

Proceedings of ICCM 2021

19th International Conference on Cognitive Modelling¹

Edited by

Terrence C. Stewart

¹ Co-located with the 54th Annual Meeting of the Society for Mathematical Psychology and held online

Preface

The International Conference on Cognitive Modelling (ICCM) is the premier conference for research on computational models and computation-based theories of human cognition. ICCM is a forum for presenting and discussing the complete spectrum of cognitive modelling approaches, including connectionism, symbolic modeling, dynamical systems, Bayesian modeling, and cognitive architectures. Research topics can range from low-level perception to high-level reasoning. In 2021, ICCM was jointly held with MathPsych – the annual meeting of the Society for Mathematical Psychology. Due to the ongoing COVID-19 pandemic, the conference was held online from July 3rd to July 9th, using a combination of prerecorded videos, live discussions, and custom software developed by the Society for Mathematical Psychology.

Acknowledgements

We would like to thank the Society for Mathematical Psychology (SMP) for their tremendous work in converting this conference to an online venue. In particular, the MathPsych Conference Chair (Joachim Vandekerckhove) guided the development of a custom website, and the officers of the SMP (especially Leslie Blaha) gave much needed logistical support.

Papers in this volume may be cited as:

Lastname, A., Lastname, B., & Lastname, C. (2021). Title of the paper. In Stewart, T. C. (Ed.). *Proceedings of the 19th International Conference on Cognitive Modelling* (pp. 6-12). University Park, PA: Applied Cognitive Science Lab, Penn State.

ISBN-13: 978-0-9985082-5-2, published by the Applied Cognitive Science Lab, Penn State.

(C) Copyright 2021 retained by the authors

Conference Committees

General and Program Chair

Terrence C. Stewart

National Research Council of Canada

Program Committee

Leslie Blaha

Air Force Research Laboratory

Edward Cranford

Carnegie Mellon University

Glenn Gunzelmann

Air Force Research Laboratory

Theodros Haile

University of Washington

Alex Kelly

Pennsylvania State University

Wouter Kruijne

University of Groningen

Othalia Larue

Wright State University

Christian Lebiere

Carnegie Mellon University

Peter Lindes

University of Michigan

Konstantinos Mitsopoulos

Carnegie Mellon University

David Peebles

University of Huddersfield

Roussel Rahman

Rensselaer Polytechnic Institute

Frank Ritter

Pennsylvania State University

Nele Russwinkel

TU Berlin

Dario Salvucci

Drexel University

Catherine Sibert

University of Washington

Sterling Somers

Carnegie Mellon University

Andrea Stocco

University of Washington

James Chapman Treyens

University of Washington

Robert L. West

Carleton University

Yuxue Yang

University of Washington

Table of Contents

Applications of information theory to perceptual independence and separability.....1 <i>Mikaela Akrenius</i>	1
How much context is helpful for noun and verb acquisition?.....8 <i>Raquel G. Alhama, Caroline Rowland, Evan Kidd</i>	8
Sequential effects in non-sequential tasks.....10 <i>Nidhi Banavar, Michael Lee, Aaron Bornstein</i>	10
Model-based explanation of feedback effects in syllogistic reasoning.....16 <i>Daniel Brand, Nicolas Oliver Riesterer, Marco Ragni</i>	16
Timing and structure of reward information influences bias in perceptual decisions as revealed by a hierarchical drift diffusion model.....23 <i>Manisha Chawla, Krishna Miyapuram</i>	23
Learning basic Python concepts via self-explanation: A preliminary python ACT-R model.....29 <i>Veronica Chiarelli</i>	29
Parameter correlations in the predictive performance equation: Implications and solutions.....36 <i>Michael Collins, Florian Sense, Michael Krusmark, Joshua Fiechter, Tiffany Jastrzembski</i>	36
Modeling phishing susceptibility as decisions from experience.....43 <i>Edward Cranford, Kuldeep Singh, Palvi Aggarwal, Cleotilde Gonzalez, Christian Lebiere</i>	43
Towards benchmarking cognitive models: A Python library for modular environment specification and partial model generation in ACT-R.....50 <i>Emmanuelle Dietz, Oliver Klaproth</i>	50
Computationally rational reinforcement learning: modeling the influence of policy and representation complexity.....57 <i>Zeming Fang, Chris R. Sims</i>	57
An ACT-R model of order effects.....64 <i>Christopher Fisher, Lorraine Borghetti, Joe Houpt, Christopher Adam Stevens</i>	64

Using cognitive agents to design dynamic task allocation systems.....	70
<i>Christopher Fisher, Mary Frame, Christopher Adam Stevens</i>	
Derivation of metric scales from ordinal data with Guttman-Goode’s scaling.....	76
<i>Vithor Franco</i>	
Criticality perception in dynamic traffic scenarios: an ACT-R model	78
<i>Noémi Földes-Cappellotto, Moritz Held, Martin Baumann, Tanja Stoll</i>	
Sampling heuristics for active function learning.....	80
<i>Rebekah Gelpi, Nayan Saxena, George Lifchits, Daphna Buchsbaum, Christopher Lucas</i>	
Simulating human periodic tapping and implications for cognitive models.....	87
<i>Pierre Gianferrara, Shawn Betts, Daniel Bothell, John Anderson</i>	
When do you buy? Predicting an individual’s decision in optimal stopping problems.....	94
<i>Manuel Guth, Marco Ragni</i>	
Inferring a cognitive architecture from multi-task neuroimaging data: A data-driven test of the common model of cognition using granger causality.....	100
<i>Holly Hake, Catherine Sibert, Andrea Stocco</i>	
Physio-cognitive modeling: explaining the effects of caffeine on fatigue.....	107
<i>Tim Halverson, Chris Myers, Jeffery Gearhart, Matthew Linakis, Glenn Gunzelmann</i>	
On disjunctions and the weak completion semantics.....	114
<i>Islam Hamada, Steffen Hölldobler</i>	
Structure learning as a mechanism for overharvesting	121
<i>Nora Harhen, Aaron Bornstein</i>	
Measuring and modelling how people learn how to plan and how people adapt their planning strategies to the structure of the environment.....	127
<i>Ruiqi He, Falk Lieder, Yash Raj Jain</i>	
Utilizing ACT-R to investigate interactions between working memory and visuospatial attention while driving.....	134
<i>Moritz Held, Jelmer Borst, Anirudh Unni, Jochem Rieger</i>	
Validating and refining cognitive process models using probabilistic graphical models....	136
<i>Laura Hiatt, Connor Brooks, Greg Trafton</i>	
Competence assessment by stimulus matching: an application of GOMS to assess chunks in memory.....	143
<i>Hadeel Ismail, Peter Cheng</i>	

Simulating proficiency and exposure effects on cross-language structural priming in simultaneous bilinguals.....	150
<i>Yung Han Khoe, Gerrit Jan Kootstra, Rob Schoonen, Stefan Frank</i>	
Toward undifferentiated cognitive models.....	157
<i>Colin Kupitz, Aaron Eberhart, Daniel Schmidt, Christopher Adam Stevens, Cogan Shimizu, Pascal Hitzler, Dario Salvucci, Benji Maruyama, Chris Myers</i>	
Exploring the decision component of the Activation-Decision-Construction-Action Theory for gain and loss facing scenarios.....	163
<i>Tei Laine, Tomi Silander</i>	
Understanding human social communication: a computational model of gossip.....	170
<i>Jeungmin Lee, Jerald D. Kralik, Jaeseung Jeong</i>	
Modelling visual decision making using a variational autoencoder.....	177
<i>Tyler Malloy, Chris R. Sims</i>	
Predicting spatial belief reasoning: comparing cognitive and AI models.....	184
<i>Johannes Mannhardt, Leandra Bucher, Daniel Brand, Marco Ragni</i>	
Diverse experience leads to improved adaptation: An experiment with a cognitive model of learning.....	191
<i>Chase McDonald, Cleotilde (Coty) Gonzalez, Leslie Blaha, Christian Lebiere, Joshua Fiechter, Erin Bugbee, Erin McCormick</i>	
Curiosity as pattern matching: simulating the effects of intrinsic rewards on the levels of processing.....	197
<i>Kazuma Nagashima, Junya Morita, Yugo Takeuchi</i>	
A predictive processing implementation of the common model of cognition.....	204
<i>Alex Ororbia, Alex Kelly</i>	
Attitudinal polarization on social networks: A cognitive architecture perspective.....	206
<i>Mark Orr, Andrea Stocco, Christian Lebiere, Don Morrison</i>	
Prediction advantage as retrieval interference: an ACT-R model of processing possessive pronouns.....	213
<i>Umesh Patil, Sol Lago</i>	
Cognitive modelling of a mental rotation task using a generalized spatial framework.....	220
<i>Kai Preuss, Nele Russwinkel</i>	
Towards precise measures of individual performance in complex tasks.....	227
<i>Roussel Rahman, Wayne D. Gray</i>	

The algebra of cognitive states: Towards modelling the serial position curve.....	234
<i>Stefan Reimann</i>	
Estimating individual differences in working memory through ACT-R modeling and resting state connectivity.....	241
<i>Patrick Rice, Andrea Stocco</i>	
Lessons learned from modelling situated cognitive agents interacting with a dynamic environment.....	248
<i>Lukas Seiling, Maximilian Plitt, Paul Schweidler, Oliver Klückmann, Lilian Befort-Trier, Robin Konczir, Severin Reuter, Kai Preuss, Sebastian Wiese, Nele Russwinkel</i>	
People are insensitive to within-category feature correlations in categorization.....	255
<i>Florian Seitz, Jana Jarecki, Jorg Rieskamp</i>	
The structured mind at rest: Evidence for the “Common Model of Cognition” in resting state fMRI.....	257
<i>Catherine Sibert, Andrea Stocco, Holly Hake</i>	
Collective intelligence as latent imagination.....	264
<i>Amit Singh</i>	
A cognitive computational model of collective search with social information.....	266
<i>Sabina J. Sloman, Robert Goldstone, Cleotilde (Coty) Gonzalez</i>	
Modeling aperture passage affordances in ACT-R 3D.....	268
<i>Sterling Somers</i>	
How good can an individual's conclusion endorsement be predicted?.....	275
<i>Sara Todorovikj, Marco Ragni</i>	
Learning reference biases from language input: a cognitive modelling approach.....	282
<i>Abigail Toth, Niels Taatgen, Jacolien van Rij, Petra Hendriks</i>	
A drift-diffusion model to explain vehicle deceleration detection of vulnerable road users	289
<i>Daniel Trommler, Claudia Ackermann, Josef Kreams</i>	
Capturing dynamic performance in a cognitive model: Estimating ACT-R memory parameters with the linear ballistic accumulator.....	295
<i>Maarten van der Velde, Florian Sense, Jelmer Borst, Hedderik van Rijn</i>	
A hidden semi-markov model classifier for strategy detection in multiplication problem solving.....	302
<i>Leendert Van Maanen, Ernö Groeneweg, Kim Archambeau</i>	
Individualizing a biomathematical fatigue model with attention data.....	309
<i>Bella Veksler, Megan Brianne Morris, Glenn Gunzelmann</i>	

Covering strategy changes: From System 1 to System 2 in syllogistic reasoning.....316
Evelyn Wiens, Alice Ping Ping Tse, Marco Ragni

Is similarity-based interference caused by lossy compression or cue-based retrieval?
A computational evaluation.....323
Himanshu Yadav, Garrett Smith, Shravan Vasishth

Individual differences in decision making strategies can be predicted by resting-state
functional connectivity.....330
Cher Yang, Andrea Stocco, Catherine Sibert

Applications of Information Theory to Perceptual Independence and Separability

Mikaela Akrenius (makreniu@indiana.edu)

Cognitive Science Program, 1001 E. 10th Street, Bloomington, IN 47405 USA

Abstract

Despite of strong historical connections between information theory and the study of perceptual independence and separability, few modern approaches take advantage of these connections. We revive Garner and Morton’s (1969) classic Mutual Uncertainty Analysis (MUA), complement it with Partial Information Decomposition (PID, Williams & Beer, 2010), and apply both to a sample of data from contemporary studies. While existing theories can dissociate between perceptual and decisional separability and identify dependencies at the level of individual stimuli, MUA and PID can provide diagnostics for identifying other types of perceptual dependencies, decompose them into their constituents, and provide a measure for their strength.

Keywords: perceptual independence; perceptual separability; information theory; mutual uncertainty analysis; partial information decomposition; general recognition theory

Introduction

Originating in studies of selective attention (Stroop, 1935) and building on Garner’s (1974) speeded classification paradigm, the study of perceptual independence and separability has become a field of its own (see Algom & Fitousi, 2016, for a review). Over the recent decades, articles and book chapters on Garner interference have come to be dominated by roughly two kinds of modeling approaches: multidimensional, signal-detection-based theories, such as general recognition theory (GRT, Ashby & Townsend, 1986), and similarity- or distance-based approaches, such as the similarity choice model (Luce, 1963; Shepard, 1957) and its further extensions (e.g. Nosofsky, 1985).

Regardless of the modeling approach used, assessments of perceptual independence are typically made based on confusion matrix data from identification experiments, and build on the assumption that the distribution of response errors is diagnostic of types of violations of perceptual independence.

As Algom and Fitousi (2016) note, despite of the strong connections that Garner’s (1962) early work on perceptual independence has to information theory, and of the usefulness of information theory in quantifying types of dependencies, it has seldom been used in the field.

To help cover this gap and to investigate whether and how information theory could be used, we will (1) re-introduce Garner & Morton’s (1969) classic mutual uncertainty analysis (MUA), along with information-theoretic preliminaries, (2) extend it with Partial Information Decomposition (PID, Williams & Beer, 2010), (3) apply both to identification experiment data from contemporary studies

and compare the results to existing, more commonly applied diagnostics (GRT), and (4) provide tentative psychological interpretations for the terms associated with PID.

Throughout the paper, we will highlight some of the formal connections between MUA, PID, and GRT. Due to limited space, this analysis will be illustrative rather than axiomatic.

Terminology

In this paper, ‘perceptual independence’ will be used to refer to the existence of statistical independence between the perceptual effects of (orthogonal) stimulus components. This is in line with Garner and Morton’s (1969) use of the term and the definition of perceptual independence used in GRT. ‘Perceptual separability’, on the other hand, will be used to refer to perceptual separability as defined by GRT.

Preliminaries: Entropy, Conditional Entropy, and Mutual Information¹

As McGill (1954) and Garner (1962) note, mutual information is an efficient tool for assessing statistical independence between two or more random variables. Unlike uncorrelation, the lack of mutual information implies statistical independence, and mutual information can capture complex (e.g. nonlinear) dependencies between variables.

Let p_i , $i \in [1, \dots, n]$, and p_j , $j \in [1, \dots, m]$, denote the probability associated with each of n , m outcomes of a discrete random variable x , y , respectively. The Shannon (1948) entropy of x is

$$U(x) = -\sum_{i=1}^n p_i \log_2(p_i), \quad (1)$$

the joint entropy of x and y is

$$U(x, y) = -\sum_{i=1}^n \sum_{j=1}^m p_{ij} \log_2(p_{ij}), \quad (2)$$

and the conditional entropy of x given y is

$$U_y(x) = U(x, y) - U(y). \quad (3)$$

The mutual information² between x and y is

$$U(x: y) = U(x) + U(y) - U(x, y), \quad (4)$$

the mutual information between x and two discrete random variables y , z , or three discrete random variables y , z , w , is

$$U(x: y, z) = U(x) + U(y, z) - U(x, y, z) \quad (5)$$

$$U(x: y, z, w) = U(x) + U(y, z, w) - U(x, y, z, w) \quad (6)$$

¹ Unless otherwise noted, the definitions used in this chapter are borrowed from Garner (1962) and McGill (1954).

² Also referred to as partial contingent uncertainty (Garner, 1962; Garner & Morton, 1969) or transmitted information (McGill, 1954).

and the conditional mutual information between x and y given z , or given z and w , is

$$U_z(x:y) = U(x:y,z) - U(x:z) \quad (7)$$

$$U_{zw}(x:y) = U(x:y,z,w) - U(x:z,w). \quad (8)$$

Mutual information is a symmetric measure of association: it is 0 if and only if x and y are statistically independent, and it can be expressed as the Kullback-Leibler (1961) divergence of the joint distribution (x, y) from the product of their marginal distributions

$$U(x:y) = \sum_{i=1}^n \sum_{j=1}^m p_{ij} \log_2 \left(\frac{p_{ij}}{p_i p_j} \right). \quad (9)$$

The mutual information between a target variable x and two source variables y, z can also be defined as

$$U(x:y,z) = U(x:y) + U(x:z) + U(xyz) \quad (10)$$

where $U(xyz)$ denotes interaction information. Interaction information is a symmetric measure

$$\begin{aligned} U(xyz) &= U_x(y:z) - U(y:z) \\ &= U_y(x:z) - U(x:z) \\ &= U_z(y:z) - U(y:z) \end{aligned} \quad (11)$$

and can be interpreted as a measure of effect size.

Mutual Uncertainty Analysis (MUA)

Garner and Morton (1969) decompose the mutual information between two stimulus components A, B (e.g. shape and color) and two response variables a, b into

$$U(a, b: A, B) = U(a:b: A: B) - U(a:b) \quad (12)$$

$$\begin{aligned} U(a:b: A: B) &= U(A: B) + U(a:A) + U(b:B) \\ &\quad + U_A(a:B) + U_B(b:A) + U_{AB}(a:b) \end{aligned} \quad (13)$$

where $U(A:B) = 0$ for orthogonally varied components, $U(a:A)$ and $U(b:B)$ measure the accuracy of responses on each component ($U(a:A) = U(A)$ and $U(b:B) = U(B)$ for maximum accuracy), and perceptual independence is violated if $U_A(a:B) \neq 0$, $U_B(b:A) \neq 0$, or $U_{AB}(a:b) \neq 0$. According to Garner and Morton, $U_A(a:B) \neq 0$ and $U_B(b:A) \neq 0$ reflect a crossing over from one perceptual channel to the other, whereas $U_{AB}(a:b) \neq 0$ measures error correlation, which can be due to perceptual or response processes. Error correlation can reflect state correlation, i.e. variation in responses across trials caused by changes in the state of the observer relative to the processing channels. $U(a:b)$, on the other hand, reflects response correlation. These terms are illustrated in Figure 1.

³ PID could also be applied to decompose the sole influence of A and B on a or b ($U(a:A, B)$ and $U(b:A, B)$), but this would yield

Partial Information Decomposition (PID)

Partial Information Decomposition (Williams & Beer, 2010) decomposes the interaction information between one target variable and two or more source variables into redundant and synergistic components, which, intuitively speaking, reflect the information shared by the sources for predicting the target (analogous to an AND gate), and unique combinations of the sources for predicting the target (analogous to a XOR gate). Formally, the information shared between target x and sources y, z can be broken into

$$\begin{aligned} U(x:y,z) &= U(x:y) + U(x:z) + U(xyz) \\ &= U(x:y) + U(x:z) + U(x:\{yz\}) - U(x:\{y\}\{z\}) \\ &= U(x:\{y\}) + U(x:\{z\}) + U(x:\{yz\}) + U(x:\{y\}\{z\}) \end{aligned} \quad (14)$$

where $U(x:\{y\})$ and $U(x:\{z\})$ denote unique information contributed by each of the sources, $U(x:\{y\}\{z\})$ denotes redundant information, and $U(x:\{yz\})$ denotes synergistic information. $U(x:\{y\}\{z\}) = U_{\min}(x:\{y,z\})$, the minimum amount of information shared by y and z for predicting x , and $U(x:y) = U(x:\{y\}) + U(x:\{y\}\{z\})$. This partitioning is illustrated in Figure 2 and can be further extended to any number of source variables. Figure 3 illustrates the case for three source variables and one target variable.

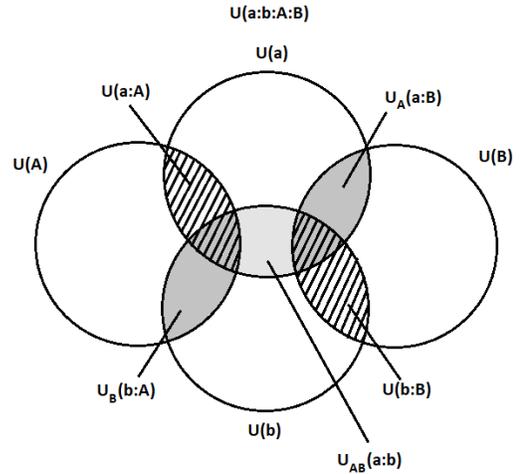


Figure 1: Illustration of the terms used by Garner and Morton (1969) in mutual uncertainty analysis.

PID for Identification Experiment Data

Identification experiment data typically involves as many response dimensions as stimulus dimensions, whereas PID has been developed to predict only one target. Due to this, PID needs to be applied separately to each response dimension. Because a majority of identification experiments consist of only two dimensions, this is relatively straightforward, and provides insight on asymmetric dependencies between the response dimensions.

Applying PID to predict response variable a yields³

less information and neglect interactions between the response dimensions.

$$\begin{aligned}
 & U(a: b, A, B) \\
 &= U(a: \{b\}) + U(a: \{A\}) + U(a: \{B\}) + \\
 & U(a: \{b\}\{A\}) + U(a: \{b\}\{B\}) + U(a: \{A\}\{B\}) + \\
 & U(a: \{b\}\{A\}\{B\}) + U(a: \{bA\}) + U(a: \{bB\}) + \\
 & U(a: \{AB\}) + U(a: \{bAB\}) + U(a: \{b\}\{AB\}) + \\
 & U(a: \{bA\}\{B\}) + U(a: \{bB\}\{A\}) + U(a: \{bA\}\{bB\}) + \\
 & U(a: \{bA\}\{AB\}) + U(a: \{bB\}\{AB\}) + U(a: \{bA\}\{bB\}\{AB\})
 \end{aligned} \quad (15)$$

which consists of all possible combinations of unique, redundant, and synergistic information contributed by each source alone or together. Applying PID to predict response variable b yields an analogous partitioning with $U(b: a, A, B)$.

Connection between MUA and PID

Using PID, the terms of MUA can be decomposed into their constituents. $U(a: A)$ (or, analogously, $U(b: B)$) can be decomposed into

$$\begin{aligned}
 U(a: A) &= U(a: \{A\}) + U(a: \{b\}\{A\}) + U(a: \{A\}\{B\}) + \\
 & U(a: \{b\}\{A\}\{B\}) + U(a: \{A\}\{bB\})
 \end{aligned} \quad (16)$$

where $U(a: \{A\})$ is indicative of unique information from A , and the remaining terms reflect the redundant information shared by A and different combinations of b and B .

Psychologically, the unique information contributed by A can be interpreted as the direct and unique perceptual influence of A on a , i.e. the part of A that is accurately reflected in a responses, not influenced by B , and not shared with b . $U(a: \{b\}\{A\})$ and $U(a: \{b\}\{A\}\{B\})$ reflect correlation between a and b that is informed by A or A and B . $U(a: \{A\}\{B\})$ measures the redundant information in A and B that is reflected in response a , which should be 0 for orthogonal stimulus dimensions. Finally, $U(a: \{A\}\{bB\})$ reflects trials in which response a correlates with A and b is informed by an interaction of A and B .

Decomposing $U_A(a: B)$ (or, analogously, $U_B(b: A)$) yields

$$\begin{aligned}
 U_A(a: B) &= U(a: \{B\}) + U(a: \{AB\}) \\
 &+ U(a: \{b\}\{B\}) + U(a: \{b\}\{AB\}) + U(a: \{bA\}\{B\}) \\
 &+ U(a: \{bB\}\{AB\}) + U(a: \{bA\}\{AB\}) \\
 &+ U(a: \{bA\}\{bB\}\{AB\})
 \end{aligned} \quad (17)$$

where $U(a: \{B\})$ is indicative of pure crossing over across perceptual channels, $U(a: \{AB\})$ reflects the synergistic influence of A and B on a , $U(a: \{b\}\{B\})$ reflects response correlation informed by B , $U(a: \{b\}\{AB\})$ reflects response correlation informed by synergistic combinations of A and B , and the remaining terms reflect types of state correlation.

The decomposition of error correlation, $U_{AB}(a: b)$, yields

$$\begin{aligned}
 U_{AB}(a: b) &= U(a: \{b\}) + U(a: \{bA\}) + U(a: \{bB\}) \\
 &+ U(a: \{bAB\}) + U(a: \{bA\}\{bB\})
 \end{aligned} \quad (18)$$

where $U(a: \{b\})$ reflects unique information shared by a and b (due to pure response correlation, e.g. bias), and the remaining terms reflect different types of state correlation: $U(a: \{bA\})$ reflects cases in which the perception of A is

enhanced (or impaired) by a certain state relative to B , $U(a: \{bB\})$ reflects cases in which B leaks into the perception of a when the observer is in a certain state relative to B , $U(a: \{bA\}\{bB\})$ reflects the redundant information shared by these cases, and $U(a: \{bAB\})$ reflects cases in which synergistic information from A and B interacts with the state of the observer, producing error correlation. Hence, under PID, $U(a: \{bAB\})$ is the term that corresponds most closely to Garner and Morton's interpretation of $U_{AB}(a: b)$.

Finally, response correlation, $U(a: b)$, can be decomposed into

$$\begin{aligned}
 U(a: b) &= U(a: \{b\}) + U(a: \{b\}\{A\}) + U(a: \{b\}\{B\}) \\
 &+ U(a: \{b\}\{AB\}) + U(a: \{b\}\{A\}\{B\})
 \end{aligned} \quad (19)$$

where $U(a: \{b\})$ is shared with $U_{AB}(a: b)$, $U(a: \{b\}\{A\})$ and $U(a: \{b\}\{A\}\{B\})$ are shared with $U(a: A)$, and $U(a: \{b\}\{B\})$ and $U(a: \{b\}\{AB\})$ are shared with $U_A(a: B)$.

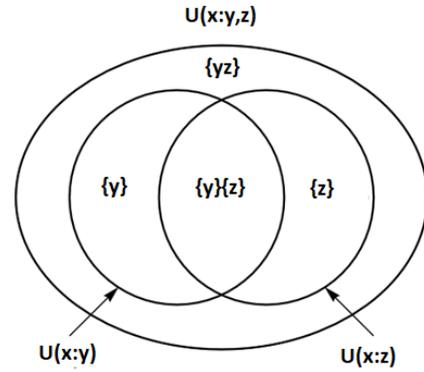


Figure 2: Partial Information Decomposition for one target variable x and two source variables y, z . Based on Figure 1 in Williams & Beer (2010).

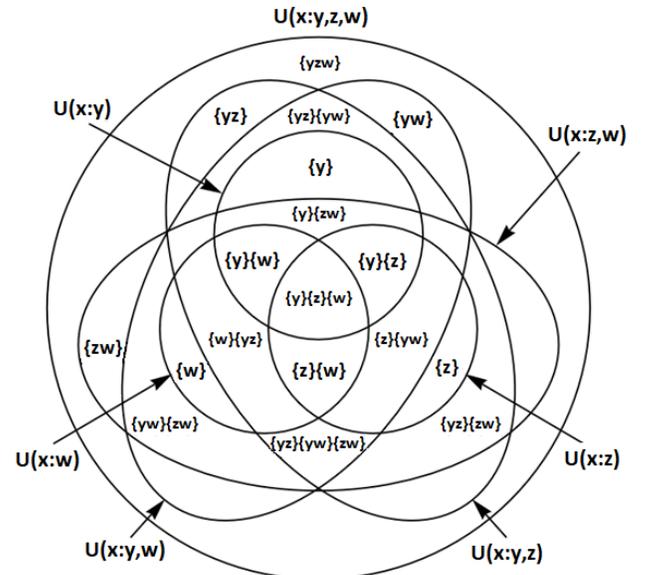


Figure 3: Partial Information Decomposition for one target variable x and three source variables y, z, w . Based on Figure S2 in Williams & Beer (2010).

Example Application: GRT

General Recognition Theory (GRT, Ashby & Townsend, 1986) is a multidimensional extension of signal detection theory, which presumes that perceptions of the stimulus components are influenced by normally distributed noise, and that response probabilities are determined by the location of choice boundaries in the perceptual space. GRT dissociates between perceptual separability, decisional separability, and perceptual independence, summarized in Figure 4, and uses a combination of probabilistic diagnostic tests to assess whether they have been violated.

According to GRT, perceptual separability holds if the perceptual distribution of a stimulus component (e.g. A_1) is uninfluenced by variation in the value of the other component (B_1 or B_2), decisional separability holds if the probability of responding e.g. a_1 given the perceptual distributions of A_1 and A_2 is uninfluenced by the value of B (i.e. the decision boundary is parallel to the B -axis), perceptual independence holds at the stimulus level if the perceptual distribution of e.g. A_1 in A_1B_1 is uncorrelated with the perceptual distribution of B_1 , and perceptual independence holds at the marginal level if it is not violated for any stimulus.

Perceptual separability

$g_{11}(x_1) = g_{12}(x_1)$ and
 $g_{21}(x_1) = g_{22}(x_1)$ for all x_1 ;
 $g_{11}(x_2) = g_{21}(x_2)$ and
 $g_{12}(x_2) = g_{22}(x_2)$ for all x_2

Decisional separability

Respond A_1 for all $x_1 \leq X_{c1}$
 Respond B_1 for all $x_2 \leq X_{c2}$
 where X_{c1} and X_{c2} are parallel to the axes

Perceptual independence

$f_{ij}(x_1, x_2) = g_{ij}(x_1)g_{ij}(x_2)$ for all x_1 and x_2

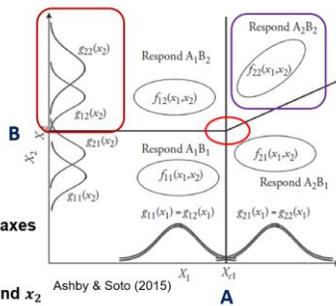


Figure 4: Postulates of GRT. Figure adapted from Ashby & Soto (2015).

Marginal Response Invariance

$$P(a_1b_1|A_1B_1) + P(a_1b_2|A_1B_1) = P(a_1b_1|A_1B_2) + P(a_1b_2|A_1B_2)$$

$$P(a_1b_1|A_1B_1) + P(a_2b_1|A_2B_1) = P(a_1b_1|A_2B_1) + P(a_2b_1|A_2B_1)$$

Sampling Independence

E.g. $P(a_1b_1|A_1B_1)$
 $= P(a_1|A_1B_1) * P(b_1|A_1B_1)$
 $= [P(a_1b_1|A_1B_1) + P(a_1b_2|A_1B_1)] * [P(a_1b_1|A_1B_1) + P(a_2b_1|A_1B_1)]$

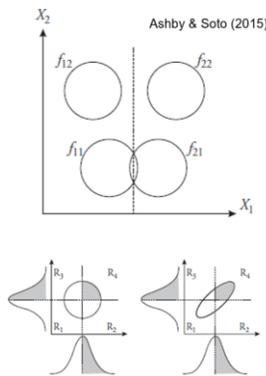


Figure 5: Some of the diagnostic tests of GRT. Figures adapted from Ashby & Soto (2015).

⁴ McGill (1954) shows that if $n_{ijm} = (n_{ij}n_{im})/n_i$ for the observed frequencies $n_i, n_j,$ and n_m of the random variables $u, v,$ and $y,$ respectively, then $U_u(v: y) = 0$. Similarly, if $n_{ijkm} =$

The diagnostic tests of GRT that are relevant for our purposes are marginal response invariance and sampling independence, summarized in Figure 5. In GRT, marginal response invariance is used as an indicator of a violation of perceptual or decisional separability, whereas sampling independence is used to assess violations of perceptual independence. GRT also employs various other statistical and signal-detection-based measures, which for the sake of space and relevance will not be reviewed here. An interested reader can consult Ashby & Soto (2015) for an illustrative review.

Formal Connections: GRT and MUA

Earlier on, it has been shown that if perceptual and decisional separability hold, marginal response invariance holds, and $U_A(a: B) = U_B(b: A) = 0$ (Theorem 6, Ashby & Townsend, 1986). Conversely, if $U_A(a: B) \neq 0$ or $U_B(b: A) \neq 0$, marginal response invariance is violated, and either perceptual or decisional separability is violated.

Analogously, it can be shown⁴ that if sampling independence holds, $U_{AB}(a: b) = 0$ (and, conversely, if $U_{AB}(a: b) \neq 0$, sampling independence is violated).

Given that Ashby & Townsend (1986, Theorem 1) show that if (and only if) decisional separability holds, a violation of sampling independence implies a violation of perceptual independence (and vice versa), this means that if decisional separability holds, $U_{AB}(a: b) \neq 0$ indicates a violation of perceptual independence at the stimulus level.

Taken together, when $U_A(a: B) \neq 0$ or $U_B(b: A) \neq 0$, perceptual or decisional separability is violated, and when $U_{AB}(a: b) \neq 0$ and decisional separability holds, perceptual independence (in the GRT sense) is violated. Hence, MUA cannot dissociate between violations of perceptual and decisional separability or prove that perceptual independence (in the GRT sense) has been violated; however, MUA can provide diagnostics for other types of violations of perceptual independence.

Empirical Results: GRT, MUA, and PID

Figure 6 shows examples of GRT, MUA, and PID applied to three kinds of identification experiment data: data from one participant in a line perception study (Townsend, Hu, & Ashby, 1981), simulated data (Ashby & Soto, 2015), and data from four participants in a facial feature perception study (Thomas, 2001b). In each data set, A and B are varied in two levels, yielding a 4x4 confusion matrix of every possible combination of A and B and their respective responses.

As predicted by the formal results, when GRT indicates a violation of perceptual or decisional separability, the respective term in MUA ($U_A(a: B)$ for A and $U_B(b: A)$ for B) deviates significantly from zero in every case except for $U_A(a: B)$ in the Townsend, Hu, and Ashby (1981) data, which is only significant at the $p < 0.10$ or $p < 0.25$ level (depending on the correction method used). Similarly, when GRT's

$(n_{ijk}n_{ikm})/n_{ik}$, where n_k refers to the observed frequencies of another random variable x , then $U_{ux}(v: y) = 0$.

perceptual independence is violated, $U_{AB}(a:b)$ deviates significantly from zero in all data sets, except for observer 4 in Thomas (2001b) where a very high accuracy for A pulls almost all interactional terms to zero. In all cases, the magnitudes of the MUA terms reflect the severity and/or number (for perceptual independence) of GRT violations.

As for the results of PID, it appears that, throughout all data sets, certain components in the decompositions of MUA terms are present very often, whereas others are seldom or never present. For example, $U(a:A)$ (and conversely $U(b:B)$) is decomposed into a nonzero $U(a:\{A\})$ (or $U(b:\{B\})$) and $U(a:\{b\}\{A\})$ ($U(b:\{a\}\{B\})$) in nearly every data set, indicating a unique contribution from A to a (B to b) and a response correlation between a and b informed by A (B). The only data sets lacking $U(a:\{A\})$ (or $U(b:\{B\})$) are Ashby and Soto (2015), where the failure of perceptual separability in A drives $U(b:\{B\})$ into $U(b:\{a\}\{B\})$ and $U(b:\{B\}\{aA\})$, i.e. all information from B is also shared with a , and observer 1 in Thomas (2001b), where $U(a:\{A\})$ is zero due to a very low accuracy in A .

Some of the data sets include a nonzero $U(a:\{b\}\{A\}\{B\})$ (or $U(a:\{b\}\{A\}\{B\})$), reflecting a crossing over in perceptual channels together with correlated responses. When both of these terms occur, the results of GRT are symmetric, whereas when only one of them occurs also GRT reflects an asymmetry in processing. For instance, in the Ashby and Soto (2015) data, only $U(a:\{b\}\{A\}\{B\})$ is nonzero and perceptual separability is violated for A (reflecting a difference in processing across levels of B), whereas for observer 1 in Thomas (2001b) only $U(a:\{b\}\{A\}\{B\})$ is again nonzero and perceptual independence fails at only one level of B . In addition, either $U(a:\{A\}\{bB\})$ or $U(b:\{B\}\{aA\})$ is nonzero in four of the six data sets, which would appear to reflect a violation of perceptual separability in the Ashby and Soto (2015) data set but is harder to explain in the Thomas (2001b, observers 2, 3, and 4) data sets. Finally, as expected with orthogonal stimulus components, $U(a:\{A\}\{B\})$ (or $U(b:\{A\}\{B\})$) is always zero.

As mentioned earlier, $U_A(a:B)$ and $U_B(b:A)$ deviate significantly from zero only in data sets in which decisional or perceptual separability is violated. In the first case (Townsend, Hu, & Ashby, 1981), $U_A(a:B)$ and $U_B(b:A)$ are decomposed into $U(a:\{AB\})$ and $U(a:\{bA\}\{AB\})$, and $U(b:\{AB\})$ and $U(b:\{aB\}\{AB\})$, whereas in the second case (Ashby & Soto, 2015), $U_A(a:B)$ consists of $U(a:\{bA\}\{AB\})$ alone. This suggests that perceptual and decisional separability could have different signatures in PID; however, the sample of data sets used here is too small to draw further conclusions on this.

In all data sets, $U_{AB}(a:b)$ is decomposed into nonzero $U(a:\{bAB\})$ and $U(b:\{aAB\})$, four of the six data sets also have nonzero $U(a:\{bA\})$ and $U(b:\{aB\})$, and one of the data sets (Ashby & Soto, 2015) has a nonzero $U(b:\{aA\}\{aB\})$. This matches with Garner and Morton's error correlation $U_{AB}(a:b)$ being primarily reflected in $U(a:\{bAB\})$ and $U(b:\{aAB\})$. The additional terms found reflect the enhanced (or impaired) perception of one dimension

depending on the state of the observer relative to the value on the other, which would (together with error correlation) appear to be reflected in GRT as a stimulus-level perceptual dependency. In one of the data sets (Thomas 2001b, observer 1), the partitioning of $U_{AB}(a:b)$ is asymmetric, with nonzero $U(a:\{bB\})$ and $U(b:\{aB\})$, possibly reflecting the fact that perceptual independence is only violated at one level of B .

Finally, as for $U(a:b)$, across all data sets only the terms shared with $U(a:A)$ or $U(b:B)$ are nonzero, indicating that response correlation always reflects information that is correct in one dimension (i.e. is never based on a relation between a and b alone, or informed by purely synergistic information from A and B).

Conclusions

To summarize, the purpose of this paper was to reintroduce MUA, to complement it with PID, to compare the results gained with MUA and PID to the results of GRT, and to provide tentative interpretations for the terms of PID. It was briefly noted that certain GRT diagnostics have MUA equivalents, and that these equivalents can be further decomposed using PID, which was illustrated in a small sample of simulated and experimental data.

The results concerning MUA and GRT are mostly in line with earlier work by Fitousi (2013), who analyzed correlations between GRT parameters and MUA terms in a simulated data set, and reanalyzed three face perception data sets (Thomas, 2001a, 2001b, and Richler et al., 2008). The novel contribution of this paper, along with formal connections between sampling independence and $U_{AB}(a:b)$, is the extension of MUA with PID and its potential psychological implications.

Suggestions for Future Work

Analogously to this paper, the results of PID could be compared to other existing approaches and extended to data sets with non-orthogonal stimulus dimensions, or to stimuli that are known to be perceptually integral. The statistical foundations underlying connections between GRT and PID would also merit further elaboration, and, like GRT, PID could be used to analyze stimulus-level information.

Methodological Notes

The MUA terms reported in this paper were computed from identification experiment data using (1) – (13) implemented in a Python program, and the PID terms presented were computed using Timme et al.'s (2014) MATLAB package. Statistical significance tests for MUA terms were executed using a chi squared approximation method described in Attneave (1959) and McGill (1954), and a correction method described in Miller and Madow (1954). The results of GRT were borrowed from the respective papers.

Acknowledgments

I would like to thank two anonymous reviewers for their comments on an earlier version of this paper.

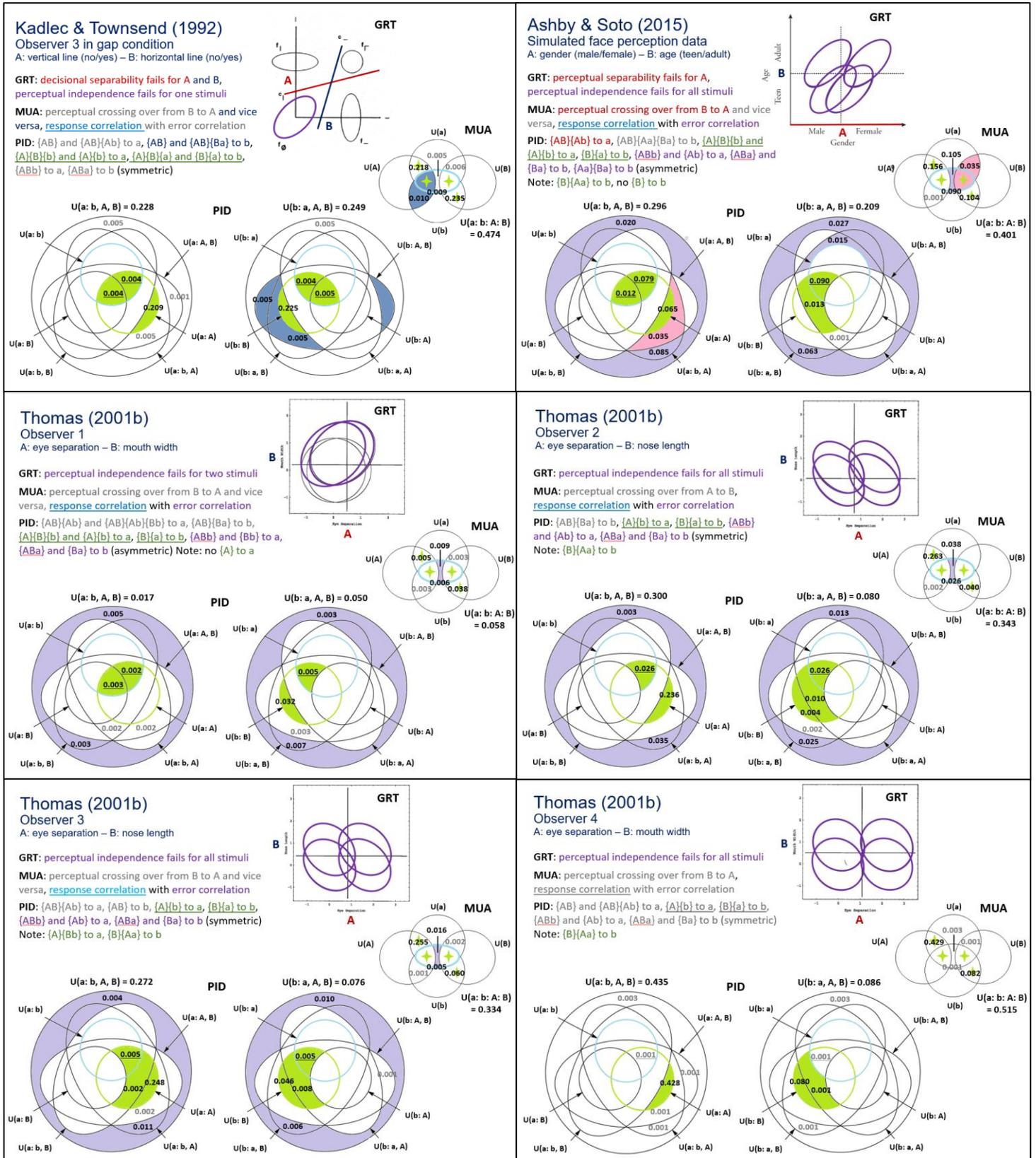


Figure 6: Examples of data sets analyzed using GRT, MUA, and PID. The figures illustrating GRT are adapted from Kadlec and Townsend (1992), Ashby & Soto (2015), and Thomas (2001b), respectively. Statistically significant MUA terms ($p < 0.05$) and their PID constituents are in black, whereas nonsignificant terms and their PID constituents are in grey. Corresponding MUA and PID terms can differ slightly due to rounding.

References

- Algom, D., & Fitousi, D. (2016). Half a century of research on Garner interference and the separability-integrality distinction. *Psychological Bulletin*, *142*, 1352–1383.
- Ashby, F. G., & Soto, F. (2015). Multidimensional signal detection theory. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.), *Oxford handbook of computational and mathematical psychology*. New York, NY: Oxford University Press.
- Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154–179.
- Attneave, F. (1959). *Applications of information theory to psychology: A summary of basic concepts, methods, and results*. New York, NY: Holt, Rinehart and Winston.
- Fitousi, D. (2013). Mutual information, perceptual independence, and holistic face perception. *Attention, Perception, & Psychophysics*, *75*, 983–1000.
- Garner, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York, NY: John Wiley & Sons.
- Garner, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum Associates.
- Garner, W. R., & McGill, W. J. (1956). The relation between information and variance analyses. *Psychometrika*, *21*, 219–228.
- Garner, W. R., & Morton, J. (1969). Perceptual independence: Definitions, models, and experimental paradigms. *Psychological Bulletin*, *72*, 233–259.
- Kadlec, H., & Townsend, J. T. (1992). Implications of marginal and conditional detection parameters for the separabilities and independence of perceptual dimensions. *Journal of Mathematical Psychology*, *36*, 325–374.
- Kullback, S., & Leibler, R.A. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, *22*, 79–86.
- Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology*. New York, NY: Wiley.
- McGill, W. J. (1954). Multivariate information transmission. *Psychometrika*, *19*, 97–116.
- Miller, G. A., & Madow, W. G. On the maximum likelihood estimate of the Shannon-Wiener measure of information. Air Force Cambridge Research Center: *Technical Report*, 54–75, August 1954.
- Richler, J. J., Gauthier, I., Wenger, M. J., & Palmeri, T. J. (2008). Holistic processing of faces: Perceptual and decisional components. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 328–342.
- Nosofsky, R. (1985). Overall similarity and the identification of separable-dimension stimuli: A choice model analysis. *Perception & Psychophysics*, *38*, 415–432.
- Shepard, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, *22*, 325–345.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.
- Timme, N., Alford, W., Flecker, B., & Beggs, J. M. (2014). Synergy, redundancy, and multivariate information measures: An experimentalist's perspective. *Journal of Computational Neuroscience*, *36*, 119–140.
- Townsend, J. T., Hu, G. G., & Ashby, F. G. (1981). Perceptual sampling of orthogonal straight line features. *Psychological Research*, *43*, 259–275.
- Thomas, R. D. (2001a). Characterizing perceptual interactions in face identification using multidimensional signal detection theory. In M. J. Wenger & J. T. Townsend (Eds.), *Computational, geometric, and process perspectives on facial cognition: Contexts and challenges*. Mahwah: Erlbaum.
- Thomas, R. D. (2001b). Perceptual interactions of facial dimensions in speeded classification and identification. *Perception & Psychophysics*, *63*, 625–650.
- Williams, P. L., & Beer, R. D. (2010). Nonnegative decomposition of multivariate information. arXiv preprint, April 14th 2010.

How Much Context is Helpful for Noun and Verb Acquisition?

Raquel G. Alhama¹

Caroline Rowland^{2,3}

Evan Kidd^{2,3,4,5}

¹Department of Cognitive Science & Artificial Intelligence, Tilburg University

²Language Development Department, Max Planck Institute for Psycholinguistics

³Donders Institute for Brain, Cognition & Behaviour, Radboud University

⁴The Australian National University

⁵ARC Centre of Excellence for the Dynamics of Language

rgalhama@tilburguniversity.edu,
{caroline.rowland, evan.kidd}@mpi.nl

Introduction

While it is widely accepted that children use distributional information to acquire multiple components of language, the underpinnings of these achievements are unclear. The goal of the current work is to investigate the role of linguistic context in the acquisition of nouns and verbs. In particular, we use a Distributional Semantic Model (DSM) to predict the age of acquisition of nouns and verbs, and we analyse the hyperparameters of the model to find out how much context is helpful for the acquisition of these words.

DSMs have been extensively evaluated against human adult ratings on semantic associations, but less so against children’s emerging semantic representations. For reasons of space, we limit our review of prior work to the most recent study that is closest to our goals. In that study, Alhama et al. (2020) propose two methods to evaluate DSMs for children’s acquisition of nouns. Their results suggest that the *Skipgram* version of *word2vec* (Mikolov et al., 2013) is most successful in predicting the Age of Acquisition (AofA) of nouns. In our work, we look more in-depth into the hyperparameters of Skipgram that best predict AofA, to find out more about the influence of context in acquisition. In addition, we extend the study to verbs.

Data

We trained the model on transcriptions of child-directed speech from CHILDES (MacWhinney, 2000), for all the English variants, for ages ranging from 0 to 60 months. To evaluate the models on AofA, we used data collected with the MacArthur-Bates Communicative Development Inventory forms (CDI). These forms contain checklists of common words that parents complete, according to whether their child *understands* or *produces* each of those words. The forms are collected at different ages, and thus can be used to estimate the AofA of words. We used the English CDIs from the Wordbank database (Frank et al., 2017) and estimate AofA as the age at which at least 50% of the children in the sample produced a given word.

How much context?

We trained Skipgram on the data described above, in order to derive vector representations for the words. We experimented with several hyperparameters of the model. We put our focus on the following:

- **Window size (win):** defined as the number of context words on each side of a target word (e.g. a window of size 1 includes a context word on each side of the target word). We explore values 1, 2, 3, 5 and 10.
- **Dynamic window size (dyn):** when this hyperparameter is enabled, the window size is dynamic, such that for each occurrence of a target word, the window size is sampled between 1 and **win**. This parameter has no practical effect when **win**=1.
- **Frequency threshold (thr):** words with frequency of occurrence below this threshold were removed, and are assumed to not be part of the vocabulary. Note that this is done after determining which words are in the context of a word, so words under the threshold are not replaced with further words in the context.

We fixed the values of the rest of hyperparameters to common default values (vector size: 100, initial learning rate: 0.025, negative sampling: off, context distribution smoothing: off, ‘dirty’ subsampling: off). Our code is available at: <https://github.com/rgalhama/public-ICCM2021>.

We then computed semantic relations between words as the cosine similarity between the corresponding vectors. As done in Alhama et al. (2020), we established a threshold *theta*, such that only words with cosine similarity larger than the threshold are considered to be neighbours. We then compute the *neighbourhood density* (ND) as the number of neighbours of each word. For reasons of space, we report results for $\theta = 0.7$, which led to highest correlations.

Figure 1 shows the results. We first focus on nouns (left graph). A very clear trend is evident for window size: given the same value of **dyn** and **thr**, a smaller window size predicts a larger correlation. Not surprisingly, the use of dynamic windows increases the fit (relative to the same fixed window size), as it decreases the amount of context available to a number of words; nevertheless, the minimum window size of 1 still performed better. We found that a small frequency threshold (**thr**=10) improves performance, indicating that even words with relatively small frequency have a role in shaping the semantic connections. In addition, the positive correlations indicate that words acquired earlier by children (i.e. smaller AofA) are those that have more semantic neighbours. This has interesting implications for language acqui-

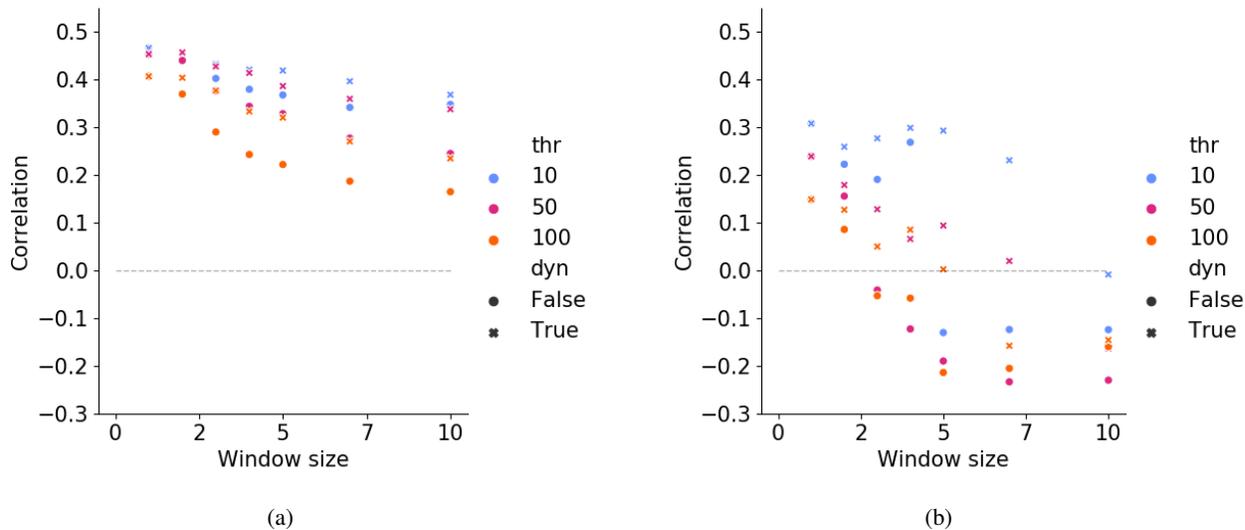


Figure 1: Correlation between AofA and ND in Skipgram, for nouns (left) and verbs (right). The hyperparameters window size (**win**), **thr** and **dyn** are defined above (in the text).

sition, as we discuss later. Overall, the results suggest that Skipgram holds promise for modelling word learning, with the best model ($\text{win}=1, \text{thr}=10$) having a correlation indicative of a medium effect size of 0.47. The results from these simulations suggest that restricting the influence of context to a very small window size consistently leads to a better fit, and that words with low frequency shape the semantic space in ways relevant to acquisition.

In order to see whether the good fit of Skipgram model extends to other syntactic categories, we evaluated its performance against AofA of verbs. As can be seen, the model shows a similar trend as for nouns, but also notable differences. For lower window sizes, results are fairly similar to nouns, albeit with smaller effect size. However, for the models with $\text{thr}=10$ (which overall performs better for verbs, as it did with nouns) there is not such a strong tendency for performance to decrease with window size, especially up to a window of size 5. As in the case of nouns, the correlations with greater effect size are positive (though this trend disappears as window size increases, specially for models with $\text{thr}>10$), indicating that having fewer semantic neighbours is beneficial for learning.

Discussion

In the case of nouns, the window size that best fits the AofA data is very small ($\text{win}=1$), suggesting that children attend to very local context, at least at an early age. Such a result makes intuitive sense in the context of children’s small verbal memory spans, which only improve as they acquire more language. The positive correlation between ND and AofA, which very consistent in the case of nouns, indicates that nouns with fewer semantic neighbours are learnt earlier. This suggests that semantic neighbours may be acting as competitors during the process of noun learning, and nouns with more

competitors are therefore less favoured.

Interestingly, we saw that the pattern of results of Skipgram is to some extent replicated for verbs, although with relevant differences. A dynamic window with a maximum size of 5 resulted in almost as good fit to the data as a window of 1 (provided $\text{thr}=10$). One potential interpretation is that larger windows allow the model to reach distant content that may include a verb’s arguments, which is likely a helpful source of information about verb meaning (Gleitman, 1990). Thus, one reason why verbs are acquired later than nouns may be the need to learn to use more distant contexts, although more simulations are needed to support this explanation (in particular, simulations with adaptive window size that depend on age and/or syntactic category). We leave this to future work.

References

Alhama, R. G., Rowland, C., & Kidd, E. (2020). Evaluating word embeddings for language acquisition. In *Proceedings of the workshop on cognitive modeling and computational linguistics* (pp. 38–42). Association for Computational Linguistics.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of child language*, 44(3), 677–694.

Gleitman, L. (1990). The structural sources of verb meanings. *Language acquisition*, 1(1), 3–55.

MacWhinney, B. (2000). *The CHILDES project: Transcription format and programs*. Lawrence Erlbaum Associates.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems* (pp. 3111–3119).

Sequential Effects in Non-Sequential Tasks

Nidhi V. Banavar (nbanavar@uci.edu)

Michael D. Lee (mdlee@uci.edu)

Aaron M. Bornstein (aaron.bornstein@uci.edu)

Department of Cognitive Sciences, University of California, Irvine
Irvine, CA 92617 USA

Abstract

In behavioral economic experiments with randomized or unstructured choice sets, trial-level sequential dependencies at the level of choice behavior or reaction time are assumed to be present only in motor or perceptual operations, but not in the cognitive valuation processes themselves. Thus, these are not explicitly accounted for. We present a flexible Bayesian hierarchical model that allows us to test for the presence or absence of linear sequential effects on cognitive, perceptual, and motor parameters of interest and subsequent choice. We apply this model to two data sets: one intertemporal choice and one risky decision making. We demonstrate sequential effects on risk tolerance inference and on the deliberative evaluation of discounted value, with many individual differences. Our results suggest that data collected in sequence cannot be treated as if it were collected independently.

Keywords: sequential decision making; intertemporal choice; risky decision making; hierarchical Bayesian modeling

Introduction

Behaviors like *temporal discounting*, how people discount value over time, and *risk tolerance*, how individuals trade off known uncertainty, are highly context dependent. While they are precisely defined within behavioral economics, or other niche fields, extant literature is rife with evidence suggesting that human behavior does not necessarily correspond to these delineations. Psychologists have demonstrated how inferences of these parameters are sensitive to many other factors including development, arousal, and cognitive capacity (Lempert & Phelps, 2015; Frey et al, 2017). Thus the question of whether these parameters can even be treated as a single (multidimensional) latent variable is a question of active philosophical and empirical research.

Importantly, researchers have also demonstrated that these inferred parameters are sensitive to the method with which they are elicited (Lempert & Phelps, 2015; Frey et al, 2017; Pedroni et al, 2017). In particular, much volatility has been observed both within and across experiments (Frey et al, 2017).

While no measure is “pure,” we must examine whether their measurement may be influenced by aspects of task structure that are unrelated, in principle, to the construct under examination. In particular, we focus on the fact that most experiments regarding intertemporal choice (ITC, infer discount factor) and risky decision making (Risk, infer risk tolerance) involve an individual making a sequence of choices, usually in one sitting. Commonly in ITC and Risk tasks, there is no ostensible structure and individuals are explicitly instructed to treat each decision independently and as if it were the only one that counts. Thus typical methods involve treating the

data as if it were independently acquired and not actually a sequence of choices.

On the other hand, empirical data and analyses from the working memory and psychophysics literature have for decades demonstrated the effect of *serial dependence*: when stimulus and choice information from previous trials influence current choice behavior and generate systematic patterns in reaction time in the absence of explicit structure in the environment or stimulus sequence (e.g., Lockhead & King, 1983; Bertelson 1961). Further, theories of intertemporal choice that involve prospecting—simulating the future (Peters & Buchel, 2010; Gabaix & Laibson 2017)—imply that already computed future values could be cached and re-used, especially if an individual has to make similar choices in sequence (Dasgupta et al, 2018). Studies have shown that episodic cues within an experiment can also influence risky decisions, suggesting a similar reliance on cognitive processes involved in simulation (Ludvig, Madan & Spetch, 2015), which could also lead to re-use.

In this paper, we develop a hierarchical Bayesian model that allows us to test for trial-level sequential influences of stimulus properties. We then apply this model to test for short-term (one-trial-back) influences of cognitive and motor perseveration in both choice behavior and response times.

Methods

Data

Inter-Temporal Choice (ITC) We model $n = 482$ adult subjects (in-person data collection, from Hunter et al, 2018) who made a sequence of 102 binary decisions between same-day monetary reward (SS: smaller sooner, range: \$1–\$85) and a larger reward in the future (LL: larger later, \$10–\$95). Delay (also indicated as T for time) between the SS and LL options ranged between 4 and 180 days. Stimuli were displayed numerically. SS and LL choices were counterbalanced to occur equally often on the Left or Right side of the computer screen. We model both choice behavior and reaction time for this data set. For brevity, however, we present results only for reaction time.

Risk We model $n = 56$ adult subjects (MTurk, from Guan et al, 2020) who made a sequence of 40 binary choices between gambles in the gain and loss domain separately, for a total of 80 trials. Each gamble was associated with two rewards and two probabilities summing to 1 (rewards range: Gain: \$1 – \$100, Loss: -\$99 – \$0, probability range: 1% – 99%). Stimuli were displayed as pie charts with labels indicating

reward amount and probability. We model and present only choice behavior for this data set, as RT was unavailable.

Choice sets were randomized for both experiments, i.e. there were no explicit trial- or task-level sequential dependencies. No outcomes were realized during the tasks (no feedback). In the following sub-section, we develop a model that tests for linear sequential effects of stimulus properties and previous choices on current choice and reaction time.

Cognitive Models: Choice Behavior in Risky Decision Making

For all models, we implement hierarchical Bayesian models in JAGS (Plummer, 2003). Unless otherwise stated, all parameters are hierarchical Normals defined with hyperpriors: $\mu \sim Normal(0, 1)$ and $\sigma \sim Normal(0, 1)_+$. Thus a hierarchical parameter X is distributed: $X \sim Normal(\mu_X, \sigma_X^2)$. We use hierarchical specifications to better capture individual differences (Lee, 2018).

Subjective Value We model the Subjective Value (SV) of a choice in accordance with Subjective Expected Utility Theory. For individuals $i = 1, \dots, n$ on trials $j = 1, \dots, J$ in conditions $c = 1, 2$:

$$SV_{(i,j,c)} = \begin{cases} \sum_{m=1}^2 P_{m(i,j)} \cdot v_{m(i,j)}^\alpha & c = 1 \text{ (Gain)} \\ \sum_{m=1}^2 P_{m(i,j)} \cdot -v_{m(i,j)}^\alpha & c = 2 \text{ (Loss)}. \end{cases}$$

On a given trial, v is the dollar reward offered for each gamble and p is the probability of reward. As each gamble is associated with two separate rewards, to compute the SV, we multiply each exponentiated reward (v^α) and probability, and then sum them. We do this separately for the left and right gamble. The exponent α is interpreted as an individual's risk tolerance (the curvature of the utility function) and is inferred at the individual, not trial, level. Note that v always refers to the objective dollar reward and v^α always refers to a subjective dollar reward. We use hyperprior $\mu_\alpha \sim Gamma(2, 1)$ for risk tolerance, with mode = 1 (risk neutrality). We further do not assume the curvature of the utility function is the same in both domains (i.e. infer $\alpha_{(i,c)}$).

Baseline. We implement a logistic choice rule to relate objective trial properties (e.g. dollar reward), subjective trial properties (e.g. SV) and choice behavior. Our baseline model includes no sequential effects. Specifically, the probability of choosing choice A vs choice B, $\theta_{A,B}$, is:

$$\theta_{A,B(i,j,c)} = \frac{1}{1 + \exp(\gamma_{(i,c)} + \beta_{(i,c)} \cdot SVD_{(i,j,c)} + \epsilon_{(i,j,c)})}$$

Here, $SVD_{(i,j,c)}$ represents the difference in $SV_{(i,j,c)}$ between the two options presented on any given trial. Then, $\gamma_{(i,c)}$ represents the shift, or bias, in a decision (towards Left or Right gamble). $\beta_{(i,c)}$ represents response variability, and we use hyperprior $\mu_\beta \sim Gamma(2, 1)$, where the mode corresponds to probability matching. Finally, $\epsilon_{(i,j,c)}$ represents effects of simple perseveration (repeat Left or Right choice). All parameters allow for variability at the individual and domain

(gain or loss) level. We pair these prior specifications with a *Bernoulli* likelihood, as no two stimuli are presented together more than once.

Sequential Effects: Properties Intuitively, we might imagine that there would be more (less) of an effect on a given parameter on sequential trials that present the subject with similar (different) values for the decision problem: e.g., if on Risk trial $j - 1$, a subject decides between a 81% chance of winning \$41 or a 55% chance of winning \$39, and the next trial j asks the subject to choose between a 80% chance of winning \$45 or a 55% chance of winning \$37, there might be little need to re-deliberate, which could thus yield an effect on either choice or response time. We consider the influence of previous (one-trial-back) and current stimulus properties and choices on representation and subsequent decision on the current trial. We compare stimulus properties by taking the absolute difference between given properties on trial j and $j - 1$. In particular, we consider the cross-trial differences in the following properties:

Property	ITC	Risk
Value (v)	$v_{LL} - v_{SS}$	$v_R - v_L$
Delay	T	
Entropy (H)		$(H_R + H_L)/2$
Composite		$EV_R - EV_L$
Heuristic (1)		$\max / \min(v \text{ or } H)$
Heuristic (2)	v and T	v and H

Table 1: Stimulus properties considered as indicator variables for the presence of sequential effects.

$H = -\sum p \log(p)$ is the Shannon Entropy of a gamble, and $EV = \sum_l p_l \cdot v_l$ is the Expected Value (assuming risk neutrality) of a gamble.

Specifically, we define all these properties as indicator variables (π), using a median split to determine whether the properties being considered in a given model (x) are large or small in difference ($x' = x_j - x_{j-1}$). Then, for every individual i on trial j and condition c :

$$\pi_{High(i,j,c)} = \begin{cases} 1, & |x_{(i,j,c)} - x_{(i,j-1,c)}| > \text{median}(\text{all } x') \\ 0, & \text{otherwise,} \end{cases}$$

and vice versa for $\pi_{Low(i,j,c)}$. For example, suppose we were interested in ITC trials with large delay differences (DD). Then, if $T_{(i,j)} = 100$ and $T_{(i,j-1)} = 6$, $\pi_{HighDD(i,j)} = |100 - 6| > 86.5 = 1$.

Sequential Effects: Model We augment our baseline model by allowing the above-mentioned properties to exert linear influences on parameters previously only inferred at the individual level. Each model considers one trial property from Table 1 (under the column Risk) at a time, but tests simultaneously for its influence on the following parameters:

Parameter	
Logistic Bias	$\Upsilon_{(i,c)}$
Logistic Slope	$\beta_{(i,c)}$
Risk Tolerance	$\alpha_{(i,c)}$
Perseveration	$\epsilon_{(i,j,c)}$

Table 2: Parameters simultaneously tested for sequential effects in the Risk task.

For example, we use:

$$\alpha'_{(i,j,c)} = \alpha_{(i,c)} + \delta_{(i,c)} \cdot \pi_{(i,j,c)}$$

instead of $\alpha_{(i,c)}$ in our $SV_{(i,j,c)}$ computation, where $\delta_{(i,c)}$ is a continuous variable representing the weight of the sequential effect. By this formulation, δ is actually a $4 \times n$ weighting matrix. Thus, the new $\alpha_{(i,c)}$ is the sequential effect adjusted risk tolerance for individual i . Our primary question of interest, then, centers around the posterior values the respective δ parameters take (in particular, zero vs non-zero).

Latent Mixture Finally, we use a latent-mixture model to allow for contaminant behavior. We assume that, for each trial, every individual belongs to one of two groups, or mixtures: task compliant or non-compliant. Specifically, if for any given trial the model infers that $\theta = 0.5$ is more likely (i.e., the subject is guessing) then that trial is considered to be non-compliant and is not included in the regular analysis. We use a *Uniform*(0, 1) prior for the base-rate of each group, paired with a *Bernoulli* likelihood.

Cognitive Models: Reaction Time in Intertemporal Choice

Researchers have used response times (RT) to improve the modeling of discount factors (Peters & D’Esposito, 2020). Previous work has also related components of the Drift Diffusion Model (DDM): both drift rate and bias to discount factor (Hunter et al., 2018). We therefore might expect that sequential effects which do not present themselves in choice outcomes might still be observable in response times.

Thus we implement a modified hierarchical Bayesian approximation of the DDM as presented in Bogacz et al. (2006). The approximation uses a shifted and scaled logistic function (tanh), and we allow for trial level variability in both the bias and drift rate terms. As with choice behavior, unless otherwise specified, all parameters can be assumed to be hierarchical and Normally distributed with independent priors. We use a *Lognormal* likelihood to fit RT.

Baseline. We present three different versions of the Bogacz approximation before considering sequential effects. First, we augment the formulation specified in the original paper with an explicit bias term (1). We assume symmetric thresholds $z'_{(i)}$ and use prior $bias \sim Normal(0, 1)T(-z'_{(i)}, z'_{(i)})$, where the bias is restricted to values that fall between $(-z_{(i,j)}, z_{(i,j)})$. A positive bias indicates a preference for LL,

while a negative bias for SS. We set $\mu_A \sim Unif(-0.9, 0.9)$ for the drift rate hyperprior.

The other two models decompose the now deterministic drift rate to incorporate stimulus properties explicitly into the model. First, we implement a simple linear regression style on decomposition modeling Subjective Value Difference (2). Here, however, Subjective Value is defined using a non-linear hyperbolic discount function, and $k_{(i)}$ is the discount factor. We also fit a model that does not include an integrated value-delay signal and instead trades off value difference and delay separately as in Hunter et al. (2018) (3). All “regression” weights have *Normal*(0, 1) prior distributions. Then, for threshold $z_{(i,j)}$, drift rate $A_{(i,j)}$ and $c^2 = 1$:

$$LL : z_{(i,j)} = z'_{(i)} - bias_{(i,j)}, \quad SS : z_{(i,j)} = z'_{(i)} + bias_{(i,j)} \quad (1)$$

$$DT_{(i,j)} = \frac{z_{(i,j)}}{A_{(i,j)}} \tanh\left(\frac{A_{(i,j)}z_{(i,j)}}{c^2}\right)$$

$$A_{(i,j)} = \beta_{0(i)} + \beta_{1(i)} \left(\frac{v_{LL}^{\alpha_{(i,j)}}}{1 + k_{(i)}T_{(i,j)}} - v_{SS}^{\alpha_{(i,j)}} \right) \quad (2)$$

$$A_{(i,j)} = \beta_{0(i)} + \beta_{1(i)}(v_{LL(i,j)} - v_{SS(i,j)}) + \beta_{2(i)} \log^{-1}(T_{(i,j)}) \quad (3)$$

$$RT \sim \log Normal(\log(DT_{(i,j)}), \sigma_{RT(i)}^2)$$

Sequential Effects As in *Choice Behavior*, we augment our baseline models by allowing the properties listed in Table 1 (under the column ITC) to exert linear influences on the parameters of interest: bias and drift. For example, we use:

$$\beta'_{0(i,j)} = \beta_{0(i)} + \delta_{(i)} \cdot \pi_{(i,j)}$$

for the intercept term in the drift rate decomposition. Here, δ becomes a 3 or $4 \times n$ matrix depending on which model was fit. The sequential effect adjusted term is the newly inferred $\beta_{0(i)}$. Again, our analysis centers around the posterior estimates of δ .

Statistical Analysis

We quantify evidence in favor of either hypothesis by using the Savage-Dickey ratio to approximate the Bayes factor as we test the two hypotheses: H_0 : no sequential effect and the alternative H_A : non-zero sequential effects. The Bayes Factor (BF) quantifies the relative strength of evidence in the data: where $BF > 3$ indicates moderate or greater evidence in favor of the hypothesis being considered (Lee & Wagenmakers, 2013). Values lower than 3 indicate that there is not enough evidence in the data to make strong statements in favor of either the null or the alternative. The Savage-Dickey ratio, then, allows us to test nested models at a particular point in the parameter space: namely 0, where there is no sequential effect. In this paper, any “evidence in favor of” a particular hypothesis reported means that the estimated Bayes Factor is greater than 3. In our analyses of hierarchical parameters, we also consider the “representative subject”, which is inferred behavior for an individual that contains all the variability of

previous experiment participants. This is distinct from the group mean, and can be thought of as answering the question “what might the next person who walks in to do the experiment look like?”

Results

Choice Behavior in Risky Decision Making

For this task, we consider sequential effects on all parameters listed in Table 2. We observed reliable sequential effects on logistic slope and risk tolerance for 7% of individuals.

Critically, and consistent with these effects being cognitively specific, these individuals only had non-zero sequential effects for specific sequences of trials: when a trial with a high difference in Expected Value between the two options (Table 1: Composite) was followed by a trial with a low difference in EV — “easy” then “difficult” in sequence — 4 individuals showed moderate to strong evidence of a negative sequential effect on risk tolerance, but only in the loss domain (see Figure 1). A negative sequential effect implies that the parameter, when inferred without sequential effects, has been underestimated. The true value, then, is greater: for example, for a specific subject, $\alpha = 1.043$ updated to $\alpha = 1.205$ when adjusted for this sensitivity. Importantly, the magnitude of α was not the only changing factor: the interpretation of the individual’s risk tolerance changed from risk *neutral* to risk *averse* in the loss domain.

Similarly, subjects demonstrated a sensitivity to sequences that were low in entropy in both domains, where the sequentially adjusted logistic slope was higher: reduced response variability than originally inferred.

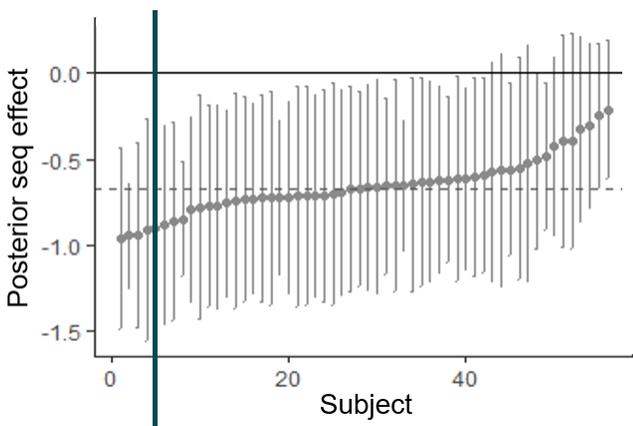


Figure 1: Sequential effects of high EV difference trials followed by low EV difference trials in the Loss domain. The dashed horizontal line is the posterior mean for the representative subject. The bold vertical line separates individuals with $BF > 3$ (left of the line) in favor of sequential effects from those with $1 < BF < 3$ (right of the line). The corresponding plot in the Gain domain not shown as there were no sequential effects in this block of the task.

We also find evidence in favor of the null for an overall

effect of motor perseveration: individuals did not systematically “stay” or “switch” their choices in both the gain and loss domain. Interestingly, this is the only result in this data set that holds at the group level and the representative subject level. All other tests for the presence of sequential effects on any parameter involved substantial individual differences: Bayes factors for representative subjects were consistently between 1 and 2, which is interpreted as anecdotal evidence in favor of the hypothesis being tested.

Finally, the model finds little evidence for guessing behavior in the data: a total 12 trials were identified as contaminant by the model (4 gain, 8 loss). This tells us that subjects were largely compliant in the task.

Given the length of the task and the size of this data set, we interpret the few subjects that do show sequential effects for specific trials as a demonstration of the model’s capability in identifying individual differences, rather than making more general claims.

Reaction Time in Intertemporal Choice

Out of $n = 482$, subjects, $n = 185$ individuals missed between 1 – 5 trials. These missed trials, and the completed trial that immediately followed a missed trial were excluded from the analysis (approximately 1% of total trials).

Baseline. We find that models that fit the deterministic drift rate decomposition and incorporate trial properties, (2) and (3), perform much better ($DIC_1 = 3967717; DIC_2 = 466889.3; DIC_3 = 365194.4$) than the model with purely stochastic drift rate (1). In particular, we find that the aggregate posterior estimates for individuals in Models 2 and 3 are similar for the primary parameters of interest: threshold, bias, and drift rate (see Table 3). Individuals in this data set, on average, appear to have a slight bias towards the Smaller Sooner option (see Table 3).

We further see that all β weights are close to zero, but with considerable individual differences. These low parameter values, however, are to be expected given how small the average inferred drift rate is.

Model 2 tests the hypothesis that as individuals accumulate information, they are considering a unified signal of value and delay which, in this case, is the difference in Subjective Value between the two options presented on the screen. Model 3, on the other hand, tests the hypothesis that individuals separately consider these properties. Our posterior estimates suggest that trading off value and delay independently may be what the subjects are doing, as Model 3 infers a very low number for the value difference parameter. On average, then, individuals are faster to make up their minds the larger the delay between the current and future options. All else held constant, this translates to lower reaction times. As such, and given superior performance in model comparison, we tested for sequential effects using the Model 3 parameterization of drift rate.

	M1	M2	M3
Parameter	Mean (95)	Mean (95)	Mean (95)
Threshold	1.98 (1.38,2.63)	1.59 (1.2,2.07)	1.56 (1.17,2.09)
Bias - S	-0.03 (-0.24,0.2)	-0.027 (-0.1,0.06)	-0.01 (-0.09,0.06)
Bias - T	-0.03 (-0.3,0.27)	-0.025 (-0.2,0.14)	-0.01 (-0.18,0.17)
Drift Rate	0.15 (-0.9,1.18)	-0.015 (-0.82,0.86)	-0.009 (-0.89,0.79)
β_0		-0.028 (-0.97,0.90)	-0.015 (-0.95,0.84)
β_1		-0.00245 (-0.45,0.4)	-1.40e-05 (-0.01,0.01)
β_2			0.008 (-0.5,0.56)
Drift Rate	0.015	-0.015	-0.009 (-0.5,0.56)
σ_{RT}	0.14 (0.04,0.3)	0.27 (0.15,0.41)	0.28 (0.14,0.4)

Table 3: Aggregate posterior estimates for DDM parameters. Bias - S is inferred bias at the subject-level, while Bias - T is the subject- and mean trial-level bias.

Sequential Effects. For the four parameters tested for sequential effects (drift rate β weights and trial-level bias), we find that 134 subjects show evidence for *non-zero* sequential effects on at *at least one parameter*. In particular, we present inferences about sequential effects driven by value, delay, or value and delay (see Table 1). This carves the stimulus space into 8 “regions” ($\pi_{(i,j)}$) of sequential effects (See Table 4).

	High Value	High Delay	Low Value	Low Delay
Low Delay	X	-	X	X
Low Value	-	X	X	
High Delay	X	X		
High Value	X			

Table 4: Specific stimulus properties that elicited sequential effects in subjects. An ‘X’ indicates a trial property or combination we explicitly modeled, and ‘-’ is undefined or a combination that has already been marked.

Of the 134, 69 subjects showed sequential effects on the bias term, 76 on the β_0 drift rate intercept term, 36 on the β_1 drift rate value term and 11 on the β_2 drift rate delay term. We note that 41 subjects have more than one non-zero sequential

effect (30 subjects with 2, 10 subjects with 3 and 1 subject with all 4 sequential terms non-zero), and again that this is across all combinations of stimulus properties. Importantly, these sequential effects were distributed roughly evenly between “main” effects driven only by differences in value or delay ($n = 83$) and “interactions” ($n = 89$), with $n = 38$ showing sequential effects for both. That is, unlike the results from modeling choice behavior in the Risk task, DDM parameters seem more susceptible to a broad range of stimulus sequences.

We also found that all 482 subjects showed evidence for *no* sequential effects on *at least one parameter* for some $\pi_{(i,j)}$. This suggests, again, that there are extensive individual differences in both the presence or absence of sequential effects, and in how and when they manifest.

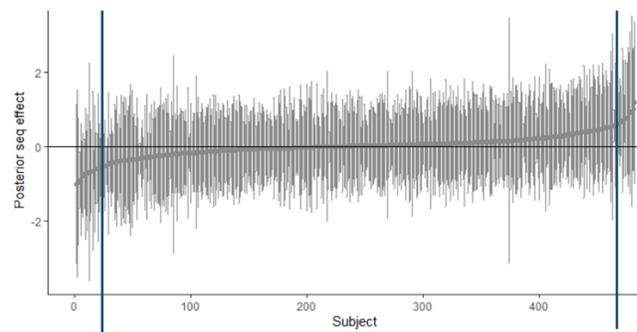


Figure 2: Sequential effects of high delay and low value difference trials on the drift rate intercept parameter. Subjects on the left and right side of the bold vertical have $BF > 3$ in favor of sequential effects. The remainder of parameters and sequential effects are not shown due to space considerations.

In sum, we find that 28% of subjects show evidence of sequential effects in DDM parameters as some function of stimulus properties value and delay. We note that the absence and presence of sequential effects are not the only conclusions we reach from the data: each subject, for some combination of stimulus properties, also had parameters where the strength of evidence was not strong enough to favor either hypothesis.

Discussion

We have introduced a flexible, generative framework to test for the presence of sequential effects on choice behavior and reaction time in explicitly non-sequential, or unstructured, environments. Our model assumes linear influences of current and previous (one-trial-back) stimulus properties on current representations, Drift Diffusion Model (DDM) parameters, and choice. Our results demonstrate evidence of stimulus-driven short-term sequential effects in both choice behavior and on reaction time related parameters in two different economic decision-making tasks. Importantly, these sequential effects were restricted to specific stimulus properties for choice behavior, but were much more widespread for parameters modeling reaction time.

The presence of such effects reinforces the sequential processing nature of the brain and adds to decades long research showing that even if stimuli in an experiment are de-correlated, they are implicitly related by time (Kiyonaga et al., 2017). This corresponds to our results, that parametric inferences even in higher order cognition can be influenced by the linear passage of time and tells us that trials completed in sequence should *not* be treated independently.

Finally, sequential effects in both choice behavior and reaction time showed overwhelming individual differences, with non-trivial changes in parameter magnitude and interpretation. For example, the interpretation of all subjects in the Risk task that presented non-zero sequential effects when “easy” trials preceded “hard” ones changed from risk neutral to risk averse in the loss domain. For the DDM parameters in intertemporal choice, we found similar changes on adjusted bias and drift rate parameters. For example, a subject whose bias term changed from positive to negative was initially interpreted as generally preferring the delayed option (and thus perhaps more patient), when, in actuality, that apparent patience was an artefact of the structure of the choice set. This is particularly important because both the magnitudes of these parameters and their resulting interpretations can be used to explain and predict real world behavior in health and clinical populations (Konova et al, 2020).

Our future directions include expanding the coding of stimulus properties to a continuous kernel: moving beyond indicator variables to continuous parameters and allowing for n-trial-back analyses. We also plan to apply this framework to larger data sets in order to establish the presence or absence of consistent stimulus driven sequential influences across individuals in economic decision making.

Acknowledgements

The authors are grateful to Catherine A. Hartley for providing the intertemporal choice data.

This work used the infrastructure for high-performance and high-throughput computing, research data storage and analysis, and scientific software tool integration built, operated, and updated by the Research Cyberinfrastructure Center (RCIC) at the University of California, Irvine (UCI). The RCIC provides cluster-based systems, application software, and scalable storage to directly support the UCI research community. <https://rcic.uci.edu>

References

Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *Quarterly Journal of Experimental Psychology*, *13*, 90-102.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*, 700--765.

Dasgupta, I., Schulz, E., Goodman, N. D., & Gershman, S. J. (2018). Remembrance of inferences past: Amortization in human hypothesis generation. *Cognition*, *178*, 67-81.

Frey, R., Mata, R., Rieskamp, J., & Hertwig, R. a. (2017). Trisk preference shares the psychometric structure of major psychological traits. *Science Advances*, *10*.

Gabaix, X., & Laibson, D. (2017). Myopia and discounting. *National Bureau of Economic Research*.

Guan M., V. J., Stokes R., & Lee, M. (2020). A cognitive modeling analysis of risk in sequential choice tasks. *Judgement and Decision Making*, *15*, 823–850.

Hunter, L. E., Bornstein, A. M., & Hartley, C. A. (2018). A common deliberative process underlies model-based planning and patient intertemporal choice.

Kiyonaga, A., Scimeca, J. M., Bliss, D. P., & Whitney, D. (2017). Serial dependence across perception, attention, and memory. *Trends in Cognitive Sciences*, *21*(7), 493-497.

Konova, A. B., Lopez-Guzman, S., Urmanche, A., Ross, S., Louie, K., Rotrosen, J., & Glimcher, P. W. (2020). Computational Markers of Risky Decision-making for Identification of Temporal Windows of Vulnerability to Opioid Use in a Real-world Clinical Setting. *JAMA Psychiatry*, *77*(4), 368-377.

Lee, M. D. (2018). Bayesian methods in cognitive modeling. In J. Wixted (Ed.), *Stevens' handbook of experimental psychology and cognitive neuroscience*. John Wiley & Sons.

Lee, M. D., & Wagenmakers, E. J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

Lempert, K. M., & Phelps, E. A. (2016). The malleability of intertemporal choice. *Trends in Cognitive Sciences*, *20*, 64–74.

Lockhead, G. R., & King, M. C. (1983). A memory model of sequential effects in scaling tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 461-473.

Ludvig, E. A., Madan, C. R., & Spetch, M. L. (2015). Priming memories of past wins induces risk seeking. *Journal of Experimental Psychology: General*, *144*, 24-29.

Pedroni, A., Frey, R., Bruhin, A., Dutilh, G., Hertwig, R., & Rieskamp, J. (2017). The risk elicitation puzzle. *Nature Human Behaviour*, *11*, 803-809.

Peters, J., & Büchel, C. (2011). The neural mechanisms of inter-temporal decision-making: Understanding variability. *Trends in Cognitive Sciences*, *15*, 227–239.

Peters, J., & D’Esposito, M. (2020). The drift diffusion model as the choice rule in inter-temporal and risky choice: A case study in medial orbitofrontal cortex lesion patients and controls. *PLOS Computational Biology*, *16*.

Plummer, M. (2003). Jags: A program for analysis of bayesian graphical models using gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing*. Vienna, Austria.

Model-Based Explanation of Feedback Effects in Syllogistic Reasoning

Daniel Brand^{1,2} (daniel.brand@cognition.uni-freiburg.de)

Nicolas Riesterer¹ (riestern@cs.uni-freiburg.de)

Marco Ragni^{1,2} (ragni@sdu.dk)

¹Cognitive Computation Lab, University of Freiburg, 79110 Freiburg, Germany

²South Denmark University, Denmark

Abstract

In the field of syllogistic reasoning research, a significant number of models aiming at describing the human inference processes were developed. There is profound work fitting the model’s parameters and analyzing each model’s ability to account for the data in order to support or disprove the underlying theories. However, the model parameters are rarely used to extract explanations and hypotheses for phenomena that go beyond the original scope of the models. In this work, we apply three state-of-the-art models, PHM, mReasoner, and TransSet, to data from reasoning experiments where participants received feedback for their conclusions. We derived hypotheses based on the models’ explanations for the feedback effect and putted these to test by conducting an experiment targeting the hypotheses. The work contributes to the field in three ways: (a) the feedback effect could be replicated and was shown to be a robust effect; (b) we demonstrate the use of the model parameters in order to derive new hypotheses; (c) we present possible explanations for the feedback effect based on existing theories.

Keywords: syllogistic reasoning; cognitive modeling; mReasoner; PHM; TransSet; feedback

Introduction

Routinely, psychological experiments are conducted to uncover robust effects and phenomena related to the latent processes of the human mind. Assumptions to shed light on the internals of the black box that constitutes the human mind are compiled into theories that are then corroborated or falsified based on comparison with experimental data.

Models, on the one hand, instantiate theories incorporating the knowledge about robust effects and phenomena that were found through observations in experiments. By providing measures to quantify the capabilities of a model to account for real world processes, this ultimately allows to test and verify the assumptions underlying the respective theories. On the other hand, models that have proven to be good accounts for their respective processes can also be transferred to different scenarios. In this way, models can be used to extract predictions even for hypothetical scenarios, which can subsequently be used to derive new hypotheses that fuel further investigations.

Consider for example the domain of human syllogistic reasoning, which will serve as the domain of interest throughout this article. Traditionally, syllogisms consist of two premises featuring one out of four quantifiers (“All”, “Some”, “No”, “Some ... not”) and two out of three categorical terms (“A”, “B”, and “C”):

All A are B.
Some B are C.

What, if anything, follows?

The goal of syllogistic reasoning is to interrelate the terms in the premises via the common *middle-term* (“B”), and derive information about the quantified relationship between the other two *end-terms* (“A”, “C”) or conclude “No Valid Conclusion” (NVC) to state that no quantified conclusion can be derived from the premises on logical grounds. For the sake of space and clarity, syllogisms are often abbreviated based on their structure. A syllogism is in one of four so-called figures, which represent the arrangement of terms:

Figure 1	Figure 2	Figure 3	Figure 4
A-B	B-A	A-B	B-A
B-C	C-B	C-B	B-C

Additionally, the quantifiers are encoded with A, E, I, O for “All”, “No”, “Some”, “Some not”, respectively (notation adopted from Khemlani & Johnson-Laird, 2012). Put together, the syllogism introduced before would be abbreviated with “AII”.

Due to the structural restrictions (exactly two premises, three terms, and four quantifiers), syllogistic reasoning is a well-defined domain with a total of 64 distinct problems and nine possible conclusion options. Because of this, syllogistic reasoning is one of the prime domains to study human deductive reasoning and explore hypotheses about the latent inferential processes of the human mind.

To date there exist at least twelve theories that try to explain the observable behavior of reasoners by drawing from a century worth of empirical investigation (Khemlani & Johnson-Laird, 2012). Models based on these theories provide sets of comprehensive and explanatory parameters to fine-tune the processes they assume to be operating the human mind. These parameters, in essence, are responsible for the explanatory value of theories as they provide the necessary information about the selection and strength of the processes that are responsible for the observable behavior. There is considerable work fitting parameters to data (Khemlani & Johnson-Laird, 2016; Riesterer, Brand, & Ragni, 2020a), which focuses on the ability of the models to account for the data. However, very little work focuses on the second use-case for these models, namely to go beyond the scope that they were

originally created for and extract explanations for new phenomena based on the interpretation of the parameters. This, however, considerably undervalues the worth of theories and models. Reflecting embodiments of insight, theories and models are capable of providing novel insight and should do so.

In this article, we attempt to use model implementations of theories to provide insight in the syllogistic reasoning process and the changes that feedback induces to these processes. Relying on a recent dataset that introduced feedback about the logical correctness of human responses as an experimental manipulation (Dames et al., 2020), we fit three prominent models (mReasoner, PHM, and TransSet) to the data. By investigating the resulting parameter distributions, we extract explanations for the effects of feedback from the theories. We then derived hypotheses that allow to experimentally test the theories explanations. At last, we conducted a study based on a modified version of the experiment by Dames et al. (2020) which featured additional questions targeting the derived hypotheses. This allowed us to replicate the feedback effect in order to ensure its robustness and test the hypotheses derived from the model’s explanations of the feedback effect.

The remainder of this paper is structured into four parts. First, we present relevant background about the theories and models for syllogistic reasoning. Second, we introduce our method of extracting explanations from the models and derive the hypotheses. Third, we describe the study and the dataset derived from it. Fourth, we present our results and discuss them with respect to the implications for the three models and the feedback effect, as well as the general implications for the field of syllogistic reasoning research.

Background

To date, syllogistic reasoning research has produced more than twelve theories attempting to explain the cognitive foundation of this form of reasoning (Khemlani & Johnson-Laird, 2012). Crucially, it was found that comparing these theories based on their ability to predict the distinctive responses of human reasoners to select an overall best explanation is difficult if possible at all (Khemlani & Johnson-Laird, 2012). Recently, however, it was found that in addition to this difficulty, predictive performances might have been overestimated due to a prevailing perspective on group analyses in the field (Riesterer, Brand, & Ragni, 2020c). If subjected to the task of predicting individual human responses instead of only the most frequently selected ones, predictive accuracies drop from above 84% (Khemlani & Johnson-Laird, 2012) to below 50% (Riesterer, Brand, & Ragni, 2020a). To see if this performance can notably be improved on—which would be clear evidence of an improved understanding of reasoning processes—or remains stuck due to high levels of noise in the data remains a crucial goal for future investigations in reasoning research.

Regardless of the questions surrounding model selection, recent results suggest that at least three accounts will play a

major role in future investigations for various reasons. First, the *Mental Models Theory* (MMT; Johnson-Laird, 1983) with its model implementation *mReasoner* (Khemlani & Johnson-Laird, 2013) is one of the most comprehensive theoretical accounts of reasoning spanning multiple domains (e.g., spatial relational, conditional, modal) and persisting for almost half a century. Second, the *Probability Heuristics Model* (Chater & Oaksford, 1999) is an instance of the probabilistic paradigm of cognitive science that adopts a stance discarding logical validity in favor of probabilistic validity. Finally, *TransSet* (Brand et al., 2020) is a recently proposed account that approaches syllogistic reasoning by focusing on a set-based interpretation of quantifiers and transitivity as its core inference rule. Currently, TransSet is the most successful model of syllogistic reasoning when judged based on predictive accuracy alone (Brand et al., 2020). In the following, the functional mechanisms of the three accounts will be introduced in greater detail.

MMT & mReasoner *MMT* approaches syllogistic reasoning via a four-step procedure (e.g., Copeland, 2006). First, a mental representation, the *mental model*, is created from the first premise. This mental model consists of a number of entities that reflect the information of the premise by being associated to the categorical terms or not. Second, the second premise is integrated into the mental model by extending the entities with information about the third term. Third, the resulting mental model is inspected to extract a conclusion candidate. In the final step, this candidate is probed by constructing alternative mental model representations that are consistent to the premises but inconsistent to the conclusion candidate. If no counterexample can be found, the conclusion is accepted as the conclusion to the syllogistic problem. Otherwise a new conclusion candidate is generated and subjected to the search for counterexamples or NVC is returned.

mReasoner is a LISP-based implementation of MMT for syllogistic reasoning that follows the four-step procedure outlined above but includes four parameters to further specify details about the model’s behavior (e.g., Khemlani & Johnson-Laird, 2016). First, λ specifies the maximum number of entities that are represented in the mental model. Second, ϵ specifies the composition of the mental model. For high values, the mental model is highly likely to exhaustively reflect the information available in the premises. For low values, it only reflects a limited canonical set of information. Third, σ reflects the propensity of the model to engage the search for counterexamples. Finally, if a counterexample is found, ω denotes the likelihood to continue the process with a weaker version of the conclusion candidate or abort the reasoning process to generate an NVC response.

PHM PHM approaches reasoning by adopting a perspective based on probabilistic validity or p-validity (Chater & Oaksford, 1999). To accomplish this without requiring computationally complex if feasible at all operations, the model is

based on a set of three generation heuristics (G1-G3) and two test heuristics (T1, T2) to approximate the p-valid behavior. To generate a conclusion, the *min-heuristic* (G1) identifies the premise with minimal informativeness (min-premise) based on the order $A > I > E > O$ and uses its quantifier as the conclusion quantifier. *p-entailment* (G2) proposes the quantifier probabilistically following from the min-heuristic result as an alternative conclusion quantifier candidate. The *attachment-heuristic* then defines the direction of the conclusion. If the min-premise begins with an end-term, it is used as the subject of the conclusion. Otherwise the end-term of the max-premise, i.e., the most informative premise in accordance to the above ranking, is used. After the conclusion is generated, the *max-heuristic* (T1) assesses a reasoner’s confidence in it by evaluating the informativeness of the max-premise. PHM assumes proportionality between confidence and max-premise informativeness. If confidence is low, NVC may be concluded instead (Copeland, 2006). Finally, the *O-heuristic* postulates that “Some ... not” conclusions should generally be avoided due to their extreme uninformativeness.

In a recent implementation of PHM (Riesterer, Brand, & Ragni, 2020a), a set of five binary parameters were used to further specify the model’s behavior. *p_ent* decides whether to use the min-heuristic or p-entailment to generate the conclusion quantifier. In addition, *A_conf*, *I_conf*, *E_conf*, *O_conf* are used to specify the confidence in the corresponding max premise quantifier.

TransSet TransSet is based on two phases: direction selection and quantifier selection (Brand et al., 2019, 2020). In *direction selection*, TransSet attempts to construct a transitive path from the premises. If this is not possible, the model returns NVC. Otherwise, it enters the *quantifier selection phase* in which the quantifier information is propagated along the transitive path. This procedure fails and leads to NVC if the first quantifier on the path is negative and the second quantifier is not all. Otherwise, the conclusion quantifier is obtained and can be combined with the direction to create the full conclusion.

TransSet uses four parameters to further specify its inferential mechanisms. First, *nvc_aversion* defines its susceptibility to the NVC aversion bias that might prevent reasoners from acknowledging the importance of this conclusion (e.g., Brand et al., 2020). In the direction selection phase, NVC aversion forces the model to create a transitive path regardless of the premises. *anchor_set* determines which term to start the transitive path from in this case. Third and fourth, *particularity* and *negativity* specify the availability of additional rules to directly derive NVC in the quantifier selection phase (Riesterer, Brand, Dames, & Ragni, 2020).

Method

Objective

The goal of the analyses presented in the following is to leverage the current understanding of human reasoning in form of

available model implementations in order to investigate the effects of feedback. By providing feedback about the logical correctness to reasoners, it is expected that the reasoning behavior changes. These changes should be reflected by different parameterizations resulting from fitting the models to the data. By interpreting the difference in parameter values, the effects of feedback on reasoning behavior can be analyzed and compared to the theoretical assumptions postulated previously (Dames et al., 2020; Riesterer, Brand, & Ragni, 2020b).

Dataset

To investigate the effects of feedback, we rely on the three datasets collected by Dames et al. (2020). First, *control* ($N = 39$) contains the control group of reasoners who were not provided with feedback about the correctness of their responses. Second, *1s* ($N = 146$) contains the group of reasoners who were presented with a feedback screen stating either “correct” or “incorrect” after each given response. Finally, *10s* ($N = 29$) contains the group of reasoners who were presented with feedback for a duration of 10s after each response. Regardless of the feedback condition, all participants were presented with the full set of 64 distinct syllogistic problems and tasked to select which of the nine possible conclusion options (including “No Valid Conclusion”) followed from the presented premises. There was a time limit for each task, which forced participants to respond within 1.5 minutes.

Performing traditional statistical (Dames et al., 2020) and data-driven modeling analyses (Riesterer, Brand, & Ragni, 2020b), feedback was shown to predominantly affect the propensity of reasoners to conclude NVC, a conclusion option that has previously been hypothesized to elicit aversion biases (Dickstein, 1976). Presenting feedback provides reasoners with the opportunity to realize the importance of the NVC response (correct in 37 out of the 64 syllogistic problems, i.e., 58%). As such, we expect models to reflect this increase in NVC usage in terms of their parameterizations. As the overall differences between the 10s condition and the 1s condition were rather small compared to the control group, we combined both feedback conditions for the following analysis.

Model Fitting

The analyses presented in the following rely on the *Cognitive Computation for Behavioral Reasoning Analysis* (CCOBRA) framework¹. CCOBRA facilitates the evaluation of computational models in a well-defined and structural manner and provides implementations for the three models considered in the analyses: *mReasoner* (Riesterer, Brand, & Ragni, 2020a), *PHM* (Riesterer, Brand, & Ragni, 2020a), and *TransSet* (Brand et al., 2020). Each model was fitted to each individual in the dataset separately. The resulting fits were then aggregated and broken down by the feedback condition.

The core results of our analysis are summarized in Figure 1a. The figure contains separate plots for each of the mod-

¹github.com/CognitiveComputationLab/ccobra

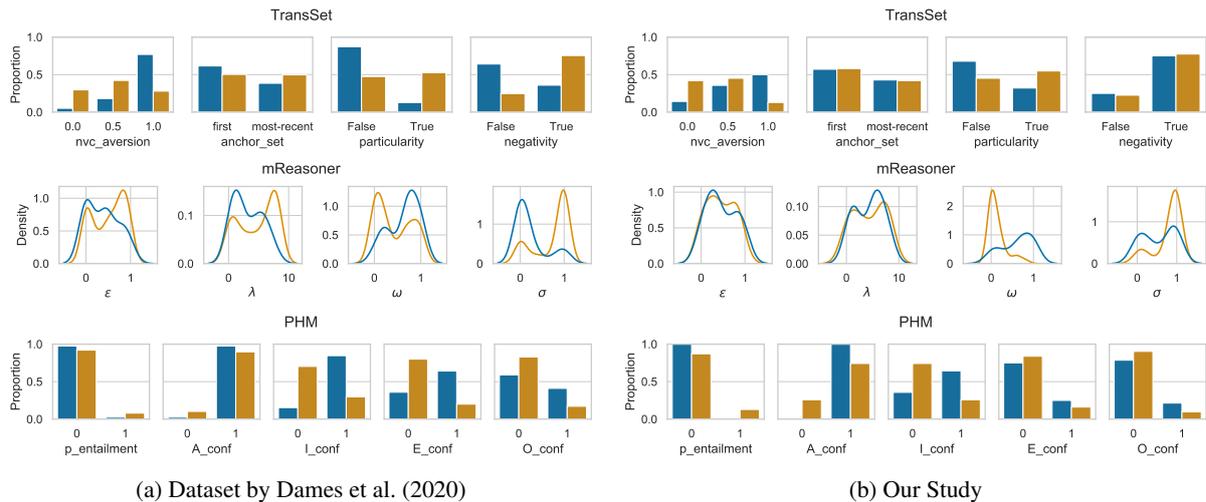


Figure 1: Parameter value distributions resulting from fitting the models to individual reasoners based on data from the original feedback study (left) and the study conducted in this work (right). Control and feedback are depicted in blue and orange, respectively.

els’ parameters. Each plot visualizes the distribution of the resulting values in terms of their proportions of occurrence (TransSet and PHM due to the parameters being discrete) or distribution (mReasoner due to being continuous). The different feedback conditions are represented by color with control and feedback in blue and orange, respectively.

On a high level, the plots reveal the obvious: the feedback manipulation of the experimental setting has an influence on human reasoning behavior that is reflected by differences in the fit results. To work out the explanatory meanings from the fits, the following sections inspect the results of each model separately and derive a hypothesis from the possible explanation.

TransSet TransSet shows distinct differences between control and the feedback condition for *nvc_aversion*, *particularity*, and *negativity*. The parameter *anchor_set* is ignored in the following due to its technical purpose and relative uniformity between the different conditions.

The value of the *nvc_aversion* parameter is substantially higher for the control condition than for feedback. In the case of *particularity*, control exhibits a strong skew in favor of *False* with the feedback condition leaning slightly towards *True*. For *negativity*, a similar skew can be observed but to a minor degree, at least for control. The feedback condition are skewed stronger towards *True*.

To summarize, TransSet attributes the effects of feedback to NVC handling, which is not surprising as it is TransSet’s main method of distinguishing individuals. The reduced value of *nvc_aversion* suggests that feedback incentivizes reasoners to accept NVC more leniently when compared to the behavior of naive reasoners (control). A similar interpretation is suggested by *particularity* and *negativity*, which control the availability of rules to abort the reasoning process in favor of

NVC. With control leaning more towards *False* and the feedback conditions to *True*, TransSet suggests that feedback allows reasoners to find and leverage heuristic rules to easily derive NVC, the response that naive reasoners (control) try to avoid.

As it is assumed that participants in the feedback condition use fast detection methods allowing them to identify NVC responses early, it is expected that the difference between the time needed for NVC responses and non-NVC responses is lower for the feedback group compared to the control group (**H1.1**). Although NVC is usually important in more difficult tasks, the NVC-specific heuristics could outweigh the difficulty and lead to overall lower times for NVC responses in the feedback condition compared to the control group (**H1.2**).

mReasoner Interestingly, mReasoner’s parameter distributions are bimodal between control and the feedback conditions. Perhaps most crucially, the σ parameter is substantially affected by feedback. As the parameter controls the propensity to engage in a search for counterexamples, which the prerequisite to derive NVC responses, this was to be expected. Less distinctly, λ and ϵ show similar behavior with control and feedback being skewed towards lower and higher values, respectively. For ω , which only plays a role within the search for counterexamples and therefore dependent on σ , feedback is mainly skewed towards the lower spectrum, while the control condition yields higher values.

To summarize, mReasoner seems to attribute the effects of feedback to a switch from a more intuitive reasoning to a more thought-out process incorporating a search for counterexamples. The propensity to rigorously evaluate the mental model via the search for counterexamples is increased (σ), and the likelihood to weaken the conclusion (which in turn allows to avoid an NVC response) is reduced. Additionally, a

Table 1: Syllogisms selected for the test phase of the study. The encoding is in line with Khemlani & Johnson-Laird (2012).

Valid	AA4, AE2, AO3, EA1, EI1, IA4, IE4, OA3
Invalid	EE2, EO3, II4, IO3, OA2, OE1, OI3, OO1

learning component can be identified: When confronted with feedback, reasoners realize the importance of correctly interpreting the premise information resulting in more comprehensive (λ) and complete (ϵ) mental models. However, it is important to note that the effects of λ and ϵ have shown to have little impact on the general behavior of mReasoner in comparison to σ (Riesterer, Brand, & Ragni, 2020a). Due to the expensive search for counterexamples, it is expected, that participants in the feedback group should be substantially slower when deriving NVC responses (**H2**).

PHM PHM’s parameterization is special because of the dependencies between the confidence parameters *A_conf*, *L_conf*, *E_conf*, and *O_conf*, which control the behavior of the max-heuristic. Since this heuristic states that confidence in the conclusion is *proportional* to the max-premise quantifier’s informativeness (Chater & Oaksford, 1999), the corresponding parameters are ordered. As soon as one parameter is set to 0, all proceeding ones must necessarily be 0 as well, indicating that confidences are so low that the conclusion is abandoned in favor of NVC.

The order of confidences is reflected by the model parameters with proportions of 1s decreasing from *A_conf* through *O_conf*. Importantly, across the board, control elicits the highest proportions of 1s with both feedback conditions eliciting similar results. *p_ entailment*, the only truly independent parameter, is dominated by 0s regardless of the condition.

To summarize, PHM suggests that feedback results in an overall decrease of confidence in conclusions potentially caused by the importance of the NVC response. Therefore, the confidence in non-NVC responses should be lower for the feedback group (**H3**).

Study design

Based on the hypotheses described above, we conducted an online-study via Prolific, in which participants were instructed to give conclusions to all 64 syllogistic problems. The study had a single-choice design, where the participants selected the conclusion by clicking on the respective button. In order to avoid a bias due to content effects, hobbies and professions were used as content for the syllogisms. The order of the response options was randomized. Participants were randomly assigned to the control condition or to the feedback condition. The experiment was divided into two parts: The first 48 syllogism (presented in random order) were regarded as a training phase, where feedback was shown for 1 second for the feedback condition. As in the original ex-

periment, the feedback only stated if the selected answer was correct. After the first 48 syllogisms, both groups received no feedback as we assumed that the feedback effect would be apparent after training. In the second phase (test phase), the participants were asked to not only select an answer, but also to estimate their confidence in the selected option by choosing values from 0% to 100% on a slider. A predefined set of syllogism was used, which featured the 8 valid and 8 invalid syllogisms that had the most differences with respect to the response behavior between feedback and non-feedback in the dataset from Dames et al. (2020), with the constraint that only unique quantifier-combinations were in the set. This was done to increase the variability and to minimize the effect of single strategies and biases (e.g., the Atmosphere effect Wetherick & Gilhooly, 1995). The selected tasks are shown in Table 1. We did not include a time limit, which allows us to disentangle the effect of feedback from the effect that the short time-frame might have had in the original study.

After excluding participants which did not take the experiment seriously (i.e., needed less than 10 minutes for all tasks, performed worse than chance, or interrupted the study for more than 5 minutes; $N = 6$), there were $N = 59$ participants, with $N = 28$ in the control group and $N = 31$ in the feedback condition. The dataset, all materials and scripts are openly available on GitHub².

Analysis

First, we compared the dataset from our study with the dataset by Dames et al. (2020). In particular, we investigated if, and to which extend, the feedback effect is apparent without a time limit. Second, we re-fitted the models to the new data in order to verify that the main predictions still hold. Subsequently, the hypotheses derived from the model’s explanations were tested based on the results of the second phase of the study. In the following section, the results are presented and discussed.

Results

The feedback effect in the original study mainly manifested in the number of NVC responses (Dames et al., 2020). This effect was also apparent in our data, as the average percentage of NVC responses (control: *mean* = 0.21, *std* = 0.17; feedback: *mean* = 0.41, *std* = 0.23) showed higher values for the feedback condition. However, the correctness (control: *mean* = 0.46, *std* = 0.50; feedback: *mean* = 0.48, *std* = 0.50) was not affected, which differs from the results by Dames et al.. This is likely the effect of the time limit, which caused the control group to perform worse (0.326 without feedback; 0.434 with feedback), while they achieved a closer result in our data.

Figure 1b shows the results of the model fits on the data from our study. Overall, the model parameters still show the same pattern, clearly showing the feedback effect. However, the effect is not as dominant as in the original study, which is

²github.com/Shadownox/iccm-feedbackexplanation

Table 2: Results of Mann-Whitney U tests (p-values and U statistic) for the hypotheses H1.1, H1.2, H2, and H3. p_{cor} shows the Bonferroni-corrected p-values.

Hypothesis	Median		U	p	p_{cor}
	Control	Feedback			
H1.1	0.92	-5.1	178.0	.006	.017
H1.2 / H2	21.26	12.06	228.0	.06	.18
H3	71.45	55.5	228.0	.001	.003

especially prominent for PHM the E_{conf} parameter, where the control condition is almost identical to the feedback condition in our data, while there was still a substantial difference in the original study. This is likely due to the missing time limit, which might have strengthened existing biases and pushed participants more towards intuitive responses. For mReasoner, this gets also apparent for the more subtle parameters λ and ϵ , which are differing substantially between the datasets with control and feedback showing almost not difference without a time limit while having distinct patterns when a time limit is present. Despite these differences, the parameter distributions between both experiments are comparable with respect to the extracted explanations, allowing a evaluation of the derived hypotheses. In the following, we compare the two conditions in order to test the hypotheses. To correct for multiple comparisons, we use the Bonferroni correction and also the corrected p-values. The results of the comparisons between control and feedback conditions for the hypotheses are shown in Table 2.

At first, we discuss the hypothesis of TransSet (H1.1). According to TransSet’s mechanism, NVC responses should be derived faster compared to non-NVC responses. As the feedback effect is explained by participants being less hesitant to derive NVC and also the utilization of NVC-specific rules, it is expected that this time difference is higher in the feedback condition compared to the control group. In fact, there is a significantly bigger difference in the feedback condition compared to the control group: participants in the feedback condition are substantially faster when deriving NVC, while the control group even needs more time for NVC responses. Additionally, the overall time of participants for NVC responses was lower in the feedback condition compared to the control group (H1.2), although significance was not reached. The hypothesis for mReasoner (H2) directly contradicts hypothesis H1.1 by predicting participants in the feedback condition to take more time, as they are more likely to engage in the expensive search for counterexamples. Since the data is even leaning towards H1.1, it is not supported by the data.

Finally, the prediction of PHM is tested. PHM predicts the confidence in non-NVC responses to be lower in the feedback group. This prediction was indeed supported by the data, showing a significantly lower confidence in the feedback condition compared to the control group. This indicates

that NVC could in fact be an option that is selected if participants have low confidence in other response options.

Discussion

The present work has three main contributions: First, we fitted three cognitive models that are state-of-the-art, PHM, mReasoner and TransSet, to each individual participant and used the resulting parameter distributions to extract explanations based on the assumed processes of the respective model. At last, these were used to derive new hypotheses that allowed us to test the models’ explanatory capabilities. Second, the feedback effect in syllogistic reasoning which was reported by Dames et al. (2020), was replicated by our study without the time limit imposed in the original study. This indicates that the effect is in fact robust and not only an interaction effect induced by the time limit. Third, by testing the hypotheses derived from the models, we were able to assess the models’ capabilities to account for the feedback effect.

Regarding the feedback effect for syllogistic reasoning, the explanations extracted from PHM and TransSet were supported by our study. Both explanations are also compatible, as it is possible that feedback at first has the effect of lowering the confidence in non-NVC-responses and later helps to develop fast and frugal detection strategies for NVC once the importance of NVC responses is realized. In contrast, the prediction of mReasoner was not supported and the data seems to even contradict its explanation. Based on our findings, we conclude that the feedback effect is best described as a heuristic process, where participants learn that NVC is a viable response option and therefore adapt their general judgment of the other response options. An explanation based on the assumption that feedback improves the reasoning process (e.g., by shifting away from intuitive responses) could not be supported. In summary, our findings indicate that PHM and TransSet are more probable accounts for the feedback effect in syllogistic reasoning. While they provide differing explanations, they might describe different parts of the same process.

Generally, our work successfully applied cognitive models for syllogistic reasoning to a new phenomenon in order to derive new hypotheses by interpreting the parameters, which is rarely done in this field. Instead, it is often the other way round: New findings were first integrated in theories and then into the respective models. While this is a valid approach to formalize the current knowledge about human reasoning into models, it does not utilize the predictive capabilities of the models. We hope that future modeling endeavors will test models more based on predictions outside their original scope, which will not only improve model selection, but also advance the field as a whole by fueling further investigations with new hypotheses.

Acknowledgements

Support to MR by the Danish Institute of Advanced Studies and the DFG (RA 1934/4-1, RA1934/8-1) is gratefully acknowledged.

References

- Brand, D., Riesterer, N., & Ragni, M. (2019). On the matter of aggregate models for syllogistic reasoning: A transitive set-based account for predicting the population. In T. Stewart (Ed.), *Proceedings of the 17th International Conference on Cognitive Modeling* (pp. 5–10). Waterloo, Canada: University of Waterloo.
- Brand, D., Riesterer, N., & Ragni, M. (2020). Extending TransSet: An individualized model for human syllogistic reasoning. In T. C. Stewart (Ed.), *Proceedings of the 18th International Conference on Cognitive Modeling* (pp. 17–22). University Park, PA: Applied Cognitive Science Lab, Penn State.
- Chater, N., & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology*, 38(2), 191–258.
- Copeland, D. E. (2006). Theories of categorical reasoning and extended syllogisms. *Thinking & Reasoning*, 12(4), 379–412.
- Dames, H., Schiebel, C., & Ragni, M. (2020). The role of feedback and post-error adaptations in reasoning. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd annual conference of the cognitive science society* (pp. 3275–3281). Cognitive Science Society.
- Dickstein, L. S. (1976). Differential difficulty of categorical syllogisms. *Bulletin of the Psychonomic Society*, 8(4), 330–332.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Harvard University Press.
- Khemlani, S. S., & Johnson-Laird, P. N. (2012). Theories of the syllogism: A meta-analysis. *Psychological Bulletin*, 138(3), 427–457.
- Khemlani, S. S., & Johnson-Laird, P. N. (2013). The processes of inference. *Argument & Computation*, 4(1), 4–20.
- Khemlani, S. S., & Johnson-Laird, P. N. (2016). How people differ in syllogistic reasoning. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 2165–2170). Cognitive Science Society.
- Riesterer, N., Brand, D., Dames, H., & Ragni, M. (2020). Modeling human syllogistic reasoning: The role of “No Valid Conclusion”. *Topics in Cognitive Science*, 12(1), 446–459.
- Riesterer, N., Brand, D., & Ragni, M. (2020a). Do models capture individuals? Evaluating parameterized models for syllogistic reasoning. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society* (pp. 3377–3383). Toronto, ON: Cognitive Science Society.
- Riesterer, N., Brand, D., & Ragni, M. (2020b). Feedback influences syllogistic strategy: An analysis based on joint nonnegative matrix factorization. In T. C. Stewart (Ed.), *Proceedings of the 18th International Conference on Cognitive Modeling* (pp. 223–228). University Park, PA: Applied Cognitive Science Lab, Penn State.
- Riesterer, N., Brand, D., & Ragni, M. (2020c). Predictive modeling of individual human cognition: Upper bounds and a new perspective on performance. *Topics in Cognitive Science*, 12(3), 960–974.
- Wetherick, N. E., & Gilhooly, K. J. (1995). ‘Atmosphere’, matching, and logic in syllogistic reasoning. *Current Psychology*, 14(3), 169–178.

Timing and Structure of Reward Information Influences Bias in Perceptual Decisions as Revealed by a Hierarchical Drift Diffusion Model

Manisha Chawla, Krishna P. Miyapuram

Centre for Cognitive and Brain Sciences, Indian Institute of Technology Gandhinagar,
Palaj, Gandhinagar 382 355, INDIA
kprasad@iitgn.ac.in

Abstract

Differential payoffs can bias simple perceptual decisions. Drift Diffusion models (DDM) have been successfully used to simultaneously model for response times (RTs) and accuracy of binary decisions. The DDM allows for identification of latent parameters that represent psychological processes underlying perceptual decisions. These parameters characterize decision making as a noisy process that accumulates evidence towards one of the two boundaries. Previous research in two alternative forced choice (2AFC) experiments has found that asymmetric payoffs result in a bias towards those decisions that result in higher payoff. We manipulate the reward structure resulting in symmetric and asymmetric payoffs for a simple orientation discrimination task and test for the differences in parameters of drift diffusion model that might relate to reward-induced bias in perceptual decisions. To understand the mechanisms of how reward information might be integrated with perceptual decisions, we altered the relative timing i.e. processing order of reward information and perceptual stimuli. Computational modelling using a hierarchical DDM revealed starting point bias towards stimuli oriented in the direction of higher rewards in asymmetric as well as symmetric rewards. The drift rate reflected the average reward expectation when reward information was presented before, but not after the perceptual stimulus. Our results suggest that integration of rewards with perceptual decisions is mediated by modulating motivation for evidence accumulation over time and prior bias in starting point.

Introduction

Computational models for Perceptual decision making describe the dynamic evolution of preferences across time until a decision is reached, rather than assuming a fixed state of preference. The Decision field theory (Bussemeyer & Townsend, 1993) is a member of a general class of sequential sampling models. Models such as drift-diffusion model (DDM: Ratcliff, 1978) suggest accumulation of varying sensory evidence that leads to a choice beyond a certain threshold. The DDM models decision-making in two-choice tasks represented by two boundaries separated by distance (represented by threshold parameter a). Lower threshold makes responding faster in general but increases the influence of noise on decision making and can hence lead to errors or impulsive choice. Higher threshold leads to more cautious responding (slower, more skewed RT distributions, but more accurate). Different studies have shown that the parameter a is sensitive to speed versus accuracy instructions (e.g., Voss et al., 2004). Additionally, there is a large body of research showing that age-related slowing in response time tasks can be partially explained by more emphasis on correct responses (e.g., Ratcliff et al., 2000, 2006, 2010, 2011). A drift-process accumulates

evidence over time with certain speed (drift-rate parameter v) until it crosses one of the two boundaries indicating the choice made. Due to noise in each trial of the drift process, the time taken to reach a particular boundary would vary across trials. If such a consistent variation is observed over different conditions the drift rate reflects task-difficulty with smaller drift rates representing more difficult tasks. In the comparison of participants, drift is a measure for individual cognitive or perceptual speed of information processing (Schmiedek et al., 2007). The DDM model also includes bias parameter z to account for starting point closer to one of the boundaries and non-decision time parameter t that encodes processes unrelated to decision making such as stimulus perception and movement (Smith & Ratcliff, 2004).

Previous research using the DDM has revealed that the effects of payoff manipulations on a perceptual decision-making task can be identified through various parameters. Dunovan et al. (2014) made a distinction between a prior bias in starting point parameter z and a dynamic bias in drift rate parameter v . The former model suggests influence of the payoffs on perceptual decisions to be only during the initial stage, while the latter suggests these influences to persist until reach the decision boundary. Van Ravenzwaaij et al. (2012) found that prior information influences starting point rather than the drift rate. Bias parameter is responsible for the starting point of response time distributions for each trial. Difference in bias parameter across conditions can reflect choices encountered with different payoff matrices. For example, Voss et al. (2004) showed that the starting point is moved toward a response threshold when the corresponding response leads to greater rewards (for a review see Voss et al., 2013). Similarly, in the domain of motivated perception, it has been found that the starting point is closer to the "positive" threshold than to the "negative" threshold in an evaluation task, even when expectancy values for both responses were symmetric (Voss et al. 2008). Diederich and Bussemeyer (2006) tested three models (1) the Bound Change Model that results in maximizing payoffs through a change in the decision threshold parameter, (2) the Drift - rate change model suggests a bias in drift rate owing to difference in payoffs, and (3) the Two-stage processing model proposed by Diederich (1997) where the decision task is separated as two evidence accumulation processes that occur sequentially. The first stage involves evidence accumulation process for the reward structure followed

by sensory evidence accumulation for the perceptual discrimination task. In a recent study, Diederich (2016) manipulated processing order of perceptual and payoff information and found further evidence in support of their multi-stage processing model. These prior studies establish a clear link for integration of contextual bias of reward information with perceptual decision making.

In the current study, we investigate how the temporal dynamics and structure of reward information bias perceptual decisions. The reward structure consists of two types of information - symmetric or asymmetric payoffs and is presented during an orientation discrimination task. As in previous studies, we presented reward information prior to the perceptual stimulus during the PreStim experiment. We conducted a second experiment, PostStim where the reward information was presented after the perceptual stimulus, but before requiring a response. Using hierarchical drift diffusion modeling, we tested separate models varying drift rate, decision threshold, and bias parameter to explain the choice distributions and response times. The models considered left and right responses as the two boundaries. Our results suggested that reward modulates perceptual decision making for both symmetric and asymmetric rewards encoded by a bias in starting point. Our results support the multistage model proposed by Diederich (2016) integrating the reward information in perceptual decisions.

Materials and Methods

Participants

PreStim experiment had 10 student volunteers (age range 19-31 years) and the PostStim experiment had 11 student volunteers (age range 19-31 years). All participants were right handed and had normal or corrected to normal vision. All participants gave written informed consent and were paid for their participation.

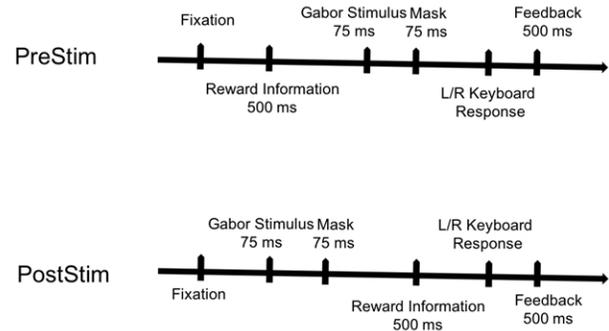
Stimuli

Each stimulus was composed of Gabor patches which were composed of a Gaussian envelope with a spatial frequency of 0.01 cycles/pixel. Maximal Michelson contrast of gratings was 0.9. Orientation of Gabor patches varied from -85 to 85 degrees with reference to vertical in step sizes of 5 degrees. A scrambled image was constructed by a combination of left and right oriented images of 45 degrees and was used for masking the gabor stimuli.

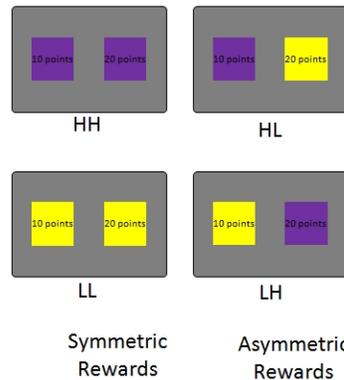
Design

The experiment was designed to test how reward information influences perceptual decisions. We manipulated the timing and the type of reward information presented that reflected the outcome (payoff) of the perceptual decisions. Therefore, reward information presented was irrelevant for performance of the perceptual task of detecting the orientation of gabor stimuli. Two types of reward information were presented - high reward magnitude (20 points) and low reward magnitude

A. Experiment Timeline



B. Reward Structure



C. Sample stimuli and Mask

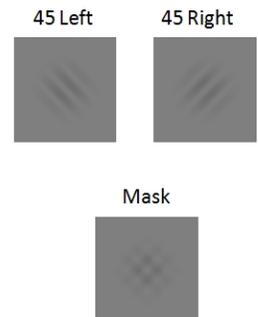


Figure 1: Experimental settings:- A. Structure of a single trial for both the PreStim and PostStim experiments. B. The four experimental conditions indicating High(H) and Low (L) rewards associated with left and right responses. Symmetric rewards are when both left and right responses would be associated with same reward. Asymmetric rewards conditions differentially reward left and right responses. C. Sample perceptual Gabor stimuli and mask stimulus. Participants were required to indicate the orientation of the Gabor stimulus.

(10 points). The reward information was displayed as the text written inside a square box and further coded in distinct colors. Reward information for the left and right oriented stimuli was presented on two sides of the screen centered vertically. The outcome of the trial was the reward displayed on the side that matched the orientation of the gabor stimulus. The reward information was manipulated in two ways - Symmetric and Asymmetric rewards. Symmetric conditions could be high rewards (HH) or low rewards (LL) in which both left and right correctly identified orientations were high or low rewarding, respectively. The Asymmetric conditions further was one of HL or LH conditions. In the HL condition, correct identification of left oriented gabor stimuli was rewarded with higher (20) points and the correct identification of right oriented gabor stimuli was rewarded with lower points (10). The LH condition was similar to the HL condition with the high and low reward contingencies being flipped.

Hierarchical drift diffusion modeling

We used an open source python toolbox for the hierarchical Bayesian estimation of the drift-diffusion model parameters (Wiecki et al., 2013). The toolbox uses Markov chain monte-carlo (MCMC) inference algorithm to estimate the joint posterior distributions of the different model parameters. We used Gelman-Rubin statistic to assess the convergence of the Markov chains by comparing the inter-chain and intra-chain variance of 5 different runs of the same model, resulting in ± 0.01 MC error suggesting 15000 samples were sufficient for convergence. For each model we generated 15000 posterior samples and discarded the first 5000 samples using burn to allow the MCMC chains to stabilize. The models were response coded with correct responses for right orientation terminating at the upper boundary and the left responses at the lower decision boundary.

To examine which model parameters are affected by the different type of reward structures and their timings we ran three different models allowing the parameters v , a and z to vary across experimental conditions, one at a time (model-V, model-A, model-Z), a composite model all three parameters to simultaneously vary between conditions, and a base model in which non of the parameters were allowed to vary across conditions. We then compared these different models using deviance information criterion (DIC) and posterior predictive checks (PPC) to find the best fitting model. DIC is a measure of relative goodness of fit for hierarchical Bayesian models (Speiegelhalter et al., 2012). DIC uses the trade-off between model fit and model complexity to compare relative goodness of the models. The best model is regarded as the one with the lowest DIC values. Difference of greater than 10 between different model DIC values is regarded as significant (Dunovan et al., 2014; Zhang & Rowe, 2014). Since, it is known that DIC is sometimes biased towards models with higher complexity we also ran Posterior Predictive Checks on group and subject data to assess the best fitting model (Michmizos & Krebs, 2014). We generated 500 simulated datasets from posterior predictive distributions of parameters corre-

sponding to the Composite model that was best-fit to the data based on lowest DIC value. We then compared the observed data distribution (empirical values from our experiment) with the simulated data generated which were found to be within 95% credible interval. Model goodness of fit is assessed using the mean standard error (MSE). Comparatively lower values of MSE for a model suggest that the model is able to reproduce observed data pattern distributions with less variability and more accuracy (Michmizos & Krebs, 2014).

The model parameters a , v and z thus estimated from the best fit model and their posterior distributions were used for statistical analysis. Our primary goal of the current research was to identify whether or not the different parameters varied across different conditions. We use posterior comparison for significance testing by calculating the proportion of overlap between the probability density of the two conditions being compared (Wiecki et al., 2013; Michmizos & Krebs, 2014). We also performed classical significance tests on mean parameter estimates as described further in Results section (Zhang & Rowe, 2014).

Results

Drift diffusion models

The drift diffusion models we considered were computed systematically allowing one parameter to vary across conditions keeping the other parameters invariant. Three models were formed to explore the modulations of parameters V , A , and Z : model-V, model-A, and model-Z, respectively. Model-V assessed for different drift-rates of evidence accumulation for left and right oriented perceptual stimuli across symmetric and asymmetric reward conditions. Similarly, model-A and model-Z assessed for decision threshold and starting point, respectively, for any biases in these parameters dependent on the reward structure. These three models were compared to a basic model in which all parameters remained invariant across conditions (BASE) using Deviance Information Criterion DIC. It was observed that all three models had lower DIC (model-V: 3426.32, model-A: 2950.80, model-Z: 3650.63) than BASE model indicating better fit to the data (3707.42). We ran a composite model that allowed for the above three parameters to vary across the four reward conditions that was found to be the best fit model (2782.35). The parameters estimated from the composite model were comparable to the independent models and future analysis are based on the estimates from the composite model. We ran a composite model including non-decision time (t) as a parameter. Mean estimates of parameter ' t ' were close to zero across reward conditions and hence are not discussed further.

Group parameter estimates from model fits of individual subjects were estimated using the hierarchical DDM. Group parameter estimates were tested for differences within the symmetric (HH and LL) and asymmetric (HL and LH) conditions using two complementary approaches (Zhang & Rowe, 2014). We compared mean parameter estimates across participants using a classical frequentist approach (i.e. t-test). We

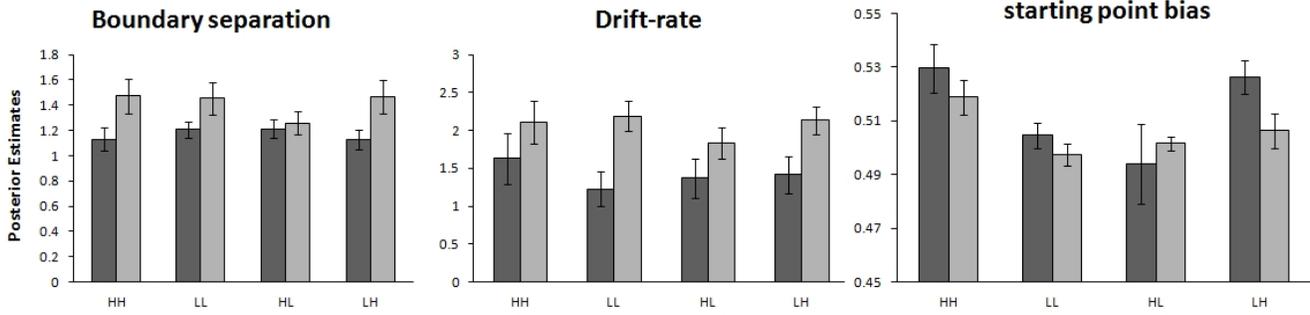


Figure 2: Group posterior estimates (y-axis) of the hierarchical drift-diffusion model parameters for the PreStim (dark gray bars) and PostStim (light gray bars). a.) Boundary separation parameter a b.) Drift-rate variability v and c.) Bias in starting point z . Error bars show standard error of mean from the posterior estimate samples.

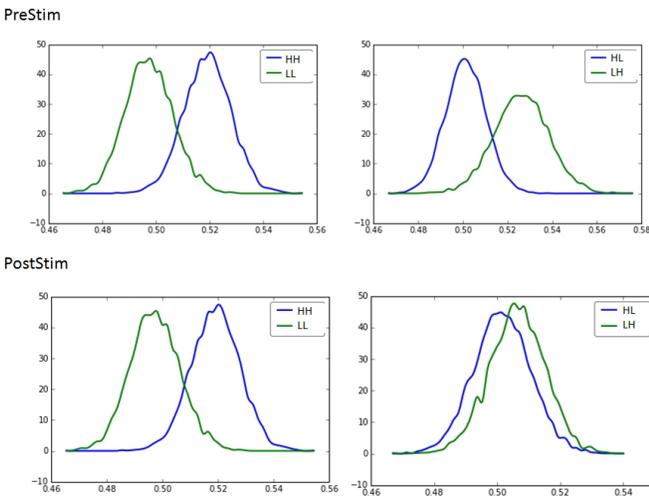


Figure 3: Posterior estimates (y-axis) of bias parameter z for a.) PreStim Symmetric conditions b.) PreStim Asymmetric conditions c.) PostStim Symmetric conditions d.) PostStim Asymmetric conditions.

compared the group posterior distributions obtained for each parameter using Bayesian approach.

Drift rate In PreStim experiment, we found bias towards high reward for drift rate corresponding to symmetric conditions ($HH > LL, t(9) = 2.47, p < 0.05; P_{Bayes} = 0.84$) but not for asymmetric conditions ($HL > LH, t(9) = -0.32, p = 0.38; P_{Bayes} = 0.46$). The drift rate for asymmetric conditions (HL, LH) was found to be intermediate to the symmetric high (HH) and low (LL) conditions possibly reflecting expectation of reward (Figure 2). In PostStim experiment we found no difference within the two symmetric ($HH > LL, t(10) = -0.17, p = 0.43; P_{Bayes} = 0.43$) and the two asymmetric reward ($HL > LH, t(10) = -1.24, p = 0.12; P_{Bayes} = 0.27$) conditions. Overall drift rates for the PostStim experiment were higher compared to the PreStim Experiment reflecting faster

response times (Figure 2). This could be due to the process of evidence accumulation being initiated during the reward information period prior to making a response.

Decision Threshold The decision threshold parameter showed no significant difference for both PreStim ($HH > LL, t(9) = -1.05, p = 0.16; P_{Bayes} = 0.18$) and ($HL > LH, t(9) = 1.37, p = 0.10; P_{Bayes} = 0.79$) and PostStim ($HH > LL, t(10) = 0.13, p = 0.45; P_{Bayes} = 0.51$ and ($HL > LH, t(10) = -1.38, p = 0.09; P_{Bayes} = 0.19$) experiments across conditions. This reflects that the boundary separation between left and right choices was not significantly different for the two symmetric and asymmetric reward conditions. The decision threshold for PostStim was observed to be greater than PreStim experiment (Figure 2) possibly due to the accumulated evidence not being allowed to reach the boundary (i.e. more conservative) while waiting for the "go" signal before the response execution.

Bias The posterior estimates of the response bias parameter for the PreStim experiment were found to be different in both the symmetric and asymmetric reward conditions. High reward condition showed relatively higher bias as compared to the low reward condition ($HH > LL, t(9) = 2.64, p < 0.05; P_{Bayes} = 0.93$). The asymmetric reward conditions had significant bias towards the boundary with high reward compared to the low reward ($HL > LH, t(9) = -2.08, p < 0.05; P_{Bayes} = 0.03$). These results reflect a prior bias for the starting point of the drift process towards the boundary with higher reward (Figure 6). For the PostStim experiment, the bias parameter was significantly different in symmetric reward conditions ($HH > LL, t(10) = 2.49, p < 0.05; P_{Bayes} = 0.96$), but was not significantly different for asymmetric rewards ($HL > LH, t(10) = -0.68, p = 0.25; P_{Bayes} = 0.36$). These results reflect absence of response bias when reward information is presented after the stimulus. On average, the bias parameters were similar for PreStim and PostStim experiments (Figure 3).

Discussion

Our results of bias parameter being dependent on reward structure supports the two-stage model proposed by Diederich and Busemeyer (2006). The timing of our experiments allows us to explicitly test support towards the two-stage model, specifically the mechanisms involved in integration of reward values during perceptual decisions. The model proposes two accumulation processes. Payoffs influence the starting point in the first stage by introducing a prior bias towards the response with higher reward. Then in the second stage, evidence accumulation is done for the perceptual stimulus. By manipulating the relative timing of presentation of reward information and the stimuli, we tested whether a dynamic bias can be induced by reward structure after the process of evidence accumulation has already begun upon stimulus presentation. Our results support the two-stage model as we find a bias in starting point for asymmetric rewards when the reward information is presented before but not after the stimulus presentation (Figure 3). The differences in starting point can therefore be attributed to the first stage of evidence accumulation process initiated by reward structure. When the reward information is presented after the stimulus, the two stage model would correspond to the second stage alone, in which evidence accumulation occurs for perceptual stimuli (Diederich & Busemeyer, 2006). Hence, we do not observe differences in starting point in PostStim experiment for asymmetric rewards.

Bias parameter encodes both symmetric and asymmetric rewards when reward information is presented before the stimulus. Further, when the reward information is presented after the stimulus, the bias parameter no longer encodes starting-point bias, rather encodes a decision bias for symmetric rewards. Similar proposal to distinguish between response-execution bias from decision bias have been made earlier (Voss et al., 2010). Reward conditions with high rewards (HH) have higher starting point relative to the reward conditions with low rewards (LL). This could possibly be due to greater motivation for perceptual discrimination in high reward conditions. This motivation-induced decision bias in decision needs to be interpreted differently from a response bias. Response bias refers to the starting point of evidence accumulation for the perceptual stimuli being biased towards the boundary corresponding to higher reward. This response bias was observed when the reward information was presented before, but not after the stimuli. The starting point in an unbiased setting would be midway of the two decision boundaries. Allowing the response bias to be estimated as a free parameter, which in turn allows us to re-interpret the bias parameter as a decision bias. Our results can be compared to previous findings (Voss et al. 2008) of motivational influences for perceptual and judgmental bias in which starting point parameter is biased towards the gain threshold. However, the current research does not explicitly dissociate the specific interpretation of the bias parameter arising from a response bias, or can be considered to be a decision bias.

The influence of reward structure on perceptual decisions can be described by two kinds of bias. The starting point reflects a prior bias, while the drift rate can encode for a dynamic bias (Dunovan et al., 2014). We found the drift rate encodes an average reward expectation for the PreStim Experiment. Our computational models estimate a single drift rate towards the two boundaries with complementary sign (v , $-v$). Hence, the finding that higher drift rate for high compared to low reward conditions (Figure 2) reflects a dynamic bias in processing the perceptual stimuli, consistent with previous research that claim that influence of payoffs persists over time, rather than only changing the starting-point (Dunovan et al., 2014; Voss et al., 2008). Together with other findings that do not find encoding of a dynamic bias in drift rate (e.g.: Mulder et al., 2012), and our manipulation of processing order, our results suggest that the reward information is encoded differently by prior bias in starting parameter and average reward expectation by dynamic bias i.e. drift rate parameter.

An intriguing result is that when the parameters from Post-Stim Experiment are compared to the PreStim experiment, we find higher drift rate and decision threshold (Figure 2). These could reflect the fact that the evidence accumulation process in support of the perceptual decision had already taken place before the reward information, following which the cue for indicating the response is given. Unlike previous studies that investigated influence of payoffs on subsequently presented perceptual stimuli, our study design separates the timing of choice execution from the perceptual decision process. Thus, we can dissociate whether the reward information influences the choice execution or the evidence accumulation mechanisms of the decision process. The results supports the notion that when stimulus is presented prior to the payoffs, evidence accumulation processes result in faster and more accurate responses, i.e. higher drift rate and more conservative decision thresholds.

Our study contributes towards understanding the mechanisms of integration of reward (value-based) information with perceptual decisions. Previous research by Rorie and colleagues (Rorie et al., 2010) demonstrated that perceptual decisions by monkeys being influenced by asymmetric but not symmetric rewards. Using computational model (DDM) analysis in the current research, we are able to identify latent parameters that correspond to influence of payoffs in perceptual decision making. The reward information pertains to value-based computations, but is unrelated to performance of the perceptual task. The reward information indicates the payoff (outcome) arising after perceptual decision. Our results demonstrated that parameters of the drift diffusion model, a model of perceptual decisions are influenced by the reward structure. This finding, though not completely novel, is further corroborated with manipulation of processing order (i.e. timing) to study the mechanisms of integration of reward values with perceptual decisions. While the behavioral results might simply suggest that it is crucial for reward information to be presented before, but not after the perceptual

stimulus, the computational modeling approach has been useful to understand the specific parameters that are encoded by the timing and structure of reward information on perceptual decisions.

In sum, our results show that symmetric and asymmetric rewards bias the starting point towards stimuli oriented in the direction of higher rewards, and also reflect the average reward expectation by the drift rate. These results can be interpreted as integration of rewards with perceptual decisions is mediated by modulating motivation for evidence accumulation over time and prior bias in starting point.

Acknowledgments

This research was supported by Cognitive Science Research Initiative, Department of Science and Technology, India - Grant No: SR/CSRI/70/2014.

References

- Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological review*, 100(3), 432.
- Chen, M. Y., Jimura, K., White, C. N., Maddox, W. T., & Poldrack, R. A. (2015). Multiple brain networks contribute to the acquisition of bias in perceptual decision-making. *Frontiers in neuroscience*, 9.
- Diederich, A. (1997). Dynamic stochastic models for decision making with time constraints. *Journal of Mathematical Psychology*, 41, 260-274.
- Diederich, A. (2016). A multistage attention-switching model account for payoff effects on perceptual decision tasks with manipulated processing order. *Decision*, 3(2), 81-114.
- Diederich, A., & Busemeyer, J. R. (2006). Modeling the effects of payoff on response bias in a perceptual discrimination task: Bound-change, drift-rate-change, or two-stage-processing hypothesis. *Attention, Perception, Psychophysics*, 68(2), 194-207.
- Dunovan, K. E., Tremel, J. J., & Wheeler, M. E. (2014). Prior probability and feature predictability interactively bias perceptual decisions. *Neuropsychologia*, 61, 210-221.
- Feng, S., Holmes, P., Rorie, A., & Newsome, W. T. (2009). Can monkeys choose optimally when faced with noisy stimuli and unequal rewards?. *PLoS Computational Biology*, 5(2), e1000284.
- Gao, J., Tortell, R., & McClelland, J. L. (2011). Dynamic integration of reward and stimulus information in perceptual decision-making. *PloS one*, 6(3), e16749.
- Gold, J. I., & Shadlen, M. N. (2002). Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*, 36(2), 299-308.
- Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature reviews neuroscience*, 9(6), 467-479.
- Leite, F. P., & Ratcliff, R. (2011). What cognitive processes drive response biases? A diffusion model analysis. *Judgment and Decision Making*, 6(7), 651.
- Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current opinion in neurobiology*, 22(6), 1027-1038.
- Michmizos, K. P., & Krebs, H. I. (2014). Reaction time in ankle movements: a diffusion model analysis. *Experimental brain research*, 232(11), 3475-3488.
- Miyapuram, K. P., & Pammi, V. S. (2013). Understanding decision neuroscience: A multidisciplinary perspective and neural substrates. *Prog Brain Res*, 202, 239-266.
- Mulder, M. J., Wagenmakers, E. J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: a diffusion model analysis of prior probability and potential payoff. *Journal of Neuroscience*, 32(7), 2335-2343.
- Philiastides, M. G., Ratcliff, R., & Sajda, P. (2006). Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. *Journal of Neuroscience*, 26(35), 8965-8975.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59-108.
- Ratcliff, R. (2006). Modeling response signal and response time data. *Cognitive psychology*, 53(3), 195-237.
- Ratcliff, R., & Childers, R. (2015). Individual differences and fitting methods for the two-choice diffusion model of decision making. *Decision*, 2(4), 237.
- Ratcliff, R., & Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *Journal of Experimental Psychology: Human perception and performance*, 26(1), 127.
- Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive psychology*, 60(3), 127-157.
- Ratcliff, R., Thapar, A., & McKoon, G. (2011). Effects of aging and IQ on item and associative memory. *Journal of Experimental Psychology: General*, 140(3), 464.
- Rorie, A. E., Gao, J., McClelland, J. L., & Newsome, W. T. (2010). Integration of sensory and reward information during perceptual decision-making in lateral intraparietal cortex (LIP) of the macaque monkey. *PloS one*, 5(2), e9308.
- Sanfey, A. G., Loewenstein, G., McClure, S. M., & Cohen, J. D. (2006). Neuroeconomics: cross-currents in research on decision-making. *Trends in cognitive sciences*, 10(3), 108-116.
- Schall, J. D. (2001). Neural basis of deciding, choosing and acting. *Nature Reviews Neuroscience*, 2(1), 33-42.
- Schmiedek, F., Oberauer, K., Wilhelm, O., Suñer, H. M., Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, 136(3), 414.
- Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of neurophysiology*, 86(4), 1916-1936.
- Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in neurosciences*, 27(3), 161-168.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583-639.
- Summerfield, C., & De Lange, F. P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. *Nature Reviews Neuroscience*, 15(11), 745-756.
- Summerfield, C., & Tsetsos, K. (2012). Building bridges between perceptual and economic decision-making: neural and computational mechanisms. *Frontiers in neuroscience*, 6.
- van Ravenzwaaij, D., Mulder, M. J., Tuerlinckx, F., & Wagenmakers, E. J. (2012). Do the dynamics of prior information depend on task context? An analysis of optimal performance and an empirical test. *Frontiers in Psychology*, 3.
- Voss, A., Rothermund, K., & Brandstaetter, J. (2008). Interpreting ambiguous stimuli: Separating perceptual and judgmental biases. *Journal of Experimental Social Psychology*, 44(4), 1048-1056.
- Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, 32(7), 1206-1220.
- Voss, A., Voss, J., & Klauer, K. C. (2010). Separating response-execution bias from decision bias: Arguments for an additional parameter in Ratcliff's diffusion model. *British journal of mathematical and statistical psychology*, 63(3), 539-555.
- Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the drift-diffusion model in Python. *Frontiers in neuroinformatics*, 7.
- Zhang, J., & Rowe, J. B. (2014). Dissociable mechanisms of speed-accuracy tradeoff during visual perceptual learning are revealed by a hierarchical drift-diffusion model. *Frontiers in neuroscience*, 8.

Learning Basic Python Concepts Via Self-Explanation: A Preliminary Python ACT-R Model

Veronica Chiarelli (veronicachiarelli@cmail.carleton.ca)

Department of Cognitive Science
Carleton University
Ottawa, Canada K1S 5B6

Abstract

This paper presents a cognitive modelling approach to investigating student learning of computer programming concepts via self-explanation. Self-explanation involves explaining instructional material to oneself by generating inferences about the material. Here, we present a preliminary Python ACT-R model of novice and experienced students studying basic Python concepts and self-explaining. Our contributions include formalizing the self-explanation process and providing a framework that can be expanded to explore and simulate more aspects of this type of student study and learning in the domain of programming.

Keywords: self-explanation; programming education; Python ACT-R; cognitive modelling

Introduction

Learning to program is difficult (Du Boulay, 1986; Duran, 2020; Robins, 2019) with high drop out and failure rates in computer science classes relative to other courses. Research has shown that novices lack in-depth knowledge of computer science concepts and instead tend to approach programming activities in a superficial way (Robins, 2019). For instance, when aiming to comprehend programs, students tend to paraphrase the code in front of them rather than provide higher-level explanations of the program's function (Biggs & Collis, 1982).

Given these consistent difficulties, some research in computer science education has explored ways to help students learn more effectively. Some techniques are specific to topics of computer science while others draw inspiration from educational tools or approaches used in other domains. One general technique is self-explanation.

Self-explanation is the process of generating explanations of instructional material to oneself (Chi et al., 1989). To illustrate, self-explanation can involve making inferences on the domain concepts needed to generate worked-out example solutions and/or by making connections between various concepts. Self-explanation is highly beneficial for learning, (Chi et al., 1989; Chi et al., 1994; Renkl, 1999). While not all students spontaneously self-explain, Chi et al. (1994) found that explanations can be elicited by simply prompting students to self-explain instructional text. This result has since been replicated in other studies (Conati & VanLehn, 2000).

To date, the utility of self-explanation has been mainly investigated in domains other than programming and so work

is needed to identify the mechanisms of self-explanation for this domain. This paper presents a preliminary model of how students learn through self-explanation of a short instructional text about the programming language Python. Before we present the model and results, we provide some background research to contextualize our findings.

Background

Chi et al. (1989) identified the self-explanation effect through their seminal study examining how students learn from instructional materials. The goal was to identify why some students learned more than others from studying activities. The student participants read a physics textbook and then studied examples and worked on problems. While they studied examples, they verbalized their thoughts, providing access to their reasoning. Student utterances were analyzed using qualitative methods and this analysis revealed that some utterances contained self-explanations while others corresponded to mere paraphrases of the instructional materials. The results showed that students who learned more produced more self-explanations than paraphrases and that their self-explanations expanded on the material and/or linked concepts or examples. The self-explanation effect has been replicated through studies in diverse domains like biology (Chi et al., 1989), math (Renkl, 1997) and programming (Recker & Pirolli, 1995). The focus on experimental work, and particularly studies aimed at characterizing the self-explanation phenomena and/or interventions to encourage this activity has meant that there is less work computationally modeling self-explanation. There are, however, notable exceptions that we now describe.

Cascade is a Prolog-based computational model of how students solve physics problems in the presence of examples (VanLehn, Jones, & Chi, 1992). Cascade can both study examples and solve problems. When studying, it "reads" an example and attempts to self-explain each solution step by deriving it using existing facts in its memory. If no appropriate rule can be found, Cascade self-explains the example using common sense and reasoning to derive a new rule. Good students modelled by Cascade use different strategies while studying examples than poor students do. Namely, good students self-explain examples that they are studying while poor students simply accept that the examples are correct without actively processing why that is the case. Running the model demonstrates the self-explanation effect.

That is, the results show that the number of simulated self-explanations is positively correlated with the number of correct problems solved by the model. This makes sense since self-explanation increases the likelihood that a student will encounter new rules or uncover impasses or gaps in knowledge that act as learning opportunities.

Jones and Fleischman (2002) investigated student learning about probability via faded examples (incomplete examples that require the students to fill in the missing material). The hypothesis was that more learning will take place if a student is challenged to complete faded examples as opposed to studying already fully worked out examples that can more easily be passively accepted as accurate. To test this hypothesis, they added a new knowledge base to Cascade to support computing probabilities. To evaluate the extended model, they conducted simulations of the model, and also compared the model actions related to learning probabilities from faded examples with a student learning from the same examples. The simulations revealed that faded examples resulted in more student self-explanations than completely worked-out examples and that faded examples exposed more impasses (knowledge gaps), thereby uncovering more learning opportunities. However, not all student learning was accurately modeled. For example, some students would learn the correct application of a rule over the course of several examples. Cascade was unable to capture that gradual progression.

Other work has focused on modelling problem solving without including example studying or self-explanation behaviours. Braithwaite and Pyke (2017) created a computational model of learning fraction arithmetic and compared it to student learning data. The model simulated problem-solving via reinforcement learning of rules, initially including various correct and incorrect rules and then increasing activation of selected rules as it progressed through example study and problem solving. The model was implemented to reinforce strategies that lead to correct problem-solving actions when applied correctly by increasing the strategies' activation. Because they were reinforced, those strategies became simultaneously more likely to be correctly selected to solve a problem and more likely to be incorrectly selected to solve a problem for which they were not appropriate. This reinforcement of strategies was implemented because the model assumed that the majority of student errors come from overgeneralizations of fraction arithmetic rules. Test runs revealed that the textbook-trained model accurately reproduces student performance data and that model simulations trained on unbiased distributions of examples and problems performed better on problems that are underrepresented in popular textbooks. The model accurately simulated student difficulties while learning and provided some evidence in support of the assumption that unrelated statistical properties also have an impact on student learning.

So far, we have described computational models that aim to simulate human problem solving and/or example studying. Other computational frameworks produce interventions to

enhance learning. For example, Conati and VanLehn (2000) extended Andes, a tutoring system for the domain of Newtonian physics, to support example studying by encouraging self-explanation through prompts and feedback. The framework doesn't simulate self-explanation but rather assesses its presence or absence using a model of the student. For example, if a student using this system spends enough time viewing an example, they will not be prompted to self-explain because the model will deem that sufficient effort was spent on the example to result in learning. Results from an evaluation of the system showed that the tutor was a beneficial tool, specifically for increasing student learning in the early stages of example study.

Recker and Pirolli (1995) also created a type of computer tutor, in this case for programming education. They investigated how students learned programming skills through self-explanation using an embedded hyper-text non-linear environment to present the instructional materials and elaborations on the text versus a typical linear instructional text. They found that high ability students (labelled as high ability based on post-test scores) benefitted from the hyper-text environment, suggesting that the self-directed learning skills of those students enabled them to use the embedded elaborations to their advantage. Conversely, low ability students did not benefit from the experimental environment, suggesting that it may have increased cognitive load for these students. Analysis of self-explanations demonstrated that good students most frequently made comments about the domain (showing that their focus was on the content) while poor students most frequently made comments related to navigation (showing that they were more focused on interface features than on the lesson content).

The challenge of understanding how students learn and how to enhance student learning is one that has been approached in various ways. We contribute to this effort and introduce a preliminary model of learning via self-explanation in the domain of programming education.

Python Self-Explanation Model

The goal of the Python self-explanation model is to simulate student learning through self-explanation in the domain of learning to program. It simulates the process of novice or experienced students self-explaining a short instructional text about Python. In its current state, the model randomly assigns a pre-test score, then self-explains each line of text by retrieving existing knowledge from memory and producing that knowledge as a self-explanation and then finally it calculates the post-test score based on the self-explanations produced. For this preliminary model, learning is simulated by an increase from pre-to post-test scores.

Theoretical Foundation

Muldner, Bursleson and Chi (2014) investigated how self-explanation helps students learn about emergent phenomena in a study where students were prompted to self-explain texts about diffusion. To investigate the impact of different kinds of self-explanations on learning, each utterance was labelled

as either a macro-level explanation, a micro-level explanation, an inter-level explanation, a paraphrase, or other. The results showed that some explanations were more strongly related to student learning than others. We extend this framework to the domain of programming, modifying the characterization of the explanations to make them suitable for the programming domain.

In the present work, micro-, macro- and inter-level self-explanations are defined for the domain of programming as follows. Micro-level self-explanations correspond to utterances about the directly visible elements in a program (such as programming syntax) or in the instructional text. Macro-level self-explanations are inferences about high-level programming concepts, such as the idea behind a given code segment. Inter-level self-explanations act as a bridge between micro-level and macro-level concepts and, as such, they explain the connection between the directly visible elements in the instructional text and their use or purpose.

For example, the line of instructional text *“It is necessary to update the condition in order to eventually break out of the loop.”* could be explained in the following ways. A micro-level self-explanation could be *“This works using conditions that can change.”* This is micro-level because it focuses uniquely on the words *condition* and *update* which are found in the line of text. A macro-level self-explanation of the same line could be *“Just like if there is no stop sign, people will keep driving.”* This self-explanation shows an understanding of the purpose and function of a while loop. An inter-level self-explanation connects micro- and macro-level ideas. *“Just like you will wash one dish at a time until there are no more dirty dishes, this will repeat until some condition happens”* is an example of an inter-level self-explanation for that same line of text. In this model, paraphrases are just restatements of the instructional text.

As in Muldner, Burleson and Chi (2014), the present model defines learning as the increase from pre- to post-test scores. Based on the learning outcomes of Muldner, Burleson and Chi (2014), in the present work, the most learning occurs with inter-level self-explanations, followed by micro-level, then macro-level, and finally the least learning occurs with paraphrases. Since the model here has not yet been developed to acquire new rules and knowledge, it uses the experimental findings about the relationship between levels of self-explanations and learning (from other domains) to calculate a post-test score, as we will describe in more detail shortly.

Model Framework, Environment and Components

Modeling Framework. ACT-R is a well-known theory of cognition which includes theories of declarative memory, procedural memory, and a chunk and buffer system (Anderson, 2007). The original computer architecture of ACT-R was implemented using Lisp. That implementation restricts modelling to directly reflect the theory, so implementing some features can implicitly have side effects on other parts of the model. To allow for flexibility, Stewart and West (2007) created Python ACT-R. This framework has the three main components from ACT-R (a chunk-based

communication system, a chunk storage system, and a pattern matching production system), but is implemented in Python. The Python code is based on the theory itself rather than being a direct translation from the original Lisp. Stewart and West (2007) thus demonstrated that the theory is separable from the code. Also, the simple module creation in of ACT-R and the ability to manually adjust more components makes it more flexible than Lisp ACT-R and promotes more extensive exploration of ACT-R theory, claims, and components. For these reasons, we used Python ACT-R as the basis for the present model. The model includes an environment and modules, described below.

Model Environment. The main component of the model environment is the instructional text. For the present work, the text describes the syntax and the concepts of “if statements” and “while loops” in Python. There are 13 lines of text in total. Each line is stored as an element in the instructional text environment and has an associated state. The lines are initially in a state of “read” to indicate that the line has yet to be read (and subsequently self-explained by the model). Each line of the text is also labelled as belonging to one of the three levels of knowledge (macro-, micro- or inter-level).

Other information stored in the environment includes a count of the number of macro-level, micro-level, and inter-level self-explanations as well as a count of paraphrases. All of these are initially set to 0 since no self-explanations have been produced before the model runs. During model execution, the counts are updated as each line of the text has been self-explained to keep track of the type of self-explanation produced. The environment also includes the pre-test score and the post-test score used to quantify learning. The calculation of these scores happens in the self-explain module (described below). Finally, the environment includes the experience level of the student being simulated by the model, either “novice” or “experienced”. This experience level influences the model execution, reflecting that novices self-explain differently and have more to learn than experienced students.

Model Modules. A key benefit of self-explanation is the integration of new and existing knowledge, meaning that students make connections between the text and what they already know. In order for the model to simulate existing knowledge, the Python ACT-R declarative memory module is initialized to model a student’s prior knowledge. Specifically, the model’s declarative memory is initialized with chunks of domain information. Each chunk specifies the knowledge itself (a piece of existing knowledge), the level of that knowledge (macro-level, micro-level, inter-level), and the topic of the knowledge (a label indicating the topic of the existing knowledge). When the model runs, the declarative memory buffer is used to retrieve existing knowledge chunks. Some noise is added to the declarative memory to account for the fact that which chunk of knowledge is retrieved is not always predictable.

When the model self-explains a line of text, it attempts to retrieve a relevant chunk of existing knowledge from

declarative memory. The relevance of a chunk is influenced by its topic and its knowledge level. A chunk is most relevant to a line of text if they have the same topic and if their knowledge level (inter-, micro-, macro-level) is similar. Knowledge levels were defined as partially similar to one another as follows. Inter-level was set to be partially similar to both macro-level and micro-level while micro-level and macro-level are set to be dissimilar. This choice was made based on the assumption that an inter-level explanation for a topic is always suitable since it serves as a bridge between the other concept levels. Meanwhile, micro-level and macro-level concepts are quite different and, therefore, it is less likely that a student would choose to produce a micro-level self-explanation for a macro-level line of text, for example.

The model also includes a self-explanation module that contains a production to change the state of a line of text from “read” to “self-explain” (to indicate that the line has been read and self-explained), a production to update the count of the different levels of self-explanations produced, and productions for updating the pre-test and post-test scores. Since this is a preliminary model, it does not yet simulate the process of taking the pre-test and the post-test. Instead, the pre-test score is randomly determined from a range of values depending on the student experience level. Novice student pre-test scores arbitrarily range from 30% to 40% based on the assumption that novices will not have the knowledge required to pass a programming pre-test. Experienced students are assumed to only have minimally more experience than novice students and, as such, their randomly chosen pre-test score will fall within the range of 45%-60%. Given that previous research has demonstrated that inter-level explanations are associated with the most learning, micro-level with slightly less learning and macro-level and paraphrases with the least learning (Muldner, Burleson & Chi, 2014), the following equation was used to determine a post-test score:

$$\text{Post-test score} = S + \left(\frac{1}{13}\right) \cdot \left(I + \left(\frac{3}{4}\right) \cdot M + \left(\frac{1}{2}\right) \cdot A + \left(\frac{1}{3}\right) \cdot P\right) \cdot (100-S)$$

where S is the pre-test score, I is the number of inter-level self-explanations produced, M is the number of micro-level self-explanations produced, A is the number of macro-level self-explanations produced, and P is the number of paraphrases produced. With this formula, all simulated students will have post-test scores higher than their pre-test scores, which makes sense since it is assumed that students will not lose any programming knowledge by self-explaining the instructional text. Further, in order to match previous findings in other domains that show the most learning is associated with inter-level explanations, a perfect score in this model is possible if all self-explanations are inter-level. (Note this is based on an assumption that the findings in other domains hold in this domain which still needs to be verified experimentally.) All other possible scores are a function of the number of each level of self-explanation produced weighted by the relative amount of learning assumed to be associated with the given level of self-explanation.

Model Execution

A run of the preliminary model begins by manually setting the model parameter for experience level as either novice or experienced. While the same Python ACT-R parameters are used to produce self-explanations when simulating either type of student, the student type influences the levels of self-explanations produced. In the programming domain, novices have been shown to be more likely to provide micro-level self-explanations than any other level of self-explanation (Robins, 2019). So, if simulating a novice student, the model rehearses micro-level knowledge in memory, thereby making it more salient in memory and strengthening its activation. In other words, novice student simulations are more likely to retrieve micro-level knowledge when self-explaining. If the experience parameter indicates previous programming experience, then the model will rehearse inter-level knowledge since, unlike novices, experienced learners are known to have more complete schemas and can therefore connect different levels of ideas (Robins 2019). This is why inter-level knowledge is more likely to be retrieved during a simulation of an experienced student’s self-explanations. Like all other levels of knowledge, macro-level knowledge still is added to declarative memory for every type of student, it is just not rehearsed and therefore is less salient and less likely to be retrieved. This is because Muldner, Burleson and Chi (2014) reported that macro-level self-explanations are least frequently produced by all students. Next, the model fires the pre-test production to randomly assign a pre-test score to the student, influenced only by the experience level.

The model then simulates self-explaining of the text. Specifically, it reads a line of instructional text and connects that text to existing knowledge. This is achieved by retrieving chunks of knowledge from declarative memory related to the topic and knowledge level of the text line (recall that all chunks in declarative memory are labelled with the topic and level). Figure 1 demonstrates a summary of the process of self-explaining a line of the instructional text. The model reads the line “*The syntax of an if statement is: if [condition]: [do something] else: [do something else]*”, retrieves a micro-level chunk from memory, and produces the chunk’s knowledge as a self-explanation “*So we have to write down the words ‘if’ and ‘else’*”.

```
-> Reading the line of text:
-line2 -> The syntax of an if statement is:
  if [condition]:
    [do something]
  else:
    [do something else]
Thinking of a self-explanation ...
[retrieve chunk from declarative memory]
[micro-level chunk retrieved]
[use knowledge of that chunk as a self-explanation]
self-explanation:
-> So we have to write down the words 'if' and 'else'
Finished self-explaining that line.
```

Figure 1: Self-explaining a line of text.

As mentioned, knowledge levels have been assigned to lines of instructional text and to chunks in declarative memory. However, given the a priori specified similarity

between knowledge levels and the effect of noise in declarative memory, the retrieved knowledge may not correspond to the knowledge level or topic of the line of instructional text. This simulates the fact that students may not always be able to retrieve related knowledge when they want to self-explain. For this model, any topic-relevant knowledge that is retrieved will lead to successful self-explanation since, for example, a macro-level line can be explained by inter-level knowledge. However, if the retrieved knowledge does not relate to the topic of the line of text, then that knowledge cannot be used to produce a self-explanation of the line resulting in the model disregarding the retrieved knowledge and, instead, just paraphrasing the line. Similarly, a failure to retrieve knowledge of any kind for a given line leads to a paraphrase of that line. After each line is self-explained, the count of each level of self-explanation is increased accordingly.

Finally, when the model has self-explained each line of the text, it calculates a post-test score using the post-test score formula previously described. The simulation displays the pre-test score, the number of self-explanations of each level produced, and the post-test score.

Results

Sample runs of the preliminary model were used to evaluate how it performed when modelling learning via self-explanation in the domain of programming. Table 1 displays results of running the Python self-explanation model 5 times as a novice programming student, and 5 times as a more experienced programming student. The table displays the pre- and post-test percentage scores and the percentage of learning gains along with the count of each level of self-explanation produced.

The results indicate that the Python self-explanation model does accurately simulate some findings of learning via self-explanation. Micro-level self-explanations were the most common type of explanation produced by novice student simulations and inter-level self-explanations were the most

common type produced by experienced student simulations. The novice student simulation reflects experimental data showing that novice programmers focus on the line-by-line details rather than the overarching concepts or connections between the syntax and the concept (Robins, 2019). This model also accurately reflects prior findings that, with experience, more complete schemas exist connecting code to concepts thereby permitting inter-level self-explanations for experienced learners. The very low number of macro-level self-explanations as compared to inter-level or micro-level self-explanations matches the observations of Muldner, Burleson and Chi (2014) and is understandable in our simulations since the model rehearses micro-level or inter-level knowledge (depending on student experience level) but not macro-level knowledge. So, while chunks of all three levels of knowledge exist in declarative memory, macro-level chunks have not been rehearsed for the reasons stated above and are therefore less salient and less likely to be retrieved for self-explanation.

The relationship between learning gains and levels of self-explanation matches the relationship described in the Muldner, Burleson and Chi (2014) data. This is built into the model as the post-test score is a weighted function of the levels of self-explanations produced. For example, the highest learning gains for a novice come from S1, and for an experienced simulation, S10, both of whom produced more inter-level self-explanations than any other simulations with their experience level. So, inter-level self-explanations resulted in the most learning.

The results accurately indicate that there were learning gains for all simulated students. Further, novice runs of the model result in more learning than experienced runs. This seems reasonable since novice students simply have more to learn. However, since the pre-test scores are randomly selected and the post-test scores are simply calculated as a function of weighted level of self-explanations produced and the pre-test score, this result is hard coded into the model rather than being determined by simulating the pre-test and post-test in full, so these results are expected.

Table 1: Results of 10 sample runs.

Student	Pre-test	Micro	Inter	Macro	Paraphrase	Post-test	Gains
N S1	37	5	6	0	2	87	50
N S2	36	6	1	0	6	72	36
N S3	38	7	2	1	3	79	41
N S4	36	5	4	0	4	80	44
N S5	32	5	3	0	5	76	44
Average (Novice)	35.8	5.6	3.2	0.2	4	78.8	43
E S6	51	3	4	1	5	82	32
E S7	55	2	3	2	6	80	25
E S8	51	3	5	1	4	85	34
E S9	52	4	3	2	4	82	30
E S10	49	1	7	0	5	85	36
Average (Experienced)	51.6	2.6	4.4	1.2	4.8	82.8	31.4

Discussion and Future Work

As described in the background section, cognitive models can provide valuable insight into how students learn and can inform effective teaching strategies and interventions. Yet, due to the complexity and intricacies involved in learning, such as individual learner differences and differences in domains, a complete model of the learning process has yet to be created. The Python self-explanation model is a preliminary step for informing on learning via self-explanation in the domain of programming.

While the current model captures some aspects of self-explanation such as drawing on existing knowledge, it does not yet simulate the acquisition of new knowledge. If the learning process were expanded to include the ability to learn new rules through commonsense and general reasoning, as is the case with models like Cascade (VanLehn, Jones, & Chi 1992), then it would be possible to also simulate the pre- and post-test activities, as opposed to randomly producing a pre-test score and then a post-test percentage calculated using weighted counts of levels of self-explanations. Supporting new rule acquisition and subsequently modelling the pre- and post-tests could provide insight into mechanisms used by students when they apply the knowledge gained from the self-explanation exercise to a problem-solving test.

There are various other avenues for future work. One extension would be to model more types of students. Robins (2019) describes that in the field of programming education, three distinct clusters of students emerge. There are “stoppers” who withdraw from or abandon the activity quickly when they encounter difficulties, “movers” who trace code and try to navigate to a correct solution when they notice an issue, and “tinkerers” who react to problems by trying different tweaks of the code somewhat haphazardly and without code tracing. These clusters all pertain to program generation and so work is needed to determine if these clusters also emerge in activities that involve reading and explaining programs and instructional materials. If similar clusters of student types exist when self-explaining, these student types’ self-explanation patterns could be modelled in addition to modelling differences in two experience levels. Alternatively, modelling the learning patterns of good and poor students as examined in Chi (1989) or of high and low ability students as in Recker and Pirolli (1995) could be an informative next step. A more sophisticated extension could include modelling individual differences on a continuum from novice to expert rather than modelling students as falling within specific experience or type categories.

Another avenue for future work involves adding capability to model student emotion during the self-explanation process. We began work for this step. Specifically, although it was not described here, our model includes a preliminary emotion module. Currently, all simulations produce states corresponding to motivated and happy at times, but also individual runs of the model will produce either frustration or boredom while self-explaining each line of text, influenced only by the student experience level. That is, runs of the

model representing experienced students produce reports of feeling bored more often than frustrated while novice runs more frequently produce frustration. This was a first step in modelling some basic emotions guided by the assumption that novice difficulties lead more often to frustration while experienced pre-existing programming knowledge makes reading a basic instructional text a more boring exercise. The model does not yet take into account text complexity (e.g., which may commonly elicit frustration across all experience levels). Also, the emotion is not yet related to the level of self-explanation produced so, the fact that a student paraphrases because they cannot retrieve relevant knowledge, for example, does not make them any more likely to feel any frustration than if they, say, successfully self-explain a line with inter-level knowledge. This leaves a lot of room for improvement in the emotion module of future versions of the model including modelling a wider range of emotions, the connection between lines of text and emotions, the relationship between successful self-explanations and emotions, or even the intricacies of the various emotions associated with more types of students (such as the stoppers, movers, tinkerers, or good and poor students suggested previously).

Additionally, most existing data comes from self-explanation studies in other domains. Confirming that the same patterns emerge within the topic of learning to program would better inform this and future models.

Acknowledgements

Thank you to Robert West for providing guidance on Python ACT-R and to Kasia Muldner for providing feedback on this work and paper.

References

- Anderson, J. (2007). 1 Cognitive Architecture. In *How Can the Human Mind Occur in the Physical Universe?* Oxford University Press.
- Biggs, J. B., & Collis, K. F. (1982). Evaluating the Quality of Learning: The SOLO Taxonomy (Structure of the Observed Learning Outcome). *New York: Academic Press.*
- Braithwaite, D. W., Pyke, A. A., & Siegler, R. S. (2017). A computational model of fraction arithmetic. *Psychological Review, 124*(5).
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*(2).
- Chi, M. et al. (1994). Eliciting self-explanations improves understanding. *Cogn. Sci. 18* (1994).
- Conati, C., & VanLehn, K. (2000). Toward computer-based support of meta-cognitive skills: A computational framework to coach self-explanation. *The International Journal of Artificial Intelligence in Education, 11*.
- Du Boulay, B. (1986). Some difficulties of learning to program. *Journal of Educational Computing Research, 2*(1).

- Duran, R. (2020). *Cognitive Complexity of Comprehending Computer Programs*. Aalto University.
- Fleischman, E., & Jones, R. (2002). Why example fading works: A qualitative analysis using cascade.
- Muldner, K., Burleson, W., & Chi, M. T. H. (2014). Learning from self-explaining emergent phenomena. *Proceedings of International Conference of the Learning Sciences, ICLS*, 2.
- Recker, M. M., & Pirolli, P. (1995). Modeling individual differences in students' learning strategies. *Journal of the Learning Sciences*, 4(1).
- Renkl, A. (1999). Learning mathematics from worked-out examples: Analyzing and fostering self-explanations. *European Journal of Psychology of Education*, 14(4).
- Robins, A. (2019). Novice programmers and introductory programming. *The Cambridge Handbook of Computing Education Research*.
- Stewart, T., & West, R. (2007). Deconstructing and reconstructing ACT-R: Exploring the architectural space. *Cognitive Systems Research*, 8(3).
- VanLehn, K., Jones, R. M., & Chi, M. T. H. (1992). A model of the self-explanation effect. *Journal of the Learning Sciences*, 2(1).
- Hill, J. A. C. (1983). A computational model of language acquisition in the two-year old. *Cognition and Brain Theory*, 6.

Parameter Correlations in the Predictive Performance Equation: Implications and Solutions

Michael Collins^{1,2,3}

Florian Sense³

Michael Krusark⁵

Joshua Fletcher⁶

Tiffany Jastrzembski¹

¹Air Force Research Laboratory ²ORISE ³Wright State University ⁴Infinite Tactics ⁵L3Harris Technologies ⁶Ball Aerospace

Abstract

Research of mathematical models of learning and retention have focused on accounting for an individual's performance across a variety of learning schedules (i.e., spaced and massed). The attempted goal of such research is to develop a model which can adequately predict human performance across a range of learning scenarios. However, little attention of this model development has focused on the interpretation of a model's best fitting parameters given the structure of a model's equations and its predicted performance values. The effect of this can lead to the development of models where the parameter values are correlated hindering a theoretical interpretation of performance. Here we examine the structure of the Predictive Performance Equation (PPE) and highlight portions of PPE's equations that lead to correlations across its free parameters. We propose a fix for these issues (Modified PPE) and conduct a formal model comparison showing the Modified PPE is simpler, has less parameter correlation and its best fitting parameters map on to identifiable aspects of an individual's performance.

Keywords: memory, learning, decay, spacing effect, mathematical modeling, model comparison, model identifiability

Introduction

Mathematical models of learning and retention are quantitative formulations of verbal psychological theories which attempt to account for and/or predict empirical data. One value of these mathematical formulations is the fact that all assumptions of a model are made explicit allowing for formal statistical evaluation. Furthermore, these mathematical models lend themselves to real-world applications, such as adaptive learning systems. Although the quantitative formulations of models have many benefits, care must be taken to ensure how these models are constructed, to ensure that a model accurately represents the assumptions of a given psychological theory.

In the domain of learning and retention, mathematical models are developed in order to represent how an individual retains knowledge based on the temporal aspects of a training schedule. Models often achieve this goal by representing three regularities of human memory: power law of learning, power law of decay, and the spacing effect. These three psychological phenomena have been represented in various mathematical models (Pavlik & Anderson, 2005; Raaijmakers, 2003; Walsh et al., 2018). The Predictive Performance Equation (PPE) is one particular mathematical model that has been found to account for a range of learning phenomena compared to other spacing effect models (Walsh, et al., 2018) and has been used to inform training applications. Each of these accomplishments was part of the explicit

purpose of PPE's development, being used as a prescriptive educational tool.

However, despite the PPE's successful applications, its current formulation limits the estimation of psychological meaningful parameter estimates due to correlation across parameters. These limitations arise not because of the underlying psychological theory PPE represents or doubt of the empirical validity of the spacing phenomena, but because of PPE's chosen mathematical representation. In this paper, I review the current formulation of the PPE, address its limitations, and offer an alternative formulation of how they might be overcome.

Predictive Performance Equation

The PPE is composed of six individual equations, containing 4 free parameters. At the center of the PPE is the Activation term M_i (Eq. 1), which is a product of the learning term (N^c) and the forgetting term (T^{-d}). The learning term is on the unit of trial exposures (N) raised to a constant learning rate (c , usually .1). While the decay term is on the scale of model time (T), raised to a decay rate (d). From Eq. 1 it can be seen that M_i is on the scale of number of exposures and model time (T).

$$M_i = N_i^c * T_i^{-d} \quad (\text{Eq. 1})$$

A novel aspect of PPE is that model time (T) is modeled as a weighted average (Eq. 2) of time since all previous presentations of an item (Eq. 3). Thus, model time (T) is on the scale of the weighted average of wall clock time (often seconds).

$$T_i = \sum_{j=1}^{n-1} w_j * t_j \quad (\text{Eq. 2})$$

$$w_j = t_j^{-x} / \sum_{j=1}^{n-1} t_j^{-x} \quad (\text{Eq. 3})$$

Additionally, PPE's decay rate (Eq. 5) dynamically changes over time, based on two free parameters, b and m , and the stability term. The stability term (Eq. 4) is a representation of the average natural logarithm of the lag_i of an item's history. Due to the fact that the natural log of the lag is taken, PPE's decay parameter is an unitless metric.

$$St_i = \left(\frac{1}{n-1} * \sum_{j=1}^{n-1} \frac{1}{\ln(lag_j + e)} \right) \quad (\text{Eq. 4})$$

$$d_i = b + m * St_i \quad (\text{Eq. 5})$$

Finally, to generate a prediction of performance, PPE's activation term (M_i) is nested within a logistic function (Eq. 6), which is controlled by two additional free parameters, τ and s , controlling the slope and intercept of the performance value. This formulation of activation value has been used in other learning contexts (Anderson, 2007)

$$P_i = \frac{1}{1 + \exp\left(\frac{\tau - M_i}{s}\right)} \quad (\text{Eq. 6})$$

Sources of correlation

As discussed in the previous section, PPE is composed of 6 equations with 4 free parameters. Equations 1 through 5 make up the PPE terms and Eq. 6 maps an unbounded activation term (M_i) onto a performance value. The free parameters are split such that they affect the PPE's decay (Eq. 5) term and the properties of the logistic function (Eq. 6). An unintended effect of this mathematical formulation is a high correlation between PPE's free parameters and an inability to compare best fitting parameters across individuals for psychometric evaluation (e.g., high versus low decay rates). The inability to compare parameters across participants is due to the fact that, since parameters correlate with each other these correlations must be taken into account before any parameter comparisons can be made across participants. Specifically, within the PPE this issue arises from two sources, (1) the PPE contains unbalanced units (i.e., M_i) and (2) the M_i term is nested within a logistic function. Each of these features have been shown in other psychological models to produce parameter correlation and issues with identifiability (Krefeld-Schwalb, Pachur, & Scheibehene, *in press*). Here, we address the origin of both these sources in the PPE and propose an alternative formulation to remedy these correlation issues.

Unbalanced Units Unbalanced units refers to instances when particular terms within an equation are combined together without the units of those terms canceling out. For example, in PPE this occurs when computing the activation term M_i (Eq. 1) when the learning term is multiplied by the decay term. PPE's learning term is on the scale of instances of exposure (N_i), while PPE's forgetting term is on the scale of model time (T_i). Combining these two terms together, leads to an activation term M_i that is on the scale of number of events and model time, which results in highly correlated parameters, due to the fact that the free parameter (d_i) within each term are dependent on that term's scale. It is this correlation of parameters that hinders PPE's parameters being able to meaningfully represent individual differences within a sample due to the fact that any parameter estimate is dependent on that term's scale. The limitations of this formulation is not unique to PPE but has been found in other psychological models. Readers interested in a more thorough explanation should see Vincent and Steward (2020) and Stewart, Scheibehene, and Pachur (2018).

Nested Equations A second source of intercorrelation within the PPE is the activation value (M_i) nested within a logistic function (Eq. 6), which is manipulated by its own free

parameters (i.e., τ and s). The nesting structure creates three difficulties with model interpretation. First, nesting the activation term within the logistic equation allows for different M_i values to have equivalent performance values. Consequently, two people with identical learning and decay terms could be predicted to exhibit different performance in the future, which suggests that PPE's parameters (especially b and m) are difficult to interpret at face value (i.e., without also knowing the values of τ and s). Second, within the logistic function, τ and s do not have independent effects on the activation term, allowing for multiple combinations of τ and s to create equivalent performance values. Third, having free parameters outside of the learning and forgetting terms obscures the interpretation of the PPE. Again, this issue is not unique to the PPE but has been noted as an issue with other psychological models Krefeld-Schwalb, Pachur, and Scheibehene (in press).

Modified Predictive Performance Equation

As discussed above, PPE consists of unbalanced units and a nested equation, each of which lead to correlation across parameters. Both issues can be fixed by making relatively minor modifications to PPE's structure keeping the remaining assumptions of human performance intact. In the rest of the document We will refer to this new equation as the Modified PPE. In this section a proposed set of modifications to the PPE to reduce the correlation of parameters and improve the parsimony of the PPE are evaluated.

One cause of PPE's intercorrelations across the model's parameters is due to the M_i term being nested within a logistic equation (Eq. 5). This nesting step is required due to the fact that M_i is not bound between 0-1. However, this formulation is not required if both the learning and decay term are bound between 0-1, which would allow the learning and decay term to be combined together to estimate performance on a 0 - 1 scale (Eq. 7). To achieve this formulation, slight modifications are made to both the learning and decay term which are outlined here.

$$Performance_i = LearningTerm_i * DecayTerm_i \quad (\text{Eq. 7})$$

Learning term In the standard PPE the learning term is produced as a power law. However, the reformulated learning term is exponential. Though there is a debate over the form of learning or forgetting term, the exponential formulation has been shown to better account for learning at an individual level performance over a power law (Heathcote, Brown, & Mewhort 2000). In this formulation the Modified PPE learning term (Eq. 8) has a learning rate m , which controls the rate at which material is acquired. When m is low information is acquired quickly, while when m increases the rate that individuals acquire information decreases. The benefit of this modified learning term is twofold. First, compared to the previous learning term (Eq. 1), it is now on a scale of 0-1. Second, due to the fact that N_i and m are used as exponents, which are unitless, the learning term can now be combined with the forgetting term .

$$\text{Learning Term}_i = 1 - (e^{-N*(1-m)}) \text{ (Eq. 8)}$$

Decay Parameter One novel component of the PPE is that its decay term dynamically changes over time based on the temporal spacing of practice. This is a result of the assumption that spacing is a result of attention moderating the spacing effect (Walsh et al. 2018). The Modified PPE retains the same assumption of the use of the stability term (Eq.4), which accounts for an item's previous temporal history of presentations. The addition of the lag term $\frac{1}{\log(\text{lag}_i)}$, which represents the most recent lag between exposures was added to PPE's decay parameter due to the fact that M_i is no longer nested in the logistics term and augmented by the τ and s parameter. The lag term is modified by the same learning rate (m) parameter as used in the learning term (Eq.8). When learning rate is low, the effect of the most recent lag is minimized, while when it is high the effect of the most recent lag is maximized. The stability term is manipulated according to the b parameter, which controls the effects that the previous temporal schedule (i.e., spaced practice) has on performance. This subtraction of the maximum value is used to format the decay term serving as a decay intercept based on the largest decay within a set of practice items (Eq. 9). The benefit of this decay term is that it is now composed of two terms (lag_i and S_t) that each represent separate aspects of the performance and each manipulated by their own free parameter.

$$d = (\frac{1}{\log(\text{lag}_i)} * m + S_t * b) - \max(\frac{1}{\log(\text{lag}_i)} * m + S_t * b) \text{ (Eq. 9)}$$

Forgetting Term For the decay term, the standard power law formulation was retained from the PPE. However, the power law is expressed as a ratio which allows the forgetting term to be expressed as a unitless metric between 0-1 (Eq. 10).

$$\text{Forgetting Term} = \frac{T_i^{-d_i}}{1+T_i^{-d_i}} \text{ (Eq.10)}$$

Summary of Changes to the PPE

Here we reviewed the PPE, a model of learning and retention which has shown great promise in accounting for both laboratory and real world findings. However, features within the PPE lead to correlation across parameters and hinder it from being used to estimate psychological constructs (i.e., learning and decay rates). To correct these limitations, we have proposed a new formulation of the PPE, decreasing the number of free parameters while retaining PPE's unique features: multiplicative performance, model time, and variable decay term.

Method

To highlight a comparison between the standard and modified PPE, a model comparison was conducted,

highlighting the correlation across free parameters and the benefits of the PPE' formulation.

Participants Sixty-one participants were recruited from a midwestern university in this paired-associate learning study. All participants completed a total of three experimental sessions spanning a three-week period.

Task Stimuli Over the course of the experiment participants memorized a set of 30 Japanese-English words. All of the words used in this study were taken from the Medical Research Council (MRC) Psycholinguistic Database manual and have been used in other previous memory studies (e.g., Pavlik & Anderson, 2005).

Experimental Design and Procedure During the experiment, an item's training schedule was manipulated according to inter-session interval (ISI) and inter-trial interval (ITI) over the course of experimental sessions. The ISI controlled the amount of time between the 1st and 2nd experimental session, with a fixed 7 day ISI between the 2nd and 3rd session across all conditions. The ISIs in this study were fixed at short (5 min), medium (7 days), and long (14 days) delay. The ITI manipulated the number of trials between presentations of the same item within a session. Two ITIs consisting of a short (every 2 trials) and long (every 11 trials) delay were embedded in each experimental session.

During the study, participants, with no knowledge of the Japanese language, were given instructions for the paired associate learning task and had an opportunity to ask any questions. Once participants began the experiment, they were shown a Japanese word (e.g., "kanboku") on the screen and asked to type the English translation (e.g., "bush") to the Japanese word. Upon first presentation of a word, participants were shown the English translation and asked to type the correct answer to ensure the item was studied. During all subsequent presentations, participants were asked to recall and type the English translation from memory. Participants were given a maximum of 7 seconds to type their answer during each trial. If a participant could not generate a response within 7 seconds, then their answer was considered incorrect. At the end of each trial participants were given feedback (correct or incorrect) and given 2 seconds to study the correct answer.

Bayesian Models

To examine the two implementations of the PPE, Bayesian hierarchical models of both the Standard and Modified PPE were implemented in JAGS (Plummer, 2012). Each model was run with 3 MCMC chains, run for 9000 iterations, with a fixed burn in period of 1000 iterations. Each models priors were chosen so that the prior predictions from each model expressed the standard learning phenomena expected from the learning schedule (i.e., slower learning in the long vs short ITI condition with more decay between the sessions in the short compared to the long ITI). Each model was fit to each of the Japanese-English word pairs across the three experimental sessions.

Results

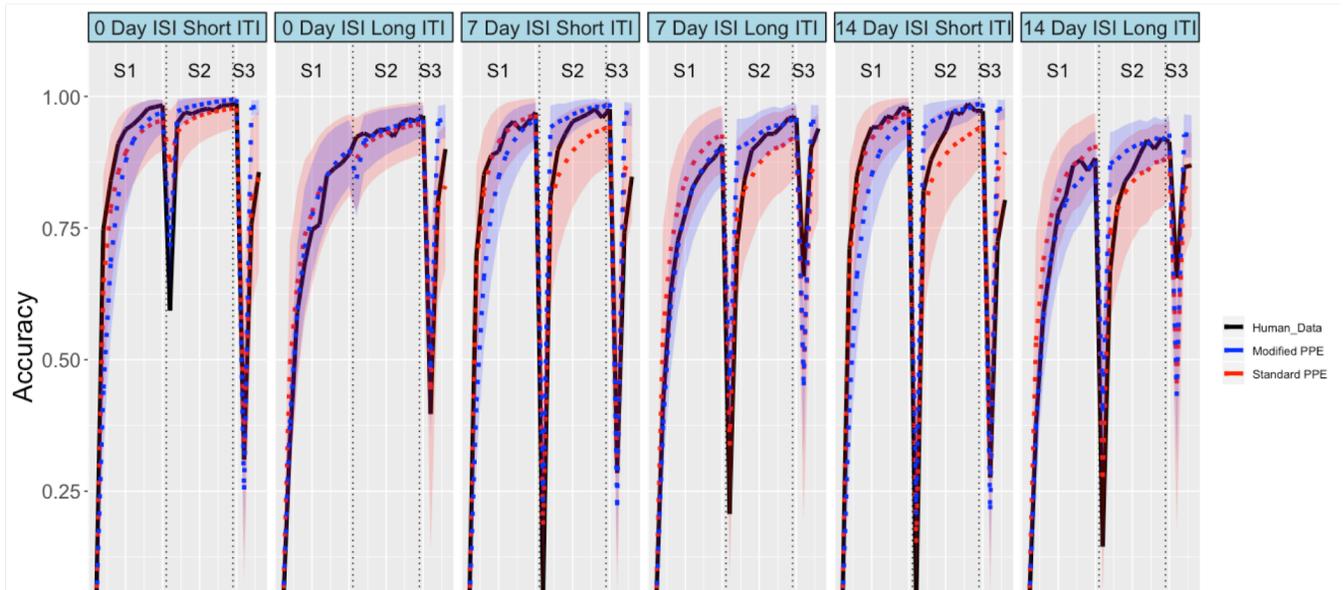


Figure 1. The average +/- 95% HDI posterior model fit of the Standard (dashed red line and ribbon) and Modified (dashed blue line and ribbon) PPE to the participants' performance (sold black line) six experimental conditions varying the three inter session interval (0, 7, and 14 days) and the two inter trial intervals (short and long).

To compare the Standard and Modified PPE equation, a comparison across three different metrics was performed. First, we examined how well each model fit to the performance of subjects across each of the learning schedules. Second, we compared the correlation between each model's free parameters. Third, the relationship between the participants' parameter estimates and learning retention were examined.

Model Fit

We first examined the average performance of participants and each model's posterior performance estimates across the six different learning schedules. An examination of the models' average fit to the participants performance reveals several interesting qualitative findings (Figure 1). First, the Modified PPE for the most part has much narrower 95% HDI compared to the Standard PPE. This difference in precision between the two models is the result of the difference in complexity. The Standard PPE has 4 free parameters, with

the activation term (M_i) being nested within a logistic equation, which gives the model additional flexibility. An example of this additional flexibility can be seen in the relearning between in 2nd and 3rd experimental sessions: the Standard PPE shows quick but attenuated relearning across sessions, while the Modified PPE shows quick relearning between sessions.

To evaluate the fit of both models to the participants' performance across the three experimental sessions, the correlation (r) and root mean squared deviation ($RMSD$) between the average accuracy and each model's posterior performance were calculated (Table 1). Both models fit the average performance of participants across all of the experimental conditions quite well, with the Standard PPE having a slightly higher correlation and lower $RMSD$ compared to the Modified PPE. However, a Bayes factor found the the Modified PPE to be strongly preferred to the Standard PPE ($BF > 30$) These results suggest strong evidence in favor of the Modified over the Standard PPE, suggesting that the Modified PPE is a more parsimonious model compared to the Standard PPE in this context.

Table 1. The correlation (r) and root mean squared deviation ($RMSD$) between the standard and modified PPE across each of the six learning schedules.

ISI	ITI	Standard PPE		Modified PPE	
		r	$RMSD$	r	$RMSD$
0	Short	0.96	0.07	0.93	0.10
0	Long	0.99	0.04	0.98	0.05
7	Short	0.99	0.05	0.93	0.11
7	Long	0.96	0.07	0.95	0.08
14	Short	0.99	0.05	0.92	0.11
14	Long	0.96	0.07	0.93	0.09

Parameter Intercorrelation

Next, we evaluated the intercorrelations between each of the models' free parameters (Figure 2). A correlation between two parameters reveals a functional interdependence, which hinders theoretical interpretations of the parameters.

Standard PPE To evaluate the comparison between the Standard PPE's free parameters, the correlation between all free parameters (b, m, s, τ) were calculated from the models

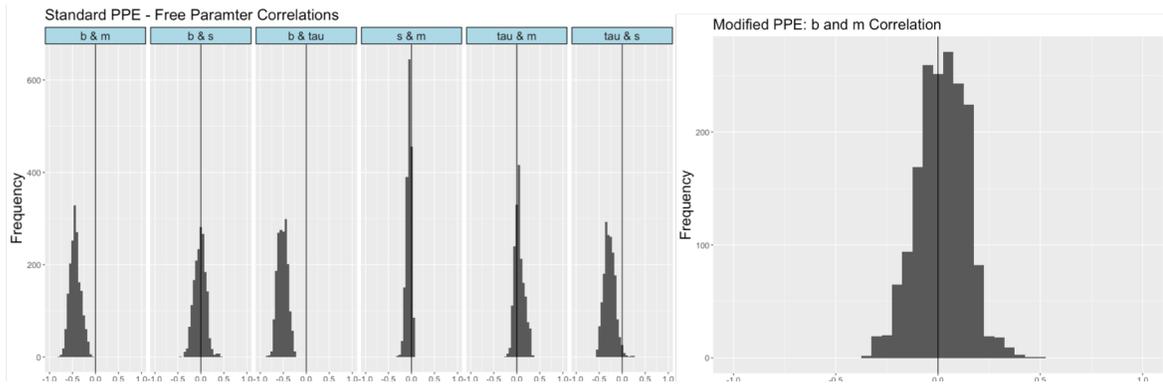


Figure 2. A histogram of each the free parameters of the Standard (left panel) and the Modified PPE (right panel) correlation with each other for each item presented over the course of the experiment.

fit for each of the Japanese – English word pairs studied by participants (Figure 2 – left panel). The results of the correlation between the Standard PPE’s free parameters are apparent and as expected from the analysis of the Standard PPE formulation. First, there was a moderate negative correlation between the b and m parameter in the standard PPE decay equation (Eq. 5). This correlation between the b and m parameters occurs due to the fact the decay parameter is structured as a linear regression with the product of the stability term and m being added to the b parameter. From this construction the same decay value can be achieved under a variety of b and m combinations. Second, the b parameter is seen to negatively correlate with τ parameter (Figure 2 – left panel). This correlation is caused by nesting the Standard PPE’s activation term (M_i) within the logistic term (Eq. 6). Due to the fact that τ affects the Standard PPE’s performance estimation outside of the activation term additional variance in the participants performance can be explained by manipulating either the τ or b parameter. Finally, a smaller negative correlation between the τ and s parameter was observed. This correlation is the result of the structure of the logistic term and unbalanced units of the Standard PPE’s activation term subtracted by the τ .

Modified PPE Compared to the standard PPE, the Modified PPE has only two parameters: b and m . Overall, the correlation between the b and m parameter is minimal compared to some of the correlations across parameters that were observed in the Standard PPE. This reduced correlation is the result of removing the logistic term from the equation, thus not having any nested terms within the equation and making sure the b and m parameters each affect only one term in the decay parameter, the stability term and the lag. The effect of each of these manipulations is that the Modified PPE is greatly simplified.

Measuring Aspects of Performance

In our final comparison between the Standard and Modified PPE, each model’s subject-level free parameters and specific aspects of the participants’ performance were

examined. The two relevant aspects of the participants’ performance chosen for this paper were the participants’ overall accuracy and retention between sessions (i.e. accuracy on the 1st trial during the 2nd and 3rd session trials 11 & 21). Ideally, a model’s free parameters should represent a latent theoretical construct, such as learning and decay, which then map on to particular measures of behavior.

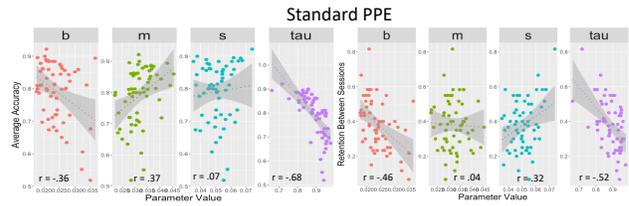


Figure 3. Scatter plot between participants’ overall average performance (left plot) and the average performance during the 1st trial during the 2nd and 3rd session and the free parameters (right plot) from the Standard PPE (left four columns).

Standard PPE As seen in Figure 3, the Standard PPE’s free subject-level parameters were seen to have a moderate correlation with both the participant’s overall average performance and their retention between sessions. However, with both measures τ was found to have the highest correlation with both overall accuracy and the retention between sessions, compared to both the b , m , and s parameter. This result highlights that the τ parameter has a disproportionate influence on the Standard PPE’s performance estimates. The predominant influence of τ can be seen as problematic due to the fact that τ modifies the activation term (M_i) and does not have any direct influence on either model time (T_i) or the stability term (St_i). One potential cause for the limited influence of the b and m parameter on accounting for performance, is the correlation between each other (i.e., b & m) and the τ parameter.

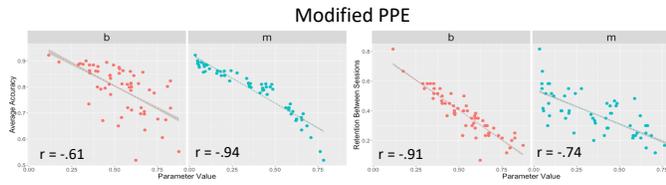


Figure 4. Scatter plot between participants’ overall average performance (left plot) and the average performance during the 1st trial during the 2nd and 3rd session and the free parameters (right plot) Modified PPE.

Modified PPE: In contrast to the Standard PPE, the free parameters in the Modified PPE are both seen to have a strong relationship with both the participants’ overall accuracy and retention between sessions (Figure 4 - Panel B). Further, the degree of each parameter’s relationship can be understood from a theoretical perspective. The *m* parameter is accounting for the participants’ ability to learn the Japanese – English word pairs during the experiment. This relationship is in line with the *m* parameter’s function within the Modified PPE, affecting the rate at which information is learned. The *b* parameter, on the other hand, which affects the model’s stability term, is seen to predominantly account for the participants’ retention between sessions. These results highlight another benefit of the Modified over the Standard PPE. The simplified structure of the Modified PPE allows for the parameters to better summarize and map on to particular aspects of the participants’ performance.

Discussion

In this paper, the formulation of the Standard PPE was examined. Two aspects of the PPE’s structure were identified as contributing to the correlation between PPE’s free parameters. Correlation between parameters increases a model’s complexity and obscures the meaning that can be attributed to particular parameter estimates.

To reduce this intercorrelation across parameters, several modifications were made to the Standard PPE’s structure, removing the activation term (M_i) from the logistic equation and modifying the forgetting and learning terms. Although these modifications to the Standard PPE changed the structure of the equations, the components unique to the PPE relative to other spacing models (i.e., variable decay rate, stability term, and model time) remained intact. A formal model comparison between the Standard and Modified PPE revealed that the Modified PPE was able to (1) account for the participants’ performance across the three experimental sessions, (2) greatly reduced the correlation across its two free parameters, and (3) parameter estimates mapped on to specific aspects of the participants’ performance.

It is important to note that the results reported in this paper do not invalidate any previous findings of the PPE, but simply address the meaning that can be attributed to its parameter estimates. PPE was initially developed as a predictive tool and to meet a set of applied criteria (i.e., assign prescriptive

scheduling, calibrate quickly to prior performance, account for relearning of spaced items after a delay; see Walsh et al., 2018 for full list). Along these criteria the Standard PPE has succeeded and has been used successfully as a predictive tool across different applied domains.

Attempting to explain data from a theoretical point of view and predicting new observations are opposing goals for scientific models (Shmueli, 2010) and neither one should be considered superior to the other. Instead, a balance between these two extremes should be found based on the pragmatic goals of the research question. If the goal is to use the PPE as a method to predict future learning and retention behavior of an individual, then the Standard PPE’s formulation is acceptable. In contrast, if the goal is to summarize an individual’s performance in terms of psychologically latent values (i.e. decay, learning) or to compare the best-fitting parameters across individuals to evaluate individual differences, then the Modified PPE proposed here is a more appropriate tool.

Several limitations need to be addressed within this paper. First, additional research needs to be conducted to further explore how well the Modified PPE can account for performance across longer and more variable learning schedules relative to the Standard PPE. Here, the Standard and Modified PPE were compared across only six unique learning conditions. Future research should compare the two models along a variety of both spaced and non-spaced learning schedules to better find where these two models differ. Second, this paper focused on reducing the parameter correlation across the Modified PPE parameters, to simplify the model and reduce its dimensionality. Future research should explore using Modified PPE for psychometric purposes, evaluating if either of its parameters correlate with particular psychological constructs such as working memory or attention.

Conclusion Mathematical models of psychological theories are useful tools for theory evaluation, development and applied technologies. For these goals to be met, care should be taken to ensure that a model’s formulation and representation are adequate and are in line with their verbal descriptions. By attending to how particular implementations of theories are represented, a balance between mathematical, statistical, and scientific validity can be found.

Acknowledgments

The authors would like to thank Joe Houpt, Christopher Fisher, and Noah Thomas for helpful discussions and comments on this work. This research was supported by MGC's participation through an appointment to the Oak Ridge Institute for Science and Education (ORISE) Student Research Participation Program. The views expressed in this paper are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government

References

- Krefeld-Schwalb, A., Scheibehenne, B., & Pachur, T. (2019, February 1). Structural Parameter Interdependencies in Cognitive Models. <https://doi.org/10.31234/osf.io/pxmnw>
- Heathcote, A., Brown, S., & Mewhort, D. J. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic bulletin & review*, 7(2), 185-207.
- Pavlik Jr, P. I., & Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science*, 29(4), 559-586.
- Plummer, M. (2012). JAGS Version 3.3. 0 user manual. *International Agency for Research on Cancer, Lyon, France*.
- Raaijmakers, J. G. (2003). Spacing and repetition effects in human memory: Application of the SAM model. *Cognitive Science*, 27(3), 431-452.
- Shmueli, G. (2010). To explain or to predict?. *Statistical science*, 25(3), 289-310.
- Vincent, B. T., & Stewart, N. (2020). The case of muddled units in temporal discounting. *Cognition*, 198, 104203.
- Walsh, M. M., Gluck, K. A., Gunzelmann, G., Jastrzembski, T., Krusmark, M., Myung, J. I., ... & Zhou, R. (2018). Mechanisms underlying the spacing effect in learning: A comparison of three computational models. *Journal of Experimental Psychology: General*, 147(9), 1325.

Modeling Phishing Susceptibility as Decisions from Experience

Edward A. Cranford (cranford@cmu.edu)

Department of Psychology, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15221 USA

Kuldeep Singh (kuldeep2@andrew.cmu.edu)

Department of Social and Decision Sciences, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15221 USA

Palvi Aggarwal (palvia@andrew.cmu.edu)

Department of Social and Decision Sciences, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15221 USA

Christian Lebiere (cl@cmu.edu)

Department of Psychology, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15221 USA

Cleotilde Gonzalez (coty@cmu.edu)

Department of Social and Decision Sciences, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15221 USA

Abstract

Traditional anti-phishing training is often non-personalized and does not typically account for human experiential learning. However, to personalize training, one requires accurate models and predictions of individual susceptibility to phishing emails. The present research is a step toward this goal. We propose an Instance-Based Learning model of phishing detection decision-making, constructed in the ACT-R cognitive architecture. We demonstrate the model's ability to predict behavior in a frequency training study, and its generality by predicting behavior in another phishing detection study. The results shed additional light on human susceptibility to phishing emails and highlight the effectiveness of modeling phishing detection as decisions from experience. We discuss the implications of these results for personalized anti-phishing training.

Keywords: phishing; cybersecurity; personalized training; decision making; instance-based learning theory; ACT-R

Introduction

Despite significant advances in security technologies, a large number of phishing emails continue to evade automated detection and are often successful because it is cognitively challenging for humans to distinguish the rare deceptive phishing message from benign emails. As such, phishing attacks remain the biggest, growing threat for cybersecurity (APWG Phishing report, 2020). While phishing attacks exploit human weaknesses using social engineering and psychological techniques (Jagatic et al., 2007), defenders typically employ technological solutions to defend against them, such as machine learning filtering of phishing emails, email authentication tools, and URL filtration/blacklisting (Prakash et al., 2010; Marchal et al., 2014; Peng, Harris, & Sawa, 2018). However, attackers are persistent and phishing emails continue

to reach their victims. Since the success of phishing attacks relies on exploiting cognitive and psychological weaknesses, it becomes essential to understand the underlying decision-making processes that influence end-user susceptibility to phishing emails (Canfield, Fischhoff, & Davis, 2016).

Recent research has shown that end-user phishing detection decisions are similar to other kinds of decisions from experience (e.g., Hakim et al., 2020; Singh et al., 2019, 2020). An individual's personal history and experience with emails can have a large influence on phishing susceptibility. Specifically, phishing decisions are influenced by the recency, frequency, and similarity of past emails to the features of the current email. For example, Singh et al. (2019) manipulated the frequency of phishing emails in an anti-phishing training study. The results showed that increasing the frequency of phishing emails during training increased the hit rate of detecting phishing emails. In other research, Singh et al. (2020) examined how the similarity of email features influenced detection accuracy. Their results showed that detection accuracy suffered the more similar the features of a phishing email were to the features of the benign emails. Lastly, Hakim et al. (2020) developed a regression model of end-user phishing susceptibility in an email rating task and revealed an effect of recency on detection decisions. In the task, end-users rated phishing and ham emails on a scale of suspiciousness. The regression model showed evidence of sequential effects of the emails, such that current ratings were positively affected by the previous rating.

Although the evidence shows that phishing decisions are influenced by experiential learning with emails, current training procedures do not take these factors into account, nor have the effects been investigated. Organizations typically

use embedded-training methods that involve sending simulated phishing emails and only provide more traditional phishing training whenever one clicks on the link in the simulated phishing message (Kumaraguru et al., 2009; Kumaraguru et al., 2007). Traditional techniques have often focused on teaching end-users to understand and identify the relevant features that distinguish phishing emails from benign ones (Kumaraguru et al., 2009; Singh et al., 2020). However, there is a deficit of effective experiential phishing training methods that directly address important underlying human cognitive processes in context. Traditional phishing training is often generic and non-personalized. That is, all end-users receive the same set of training emails, the non-phishing emails lack the familiar context that personal ham emails tend to have (e.g., from senders who are familiar to the end-user), and the phishing emails are sent without consideration of the individual's history. Consequently, in many phishing-detection tasks, end-users have trouble distinguishing the phishing emails from the ham, and due to the generic nature of training, the effects of training vary considerably between individuals. In addition, different types of phishing emails have had varied effectiveness across individuals, further emphasizing the need to personalize anti-phishing training (Lin et al., 2019; Oliveira et al., 2017).

Personalized training interventions could prove immensely useful for improving anti-phishing detection, but such methods require models that can be tailored to individuals and that can make accurate decision predictions for a specific phishing email presented at a specific time. Therefore, as a first step toward this goal, we propose a cognitive model that leverages the influence of individual experience on phishing detection decisions, specifically turning to a memory-based theory of experiential learning called instance-based learning theory (IBLT; Gonzalez, Lerch, & Lebiere, 2003). According to IBLT, decisions are made by generalizing across past experiences, or instances, that are similar to the current situation. Typically, instances represent the features of the decision, the action taken, and the outcome of that decision. However, for emails, there is usually a dissociation between the actions taken and feedback regarding whether the email was ultimately malicious. For a given email, IBLT suggests that end-users make decisions by retrieving a classification from memory based on the similarity of features of the current email to features of past emails. Thus, decisions are influenced by typical memory effects such as recency and frequency of past instances and are susceptible to cognitive biases that emerge from these memory processes (e.g., confirmation bias; Lebiere et al., 2013).

General cognitive theories of decisions from experience indicate that the low frequency of phishing emails (compared to benign emails) could be a major issue in the success of detection decisions if end-users underweight the probability of these rare events (Gonzalez et al., 2003; Gonzalez & Dutt, 2011). Additionally, phishing emails often mimic quite well the benign (i.e., ham) emails that regularly flood our inboxes. In other words, phishing emails are similar to the highly frequent and usually recent benign emails that we receive

regularly, and phishing decisions are susceptible to effects of frequency, recency, and similarity of features.

Our cognitive model builds upon that proposed by Cranford et al. (2019). In this paper, we first extend and improve upon that model to explore the effects of frequency on phishing detection training by modeling the Phishing Training Task (PTT) in Singh et al. (2019; 2020). We then demonstrate the model's generality by running it through the task in Hakim et al. (2019), the Phishing Email Suspicion Test (PEST), which tests on a different database of emails. Finally, we discuss the implications of the model for future research towards personalized, adaptive anti-phishing training interventions.

Modeling the Phishing Training Task

The PTT (Singh et al., 2019) was designed to examine the impact of learning factors (e.g., frequency effects) on phishing detection decisions. The task is based on the design in Canfield et al. (2016) in which participants are presented a series of email messages and are requested to make classification decisions. In the PTT, participants make three responses to each email: a classification decision of whether the email was a phishing email or not, a confidence rating of their decision (from 50, "not confident at all", to 100, "fully confident"), and the action they would take in response to each email (selected from a 6-point, Likert-type scale ranging from "Respond to this email" to "Report this Email"). For the present model, we focused on the first classification decision.

The PTT consists of three phases: pre-test, training, and post-test. During the pre- and post-test phases, end-users are presented with 10 emails, two of which are phishing emails, and the remaining are benign, ham emails. During the training phase, end-users are presented 40 emails of which 10, 20, or 30 are phishing emails. End-users are randomly assigned to one of the three phishing frequency conditions. Feedback about decision accuracy is provided after each trial during the training phase but not during either testing phase.

IBL Model Description

The IBL model was adapted from Cranford et al. (2019) and constructed in the ACT-R cognitive architecture (Anderson & Lebiere, 1998). The modifications made to the model were few, but important, and provided substantial improvement to predicting human behavior in the PTT. These will be discussed below, after presenting the model results.

The model performs the PTT in the same way as humans, processing one email at a time, judging whether each is phishing or ham. For each email, the model takes the content of the email as input and generates a classification by retrieving from similar past instances. For the PTT, the elements of an email include the sender's email address, subject line, email body, link text, and underlying link URL. The classification (i.e., decision) is either phishing or ham. In ACT-R, the retrieval of past instances is based on the activation strength of the relevant chunk in memory and its similarity to each of the elements of the current situation. The activation A_i of a chunk i is determined by the following equation:

$$A_i = \ln \sum_{j=1}^n t_j^{-d} + MP * \sum_k Sim(v_k, c_k) + \varepsilon_i \quad (1)$$

The first term provides the power law of practice and forgetting, where t_j is the time since the j th occurrence of chunk i and d is the decay rate of each occurrence. The second term reflects a partial matching process, where $Sim(v_k, c_k)$ is the similarity between the actual memory value and the corresponding element for chunk slot k , and is scaled by the mismatch penalty (MP, which was set at 2.0; discussed below). The term ε_i represents transient noise, a random value from a logistic distribution with a mean of zero and variance parameter s of 0.25 (common ACT-R value, e.g., Lebiere, 1999), and introduces stochasticity in retrieval.

The probability of retrieving a particular instance is determined according to the SoftMax equation (i.e., the Boltzmann equation), reflecting the ratio of an instance’s activation A_i and the temperature t (which was set to the default value which scales to the noise parameter, $\sqrt{2} * s$):

$$P_i = \frac{e^{A_i/t}}{\sum_j e^{A_j/t}} \quad (2)$$

The model uses ACT-R’s *blending* mechanism (Lebiere, 1999, Gonzalez et al., 2003) to generate a classification based on the similarity to past instances. Blending is a memory retrieval mechanism that returns a consensus value across all memories with similar elements, rather than from a specific memory, and is computed by the following equation:

$$\underset{V}{\operatorname{argmin}} \sum_i P_i \times (1 - Sim(V, V_i))^2 \quad (3)$$

The value V is the one that best satisfies the constraints among actual values V_i in the matching chunks i weighted by their probability of retrieval P_i . Satisficing is defined as minimizing the dissimilarity between the consensus value V and the actual answer V_i contained in chunk i .

In summary, the model matches memories to the current email content and uses blending to generate the classification decision. After generating a classification, the experience (email content plus decision) is saved in declarative memory as a new instance, which affects future decisions. During the training phase, the classification slot is first updated to match the feedback prior to being saved to memory.

While prior research has identified relevant features for detecting phishing emails (Kumaraguru et al., 2009; Singh et al., 2020), and training tools have attempted to teach end-users to identify such features, the current model relies solely on the semantic features of the email to make classifications. At the lowest level, an end-user that has not undergone training to identify expert features would likely rely on the semantics of an email to make a classification. An email that is semantically similar to past known phishing emails would more likely be classified as phishing. Additionally, by relying on only the semantics of the email, the model does not need to identify expert features in a body of text (which is a difficult task to automate by any natural language processing, NLP, standards). In contrast, current NLP techniques are quite efficient at computing the semantic similarity between texts and can

therefore feasibly be used to generate the similarities between emails required for blending computations.

A novel feature of the model, therefore, is how similarities are computed between slot values. Typically, similarities between numeric values are computed using a linear function scaled between 0 and 1.0, where 1.0 is a perfect match and 0 is maximally dissimilar. However, for non-numeric information, unless a value is specified for a relation, they are either maximally similar or maximally different. For emails, the content is non-numeric, often several words to paragraphs in length. Because two texts that are semantically similar should have higher similarity values (closer to 1.0) compared to texts that are semantically very dissimilar, it is possible to compute individual similarities between semantic content.

To compute similarities between textual information, we used the University of Maryland Baltimore County’s semantic-textual-similarity tool (UMBC; Han et al., 2013). The tool uses a combination of latent semantic analysis (LSA) and WordNet to produce semantic similarity values between two texts. The two input texts can be of any word-length and it produces a value between 0.0 and 1.0, with 1.0 being maximally similar in meaning. For example, the similarity between “happy dog” and “joyful puppy” is 0.65, whereas “happy dog” and “sad feline” is 0.34, and “happy dog” and “hot tea” is 0.0. This technique has proven useful for producing meaningful similarity values between textual content.

Model Results

To generate stable estimates of performance compared to that of humans, the model was run 10 times per participant and given the same sequence of emails presented to the participant. Therefore, in the analyses below, we compare 2980 model runs to 298 humans. Before beginning the task, the model must first be initialized with a set of instances to be able to retrieve a classification. Therefore, the model was initialized with 10 instances that include the email content and ground-truth classification, five of which were phishing emails and five were ham. The initialized instances were sampled from the remaining emails that were not presented during the task.

To examine the model performance compared to that of humans, we computed signal detection measures and plotted the receiver operating characteristic (ROC) curve for each phase and frequency condition of the task. We plotted the mean True Positive Rate (TPR; or Sensitivity) on the y-axis and the False Positive Rate (FPR; or 1-Specificity) on the x-axis. The TPR is equivalent to the hit rate of classifying phishing emails as phishing. The FPR is equivalent to the false-alarm rate of classifying ham emails as phishing. Therefore, in ROC space, points closer to the top left of the graph indicate greater discriminability while points toward the middle indicate less discriminability. Meanwhile, points toward the top right or bottom left indicate greater overall bias toward responding phishing or ham, respectively.

Figure 1 shows the mean ROC curves for the humans (black) compared to the model (gray). As can be seen, the model generates very accurate predictions of human behavior

across phases and frequency conditions. Like humans, the model does not perform perfectly, highlighting the difficulty of the task in discriminating phishing from ham emails. As observed in humans, the frequency of phishing emails observed during training (Phase 2) had a direct impact on discriminability in the post-test phase (Phase 3), such that greater increases in frequency during Phase 2 led to greater increases in TPR, but also FPR, in Phase 3 compared to Phase 1. However, as can be observed, the model is slightly more sensitive to frequency effects than are humans. When the base rate is 25% (10 phishing, 30 ham) the model tends to underpredict human performance at post-test and classifies more of the phishing emails as ham. When the base rate is 50% phishing emails or more, then the model tends to more accurately classify the phishing emails compared to humans. The model demonstrates that a greater frequency of experience with phishing emails leads to more cautious decisions with future emails. This is because the greater number of phishing instances in memory the greater influence they have on retrieval (i.e., a greater probability of retrieving a phishing classification from memory).

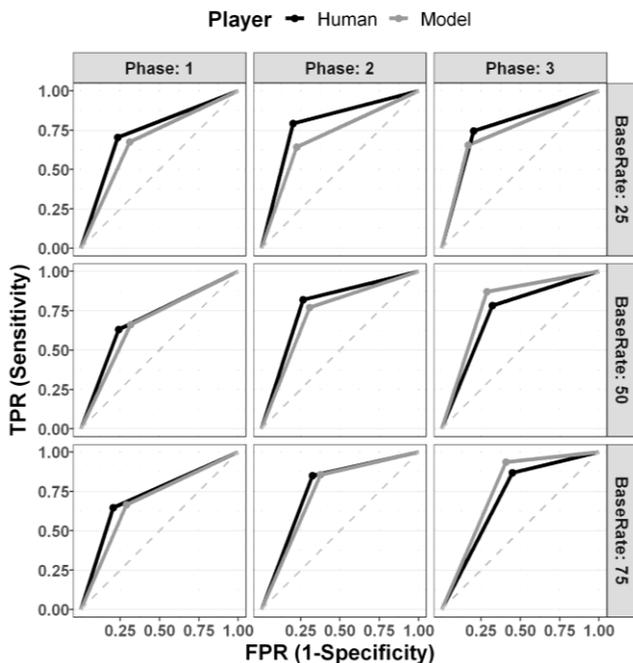


Figure 1: ROC curves of phishing decision accuracy across three phases of the PTT and three frequency conditions, for humans (black) compared to the model (gray).

Discussion

That the model generated highly accurate predictions of human behavior is good news towards developing personalized anti-phishing training interventions. The model is able to rely on experience, through interaction with the environment, and the dynamics of memory to generate a range of behavior. The modifications we made to the original

Cranford et al (2019) model helped to provide a better understanding of end-user susceptibility to phishing emails.

For the original model, the important parameter values for activation and blending were left at their default values. These include, decay rate d , mismatch penalty MP , transient noise s , and temperature t . For the decay rate, the default value is 0.5. Decay rate is related to forgetting and influences recency effects such that the higher the value the less of an impact older instances will have in retrieval, and thus more recent instances will have a greater impact. At the default value, the model tended classify emails as ham due to the greater frequency of ham emails during the pre-test phase. In the current model, we set this parameter to 0. This allowed all instances to play a more equal role in retrieval and reduced excessive recency effects. Studies have shown that the default decay rate of 0.5 is effective for modeling the typical laboratory task that is short in duration and involves novel stimuli, however, for longer duration tasks this value is less useful at representing retrieval effects (Pavlik & Anderson, 2005). For the present task, reading emails is a task with which humans come to the experiment with vast amounts of prior experience. It is presumable then that these past experiences play a role in retrieval and have decayed to a steady level at experimentation. With such a vast memory base, spreading activation more evenly across new instances by setting decay to zero is a suitable solution for representing this memory phenomenon.

Ideally, we would use a portion of the end-user’s actual history of emails to initialize the model, but using examples from the database was a reasonable alternative. Interestingly, the model was initialized with equal numbers of ham and phishing emails, whereas in reality, humans see many more ham emails than phishing. Initializing the model with comparatively more ham emails resulted in too many ham classifications. It could be that phishing emails are inherently more emotionally valent and thus more salient in memory. Or it could be an experimental effect of end-users expecting some of the emails to be phishing. In any case, these results reveal a bias to respond phishing in the task that was accounted for through initialization of instances.

One issue with the original model was that the UMBC semantic similarity tool produces a compressed range of values, which in turn compresses the range of differences between emails and makes it more difficult to discriminate stimuli. During retrieval, the instances are more evenly weighted. To alleviate this constraint, we modified the temperature and mismatch penalty parameters. The noise s was left at its default value of 0.25 which provides a reasonable amount of stochasticity in retrieval. Increasing this value resulted in overly varied responses, and reduced discriminability. However, lowering this value did not produce enough stochasticity between model runs. The temperature on the other hand was reduced from a neutral value of 1.0 to the default value of $\sqrt{2} * s$, which equals approximately 0.35 given the current value of s . Temperatures of 1.0 reflect an unbiased retrieval given the historical frequency distribution of instances. This means that retrieval is

more evenly distributed across instances. Increasing this value tends toward randomness, making discrimination more difficult. Therefore, lowering the temperature by reverting to the default ACT-R value resulted in greater discriminability where more weight is given to instances with higher activation values. This in turn rewards those instances that are more semantically similar to the current instance. Finally, we increased the mismatch penalty MP , from 1.0 to 2.0. The mismatch penalty directly influences the model's discriminability because it scales the dissimilarity between instances. Therefore, increasing this value enhanced the differences between different emails while simultaneously strengthening the similarities between similar emails, effectively decompressing the range of similarities produced by the UMBC semantic-similarity tool. The result was an increase in overall discriminability of the model.

The current model predicts human performance well, but even still, there is room for improvement. We did not perform any detailed parameterization of the model, but instead settled on reasonable and justifiable values through strategic exploration. Therefore, the model has potential to be further refined. Additionally, relying on the semantic similarity of features of an email, generated through NLP techniques, instead of attempting to extract the presence of features within the email text, allowed us to create a model that can more easily generalize to novel environments (i.e., with different emails). Relying on the semantic content means we do not have to preprocess new emails, manually or through automated means, to identify relevant features. To test the generality of the model, we ran the model through another task that used a different database of emails, the Phishing Email Suspicion Test (Hakim et al., 2019).

Modeling the Phishing Email Suspicion Test

Hakim et al. (2019) used the PEST task to assess the relationship between real and simulated phishing and ham emails and to examine the efficacy of using the simulated phishing emails for anti-phishing training against real-world phishing attempts. In the PEST task, 97 participants rated a total of 160 emails each on a Likert-type suspiciousness scale from 1 "Definitely Safe" to 4 "Definitely Suspicious". Participants were presented 40 of each type of email: real-ham, simulated-ham, real-phishing, and simulated-phishing. The emails were presented in random order and selected randomly from the database of emails.

The IBL model described above was tasked to perform the PEST. Because the PEST included four types of emails, the model was initialized with a total of 20 emails (five of each type). However, to model the individual differences observed in the human PEST data, we introduced stochasticity in the initialization. That is, the model was initialized with three to six ham emails of each type, randomly selected from a uniform distribution, and the remaining were phishing. This also introduced varied initial biases between model runs, where some runs were initially biased toward ham and other runs more biased toward phishing, thus resembling a human population, but with a skew toward phishing. However, as

will be discussed in more detail below, to produce the following results, each run was initialized with an extra 2 simulated-phishing emails. This is consistent with the finding of Hakim et al. (2019) that participants displayed a bias to respond phishing, and with the phishing bias observed in the PTT model. Increasing the number of phishing emails was required to drive such a bias in the PEST model.

Since the PEST database of emails includes only 40 examples of real-ham emails, to ensure no initialized real-ham emails were presented during the test, we reduced the number of emails of each type presented during testing from 40 to 30. Therefore, the model experiences 120 total emails per run, still allowing for ample observations. For the PEST, emails did not show the underlying link URL if hovered over with a mouse, so the URL feature was removed from the instance representation and only the link text was compared in retrievals. Instead of generating a classification, the model takes the semantic features of the email as input and generates, via blending, a rating score between 1 and 4. The retrieved value is a continuous value that is rounded to the nearest whole number to provide a discrete rating. The blended rating value is replaced with the discrete rating value before saving the instance to declarative memory at the end of each trial. Like humans, the model does not receive feedback regarding the accuracy of its decisions.

Model Results

To generate stable estimates of performance, the model ran through the task 200 times, with initialized and tested emails selected randomly for each run. Therefore, in the analyses below, we compare 200 model runs to 97 humans.

Because the PEST requires a rating response as opposed to a classification response, to analyze the model performance compared to humans, we examined the mean suspicion scores per email type as well as the subject-level correlation between ratings for simulated and real emails, separately for phishing and ham emails. These combined measures provide accounts for the mean as well as the full range of human behavior.

Figure 2 shows a boxplot of the mean suspicion score per email type. The results align very well with the human data, closely accounting for the mean behavior as well as the variance between individuals. The real-ham emails were rated the lowest at approximately 2, while the simulated- and real-phishing emails were rated highest at almost 3. Meanwhile, the simulated-ham emails were rated at near 2.5.

Figure 3 shows the correlation between simulated and real emails for ham and phishing emails separately. These results highlight the model's ability to account for both the within- and between-subject variances in performance. As will be discussed further, a key contributor to the model's performance is the randomization of initial instances.

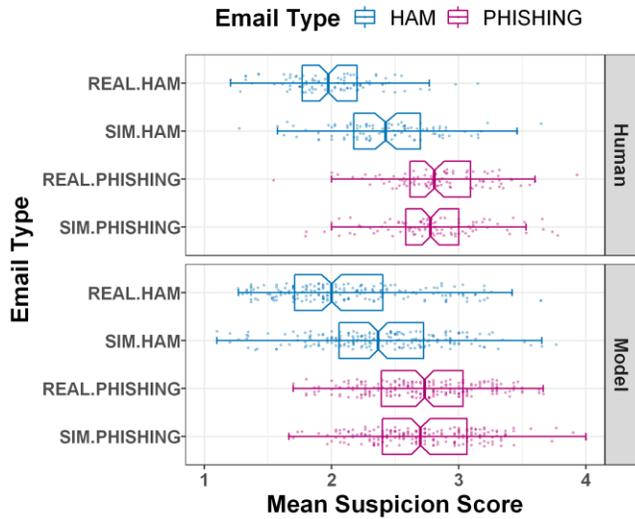


Figure 2: Boxplot of mean suspicion scores for each type of email in the PEST for humans compared to the model.

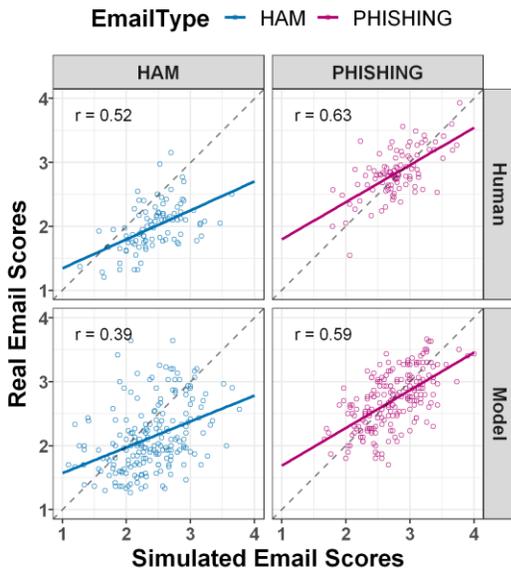


Figure 3: Scatterplots showing the correlation between real emails and simulated emails for ham emails and phishing emails separately, comparing humans to the model.

Discussion

The IBL phishing susceptibility model was able to successfully generalize to other environmental conditions with a different pool of participants performing a slightly different task with different stimuli. To produce the level of accuracy in predicting human behavior in the PEST, the model required stochasticity in initialization and additional initialized phishing instances. The result was an increase in correlations between real and simulated emails and an overall bias toward rating emails more suspicious. In fact, using a static initialization considerably reduced the observed correlations in Figure 3.

In exploring an appropriate initialization for the model, the results revealed a relationship between simulated-ham and -phishing emails. For example, increases in simulated-phishing emails had a positive impact on simulated-ham emails. These results suggest there is large semantic overlap between simulated emails, which is consistent with how the simulated-phishing and -ham emails were constructed. The simulated-ham emails were modified versions of simulated-phishing emails made to seem less suspicious. The model picks up on this semantic overlap which results in simulated-ham emails having a higher match to simulated-phishing emails, producing inflated ratings for simulated-ham emails.

Conclusion

Our IBL model highlights the role of experiential learning for end-user phishing detection decisions. The major influences in generating accurate predictions of human susceptibility to phishing emails were initialization of instances and similarity between email features. A phishing bias was accounted for by adding disproportionately more phishing emails than ham emails compared to real-world frequencies. Adding stochasticity in initialization accounted for individual differences in behavior. Humans have distinct experiences that influence decisions and capturing this background knowledge is essential to building models that not only predict a range of human behavior, but also that can predict a specific individual’s behavior. Because the model is expected to generate better predictions of an individual the more the model’s memory aligns with the human’s, model-tracing techniques should prove useful in developing personalized anti-phishing training interventions (Anderson et al., 1995). Future research is aimed at further exploring the effects of initialization, with an emphasis on generality and in exploring ways to decompress the range of semantic similarities or even representing alternative features.

While the current model uses only the semantics of the email to make decisions, current training methods teach end-users to identify so-called “expert” features (e.g., a request for personal information; Singh et al., 2020). Using only semantic features of an email produces human-like, albeit fairly poor discriminability in the experimental tasks. A goal for future research is developing a model that can learn to identify expert features so that we can use the model to help train end-users to detect such features. For now, the current model proved a successful first step toward personalized anti-phishing training.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant Number 2026148, the Carnegie Mellon University CyLab Security and Privacy Institute, and the Army Research Office under MURI Grant Number W911NF-17-1-0370.

References

- Anderson, J. R., Corbett, A. T., Koedinger, K., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of Learning Sciences*, 4, 167–207.
- Anderson, J. R., & Lebiere, C. (1998). *The Atomic Components of Thought*. Mahwah, NJ: Erlbaum.
- Anti-Phishing Working Group (2021). *Phishing Activity Trends Report: 4th Quarter 2020*. Retrieved from https://docs.apwg.org/reports/apwg_trends_report_q4_2020.pdf. APWG.
- Canfield, C. I., Fischhoff, B., & Davis, A. (2016). Quantifying phishing susceptibility for detection and behavior decisions. *Human factors*, 58(8), 1158-1172.
- Cranford, E. A., Lebiere, C., Rajivan, P., Aggarwal, P., & Gonzalez, C. (2019). Modeling cognitive dynamics in end-user response to phishing emails. In *Proceedings of the 17th Annual Meeting of the International Conference on Cognitive Modelling*. Montreal, CA.
- Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological review*, 118 (4), 523.
- Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance based learning in dynamic decision making. *Cognitive Science*, 27(4), 591-635.
- Hakim, Z.M., Ebner, N.C., Oliveira, D.S. *et al.* (2020). The Phishing Email Suspicion Test (PEST) a lab-based task for evaluating the cognitive mechanisms of phishing detection. *Behavioral Research Methods*.
- Han, L., Kashyap, A. L., Finin, T., Mayfield, J., & Weese, J. (2013). UMBC_EBIQUITY-CORE: Semantic Textual Similarity Systems. In *Proceedings of the 2nd JCLCS* (pp. 44-52). Atlanta, GA.
- Jagatic, T. N., Johnson, N. A., Jakobsson, M., & Menczer, F. (2007). Social phishing. *Communications of the ACM*, 50(10), 94–100
- Kumaraguru, P., Cranshaw, J., Acquisti, A., Cranor, L., Hong, J., Blair, M. A., & Pham, T. (2009). School of phish: a real-world evaluation of anti-phishing training. In *Proceedings of the 5th Symposium on Usable Privacy and Security* (p. 3). Mountain View, CA.
- Kumaraguru, P., Rhee, Y., Sheng, S., Hasan, S., Acquisti, A., Cranor, L. F., & Hong, J. (2007). Getting users to pay attention to anti-phishing education: evaluation of retention and transfer. In *Proceedings of the anti-phishing working groups 2nd annual ecrime researchers summit* (pp. 70-81).
- Lebiere, C. (1999). A blending process for aggregate retrievals. In *Proceedings of the 6th ACT-R Workshop*. George Mason University, Fairfax, Va.
- Lebiere, C., Pirolli, P., Thomson, R., Paik, J., Rutledge-Taylor, M., Staszewski, J., & Anderson, J. R. (2013). A Functional Model of Sensemaking in a Neurocognitive Architecture. *Computational Intelligence and Neuroscience*.
- Lin, T., Capecci, D. E., Ellis, D. M., Rocha, H. A., Dommaraju, S., Oliveira, D. S., & Ebner, N. C. (2019). Susceptibility to Spear-Phishing Emails: Effects of Internet User Demographics and Email Content. *ACM Trans. Comput.-Hum. Interact.* 26, 5, Article 32 (July 2019), 28 pages.
- Marchal, S., François, J., State, R., & Engel, T. (2014). Phishstorm: Detecting phishing with streaming analytics. *IEEE Transactions on Network and Service Management*, 11(4), 458-471.
- Oliveira, D., Rocha, H., Yang, H., Ellis, D., Dommaraju, S., Muradoglu, M., Weir, D., Soliman, A., Lin, T., & Ebner, N. (2017). Dissecting Spear Phishing Emails for Older vs Young Adults: On the Interplay of Weapons of Influence and Life Domains in Predicting Susceptibility to Phishing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6412–6424.
- Pavlik Jr, P. I., & Anderson, J. R. (2005) Practice and forgetting effects on vocabulary memory: an activation-based model of the spacing effect. *Cognitive Science*, 29(4), 559–586.
- Peng, T., Harris, I., & Sawa, Y. (2018). Detecting phishing attacks using natural language processing and machine learning. In *Proceedings of the IEEE 12th international conference on semantic computing* (pp. 300–301).
- Prakash, P., Kumar, M., Kompella, R. R., & Gupta, M. (2010). Phishnet: predictive blacklisting to detect phishing attacks. In *2010 Proceedings IEEE INFOCOM* (pp. 1-5). San Diego, CA.
- Singh, K., Aggarwal, P., Rajivan, P., & Gonzalez C. (2019). Training to detect phishing emails: Effect of the frequency of experienced phishing emails. In *Proceeding of the 63rd International Annual Meeting of the HFES*. Seattle, WA.
- Singh, K., Aggarwal, P., Rajivan, P., & Gonzalez, C. (2020). What makes phishing emails hard for humans to detect? In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 64(1). Chicago, IL.
- Vishwanath, A., Harrison, B., & Ng, Y. J. (2018). Suspicion, Cognition, and Automaticity Model of Phishing Susceptibility. *Comm. Research*, 45(8), 1146–1166.

Towards Benchmarking Cognitive Models: A Python Library for Modular Environment Specification and Partial Model Generation in ACT-R

Emmanuelle Dietz (emmanuelle.dietz@airbus.com) and Oliver W. Klaproth (oliver.klaproth@airbus.com)

Airbus Central Research & Technology, Hein-Sass-Weg 22
21129 Hamburg, Germany

Abstract

In this paper we present the cognitive modeling library *txt2actr*, which facilitates the specification of an ACT-R environment through simple text files and partially automates the construction of certain components within a cognitive model. Our general purpose goes beyond this library and aims at promoting the modular construction and evaluation of cognitive models. In particular, we suggest to establish benchmarks that allow (i) the competition among models with respect to classical tasks in experimental psychology, and (ii) the evaluation of possibly new or more applied tasks with respect to benchmark models. Such benchmarking proposals can be found in various other disciplines and usually serve as an incentive to improve existing theories and eventually converge towards a common language. Yet, *txt2actr* is far from providing a solution to the associated challenges. It rather serves as a proof of concept by illustrating how two model components for very specific cognitive phenomena in situation awareness can be applied in three different environments.

Introduction

After 50 years, Newell's criticism that the scientific community does not *seem in the experimental literature to put the results of all the experiments together* (Newell, 1973, p. 298) still seems to hold. Interestingly, this problem persists independent of the *bands of cognition* (Newell, 1990), that is the proposed models that deal with *lower levels* such as biological processes or the ones which address *higher levels*, such as human reasoning and decision making. As Khemlani and Johnson-Laird (2012) observed for the particular case of modeling human syllogistic reasoning, *the existence of 12 theories of any scientific domain is a small disaster*. Or, as Taatgen and Anderson (2010) put it, *multiple possible models is not just a problem for cognitive architectures but for any scientific theory*.

One bottleneck might be that most cognitive theories are not formalized and therefore ambiguous: There is no commonly accepted language which allows to compare and thus evaluate their cognitive plausibility on a set of benchmark tasks. As noted by Marewski and Mehlhorn (2011), this leads, among others, to the specification problem (Lewandowsky, 1993), how to translate an under-specified hypothesis into a detailed model, and the identification problem (Anderson, 1976), what to do with many different models which are equally capable of reproducing and explaining data.

During the past decades, cognitive architectures (CA) have been proposed (e.g. ACT-R (Anderson 2007), SOAR (Laird,

2012)) to *address under-specified process hypotheses and provide a falsifiable methodology* (Thomson, Lebiere, Anderson, & Staszewski, 2015). These architectures contributed largely to the development of the field and the comparability of cognitive modeling. However, two main issues still persist: Firstly, it requires substantial intellectual commitment to learn, understand and construct models within these architectures (Taatgen & Anderson, 2010). Secondly, these architectures are highly parametrized, which on the one hand provides a great amount of modeling freedom, but on the other hand leads to models which rather capture *the intuitions of the designers* (Thomson et al., 2015). Laird, Lebiere, and Rosenbloom (2017) proposed a common model of cognition as an abstracted framework depicting the *best consensus given the community's current understanding of the mind* on the architectural level. Such a standard model could then be used as a common language for the community and guide researchers by enabling them to include or extend other components and evaluate or develop psychological experiments.

According to Taatgen and Anderson (2010), a good model is characterized as being applicable to various tasks, as simple as possible and able to predict outcomes of new tasks. That means that the metric for a good model can then be specified by its generalizability (Thomson et al., 2015) and its predictability, including predictions of yet untested cognitive phenomena (Ragni, 2020). These models should be built out of components (Taatgen & Anderson, 2010) and the applicable strategies and heuristics should be rather *selected by the model than by the designer* (Thomson et al., 2015).

According to Ragni (2020), another difficulty is how to *identify the relevant problems* (or tasks) that a model should account for. He suggests to establish generally accepted benchmarks, similar to the PRECORE Challenge (Ragni, Riesterer, & Khemlani, 2019) for human reasoning tasks. The evaluation of this challenge was done with the benchmarking tool Cognitive COmputation for Behavioral Reasoning Analysis (CCOBRA) framework,¹ which was thereafter again applied for new prediction mechanisms in individual human syllogistic reasoning (Dietz Saldanha & Schambach, 2020). The success of establishing benchmarks and developing competitions can be observed in other disciplines (e.g. SAT,²

¹<https://github.com/CognitiveComputationLab/ccobra>

²<https://satcompetition.github.io/2021/>

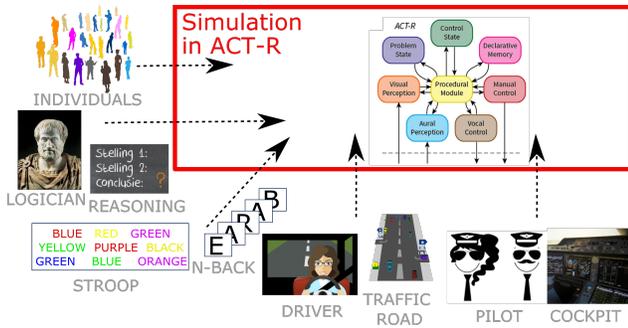


Figure 1: Currently, a majority of model and environment specifications in ACT-R are manually implemented.³

ASP (Gebser, Maratea, & Ricca, 2020)), which have the beneficial side effect to improve existing approaches, and more importantly, motivating the scientific community to agree on a common language. Furthermore, contrasting cognitive models with benchmark models (such as statistical baselines or data-driven neural networks) will determine the current empirical upper bound of the models’ performance (Riesterer, Brand, & Ragni, 2020). At the same time these upper bounds can serve as new incentives for future models to outperform the benchmark models. However, the previously mentioned criteria for good models should be the main focus. In particular even if machine learning techniques might have better predictive overall performances, the generalization of models across a range of paradigms and conditions can be more powerful. As (Lebiere et al., 2013) stated, approaches based on cognitive modeling require less data and fewer domain-specific assumptions to be parametrized as they can be guided by cognitive constraints. Furthermore, they have the advantage to combine symbolic structures and statistical parameters. Taking these proposals as a starting point, we suggest an environmental setting where cognitive models can be benchmarked according to their performance with a set of tasks. In the ideal case, if results are openly shared the ones who have the experimental data but are missing the best predicting models can benefit from the ones who have the cognitive models and vice versa. In this paper, we present the software library *txt2actr* in which the task, the environment, and the model can partially be specified through text files. With this library we aim at developing a modular task design through an ACT-R interface and the parametrization of ACT-R models by a modular and guided production and chunk engineering process. *txt2actr* serves as a proof of concept to address some of the challenges discussed above, and is far from providing a solution or a complete benchmark of tasks or models.

Related Work

The approaches we briefly discuss here are related in the sense that they emphasize the importance of generalizing or benchmarking models, which we consider highly relevant for

the challenges we intend to address. Interestingly they all address this issue on a different level of cognition.

Salvucci (2013) proposes a single model of cognitive skill acquisition in ACT-R by reusing component skills across 7 different task domains. The results are a step towards a more unified account of skill learning and demonstrate that a model can reuse knowledge by transferring it to various tasks.

One of the goals suggested by Taatgen (2013) is to reuse cognitive processes and structure them in a way so that they can be applied in many different combinations, similarly to a construction kit that should be deployable throughout all tasks. As almost all cognitive models suffer from the problem of prior knowledge, transfer in cognitive control, that is the process by which goals or plans influence behavior, might be a promising approach to address this issue, where processes of cognitive control are based on skills.

Marewski and Mehlhorn (2011) specify 39 different process models, which should not only predict what decision a person will make, but also how the information used to make the decision will be processed. In particular, they focus on a class of models that makes decisions by exploiting the accessibility of memory contents. For this purpose they choose to model a task for recognition heuristic. These 39 models either differ in very small aspects or very fundamental assumptions about processing. The main purpose of Marewski and Mehlhorn is not to advocate any particular process model for the task in consideration but rather using the debate as a case study to provide a methodological primer on how architectures like ACT-R can be used to lend precision to theorize decision processes. By implementing models of different levels of description and specificity in one architectural modeling framework, they make the models and their predictions comparable providing a basis for future model tests.

GOMS (Card, Moran, & Newell, 1983) contains hierarchical methods, visual and memory stores, and control constructs and aims at explaining expert routine behaviors and reduce the effort for detailed task analysis and cognitive modeling techniques. Amant and Ritter (2004) provide an automatic generation of GOMS models into ACT-R models. However, it suffers from under-constraints in many areas, for example visual processing (Amant, McBride, & Ritter, 2006). The extension SGOMS (West & Nagy, 2007; West & Somers, 2011) assumes that cognitive modeling at the level of psychological experiments (micro cognition) can scale up to higher level task, such as dealing with task interruptions, by an additional higher-level control structure and multi-tasking.

Cognitive Modeling with *txt2actr*

The python library *text to ACT-R*, abbreviated as *txt2actr*, provides (i) an interface between text files that describe and dynamically change the environment which the cognitive model interacts with during the ACT-R simulation and (ii) a partial and modular construction of the cognitive model. *txt2actr* is publicly available on github.⁴

³The image in the red box is from Anderson and Borst (2017).

⁴<https://github.com/eadietz/txt2actr>

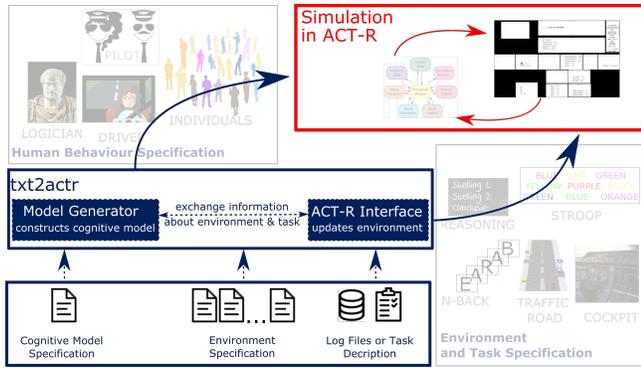


Figure 2: Overview of the two main components in *txt2actr*.

Figure 2 shows the two main components of *txt2actr* in the blue box (labeled *txt2actr*), where the *ACT-R interface* (on the right) and the (*partial*) *model generator* (on the left) interact and depend on each others specifications. The necessary environment specifications and cognitive model specifications can be defined in the respective text files that are in the blue box below the labeled *txt2actr* box. The task descriptions or the log files about the environment and the cognitive model file replace the specifications of the individual environment and tasks (on the lower right corner) and the individual human behavior (upper left corner). The decision to structure *txt2actr* this way is driven by the idea that ideally, a set of cognitive models could then be systematically evaluated with respect to a set of various tasks. Whether such a generic parametrization of tasks or models will eventually be possible needs to be further investigated. Additionally, we aim at a lower inhibition threshold for cognitive modeling, which should be usable by the ones with simulation needs but with little or no experience with cognitive modeling.

We chose the cognitive architecture ACT-R as basis our purpose as it provides a wide range of functionality, it is well established within the community and has a very well documented manual (Bothell, 2020) including an extensive tutorial. ACT-R (Anderson, 2007) is a theory about how human cognition works. It allows to get a better understanding human cognition by simulating different cognitive functions. Each function is represented by a particular (and independent) module that communicates with other modules through buffers. Knowledge in ACT-R is either encoded as declarative memory or procedural rules. Cognitive architectures are also used for tasks within the real world, such as aircraft cockpit or car driving environments (Salvucci, 2006).

We heavily rely on the already existing and publicly available python module *actr.py* which allows a direct interaction within the ACT-R environment in Lisp through python.⁵

ACT-R Environment Interface

We will briefly introduce the two most important aspects that affect the dynamics of the (ACT-R) environment and can be

⁵The tutorial and *actr.py* can be found here <http://act-r.psy.cmu.edu/software/>

vision	description	y position	x position	width	length	fontsize	x text position	y text position	header labels
leading_car	showing red when leading car brakes	255	410	550	250	12	25	25	car_normal.car_brake.car_hold
own	showing own behavior	590	410	250	150	12	50	50	clock.brake.react_emg.car_collision
clock	showing the time	590	710	160	80	12	15	15	4 time

image	description	y start	x start	y end	x end	color	window	label
backlight_1	should appear when leading car brakes	90	10	180	100	back-light.gif	leading_car	car_brake
backlight_2	should appear when leading car brakes	90	400	180	100	back-light.gif	leading_car	car_brake

button	y position	x position	height	width	action	window	label
button_1	10	10	60	60	own_display	brakes	

Figure 3: The specifications in *txt2actr* for the positions of the windows, the images and the buttons in ACT-R for a driving environment use case. Further specifications such as which items should appear when and where (depending on the log files) are also possible.

modified by *txt2actr* during the simulation:

Visual Scene Everything that can be perceived by the cognitive model through its vision module, such as numbers, text, geometric figures, buttons or images from external sources can be specified by their sizes, location and colors (if applicable).

Audio Scene Everything that can be perceived by the cognitive model through its audio module, such as tones, (spoken) words or digits can be specified by their volume, type of tone or duration.

In order to make sure that the information about the environment is displayed at the intended time, the ACT-R interface in *txt2actr* uses the *schedule_time* function provided by *actr.py*. As illustration consider a cockpit environment in an aircraft: In flight, some of the values shown in the cockpit displays in the visual scene will permanently change according to the specifications of a given log file. Additionally, some tone or sound in the audio scene might occur as well. All of the above described components can be specified within the respective text files in *txt2actr*.

Figure 3 shows a screenshot of the environment specification files from a very simplified car driving use case. The full environment specification of this use case can be found in the *use-cases/driving-task* folder in *txt2actr*.

So far we have not discussed how the model perceives its environment. Thus, possibly even when a tone appears in the model’s acoustic scene or a value is shown in one of the windows at a certain time, it might well be that the cognitive model does not perceive it. The model behavior depends on its specification which we will discuss in the next section.

Partial Model Generator

Before we lay out our understanding of the model generator, we introduce the concept of cognitive principles, as they form the foundation for the generation of models.

Cognitive Principles are *cognitively plausible explanations* for some episodes of human behavior, which can be anything from biases, heuristics, judgments or even decision making and reasoning. In particular, cognitive principles are

```
(add_dm
;; The location specification for each item (label) value
(car_normal-info isa display-info name car_normal screen-x 442 screen-y 334)
(car_brake-info isa display-info name car_brake screen-x 442 screen-y 310)
(car_hold-info isa display-info name car_hold screen-x 442 screen-y 286)
(backlight_1-info isa image-info name backlight_1 screen-x 470 screen-y 435)
(backlight_2-info isa image-info name backlight_2 screen-x 860 screen-y 435)
(car_collision-info isa display-info name car_collision screen-x 467 screen-y 694)
(react_emg-info isa display-info name react_emg screen-x 467 screen-y 670)
(time-info isa display-info name time screen-x 732 screen-y 648)
;; the list of items that are to be attended in a routine loop
;; in case of (visual) goal-driven information processing
(car_normal-0 ISA list-info current-on-list car_normal next-on-list car_brake)
(car_brake-1 ISA list-info current-on-list car_brake next-on-list car_normal)
)
```

Figure 4: The environment specifications of Figure 3 reappear as declarative knowledge in the cognitive model.

not necessarily in line with *rational* or (*classical*) *logical* reasoning, but rather demonstrate naturalistic thinking in everyday life. To some extent, the identification of these psychological phenomena are one of the main motivators in the field of Cognitive Science. As many decades of research show, these psychological phenomena are very insightful for the understanding of the human mind, but at the same time extremely difficult to specify unambiguously during the development of the corresponding model. Cognitive principles are modular and formalized approximations of these phenomena. It is likely that a psychological phenomena can be formalized in various ways, and that there is no agreement on their formal representations. In this case, each of the formalization is an instance of them. We intend to specify a catalogue of cognitive principles, each of them as a module, such that they apply independently of each other. We are aware that it various cognitive phenomena influence each other, or only have an effect when applied together. However, for the goal of benchmarking, i.e. construct a benchmark of cognitive models out of these principles, it is helpful to consider them separately, similarly has been done for the case of human syllogistic reasoning in (Dietz Saldanha & Schambach, 2020). Different to the approaches from the previous section, we do not try to identify new cognitive phenomena but rather better understand, formalize and classify already well-established ones. In the first step, cognitive principles are formalized as abstracted and modular entities, and in the second step, during the model construction, they are instantiated with respect to a given task or environment. In the ideal case, a model can be simply specified through the underlying cognitive principles that it assumes.

Modular chunk and production engineering takes place in two steps:

1. Modular specification of the model’s properties through cognitive principles.
2. Model construction by instantiating the specification with respect to the environment and the task.

Fortunately, the ACT-R architecture itself consists of a set of modules, which allow us to naturally specify the different components in a modular way. As already mentioned, knowledge can be either represented declaratively, by means

of chunks that belong to certain chunk types with a set of slot configurations. For instance,

(chunk-type display-info name screen-x screen-y)

is a *display-info* that might have the slots *name*, *screen-x*, and *screen-y*. A chunk can then be understood as an instantiation thereof, such as

**(ALTITUDE-info isa display-info
name ALTITUDE screen-x 389 screen-y 514)**

tells us at which coordinates we can find the current altitude. The automatic construction of the initial chunk types and chunks is based on our assumption that the model has some basic knowledge about the environment it interacts with. By default, there will be four different chunk types: *display-info*, *button-info*, *image-info* and *sound-info*. These chunks are derived from the task specification and can be used to model familiarity with a task environment such as knowledge about positions and functions of buttons in an aircraft cockpit. Consider again the environment specification files in Figure 3: If not specified otherwise in the cognitive model specification file, then *txt2actr* will automatically include these items and their coordinates into the declarative knowledge of the cognitive model.

Figure 4 shows a screenshot of the declarative knowledge from the very simple car driving use case.

Two Cognitive Processes in three Environments

txt2actr also allows to specify initial *procedural knowledge*, which in ACT-R is done by means of production rules. These production rules can be read as conditional statements, where, in case the condition holds, the consequence will be executed. The automated construction of an initial set of production rules is more complex, in particular if aiming at constructing them independent of the task. Therefore, we will illustrate the modularity of *text2actr* by implementing two cognitive processes from information processing and test them in three different environments.

Situation Awareness We will consider two essential processes for modeling situation awareness (Freiman, Myers, Caisse, Halverson, & Ball, 2019). Situation awareness describes to which extent someone has perceived and understood vital elements of a situation for completing the task at hand (Endsley, 2015). According to Endsley (2015) and Freiman et al (2019), a good model of situation awareness needs to account, among others, for the alternation between data-driven and goal-driven information processing, which both can be understood as two distinctive cognitive principles: In the case of *data-driven information processing*, visual attention should be guided by changes in the environment, while for *goal-driven information processing* the model actively engages in search of specific information by means of coordinates. In both cases, the model keeps an updated representation of the values of the attended items. Figure 5 illustrates this idea by the cockpit environment: Goal driven

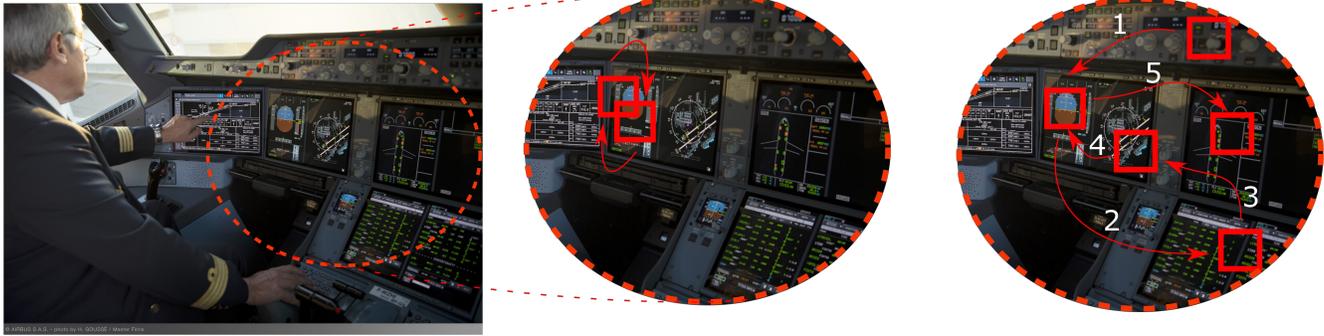


Figure 5: A real cockpit (left), and two ways on how visual information could be attended (middle and right), where the red squares and the arrows denote a possible sequence in which locations are attended.

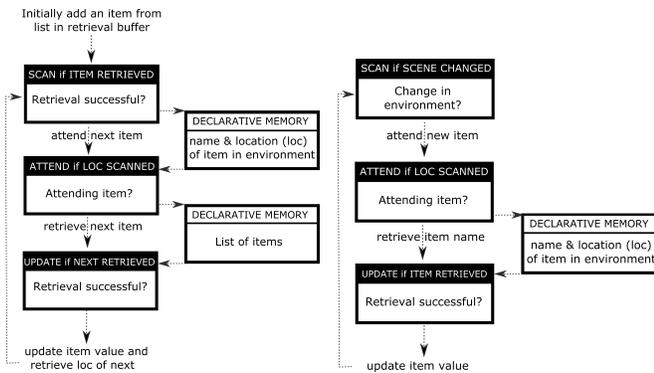


Figure 6: The goal-driven (left) and the data-driven information processing (right) components in ACT-R. Each box with a black header represents a production rule.

information processing (middle) could be where the visual attention alternates between two locations. An example of data-driven (or event-driven) information processing (right) would be where the visual attention is determined by any newly appearing item on the displays (the arrows denote a possible sequence of attention).

Both cases of information processing are implemented as abstracted modular components in *txt2actr*, each represented by a set of general production rules. Additionally, for the goal-driven case, the list of the items to be attended (based on the text files) is automatically generated and added to the declarative memory. An example of such a list for the car driving environment is shown in the last two lines of Figure 4. In case one or both components are chosen to be part of the model, *txt2actr* will create a model with these component(s) instantiated with respect to the specified environment.

The lisp files of both model components in ACT-R with the respective names (data-driven.lisp and goal-driven.lisp) are part of *txt2actr*.⁶ Figure 6 shows a description of the processes: Initially, for the goal-driven information processing

(left), an item from the list is set into the retrieval buffer. When the retrieval of the current item’s location (first production rule) was successful, then this location is attended and the next item on the list is retrieved. Finally, the current item’s value is updated, and process starts again by retrieving the location of the next item. In case of data-driven information processing (right), the first rule only fires when the scene changes, for example when new values appear in the visicon. The new item is attended and based on its location, its name is retrieved. Finally, similar to the previous case, this item’s value is updated in the model’s memory. These model components are not use case specific and thus generally applicable.

For testing whether both model components would behave as intended, both individually and together, we have chosen and specified three very simple but different environments. These three environments were specified exclusively through the provided text files in *txt2actr* (which are contained in the folder environment-specification of each use case). The first one is the *paired associates task* (Anderson, 1981) where the environment and model specification are adapted from the ACT-R tutorial.⁵ Note that we intentionally chose for a task in which the model has a different purpose in order to observe how the model behaves. The second and third examples are about the simulation of real-world scenarios: A driving environment and a cockpit environment. Different than in the first example, the values in their environment continuously change according to the log files. Furthermore, the model does not have any other task to accomplish except of updating values in its own memory according to the two cases of the above described visual information processing. The driving environment is built on data from an empirical study originally by Haufe et al (2011) and takes processed extracts of datasets from BNCI Horizon as log file input.⁷ The cockpit environment takes as log files extracts from Dashlink.^{8,9}

⁵The used dataset (VPae.mat) can be found here: <http://bnci-horizon-2020.eu/database/data-sets> (30.4.21)

⁸<https://c3.nasa.gov/dashlink/projects/85/> (30.4.21)

⁹The dataset can be found here <https://c3.nasa.gov/>

⁶<https://github.com/eadietz/txt2actr/tree/master/benchmarks/model-components>



Figure 7: Simulation of paired-associates task (left), driving environment (middle) and cockpit environment (right) in ACT-R. The red dot in the middle and right image shows the model’s visual focus.

Observations Figure 7 shows a screenshot of the simulation of each environment. In the driving environment, different values change continuously, while in the paired-associates task, only one value changes, and this happens only occasionally. Therefore, naturally the data-driven information processing is fired more often in the driving environment than in the paired-associates task and even more often in the cockpit environment. Interestingly, in the compound model consisting of both visual processing components the production rules of the data-driven information processing do not apply anymore. Only by the specification of high utilities for this component the production rules of both components apply. This leads us to the more general question of how such a compound model of visual processing should behave. On the one hand the model should be able to pursue goal-driven visual behavior while being sensitive to new stimuli that compete for visual attention. When goal-driven behavior does not occupy all buffers, these buffers are available to being used by salient, not goal-related stimuli which can lead to distractions and mind wandering (Taatgen et al., 2021). While utility functions can help modeling commitment to goals or susceptibility to distraction, we believe that modeling of attentional control should recognize the interplay of cognitive resources and the environmental factors such as the salience of stimuli (e.g., alert sounds in the cockpit). It is very likely, that other (ACT-R) modelers would have implemented the above components differently in ACT-R or even diverge from the processes shown in Figure 6. The novelty of our approach is not to demonstrate that the proposed components are the most cognitively plausible ones, but rather that we can build abstracted modular entities of these components, which can be instantiated with respect to different environments.

Conclusions

This paper proposes a possible path for benchmarking cognitive models. Yet, we are far from providing a solution but rather show how a very specific cognitive phenomenon might be applicable to different environments. Already modeling concrete aspects of visual information processing in ACT-R leads to plenty of choice points on the implementation side that are not specified in the theories of situation awareness.

dashlink/static/media/dataset/Tail.687.9.zip (30.4.21)

However, we believe that taking the effort to approximate theories by implementing instances thereof as models could be beneficial to identify under specifications that might not be immediately obvious.

Our contributions with *txt2actr* are two-fold: First, the specification of an ACT-R environment can be done through text files. Second, we have shown that it is possible to formulate abstracted entities of cognitive phenomena from which model components can be automatically generated. However, we are very aware that this process needs to be built with care and based on more objective criteria for cognitive plausibility or consensus. Therefore, we also need to find a more accessible way of individually assembling cognitive models, possibly guided by a catalog of cognitive principles usable for modular and guided model construction. Currently, we only consider the (partial) automation of initial chunk types, initial chunks and initial productions. The general parameter settings, additional chunk types, chunks, productions or other commands can be specified manually via a text file. For the future, a systematic account on producing general parameter settings might be considered as well. We also argue that some choices (or strategies) do not need to be taken before modeling a task, but can be taken at a later stage (e.g. evaluation of different strategies by instantiating different models for the same task), should be able to include different theories and allow a systematic comparison between modeling approaches.

Finally, we believe that establishing benchmarks will promote (i) the competition among models with respect to the most typical tasks in cognitive psychology, and (ii) the evaluation of (possibly new) tasks with respect to benchmark models. This might help the community to address the previously mentioned issues and eventually unify the field.

References

Amant, R. S., McBride, S. P., & Ritter, F. E. (2006). AI support for building cognitive models. In *Proceedings of the twenty-first national conf. on artificial intelligence and the eighteenth innovative applications of artificial intelligence conf.* (pp. 1663–1666). AAAI Press.

Amant, R. S., & Ritter, F. E. (2004). Automated goms-to-act-r model generation. In *Proceedings of the sixth international conf. on cognitive modeling (iccm)* (pp. 28–34).

- Anderson, J. R. (1976). *Language, memory, and thought / john r. anderson*. L. Erlbaum Associates ; distributed by the Halsted Press Division of Wiley Hillsdale, N.J. : New York.
- Anderson, J. R. (1981). Interference: The relationship between response latency and response accuracy. *Experimental Psychology: Human Learning and Memory*, 7, 326-343.
- Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* Oxford University Press.
- Borst, J. P., & Anderson, J. R. (2017). A step-by-step tutorial on using the cognitive architecture act-r in combination with fmri data. *Mathematical Psychology*, 76, 94–103.
- Bothell, D. (2020). *ACT-R 7.21+ Reference Manual*. (<http://act-r.psy.cmu.edu/actr7.x/reference-manual.pdf>, accessed on 22.4.21)
- Card, S., Moran, T., & Newell, A. (1983). *The psychology of human-computer interaction*. CRC.
- Dietz Saldanha, E., & Schambach, R. (2020). A computational approach for predicting individuals' response patterns in human syllogistic reasoning. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42th annual meeting of the cognitive science society*. cognitivesciencesociety.org.
- Endsley, M. R. (2015). Situation awareness misconceptions and misunderstandings. *Cognitive Engineering and Decision Making*, 9(1), 4-32.
- Freiman, M., Myers, C. W., Caisse, M., Halverson, T., & Ball, J. T. (2019). Assessing cognitive fidelity in a situation awareness process model. In G. L. Rogova, N. M. McGeorge, O. E. Gundersen, K. Rein, & M. D. Freiman (Eds.), *IEEE conf on cognitive and computational aspects of situation management, cogsima 2019* (pp. 100–106). IEEE.
- Gebser, M., Maratea, M., & Ricca, F. (2020). The seventh answer set programming competition: Design and results. *Theory and Practice of Logic Programming*, 20(2), 176–204.
- Haufe, S., Treder, M. S., Gugler, M. F., Sagebaum, M., Curio, G., & Blankertz, B. (2011, jul). EEG potentials predict upcoming emergency brakings during simulated driving. *Neural Engineering*, 8(5), 056001.
- Khemlani, S., & Johnson-Laird, P. (2012). Theories of the syllogism: A meta-analysis. *Psychological bulletin*, 138 3, 427-57.
- Laird, J. E. (2012). *The soar cognitive architecture*. The MIT Press.
- Laird, J. E., Lebiere, C., & Rosenbloom, P. S. (2017, December). A Standard Model of the Mind: Toward a Common Computational Framework across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics. *AI Magazine*, 38(4), 13.
- Lebiere, C., Pirolli, P., Thomson, R., Paik, J., Rutledge-Taylor, M., Staszewski, J., & Anderson, J. R. (2013). A functional model of sensemaking in a neurocognitive architecture. *Computational Intelligence and Neuroscience*, 2013, 921695:1–921695:29.
- Lewandowsky, S. (1993). The rewards and hazards of computer simulations. *Psychological Science*, 4(4), 236-243.
- Marewski, J. N., & Mehlhorn, K. (2011). Using the act-r architecture to specify 39 quantitative process models of decision making. *Judgment and Decision Making*, 6(6), 439-519.
- The nature and transfer of cognitive skills. (2013, 06). *Psychological review*, 120, 439-471. doi: 10.1037/a0033138
- Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In *Visual information*. New York: Academic Press.
- Newell, A. (1990). *Unified theories of cognition*. USA: Harvard University Press.
- Ragni, M. (2020). Artificial intelligence and high-level cognition. In P. Marquis, O. Papini, & H. Prade (Eds.), *A guided tour of artificial intelligence research* (Vol. 3, pp. 457–486). Germany: Springer.
- Ragni, M., Riesterer, N., & Khemlani, S. (2019). Predicting individual human reasoning: The precore-challenge. In A. K. Goel, C. M. Seifert, & C. Freksa (Eds.), *Proceedings of the 41th annual meeting of the cognitive science society, cogsci 2019: Creativity + cognition + computation, montreal, canada, july 24-27, 2019* (pp. 9–10). cognitivesciencesociety.org.
- Riesterer, N., Brand, D., & Ragni, M. (2020). Predictive modeling of individual human cognition: Upper bounds and a new perspective on performance. *Topics in Cognitive Science*, 12(3), 960-974.
- Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors*, 48(2), 362-380.
- Salvucci, D. D. (2013). Integration and reuse in cognitive skill acquisition. *Cognitive Science*, 37(5), 829–860.
- Taatgen, N., & Anderson, J. R. (2010). The past, present, and future of cognitive architectures. *Topics in Cognitive Science*, 2(4), 693-704.
- Taatgen, N., van Vugt, M. K., Daamen, J., Katidioti, I., Huijser, S., & Borst, J. P. (2021). The resource-availability model of distraction and mind-wandering. *Cognitive Systems Research*, 68, 84-104.
- Thomson, R., Lebiere, C., Anderson, J. R., & Staszewski, J. (2015). A general instance-based learning framework for studying intuitive decision-making in a cognitive architecture. *Applied Research in Memory and Cognition*, 4(3), 180-190. (Modeling and Aiding Intuition in Organizational Decision Making)
- West, R., & Nagy, G. (2007). Using goms for modeling routine tasks within complex sociotechnical systems: Connecting macrocognitive models to microcognition. *Cognitive Engineering and Decision Making*, 1(2), 186-211.
- West, R., & Somers, S. (2011). Scaling up from micro cognition to macro cognition: Using sgoms to build macro cognitive models of sociotechnical work in act-r. *Cognitive Science*, 33.

Computationally Rational Reinforcement Learning: Modeling the Influence of Policy and Representation Complexity

Zeming Fang (fangz5@rpi.edu) & Chris R. Sims (simsc3@rpi.edu)

Department of Cognitive Science, Rensselaer Polytechnic Institute
110 8th St, Troy, NY 12180 USA

Abstract

In recent years, several models of human reinforcement learning have been proposed that balance rationality (maximizing expected utility) against cognitive costs. Lai and Gershman (2021) proposed a model in which the cognitive cost was assumed to be the policy complexity, defined in terms of information theory as the mutual information between the sensory input and behavioral response. Here, using evidence from a published data set (Collins & Frank, 2012), we show that this model fails to account for the “set size effect” in learning: humans’ learning efficiency decreases when the number of the presented stimuli increases. We therefore propose an alternative computational model, in which cognitive cost constitutes not only the policy complexity, but also the representation complexity—the amount of information conveyed from sensory inputs to internal representations. We quantify information processing cost as the combination of representation complexity and policy complexity. The resulting model captures the set size effect in an instrumental learning paradigm.

Keywords: Computationally rational; Reinforcement Learning; Information theory; Set size effect

Introduction

Human working memory is known to be capacity limited. A well-established consequence of this is the *set size effect*—namely, humans’ memory performance systematically decreases as the number of items to be stored in memory increases (Ma, Husain, and Bays (2014)). Much existing work has sought to quantify what is meant by working memory capacity and explain the set size effect. One example is the work by Sims (2016), who formalized working memory capacity as a limited pool of information quantity that enables a cognitive function (e.g., store a stimulus) or process (e.g. making decisions). The information resource can be subdivided into portions, and more items to be stored implies less resource allocated to encode each item, resulting in lower recall precision per item. Up until now, however, most research on working memory has not examined how these limits might impact other cognitive systems.

Collins and Frank (2012) and Collins, Brown, Gold, Waltz, and Frank (2014) studied how working memory limits impact humans’ reinforcement learning (RL). They reported an analog of the set size effect in an instrumental learning paradigm, and showed that a standard RL model (M^{RL} model in this article) cannot capture this phenomenon.

Gershman and Lai (2020) reexamined Collins et al. (2014), and proposed a computationally rational (Gershman, Horvitz, & Tenenbaum, 2015) account of humans’ suboptimal learning performance. The mathematical framework they used is

known as *rate distortion theory* (Berger, 1971). This framework provides the tools for predicting the highest achievable performance under a given information capacity constraint, and hence is directly applicable to explaining human learning performance under a limited pool of (information-theoretic) resource. They considered the capacity constraint as *policy complexity* (Tishby & Polani, 2011; Still & Precup, 2012; Lerch & Sims, 2018), which measures the rate of information extracted from states and transmitted to actions. They concluded that in general human participants optimized this reward-policy complexity trade-off, and humans’ suboptimal performance can be understood as a compromise to limited policy complexity.

While interpreting humans’ suboptimal performance, Gershman and Lai (2020) did not explicitly address how the set size effect emerges in human learning. This article seeks to fill this gap. Intuitively, one expects that when learning in larger set size conditions, the overall cognitive cost is higher, and hence humans will rationally trade task performance against rising cognitive costs. By analyzing the data set in Collins and Frank (2012), we show that policy complexity does not suffice as an explanation for human behavior: the policy complexity to reach optimal performance does not necessarily increase with the set size. This observation also violates humans’ experience that the larger set size task is more difficult. This implies that the policy complexity is not sufficient for cognitive cost, other complementary constitutions are needed. We then considered another information notion, *representation complexity* (Tishby & Polani, 2011; Genewein, Leibfried, Grau-Moya, & Braun, 2015; Zenon, Solopchuk, & Pezzulo, 2019), measuring information transmitted about environmental state to an agent’s internal representation.

Directly measuring internal representations is a notoriously difficult problem because they are latent constructs. In this article, we resort to a model-based analysis to understand how the set size impacts human learning performance. We compare three classes of models: a standard RL model as a benchmark, two RL models with policy complexity adopted from Lai and Gershman (2021) to show the failure of policy complexity, and another two that explains cognitive cost as the summation of both representation complexity and policy complexity to interpret the set size effect.

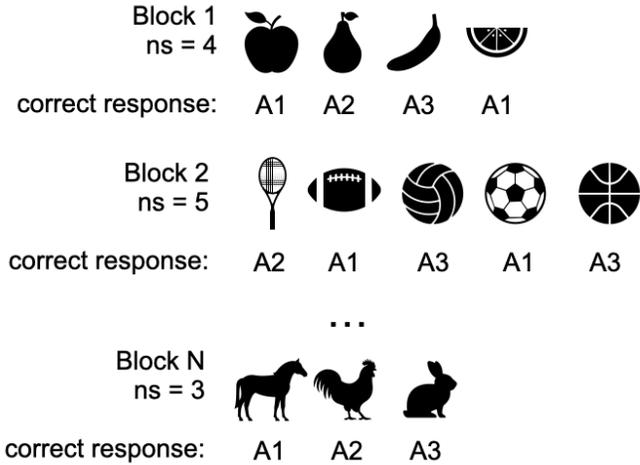


Figure 1: Schematic of experimental task studied by Collins and Frank (2012). On each trial, subjects were shown one single stimulus and were instructed to choose one of three actions. Each stimulus corresponded to one correct action and the number of stimuli varied across blocks. Note that the stimuli shown here are for illustrative purposes and are not the actual stimuli used in the experiment.

Methods

Data set

We tested a series of models on the data set reported in Collins and Frank (2012). This data set consists of 78 subjects' learning performance in a multi-armed bandit task. On each trial, subjects were shown one single visual stimulus (drawn from categories such as sports, fruits, etc.) and were instructed to quickly choose a key among three alternatives. Each response was followed with a binary outcome, either 1 (reward) or 0 (no reward). For each stimulus, the reward was deterministically associated with only one of the three responses. All stimuli were repeated 9-15 times within a block, and did not appear across blocks. The set size ns (the number of different stimuli within a block) systematically varied across blocks, ranging from 2 to 6 (Figure 1). Each subject completed 19 blocks, six in which with $ns = 2$, four with $ns = 3$, three blocks each with $ns = 4, 5$, or 6. See Collins and Frank (2012) for complete details.

Computational rationale of policy complexity

In the standard RL scenario, decision-making involves two variables: *environmental state* S and the *action* A . In the language of information theory, we can think of this cognitive process as an information channel: a policy $\pi(a|s)$ that maps the environmental states S onto a probability distribution over actions A . According to information theory, the average computational demands necessary to convey information over this 'policy channel' is equal to the *mutual information* between

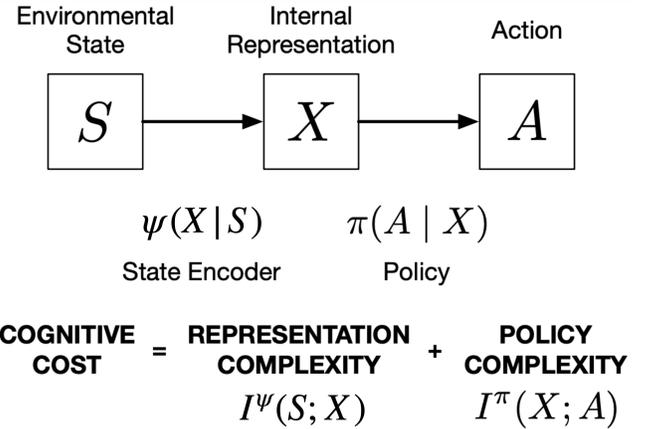


Figure 2: Schematic of cognition process. The biological sensory signal of the input stimuli S are encoded to internal mental representation X , and based on which human make decisions A .

state and action, the general equation of which is:

$$I(Y; Z) = \sum_i p_y(y_i) \sum_j p_{z|y}(z_j|y_i) \log \frac{p_{z|y}(z_j|y_i)}{p_z(z_j)} \quad (1)$$

where Y means the sender and Z , the receiver. The calculation of mutual information requires us to know the marginal distribution of both variables, p_y and p_z , as well as the channel statistics, $p_{z|y}$.

Gershman and Lai (2020) and Lai and Gershman (2021) considered the mutual information $I^\pi(S; A)$ as policy complexity. This is correct with the implicit assumptions that humans have full access to the environmental state (Tishby & Polani, 2011) and that they do not rely on internal representations of stimuli. Under these two assumptions, human decision-making is much like the stimulus-response (S-R) mapping in classic behaviorism.

Computational rationale of representation complexity

Instead of considering humans' decision-making as an S-R process, we introduce a third construct: the encoded *internal representation* of the state, X . Humans may now respond A to the given representation X . We may now consider the whole decision process as a cascade information channel (Figure 2): a state encoder $\psi(x|s)$ that maps the environmental states S onto a probability distribution over internal representations X , followed by a policy $\pi(a|x)$ that maps the mental states X to a distribution over actions A . The mutual information $I^\psi(S; X)$ is considered as representation complexity and the policy complexity is now $I^\pi(X; A)$.

The advantage of introducing the representation is to allow the emergence of abstractions, which is thought of as a hallmark of intelligence (Kemp, Perfors, & Tenenbaum, 2007;

Gershman & Niv, 2010). When the environment is very complicated with an unaffordable information cost, an adaptive agent can cluster environmental states with a similar policy to lower the information cost during the state encoding stage (Genewein et al., 2015). However, the goal of this article is to identify what constitutes the cognitive cost, and the formation of adaptive representations is beyond our focus. In the present paper, we implemented a simple fixed state encoder ψ (see Models section for details).

Models

RL baseline: M^{RL} We use the RL baseline from Collins and Frank (2012). The computational goal of the RL baseline model is to find a policy that maximizes the expected total reward over all trials within a block,

$$\max_{\pi} E[r_t | p_s, \pi] \quad (2)$$

where p_s represents the prior knowledge about the state distribution and it is a uniform distribution in this experiment, in keeping with the experiment design where stimuli are uniformly sampled. r_t is the reward subjects received at trial t .

To achieve this, the model learns a state-action value $Q(s, a)$ and a policy $\pi(a|s)$ to guide action selection. Both the Q function and the policy are updated after each trial t . The update of the Q function follows:

$$Q^t(s_t, a_t) = Q^{t-1}(s_t, a_t) + \alpha_q [r_t - Q^{t-1}(s_t, a_t)] \quad (3)$$

where α_q is the learning rate. The s_t and a_t are the observed current state and action. We use the superscript t to note temporally changing variables.

To balance exploration and exploitation in the RL baseline model, the policy is formalized as the output of the softmax function of the most recent Q value,

$$\pi^t(s, a) = \frac{\exp[\beta Q^t(s, a)]}{\sum_j \exp[\beta Q^t(s, a_j)]} \quad (4)$$

where $\beta \geq 0$ is the inverse temperature parameter that controls the degree of stochasticity in the policy (Sutton & Barto, 2018).

The only parameters for the RL baseline are the learning rate α_q and the inverse temperature β for the policy. To apply the model to behavioral data, we fit both parameters via maximum likelihood estimation.

Policy complexity: $M_{(1)}^{\pi}$ For the model that considers policy complexity, the computational goal is to maximize expected utility while ensuring that the policy complexity does not exceed a fixed capacity limit:

$$\max_{\pi} E[r_t | p_s, p_a, \pi] \quad s.t. \quad I^{\pi}(S; A) \leq C \quad (5)$$

where p_a is the marginal action distribution and C denotes the channel capacity—the maximum available cognitive resource. Equation 12 can be rewritten in a Lagrangian form:

$$\max_{\pi} \beta E[r_t | p_s, p_a, \pi] - I^{\pi}(S; A) \quad (6)$$

where $\beta \geq 0$ regulates the tradeoff between external reward and policy complexity. When $\beta \rightarrow \infty$, the agent can be considered fully rational; when $\beta \rightarrow 0$, the agent sticks with its prior policy p_a .

To solve equation 6, we use the gradient-based process model developed in Lai and Gershman (2021). For more details, see (Lai & Gershman, 2021, appendix). This is an ‘‘actor-critic’’ model using the ‘‘policy gradient’’ algorithm (Sutton & Barto, 2018) to incrementally update the parameterized policy π_{θ} (the parameters of which are θ) and value function V_w (the parameters of which are w). In the original paper, all parameters are initialized as 0. while in this article we initialized the value parameters w as 1 as it provided a better fit to the data.

In each timestep t , the model first estimates the value of the current state s_t and the current policy of the state:

$$\hat{V}_w(s_t) = w^{t-1} \cdot \mathbb{I}(s_t) \quad (7)$$

and

$$\hat{\pi}_{\theta}(a|s_t) = \exp(\beta \theta^{t-1} \cdot \mathbb{I}(s_t) + \log p_a(a)) \quad (8)$$

where $\mathbb{I}(\cdot)$ is the indicator function that returns an one-hot encoding of the input. Note that $\hat{\pi}_{\theta}(a|s_t)$ is a distribution over action a .

The model update the critic using :

$$w^t = w^{t-1} + \alpha_w \mathbb{I}(s_t) \delta \quad (9)$$

where α_w means the learning rate of the parameters of the value function, and $\delta = \beta r_t - \log \frac{\hat{\pi}_{\theta}(a_t|s_t)}{p_a^{t-1}(a_t)} - \hat{V}_w(s_t)$ is the prediction error.

The update of the critic is divided into two sub-steps. The first step is to update the policy:

$$\theta_{a_t}^t = \theta_{a_t}^{t-1} + \alpha_{\theta} \mathbb{I}(s_t) \beta [1 - \hat{\pi}_{\theta}(a_t|s_t)] \quad (10)$$

where α_{θ} is the learning rate of the policy parameter and $\theta_{a_t}^t$ means the parameters for action a_t . The update of the policy is followed by the update of the marginal action distribution:

$$p_a^t(a) = p_a^{t-1}(a) + \alpha_a [\hat{\pi}_{\theta}(a|s_t) - p_a^{t-1}(a)] \quad (11)$$

To use this model, we need to fit four hyperparameters $\{\alpha_w, \alpha_{\theta}, \alpha_a, \beta\}$. Note this model is a general case of the RL baseline. When $\alpha_a = 0$ and p_a is a uniform distribution, this model collapses to a policy gradient variant of RL baseline.

Policy complexity: $M_{(2)}^{\pi}$ As shown in the results section below, the basic policy complexity model ($M_{(1)}^{\pi}$) does not predict a set size effect in human learning. The limitation stems from assuming a single tradeoff parameter (β) for all set sizes. To avoid this limitation we can fit a specific β to set size ns .

Thus, $M_{(2)}^{\pi}$ is exactly the same as $M_{(1)}^{\pi}$ except it has eight hyperparameters $\{\alpha_w, \alpha_{\theta}, \alpha_a, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6\}$. The subscript of the β represents the set size the β is fit to.

Representation + policy complexity: $M_{(1)}^{\psi+\pi}$ This model considers the cognitive cost as summation of both representation complexity and policy complexity, as illustrated in Figure 2. The resulting objective is

$$\max_{\pi} E[r_t | p_s, p_x, p_a, \psi, \pi] \quad s.t. \quad I^{\psi}(S;X) + I^{\pi}(X;A) \leq C \quad (12)$$

and the corresponding Lagrangian form is,

$$\max_{\pi} \beta_{ns} E[r_t | p_s, p_x, p_a, \psi, \pi] - I^{\psi}(S;X) - I^{\pi}(X;A) \quad (13)$$

where p_x means the prior belief about the internal representations and is assumed as a uniform distribution. Representations X are generated probabilistically according to the state encoder $\psi(x|s)$. For example, "apple" and "orange" may evoke very distinct sensory representations, but are mapped to one latent representation because they both have the same optimal response (and hence are functionally, if not perceptually, equivalent). In this paper, we implemented a simple (non-adaptive) model for the state encoder ψ inspired by the ϵ -greedy policy in RL (Sutton & Barto, 2018). An environmental state s has $1 - \epsilon$ probability to be recognized as s and has $\frac{\epsilon}{|S|-1}$ probability to be recognized as any of stimuli other than s . Increasing ϵ increases the "noise" in the state encoder, and hence reduces its information-theoretic channel capacity. The motivation behind this design is that we need a noisy categorical distribution (environmental state s is recorded as a categorical variable in the data) that may collapse to a one-hot encoding (the indicator function $\mathbb{I}(\cdot)$ in equation 8, assuming humans participants had full access to the environmental state) if humans are really optimal state encoders. If the fitted ϵ is 0, we may conclude humans develop perfect representations for the external stimuli in this simple experiment paradigm.

To implement a gradient-based RL model with a state encoder ψ , we only need to change the indicator function of the state indicator function $\mathbb{I}(s_t)$ to $\psi(x|s_t)$. The nine hyperparameters of this model are $\{\alpha_w, \alpha_{\theta}, \alpha_a, \epsilon, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6\}$. When $\epsilon = 0$, $M_{(1)}^{\psi+\pi}$ collapses to $M_{(1)}^{\pi}$.

Representation + policy complexity: $M_{(2)}^{\psi+\pi}$ The previous model utilized a gradient-based optimization procedure to achieve the learning objective. We also tested a gradient-free normative model based on Tishby and Polani (2011) and Genewein et al. (2015). The model is built upon RL baseline M^{RL} with a same critic formulation and update rule.

The actor component of the model is conditional on the internal representation x and action a . Since we have no access to the latent representation in the observed data, we can only infer the representation-action value function $Q_{bel}(x, a)$ following (Genewein et al., 2015),

$$Q_{bel}(x, a) = \sum_x p(s|x) Q^t(s, a) \quad (14)$$

where $p(s|x) = p_s(s)\psi(x|s)/p_x(x)$ is the Bayesian posterior over s given x and ψ follows the same design with $M_{(1)}^{\psi+\pi}$.

With the representation-action function $Q_{bel}(x, a)$, we can formulate the optimal update of the actor as,

$$\pi^t(a|x) = \frac{\exp[\beta_{ns} Q_{bel}^t(x, a) + \log p_a^{t-1}(a)]}{\sum_j \exp[\beta_{ns} Q_{bel}^t(x, a_j) + \log p_a^{t-1}(a_j)]} \quad (15)$$

This is the optimal policy update for a given value function (Tishby & Polani, 2011). In contrast to the gradient-based update of the previous model, this model would be expected to learn more quickly.

The update of marginal policy p_a follows equation 11. The hyperparameters of $M_{(2)}^{\psi+\pi}$ are $\{\alpha_q, \alpha_a, \epsilon, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6\}$.

Optimal policy As a benchmark for evaluating our models, we also determined the optimal policy for a learning agent. To achieve the optimal solution, we can simply use the RL baseline model with $\alpha_q = 1$ (high learning rate) and $\frac{1}{\beta} = 0$ (no exploration). This is a consequence of the particular task environment, as there is exactly one action that is deterministically rewarded for each stimulus.

Results

Model fits and the set size effect

Figure 3 compares human and model learning curves. As expected, the RL baseline (M^{RL}) does not reproduce the set size effect. More surprisingly, a model incorporating policy complexity ($M_{(1)}^{\pi}$) also fails to account for this effect. This model utilizes a fixed utility-complexity tradeoff parameter (β) for all set sizes. Model $M_{(2)}^{\pi}$ fits separate parameters for each set size, but offers no explanation as to why this parameter should differ according to set size. The models that incorporate both policy complexity and representation complexity were able to demonstrate the set size effect.

Table 1: Models' goodness-of-fit.

-	NLL	SSE
M^{RL}	28135.358	0.463
$M_{(1)}^{\pi}$	26299.911	0.256
$M_{(2)}^{\pi}$	25889.932	0.083
$M_{(1)}^{\psi+\pi}$	25784.780	0.078
$M_{(2)}^{\psi+\pi}$	26347.952	0.089

Table 1 summarizes the negative log-likelihood (NLL) and sum-of-squared-error (SSE) for all models. NLL evaluates how well the model accounts for the experimental data, and SSE measures the degree of similarity between the model's predictive learning curves and that of humans. In terms of these two criteria, $M_{(1)}^{\psi+\pi}$ accounts best for subjects' behaviors. However, $M_{(1)}^{\psi+\pi}$ fails to capture one observation in human data: human follows a nearly optimal learning curve in set size 2 and 3. This phenomenon is only captured by $M_{(2)}^{\psi+\pi}$.

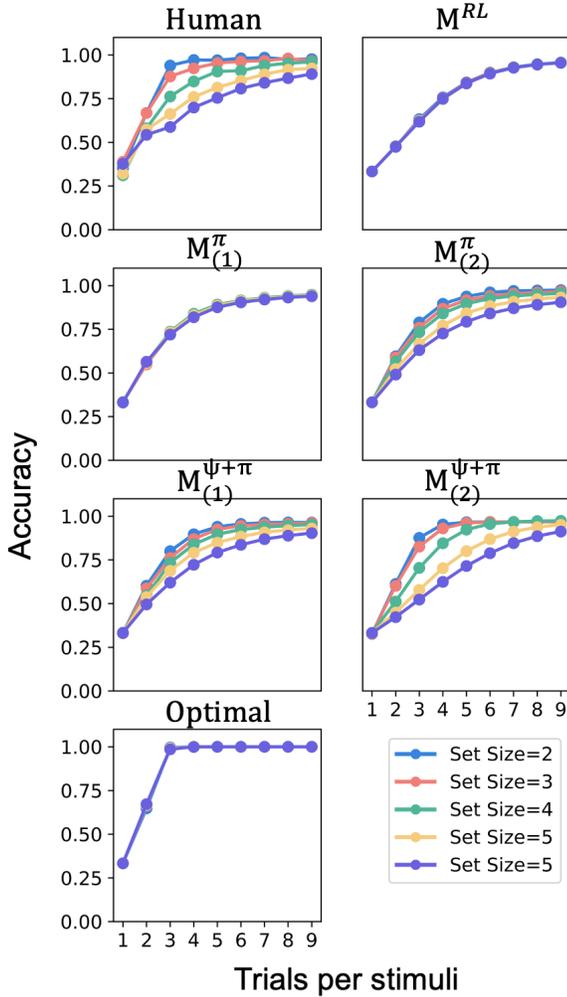


Figure 3: Model results. Learning curves generated using the fit parameters for each set size. Accuracy indicates proportion of responses that were rewarded.

Policy complexity does not account for the set size effect

The key assumption of our information-theoretic approach is that humans behave rationally subject to a fixed cognitive resource (the information constraint), and consequently their suboptimal task performance is explainable via this constraint. In this sense, if we estimate the cognitive cost of the optimal policy, the amount of information (measured in nats) to encode this policy may increase monotonically with the set size, whereas the cognitive cost for the empirical human policy should saturate to a certain value. We may consider this asymptotic value as the effective constraint on cognitive cost for human participants.

Figure 4 shows the model-based cognitive cost estimation. Details for estimating policy complexity and representation complexity are given in the Methods section above. The left column shows the estimation of policy complexity reveal-

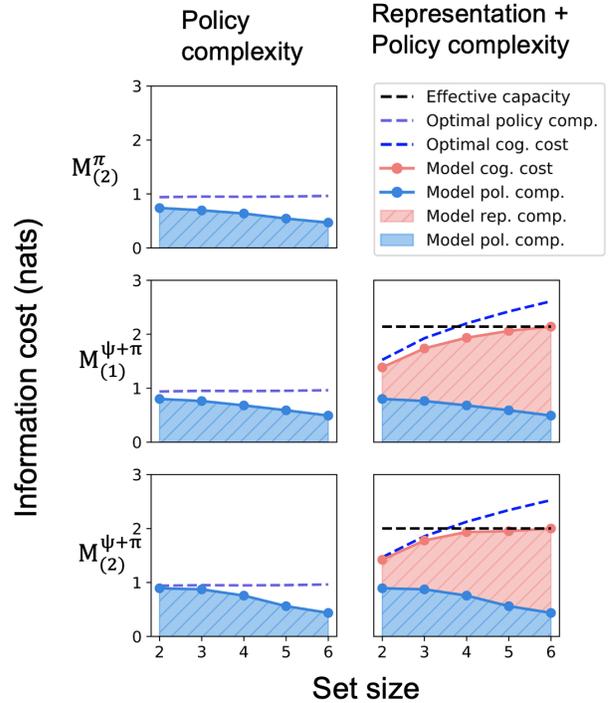


Figure 4: Cognitive cost estimation. The left column shows the complexity of both models’ policy (blue shaded region) and the optimal policy complexity (purple dashed line). The right shows the total cognitive cost (red solid line), constituted of policy complexity (blue shaded region) and representation complexity (red shaded region) for both models. The working memory capacity (black dashed line) equals the maximum of models’ cognitive cost. The blue dashed line indicates the total cognitive cost for the optimal policy. All quantities are measured in nats

ing two salient features: 1) the optimal policy complexity $I^{opt}(X;A)$, calculated using equation 1, is almost constant over set sizes instead of monotonically increasing (purple dashed line), hence larger set sizes do not appear to be more cognitively demanding according to this model; 2) the empirical policy complexity monotonically decreases instead of saturating at a fixed channel capacity (blue shaded region). A problematic question therefore arises: if the tasks in all set sizes are equally complex from an information-theoretic perspective, why do human participants adopt simpler policies in a larger set size conditions?

Neither of these properties is readily explainable from the perspective of computational rationality. We, therefore, conclude that policy complexity alone does not adequately explain human cognitive costs in this experiment.

Representation complexity plus policy complexity captures the set size effect

The right column of figure 4 displays the estimation for both policy complexity and representation complexity (cognitive

cost). The number of nats required to encode the combined representation and optimal policy ($I^\Psi(S;X) + I^{\text{opt}}(X;A)$) increases monotonically with set size (blue dashed line). However, whereas the task demands for optimal performance grow monotonically with set size, the empirically estimated cognitive costs (red solid line) appear to grow much slower in $M_{(1)}^{\Psi+\pi}$ and reach an asymptote at ~ 1.998 nats in $M_{(2)}^{\Psi+\pi}$. Consistent with our expectation, both properties imply the existence of an upper limit on the cognitive capacity that is the sum of both representational complexity and policy complexity.

This formulation of cognitive cost captures and quantifies the subjective experience that increasing the set size increases the cognitive difficulty of the task. In addition, while the estimated policy complexity saturates (or tends to saturate), Figure 4 also shows that rising representation complexity (red shaded region) imposes extra constraints on policy complexity $I^\pi(X;A)$ as the set size increases, answering the question we asked in the last paragraph. According to this model, for set size $ns = 2, 3$ conditions, human decision-makers are able to perform near-optimally because the total cognitive cost is below the available capacity. However, when $ns = 4, 5, 6$, as the state representation complexity grows, human decision makers must resort to an increasingly suboptimal policy to prevent total cognitive cost from exceeding a maximum limit.

Conclusions

In this article, we proposed a new model that optimizes the resource-rational computational goal. Comparing with a similar model published (Griffiths, Lieder, & Goodman, 2015), a large improvement has been made in predicting a deterministic reinforcement learning task. The empirical results indicated that the progress was made because of refining three modeling assumptions: (i) constructing the cost as the sum of representation and policy complexity, (ii) estimating the complexities using a wrong prior, and (iii) updating the model in terms of distribution.

Many suboptimal decisions can be explained as a trade-off between maximizing utility and minimizing costs or constraints imposed by limited cognitive resources (Sims, 2016; Lerch & Sims, 2018; Gershman, 2020). We contribute to this line of thought by arguing that there are two separate sources of cognitive demand in a reinforcement learning setting: representation complexity, and policy complexity. Through a model-based analysis, we showed that the total cognitive cost incorporating both of these constructs appears to saturate to an upper limit in human reinforcement learning. This tentatively suggests the existence of a fixed cognitive resource that can be allocated to a learning task.

Based on this conclusion, we made one further step to interpret how the set size leads to humans' suboptimal performance in the (Collins & Frank, 2012) experiment. Although a larger set size is not necessarily more complicated in terms of policy complexity, it does require the human subjects to hold more representations of the world stimuli. Humans, thus,

have to seek a simpler policy to balance the rising cognitive cost.

In future work, we seek to increase the quality of our model-based analysis by developing more accurate models that better describe humans' learning and decision-making under limited cognitive resources. We expect the following properties from a better model: First, instead of fitting the tradeoff parameters β_{ns} for each set size to describe humans' policy, we can model the principle humans may follow in balancing the reward and the resource. Also, our models assume that human sensory processing is fixed. However, substantial evidence supports that human sensory channel might be adaptive (see review Orhan, Sims, Jacobs, and Knill (2014)). Perhaps, human subjects start with a less resource-efficient sensory code but end up with more efficient coding, allowing humans to learn a more rewarding but complicated policy. This change might be observed only after extensive training because the update of the sensory channel should follow an extremely small learning rate due to it is hardwired in the human neural system.

References

- Berger, T. (1971). The source coding game. *IEEE Transactions on Information Theory*, 17(1), 71–76.
- Collins, A. G., Brown, J. K., Gold, J. M., Waltz, J. A., & Frank, M. J. (2014). Working memory contributions to reinforcement learning impairments in schizophrenia. *Journal of Neuroscience*, 34(41), 13747–13756.
- Collins, A. G., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? a behavioral, computational, and neurogenetic analysis. *European Journal of Neuroscience*, 35(7), 1024–1035.
- Genewein, T., Leibfried, F., Grau-Moya, J., & Braun, D. A. (2015). Bounded rationality, abstraction, and hierarchical decision-making: An information-theoretic optimality principle. *Frontiers in Robotics and AI*, 2, 27.
- Gershman, S. J. (2020). Origin of perseveration in the trade-off between reward and complexity. *Cognition*, 204, 104394.
- Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, 349(6245), 273–278.
- Gershman, S. J., & Lai, L. (2020). The reward-complexity trade-off in schizophrenia. *bioRxiv*.
- Gershman, S. J., & Niv, Y. (2010). Learning latent structure: carving nature at its joints. *Current opinion in neurobiology*, 20(2), 251–256.
- Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in cognitive science*, 7(2), 217–229.
- Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical bayesian models. *Devel-*

- opmental science*, 10(3), 307–321.
- Lai, L., & Gershman, S. J. (2021). Policy compression: An information bottleneck in action selection. *Psychology of Learning and Motivation S*.
- Lerch, R. A., & Sims, C. R. (2018). Policy generalization in capacity-limited reinforcement learning.
- Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature neuroscience*, 17(3), 347.
- Orhan, A. E., Sims, C. R., Jacobs, R. A., & Knill, D. C. (2014). The adaptive nature of visual working memory. *Current directions in psychological science*, 23(3), 164–170.
- Sims, C. R. (2016). Rate–distortion theory and human perception. *Cognition*, 152, 181–198.
- Still, S., & Precup, D. (2012). An information-theoretic approach to curiosity-driven reinforcement learning. *Theory in Biosciences*, 131(3), 139–148.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.
- Tishby, N., & Polani, D. (2011). Information theory of decisions and actions. In *Perception-action cycle* (pp. 601–636). Springer.
- Zenon, A., Solopchuk, O., & Pezzulo, G. (2019). An information-theoretic perspective on the costs of cognition. *Neuropsychologia*, 123, 5–18.

An ACT-R Model of Order Effects

Christopher R. Fisher (christopher.fisher.27.ctr@us.af.mil)

Cubic Defense
Beavercreek, OH 45324 USA

Lorraine Borghetti (lorraine.borghetti.1.ctr@us.af.mil)

Air Force Research Laboratory
Wright Patterson AFB, OH USA

Joseph W. Houtp (joseph.houtp@utsa.edu)

University of Texas at San Antonio
San Antonio, TX USA

Christopher Stevens (christopher.stevens.28@us.af.mil)

Air Force Research Laboratory
Wright Patterson, OH USA

Abstract

Models based on classical probability theory have difficulty accounting for order effects, which occur when the order of question presentation affects response probabilities. Recently, quantum models have garnered support as an account of order effects. In particular, the pattern of order effects is consistent with a critical property of the quantum model called the QQ equality. We investigate whether the ACT-R cognitive architecture can produce order effects and satisfy the QQ equality based on memory retrieval mechanisms. In the ACT-R model, the answer to the first question creates a new context through which spreading activation influences retrieval probabilities for the second answer. Our analysis shows that spreading activation can produce order effects and satisfy the QQ equality, depending on the composition of declarative memory. Across a wide range of conditions, violations of the QQ equality are typically small, but moderate to large in a smaller set of cases.

Keywords: ACT-R; Quantum cognition; Order effects

Introduction

An order effect occurs when a response depends on the order in which stimuli are presented. In cognitive science, order effects are commonly treated as a nuisance factor in experimental design and data analysis. Typically, stimulus order is counter-balanced, marginalized out, and subsequently ignored. Recently, however, there has been growing interest in developing theoretical accounts of order effects (Trueblood & Busemeyer, 2011; Jones, Curran, Mozer, & Wilder, 2013; Wang, Solloway, Shiffrin, & Busemeyer, 2014). What makes order effects interesting is that they are difficult to account for using models based explicitly or implicitly on the foundation of classical probability theory. In particular, order effects violate the commutative law of classical probability theory according to which $\Pr(A \wedge B) = \Pr(B \wedge A)$. Furthermore, order effects are interesting because they highlight the context-dependent nature of cognition.

One example of order effects comes from a national poll asking respondents about the trustworthiness of former President Clinton and former Vice President Gore in two separate questions. One set of respondents judged “yes” or

“no” whether Clinton is generally trustworthy followed by the same judgment about Gore, whereas the other set of respondents made the same judgments in the opposite order. In Table 1, the results for both question orders are transcribed from (Wang et al., 2014). The third sub-table shows the order effect, calculated as the difference between corresponding cells of the Gore-Clinton table and the Clinton-Gore table. A clear violation of the commutative law can be seen, as the values differ from zero.

During the 20th century, quantum probability theory was developed to account for order-dependence of measurement in physics. More recently, it has been adopted in cognitive science in response to similar violations of classical probability theory found in human cognition (Busemeyer, Pothos, Franco, & Trueblood, 2011; Atmanspacher, Römer, & Walach, 2002). Quantum probability theory is based on an alternative set of axioms which violate the commutative law under specific conditions. Several lines of research provide strong support for quantum probability account of order effects. For example, a quantum model provided a superior account of sequential belief updating compared to several competing models (Trueblood & Busemeyer, 2011). Perhaps the most compelling line of evidence is that the results from a large corpus of surveys and several experiments were consistent with a critical property of quantum probability theory called the Quantum Question (QQ) equality (Wang et al., 2014). The QQ equality is a structural property of the quantum model that constrains the possible patterns of order effects.

Very few alternative accounts of order effects have been proposed to date—perhaps reflecting the inherent challenge of the task. From a purely mathematical perspective, a Bayesian updating model can produce order effects with the inclusion of order-dependent events (Trueblood & Busemeyer, 2011). However, such a model is problematic because it is saturated, and thus merely re-describes any pattern of data without providing a principled way to assign values

to its numerous terms. Recently, a multinomial processing tree called the repeat-choice model was developed to account for order effects and the QQ equality (Kellen, Singmann, & Batchelder, 2018). Multinomial processing trees use a tree-like structure of processing stages to describe choice behavior. According to the repeat-choice model, there is a probability distribution of over preference states, such as the preference to respond yes to both questions. With some probability, the responses will be made based only on the preference state. However, additional information will be considered with the complementary probability. If additional information is considered, there is some probability that the second response is the same as the first (assimilation effect) and the complementary probability that the second response will differ from the first (contrast effect). Some variations of the repeat-choice model provided a similar fit to the data as the quantum model, thus demonstrating that model based on classical probability theory can account for the data.

One question that remains is whether a cognitive architecture can account for order effects. A cognitive architecture is a framework for simulating and developing unified theories of cognition (Newell, 1990). The primary goal of cognitive architectures is to provide a broad account of human cognition, spanning areas such as memory, multi-tasking, and perception among others. For this reason, order effects provide a new and important benchmark for testing cognitive architectures. Our goal in the present paper is to develop a model of order effects based on the ACT-R cognitive architecture (Anderson et al., 2004) and outline its predictions. Order effects present an interesting challenge for ACT-R because, unlike the quantum model—which was developed in physics to account for order dependence in measurement—it was not developed specifically to account for order effects. Instead, order effects must emerge from existing cognitive processes and mechanisms postulated by ACT-R. In what follows, we will demonstrate that ACT-R can produce order effects using memory retrieval mechanisms. In some cases, the model satisfies the QQ equality, while in other cases it violates the QQ-equality to varying degrees. The pattern of predictions depends critically upon the composition of declarative memory.

Overview

The remainder of the paper is organized as follows. We will begin with a brief overview of the quantum model and introduce the QQ equality. Next, we will provide a formal description of the ACT-R model of order effects. In the following section, we will describe its properties in terms of order effects and the QQ equality. Finally, we will discuss directions for future research.

Quantum Model

Quantum probability theory is based on an alternative set of axioms which allows non-commutative behavior (Busemeyer et al., 2011). Unlike classical probability theory in which events are subsets of a universal set, quantum probability

Table 1: Joint probability table for Clinton-Gore order, the Gore-Clinton order, and the order effect. Column and row labels C and G correspond to Clinton and Gore. Subscripts y and n correspond to yes and no.

		Clinton-Gore		Gore-Clinton	
		G_y	G_n	G_y	G_n
C_y		0.4899	0.0447	C_y 0.5625	0.0255
C_n		0.1767	0.2886	C_n 0.1991	0.213
Order Effect					
		G_y	G_n		
C_y		-0.0726	0.0192		
C_n		-0.0224	0.0756		

theory is based on a geometric representation of uncertainty. Events are sub-spaces within an n dimensional vector space called a Hilbert space. A cognitive state represented as a state vector which is a linear combination of basis vectors that define the Hilbert space. In the quantum model, probabilities are formed by projecting the state vector onto a target subspace and computing the squared magnitude of the projection. A key distinction between classical and quantum probability theory is the concept of compatibility. Compatible events can be evaluated with respect to the same basis vectors, in which case quantum and classical probability theory make the same predictions. By contrast, incompatible events cannot be evaluated with respect to the same basis vectors. Instead, the basis vectors are rotated to create new set of basis vectors for the incompatible events. In other words, incompatible events cannot be evaluated simultaneously. Importantly, rotation leads to non-commutative behavior and other violations of classical probability theory. At a psychological level, rotation of the basis vectors represents a change of perspective.

Although rotation of the vector space provides the quantum model with the flexibility to produce order effects, the range of behavior is highly constrained by a critical property known as the QQ equality (Wang et al., 2014). The QQ equality imposes a symmetrical relationship on the order effects in which both the diagonal elements and off-diagonal elements of the difference table must sum to zero. For example, in the third sub-table of Table 1, the diagonal elements $-0.0726 + 0.0756 \approx 0$ and the off-diagonal elements $-0.0224 + 0.0192 \approx 0$. Importantly, the QQ equality holds regardless of the initial state vector, the degree of rotation of the basis vectors, and is preserved in aggregated data (Wang et al., 2014). Formally, the QQ equality is defined by the following two statements:

$$\begin{aligned}
 q_1 &= \Pr(Y_c = \text{yes} \wedge Y_g = \text{yes}) - \Pr(Y_g = \text{yes} \wedge Y_c = \text{yes}) + \\
 &\Pr(Y_c = \text{no} \wedge Y_g = \text{no}) - \Pr(Y_g = \text{no} \wedge Y_c = \text{no}) = 0
 \end{aligned}
 \tag{1}$$

$$\begin{aligned}
 q_2 &= \Pr(Y_c = \text{yes} \wedge Y_g = \text{no}) - \Pr(Y_g = \text{no} \wedge Y_c = \text{yes}) + \\
 &\Pr(Y_c = \text{no} \wedge Y_g = \text{yes}) - \Pr(Y_g = \text{yes} \wedge Y_c = \text{no}) = 0
 \end{aligned} \tag{2}$$

where Y_p represents the response to question about person $p \in \{c, g\}$, and c denotes Clinton and g denotes Gore. Throughout, we will designate Y_p as a random variable and y_p as a specific realization of Y_p .

ACT-R Model

We developed a memory-based model of order effects within the ACT-R cognitive architecture (Anderson et al., 2004). ACT-R operates as a production system and is organized as a set of specialized processing modules which includes memory, visual/auditory perception and motor execution. Each module can process only one request at a given time and contains a buffer that holds a maximum of one chunk of declarative knowledge. For our present purposes, we will focus primarily on the declarative memory module. Although we will frame the model in terms of the Clinton-Gore example above, the model is applicable to many other cases in which responses are based on memory.

The model assumes that declarative memory contains chunks which represent true or false statements made by Clinton or Gore. When a question about a person is posed, a retrieval request is issued to declarative memory where the most active chunk about the target person is returned. The answer to the question is yes if the chunk contains a true statement. By contrast, the answer to the question is no if the retrieved chunk contains a false statement. During the retrieval of the first answer, there is no influence of contextual information. However, the answer to the first question creates a new context for answering the second question. In particular, the chunk for the first answer is stored in the imaginal buffer where activation spreads to chunks in declarative memory that share the same truth value, resulting in order effects under specific conditions.

Declarative Memory

Within ACT-R, a chunk is a basic unit of declarative knowledge given by a collection of slot-value pairs. For example a memory chunk could contain the slot ‘name’ with value ‘Sigma’ and the slot ‘animal-type’ with value ‘dog’. For the remainder of the paper, we use the following notation for chunks: We use c_m to indicate a chunk in memory, where m is an index that ranges over all of the chunks in memory. We write the relationship between the slot s and value v for chunk m as $c_m(s) = v$. We will also need to reference the slots in the chunk for which a value is defined, which we denote Q_m . Note that all chunks do not necessarily (and generally do not) have the same slots, so to maintain the generality of the notation, we assume $c_m(s) = \emptyset$ for any s that is not a slot in chunk c_m i.e., for all $s \notin Q_m, c_m(s) = \emptyset$.

Like chunks, retrieval requests in the ACT-R architecture are collections of slot-value pairs, which we designate using set notation, $\mathbf{r} = \{(s_i, v_i)\}_{i \in I}$.

Knowledge Representation

Each chunk c_m in declarative memory contains a name slot, a statement slot, and a truth slot: $Q = \{\text{name, statement, truth}\}$. The name slot contains the name of the person, the statement slot contains the content of the statement and the truth slot contains the truth value of the statement.

Activation

The probability of retrieving a chunk increases monotonically with its activation value. Activation for chunk m is the sum of the following three components:

$$a_m = \beta + S_m + \epsilon_m \tag{3}$$

where β is the base level constant, which scales activation up or down, S_m is spreading activation, and $\epsilon_m \sim \text{Normal}(0, \sigma)$ is normally distributed noise. Spreading activation reflects the influence of context whereby active information in the architecture facilitates the retrieval of chunks containing the same values. Spreading activation has been used to explain the fan effect whereby concepts associated with more facts require more time to retrieve (Anderson, 1974). In the model, activation can only spread from the truth value of the chunk in the imaginal buffer to chunks in declarative memory. This follows from the simplifying assumption that statements are unique, and thus do not contribute to spreading activation. Given these simplifications, we can express spreading activation as:

$$S_m(x) = \frac{\gamma + \log\left(\frac{1}{1+x}\right)}{|Q|}$$

where x is the number of chunks in declarative memory with a same truth value as $c_{r,\text{imaginal}}(\text{truth})$, γ is the maximum association parameter, and $|Q|$ is the number of slots in each chunk, which is 3 in this case.

Retrieval Process

After the first question is encoded, a retrieval request $\mathbf{r} = \{(\text{name}, v_1)\}$ is submitted to declarative memory where a set of matching chunks $R = \{c_m \in M : (\text{name}, c_m(\text{name})) \in \mathbf{r}\}$ compete for retrieval and the chunk with maximum activation $c_r \in R$ is retrieved. During the retrieval of the first answer, there is no influence of spreading activation because the imaginal buffer is empty. The retrieved chunk is placed into the imaginal buffer and becomes $c_{r,\text{imaginal}} = c_r$ where it will influence memory retrieval for the second answer through spreading activation.

Response Mapping

The mapping between the truth value of the retrieved chunk and the response y for person p is given by:

$$y_p = \begin{cases} \text{yes} & c_r(\text{truth}) = \text{true} \\ \text{no} & c_r(\text{truth}) = \text{false} \end{cases}$$

Retrieval Probability

The retrieval probability is found by comparing the response set $W_x \subset R$ to the retrieval set R . The response set is the subset of chunks in the retrieval set that map to the observed response y_p . The response set for yes and no are defined as:

$$W_{\text{yes}} = \{c_m \in R : c_m(\text{truth}) = \text{true}\}$$

$$W_{\text{no}} = \{c_m \in R : c_m(\text{truth}) = \text{false}\}$$

The probability of responding x on the first question is given by the following softmax function:

$$\Pr(Y_p = x) = \frac{\sum_{c_m \in W_x} e^{\frac{\mu_m}{\sigma}}}{\sum_{c_k \in R} e^{\frac{\mu_k}{\sigma}}} = \frac{e^{\frac{\beta}{\sigma}} \sum_{c_m \in W_x} e^0}{e^{\frac{\beta}{\sigma}} \sum_{c_k \in R} e^0} = \frac{|W_x|}{|R|}$$

where μ is mean activation, $\sigma = s\sqrt{2}$ controls activation noise and s is the logistic scale parameter. The expression simplifies to the ratio of chunks leading to response x over all chunks that match the retrieval request because $e^{\beta/\sigma}$ can be factored out of each term. We can rewrite the expression in terms of the number of true and false statements for Clinton and Gore. Let $T = T_c + T_g$ be the total number of chunks containing a true statement and $F = F_c + F_g$ be the total number of chunks containing a false statement, where subscript c represents Clinton and subscript g represents Gore. For example, the probability of responding yes to Clinton on the first question is defined as

$$\Pr(Y_c = \text{yes}) = \frac{T_c}{T_c + F_c}$$

which is simply the ratio true statements made by Clinton compared to all statements made by Clinton. The expression for the second question includes a term for spreading activation, which can be simplified as: $z \left(\frac{1}{x+1}\right)^h$ where $h = \frac{1}{|\mathcal{Q}| \sigma}$ and $z = e^{\eta}$. For example, the probability of responding yes to Gore on the second question given a response of yes to Clinton on the first question is defined as:

$$\Pr(Y_g = \text{yes} | Y_c = \text{yes}) = \frac{T_g \cdot z \cdot \left(\frac{1}{T+1}\right)^h}{T_g \cdot z \cdot \left(\frac{1}{T+1}\right)^h + F_g}$$

In this example, spreading activation increases the probability of responding yes to the question about Gore. The full set of equations can be found in Table 2 for the Clinton-Gore order and Table 3 for the Gore-Clinton order. Note that each joint probability table sums to 1 as required by classical probability theory. Under certain conditions, however, spreading activation causes the probability mass to shift to different cells in each table, producing two different joint probability distributions. In some sense, this is similar to using a different set of basis vectors to define events in the quantum model. In the ACT-R model, the table for each order is consistent with classical probability theory. Similarly, in the quantum model,

probabilities based on projection any of a set of orthonormal basis vectors are consistent with classical probability theory. However, just as the ACT-R model is not necessarily consistent with classical probability theory across tables, the quantum model is not necessarily consistent with classical probability theory across different rotations of the basis vectors.

Predictions

In what follows, we describe the predictions of the ACT-R model for order effects and the QQ equality. Although we have proved the following properties, the proofs are omitted due to space limitations.

Table 2: Predictions of the ACT-R order model for the Clinton-Gore order. Column and row labels C and G correspond to Clinton and Gore. Subscripts y and n correspond to yes and no.

	G_y	G_n
C_y	$\frac{T_c}{T_c + F_c} \cdot \frac{T_g \cdot z \cdot \left(\frac{1}{T+1}\right)^h}{T_g \cdot z \cdot \left(\frac{1}{T+1}\right)^h + F_g}$	$\frac{T_c}{T_c + F_c} \cdot \frac{F_g}{T_g \cdot z \cdot \left(\frac{1}{T+1}\right)^h + F_g}$
C_n	$\frac{F_c}{T_c + F_c} \cdot \frac{T_g}{T_g + F_g \cdot z \cdot \left(\frac{1}{T+1}\right)^h}$	$\frac{F_c}{T_c + F_c} \cdot \frac{F_g \cdot z \cdot \left(\frac{1}{T+1}\right)^h}{T_g + F_g \cdot z \cdot \left(\frac{1}{T+1}\right)^h}$

Table 3: Predictions of the ACT-R order model for the Gore-Clinton order. Column and row labels C and G correspond to Clinton and Gore. Subscripts y and n correspond to yes and no.

	G_y	G_n
C_y	$\frac{T_g}{T_g + F_g} \cdot \frac{T_c \cdot z \cdot \left(\frac{1}{T+1}\right)^h}{T_c \cdot z \cdot \left(\frac{1}{T+1}\right)^h + F_c}$	$\frac{F_g}{T_g + F_g} \cdot \frac{T_c}{T_c + F_c \cdot z \cdot \left(\frac{1}{T+1}\right)^h}$
C_n	$\frac{T_g}{T_g + F_g} \cdot \frac{F_c}{T_c \cdot z \cdot \left(\frac{1}{T+1}\right)^h + F_c}$	$\frac{F_g}{T_g + F_g} \cdot \frac{F_c \cdot z \cdot \left(\frac{1}{T+1}\right)^h}{T_c + F_c \cdot z \cdot \left(\frac{1}{T+1}\right)^h}$

Order Effects

According to the model, order effects depend on the ratio of true to false statements for each person. Four order effects can be obtained by subtracting the corresponding cells of Gore-Clinton and Clinton-Gore joint probability tables. Figure 1 shows that the predicted order effect depends on the ratio of true to false statements for each person. For example, in the top left panel, T_c is varied from 0 to 6 while $F_c = T_g = F_g = 1$. The order effect is large when $T_c = 0$ and small to moderate for all other values. Similar patterns can be found in the remaining panels.

Equal Ratios The model predicts no order effects for matching responses (e.g. yes, yes) when the ratios are equal. However, the order effect for a yes, no response is the negative of the order effect for a no, yes response, which ranges over negative and positive values. More formally, we define O_{y_c, y_g} as the order effect for response y_c to the Clinton question and y_g to the Gore question. With this notation,

we can write: $O_{yes,yes} = O_{no,no} = 0$ and $O_{yes,no} = -O_{no,yes}$ if $T_c = v \cdot T_g$, $F_c = v \cdot F_g$. As a special case, if $T_c = F_c$ and $T_g = F_g$, then $O_{yes,no} = O_{no,yes} = 0$. Intuitively, this means that the effects of spreading activation in both question orders cancel out because the ratios are 50-50, thus eliminating the order effect. This can be seen in Figure 1 where the values on the x-axis equal 1.

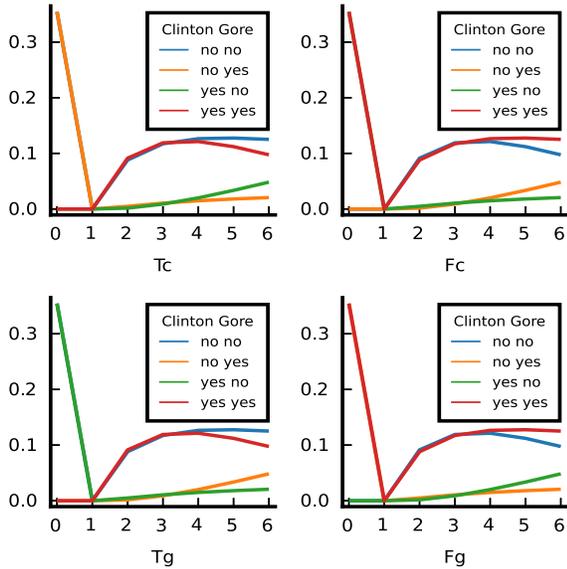


Figure 1: Order effect predictions. Each plot shows the absolute order effect along the y-axis for the number of chunks of the type specified in the x-axis while the chunk types are hold constant at a value of 1.

QQ Equality

The ACT-R model follows the basic mathematical constraint whereby $q_1 = -q_2$, but satisfies the QQ equality ($q_1 = q_2 = 0$) under specific conditions that depend on the ratio of true to false statements for each person. Figure 2 shows the q values as a function of the value on the x-axis while the other three values are fixed at 1. When value on the x-axis is zero, a large violation of the QQ equality is predicted. In other cases, the predicted violation of the QQ equality is small or zero.

Equal Ratios ACT-R satisfies the QQ equality when the ratios of true to false statements are equal for each person. More formally, $q_1 = q_2 = 0$ if $T_c = v \cdot T_g$, and $F_c = v \cdot F_g$. This occurs in Figure 2 where the value on the x-axis is equal to 1.

Complementary Ratios ACT-R satisfies the QQ equality when the ratios of true to false statements are complementary for each person, but the number of chunks for each person are equal. More formally, $q_1 = q_2 = 0$ if $T_c = F_g$ and $T_g = F_c$.

Q Distribution To investigate the extent to which the ACT-R model violates the QQ equality, we computed q values across all permutations of memory composition for values from 0 to 6. As shown in Figure 3, most of the density is cen-

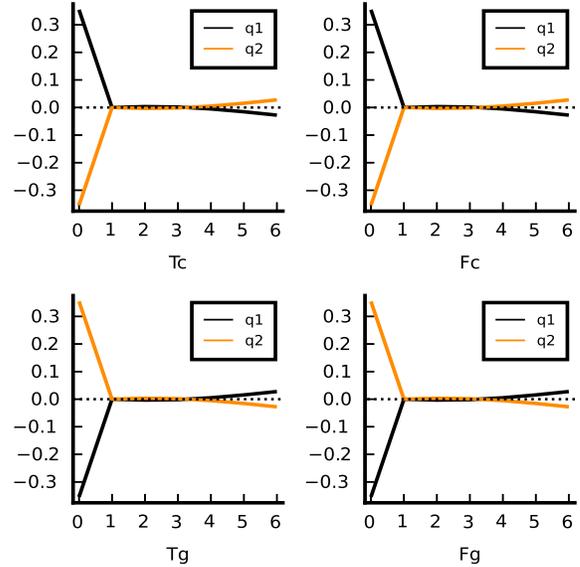


Figure 2: q-value predictions. Each plot shows the q values for the number of chunks of the type specified in the x-axis while the other chunk types are hold constant at a value of 1.

tered near zero. In particular, approximately 77% of q values range between -.1 and .1 and approximately 68% range between -.05 and .05. The remaining portion of the distribution extends towards -.4 and .4 in a roughly uniform manner. Collectively, these results suggest that violations of the QQ equality are typically small, but can be quite large under some circumstances.

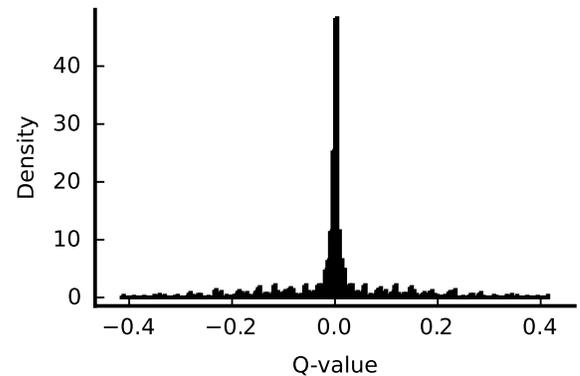


Figure 3: Distribution of q values across all $7^4 = 2,401$ permutations of T_c , F_c , T_g , and F_g for values 0 to 6.

Discussion

In the present paper, we developed a memory-based model of order effects within the ACT-R cognitive architecture and outlined many of its predictions and properties. Our analysis reveals that ACT-R can produce order effects and satisfy the QQ equality depending on the composition of declarative

memory. Across a large range of memory sets, the model produces q values that either satisfy the QQ equality or violate it only by a small degree. In other cases, there is a clear divergence from the QQ equality.

Some points of similarity and difference between the quantum model and the ACT-R model are worth noting. One point of similarity is that context is an important determinant of order effects in both models. In the quantum model, order effects arise from non-commutative evaluation processes when events are incompatible. In the ACT-R model, the answer to the first question creates a new context through which spreading activation modulates the retrieval probabilities for the second answer. The models differ in several important regards. One difference is the distinction between memory-based vs. online judgments (Hastie & Park, 1986). In ACT-R, judgments are formed from a set of experienced events stored in memory, whereas in the quantum model, judgments are constructed online through comparison processes and do not require a definite reference class. One direction for future research is to determine whether ACT-R can perform online judgments by comparing chunks in different buffers.

Although we have demonstrated as a proof of concept that ACT-R can produce order effects and satisfy the QQ equality under specific conditions, the model has not been tested against empirical data. For this reason, it remains unclear how it compares to the quantum model in terms of empirical support. As a memory-based model, ACT-R requires a well-controlled experiment in which the composition of declarative memory is manipulated to test the properties outlined above. Existing data sets are not suitable for testing the ACT-R model because factors influencing memory were not controlled or measured. For example, respondents in the national survey likely differed in terms of political knowledge and information sources, which, in turn, could introduce heterogeneity in judgments about Clinton and Gore. Uncertainty and heterogeneity in memory composition would render the results uninterpretable from the standpoint of the ACT-R model. In future research, we plan to design a memory based experiment to test the predictions outlined above.

Our analysis shows that the predictions hold for different values of maximum associative strength and activation noise so long as maximum association strength is sufficiently large to produce a positive spreading activation term. One may wonder how the predictions might change when certain assumptions of the model are relaxed. For example, relaxing the assumption that β is equal across chunks leads to a somewhat more complex model, but the predictions ultimately depend on the ratio of activation of true and false statements for each person rather than the ratio of chunks.

Conclusion

Order effects are an interesting benchmark for testing the ACT-R cognitive architecture because it was not developed to account for such effects. Nonetheless, we demonstrated that ACT-R can produce order effects using existing memory

retrieval mechanisms and can satisfy the QQ equality under some conditions. Although more work is required to test the model, we regard this proof of concept as an important first step towards stress testing the architecture against new benchmark phenomena.

Acknowledgments

The opinions expressed herein are solely those of the authors and do not necessarily represent the opinions of the United States Government, the U.S. Department of Defense, the U.S. Air Force, or any of their subsidiaries or employees. This research was supported by an AFOSR LRIR. The contents have been reviewed and deemed Distribution A. Approved for public release. Case number: RH-21-121922.

References

- Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, 6(4), 451–474.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological review*, 111(4), 1036.
- Atmanspacher, H., Römer, H., & Walach, H. (2002). Weak quantum theory: Complementarity and entanglement in physics and beyond. *Foundations of Physics*, 32(3), 379–406.
- Busemeyer, J. R., Pothos, E. M., Franco, R., & Trueblood, J. S. (2011). A quantum theoretical explanation for probability judgment errors. *Psychological Review*, 118(2), 193.
- Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review*, 93(3), 258.
- Jones, M., Curran, T., Mozer, M. C., & Wilder, M. H. (2013). Sequential effects in response time reveal learning mechanisms and event representations. *Psychological Review*, 120(3), 628.
- Kellen, D., Singmann, H., & Batchelder, W. H. (2018). Classic-probability accounts of mirrored (quantum-like) order effects in human judgments. *Decision*, 5(4), 323.
- Newell, A. (1990). *Unified theories of cognition*. Harvard University Press.
- Trueblood, J. S., & Busemeyer, J. R. (2011). A quantum probability account of order effects in inference. *Cognitive Science*, 35(8), 1518–1552.
- Wang, Z., Solloway, T., Shiffrin, R. M., & Busemeyer, J. R. (2014). Context effects produced by question orders reveal quantum nature of human judgments. *Proceedings of the National Academy of Sciences*, 111(26), 9431–9436.

Using Cognitive Agents to Design Dynamic Task Allocation Systems

Christopher R. Fisher (christopher.fisher.27.ctr@us.af.mil)

Cubic Defense
Beavercreek, OH 45324 USA

Mary E. Frame (mary.frame@parallaxresearch.org)

Parallax Advanced Research
4035 Colonel Glenn Hwy, Beavercreek, OH 45431

Christopher Stevens (christopher.stevens.28@us.af.mil)

Air Force Research Laboratory
Wright Patterson, OH USA

Abstract

Although cognitive models are primarily used to formalize theories of cognition, they could be applied in artificial intelligence (AI) systems, such as autonomous managers (AMs) which optimize team performance through dynamic task allocation. Cognitive models can be incorporated into the AM's decision system to understand the implications of alternative task distributions. They can also be used as simulated agents to stress test AMs under a wide range of conditions. In a simulation study, we varied the cognitive model used in the AM's decision system and the cognitive model performing a task to explore the design space of AMs. We found a trade-off between optimality and robustness in which complex models performed the best when assumptions were met, but were not robust to violation of assumptions. These results highlight the importance of considering simple models when assumptions could be violated and showcase the utility of cognitive models in AI systems.

Keywords: task allocation; cognitive agents;

Introduction

Cognitive models have a multitude of uses ranging from formalizing theories of cognition and sharpening research questions (McClelland, 2009), to measuring individual differences in cognition among clinical and non-clinical populations (Riefer, Knapp, Batchelder, Bamber, & Manifold, 2002; Yechiam, Busemeyer, Stout, & Bechara, 2005). One of the promises of cognitive architectures—and perhaps cognitive models more generally—is the ability to scale up to complex tasks where the limits of theory and practical application can be pushed (Newell, 1990). Some complex tasks in which cognitive models have been applied include training teams involving synthetic teammates (McNeese, Demir, Cooke, & Myers, 2018) and driving (Salvucci, 2006).

One burgeoning area in which cognitive models could be informative is artificial intelligence (AI). Deep learning in particular is a remarkably flexible function learning algorithm. According to the Universal Approximation Theorem, a deep neural network containing a sufficient number of layers can approximate any continuous function with sufficient training data (Zhou, 2020). However, this flexibility comes at a high cost: copious amounts of data are required for training in order to compensate for the lack of predefined structure. In addition, deep learning and similar approaches have been criticized for being opaque, brittle, and vulnerable to sabotage

(Nguyen, Yosinski, & Clune, 2015). One solution is to incorporate scientific models into AI systems in order to provide a structure that reduces demanding data requirements. Currently, there are efforts to integrate neural networks with differential equation models commonly used in physics, biology and pharmacology (Rackauckas et al., 2020) to achieve a better balance between flexibility and data requirements. Along similar lines, cognitive models could be integrated into AI systems in situations that require interacting with or reasoning about humans.

We argue that cognitive models can be integrated into autonomous managers (AMs) designed to optimize performance in team-based work environments. An AM monitors performance of a team and dynamically allocates tasks between workers in order to improve performance of the team. An autonomous task manager can draw upon several sources of information and methods to inform a task allocation decision, including performance history of workers, AI, and mathematical optimization. Cognitive models can be integrated into an AM in at least two ways. First, an AM could use a cognitive model to predict and understand the implications of alternative task allocations. Second, a wide range of cognitive models can perform a simulated task environment in order to investigate the robustness of the design space of AMs.

We performed a series of simulations to explore the design space of AMs that incorporate different cognitive models into the decision process. Our primary research goal was to identify trade-offs between optimality and robustness in the design space. For example, do some designs perform optimally when model assumptions are satisfied, but perform poorly when assumptions are violated? In our simulations, an AM dynamically allocates sub-tasks of a larger, more complex task to cognitive agents (i.e. instances of a cognitive model that perform the task). We varied both the type of cognitive agent that performed the task and the cognitive model that the AM used to inform task allocation decisions. The cognitive models/agents varied according to cognitive processing constraints and the relationship between workload and performance. In some cases, the cognitive agent and the model used in the AM were the same; in other cases, they differed.

Overview

The remainder of the paper is organized as follows. First, we discuss the logic behind dynamic task allocation and the need for automated task allocation systems. Next, we describe the novel complex task environment used to test the effectiveness of AMs. Then, we describe the cognitive agents and AMs used in our simulation. Finally, we present and discuss the results of the simulation study. To preview a key result, we found a trade-off between optimality and robustness.

Dynamic Task Allocation

The performance of a team may vary according to numerous factors, including differences in skill specialization, sensitivity to workload, and temporal dynamics associated with fatigue. In some cases, team performance could be improved simply by allocating tasks to workers with the appropriate specialized skills. However, optimizing team performance is rarely this simple. For example, a person who skillfully performs two tasks in isolation may struggle to perform both together if she or he is sensitive to changes in workload level. Performance may also vary randomly from day to day due to unknown factors or could vary with fluctuations in task demands or fatigue. Dynamic task allocation is necessary in order to deal with uncertainty and to adapt to changes in performance across time. Collectively, these factors present a challenge for optimizing team performance.

Some work environments could benefit from an AM because it is difficult and costly for a human supervisor to manually monitor and allocate tasks. In addition, some research indicates that delegating work distribution decisions to workers can be disruptive due to the additional workload imposed by monitoring the performance oneself and others (Katidioti, Borst, van Vugt, & Taatgen, 2016; Won, Condon, Landon, Wang, & Hannon, 2011). One important research question that remains unanswered is how to design an AM that can adapt to dynamic situations and is robust to individual differences. For example, do more complex models provide a large performance gain compared to simpler models? Are more complex models less robust to violation of assumptions? We attempt to address these questions in the present research.

ISR-MATB

We developed a complex task environment called the Intelligence Surveillance and Reconnaissance Multi-Attribute Task Battery (ISR-MATB) to induce task demands similar to what is found in ISR operations. The ISR-MATB is a variation of the Multi-task Attribute Task Battery (Santiago-Espada, Myer, Latorella, & Comstock Jr, 2011) which was designed to emulate task demands in aviation. Whereas the MATB focuses on performing multiple tasks concurrently, the ISR-MATB focuses on goal switching, information search, interdependence between operational procedures, and synthesis of information into actionable decisions. The ISR-MATB uses variations of standard cognitive tasks to tap into each of these cognitive demands. Although each sub-task is relatively sim-

ple in isolation, they combine into a more complex whole due to inter-dependencies between sub-tasks. In what follows, we will describe each sub-task and then explain how they fit into an inter-dependent task flow.

Psychomotor Vigilance Task

The ISR-MATB uses a modified Psychomotor Vigilance Task (PVT) (Dinges & Powell, 1985) to emulate unpredictable changes in goals that may occur in ISR operations. The PVT is commonly used to measure fatigue and sustained attention. In a standard PVT, a millisecond counter appears after a uniformly distributed inter-stimulus interval of 2 to 10 seconds. Participants are instructed to respond as quickly as possible when the stimulus is presented. Upon responding, the millisecond counter stops and is displayed for 1 second as feedback. In the modified PVT, a target for the current trial is randomly selected from the stimulus set {grey Q, grey O, black Q, black O} and presented. As explained further below, the target is used to complete the visual and auditory search tasks.

Visual Search Task

The ISR-MATB uses the classic conjunctive visual search task (Treisman & Gelade, 1980) to emulate visually demanding search tasks in ISR, such as searching through a visually dense video feed for a target. The conjunctive search task requires participants to search for a target among an array of scattered distractors. Each stimulus has two dimensions: color and letter. A stimulus is considered a target if it matches on both dimensions (e.g. black Q). A distractor matches on one dimension but differs on the other (e.g. grey Q, black O). On half of the trials the target is present and on the other half of trials the target is absent.

Auditory Search Task

The ISR-MATB uses an auditory search task to emulate similar search tasks in ISR operations. For example, an operator might be required to search for keywords and phrases in communication channels and audio recordings where audio signals can be degraded or embedded in background noise. In the auditory search task, a participant is instructed to scan up to four radio channels containing background white noise for the search target (e.g. an audio recording of the words "black Q"). Difficulty is manipulated by changing the number of radio channels and the signal to noise ratio.

Decision Task

In ISR operations, information from multiple sources must be synthesized into a decision to act or refrain from taking action. We capture this aspect of ISR operations with the multiple-cue decision task inspired by similar tasks in the literature (Sieck & Yates, 2001). As summarized in Table 1, decisions are based on two binary cues: (1) whether the target state (i.e. present or absent) is the same or different in the visual and auditory search sub-tasks, and (2) whether confidence in the accuracy of the information is low or high. For

example, if confidence is low and the target is present in the visual and auditory search tasks, the correct action is to refrain. The base rate of cue values is 50-50, meaning it is not possible to perform better than chance with incomplete information.

Table 1: A decision matrix of four rules based on whether the visual and auditory target states are the same (present in both or absent in both) or different and the confidence in the information accuracy.

Confidence	Target State	Correct Action
Low	Same	Refrain
Low	Different	Act
High	Same	Act
High	Different	Refrain

Task Flow

The ISR-MATB features an inter-dependent task flow in which information must be acquired and integrated into a decision to act or refrain. At the beginning of each trial, the target must be acquired in the PVT before other sub-tasks can be performed. Once the target has been acquired, it is used to perform the visual and auditory search tasks. Similarly, the visual and auditory search tasks must be completed in order to ascertain the the first binary cue for the decision task. The second binary cue is acquired by clicking on a designated button, which reveals whether confidence in the information is low or high. Once the cues are acquired and the correct rule is retrieved (see Table 1), the participant can decide the correct course of action.

Cognitive Agents

As illustrated in Figure 1, we developed five types of cognitive agents with performance profiles that differ as a function of workload. Although the cognitive agents are not based on high-fidelity cognitive models, they must operate with realistic cognitive constraints on performance. Importantly, this set of cognitive models provides a wide range of performance patterns against which the AMs can be stress tested.

Before proceeding, we note some common notation and characteristics across agents: Define $\theta_{s,j}$ as the accuracy of agent j on sub-task $s \in S = \{p, v, a, d\}$, which corresponds to the PVT, the visual search task, the auditory search task and the decision task, respectively. All cognitive agents guess on the visual and auditory search task if no response is provided to the PVT, which we denote as $y_p = 0$.

Constant

As the name implies, the accuracy of the Constant agent does not vary according to workload level. However, performance can differ between sub-tasks. See (Frame, Lopez, & Boydston, 2019a) for a similar approach. Each parameter value $\theta_{s,j}$ is randomly sampled from the following distribution: Uniform(.50, 1). Once a parameter is selected, it is fixed for

the duration of the simulation, making the expected accuracy constant.

Random-Dynamic

As a stress test for the AM, we developed a Random-Dynamic agent which changes on randomly selected sub-tasks after a set of 30 trials have been completed. Initial parameter values $\theta_{s,j} \sim \text{Uniform}(.50, 1)$. After a block of 30 trials has been completed, a new value for each accuracy $\theta_{s,j}$ is re-sampled with probability $p_{\text{change}} = .20$. Otherwise, the accuracy parameter remains the same for the next block.

Capacity-Limited

Performance of the Capacity-Limited Agent decreases as a function of workload—defined here simply as the number of tasks assigned to the cognitive agent. The probability of a correct response on sub-task s is given by the following piecewise equation:

$$\theta_{s,j} = \begin{cases} \frac{1}{1 + e^{-(\beta_{0,j} + \beta_{1,j} \times w)}} \\ .5 \end{cases} \quad \text{if } y_p = 0 \text{ and } s \in \{v, a\}$$

where $\beta_{0,j}$ is the intercept and $\beta_{1,j}$ the slope for agent j , and w is workload level. The slope $\beta_{1,j}$ represents the sensitivity of accuracy to changes in workload, such that negative values of $\beta_{1,j}$ lead to a decrease in accuracy with increasing levels of workload. The second piece of the equation above indicates the model guesses on the visual and auditory search tasks if no response on the PVT is provided within the response deadline. Parameter values are initialized such that $\beta_{0,j} \sim \text{Uniform}(0, 3)$ and $\beta_{1,j} \sim \text{Uniform}(-1, 0)$, subject to the constraint that $\theta_{s,j} \geq .5$ under maximum workload to ensure that performance cannot drop below chance levels.

Yerkes-Dodson

Accuracy for the Yerkes-Dodson agent follows a parabolic (e.g. inverse U-shaped) relationship with workload in which optimal performance is achieved with moderate levels of workload. Some evidence indicates that the relationship between arousal and performance might be parabolic under some circumstances (Yerkes & Dodson, 1908). Following this logic, if low levels of workload induce boredom or mind-wandering, and high levels of workload are overwhelming, then optimal performance for an agent might be achieved at a moderate level of workload. We make one small modification to the Capacity-Limited agent to incorporate this assumption:

$$\theta_{s,j} = \begin{cases} \frac{1}{1 + e^{-(\beta_{0,j} + \beta_{1,j} \times (w - 2.5)^2)}} \\ .5 \end{cases} \quad \text{if } y_p = 0 \text{ and } s \in \{v, a\}$$

The primary difference is that w is replaced with $(w - 2.5)^2$. Subtracting 2.5 from w places the maximum at the midpoint between one and four tasks and the exponent of 2 produces the parabolic relationship. Parameter values are sampled from the same distributions used for the Capacity-Limited agent.

Fatigue-Dynamic

The performance of the Fatigue-Dynamic agent is based on a dynamical system composed of two opposing processes. See (Patterson, Lochtefeld, Larson, & Christensen-Salem, 2019) for a related model. One process represents the gradual depletion of cognitive resources due to fatigue which is modulated by the instantaneous level of workload. An opposing recovery process replenishes the cognitive resource during periods of sufficiently low workload. If the net effect of the opposing processes is zero, the system achieves a state of equilibrium in which no change occurs. We approximate this dynamical process with the following logistic difference equation:

$$v_t = v_{t-1} + \Delta \times (\beta_{1,j} \times w_{t-1} + \beta_{2,j}) \times (v_{t-1} - v_{\min}) \times \left(1 - \frac{v_{t-1} - v_{\min}}{v_{\max,j} - v_{\min}}\right)$$

where v_{\min} and $v_{\max,j}$ are the lower and upper asymptotes of accuracy, respectively, t indexes the time step, $\Delta = 1$ (seconds) is the change in time per time step, $\beta_{1,j}$ the slope the fatigue decrement, w_{t-1} is the workload level at time step $t - 1$, and $\beta_{2,j}$ is the slope for the recovery process. Accuracy is defined as:

$$\theta_{s,j} = \begin{cases} v_t & \\ .5 & \text{if } y_p = 0 \text{ and } s \in \{v, a\} \end{cases}$$

At the beginning of each simulation, parameters were initialized as follows: $v_{\max,j} \sim \text{Uniform}(.85, 1)$, $\beta_{1,j} \sim \text{Uniform}(-.0015, -.0005)$, and $\beta_{2,j} \sim \text{Uniform}(|\beta_{1,j}|, 2 \times |\beta_{1,j}|)$. The purpose of constraining $\beta_{2,j}$ in terms of $\beta_{1,j}$ is to ensure that the neither the fatigue nor the recovery process are dominant at all levels of workload. We fixed $v_{\min} = .5$ to ensure that performance cannot fall below chance. We set initial accuracy to $v_0 = .9 \times v_{\max,j}$ under the assumption that initial accuracy is near the maximum.

Autonomous Managers

We developed five autonomous managers (AM) that base task allocation decisions on different cognitive models¹. Each AM has the following in common: First, each AM learns the performance profile of each cognitive agent from observed data. Second, each AM monitors ongoing performance and may change the task allocation after each block of 10 trials. Third, unless otherwise noted, sub-tasks are randomly allocated to cognitive agents on the first block to avoid bias in initial conditions.

Recent-Maximum

The Recent Maximum AM takes a data-driven approach, using the most recent block of trials as the best estimate of a performance of a cognitive agent. In other words, it makes no

¹We did not include an AM that uses a Fatigue-Dynamic agent as a model because parameters could not be reliably estimated with optimizers freely available in Java.

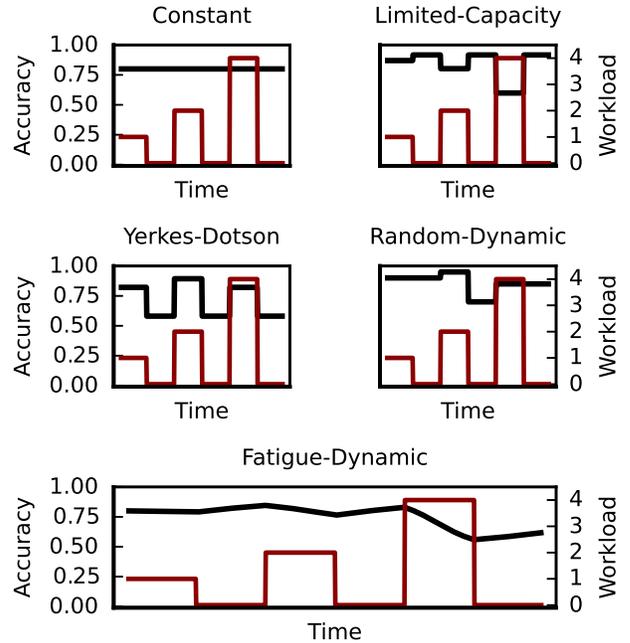


Figure 1: An illustration performance for each agent type as a function of workload. Black: agent performance on a single sub-task. Red: Agent workload.

assumptions about the relationship between workload and accuracy. After each block, the AM allocates the sub-task to the cognitive agent whose most recent block of data for that sub-task is the highest. In order to promote exploration in early blocks, the AM initializes each agent's performance history with a high accuracy of .90 for each sub-task.

Constant

The Constant AM assumes that the performance of each cognitive agent may differ by sub-task, but is otherwise constant across time and does not vary according to workload level. As such, the constant AM is similar to the Recent-Maximum AM, except it use all blocks of trials to estimate accuracy of each sub-task.

Capacity-Limited

The Capacity-Limited AM assumes that all cognitive agents are Capacity-Limited agents. After each block of trials, the Capacity-Limited AM estimates the parameters $\beta_{0,j}$ and $\beta_{1,j}$ from each agent's entire history of data. Using the maximum likelihood estimates, the Capacity-Limited AM iterates through all possible sub-task allocations and selects the allocation that maximizes expected accuracy on the decision task. Rather than allocating sub-tasks to agents randomly, the Capacity-Limited AM completes two exploration blocks to improve parameter estimation. During the first exploration block, one sub-task is allocated to the first agent, and three sub-tasks are allocated to the other agent. On the next exploration block, the sub-tasks are swapped between agents.

Yerkes-Dodson

The Yerkes-Dodson AM is identical to the Capacity-Limited AM, except it assumes the relationship between workload and accuracy is parabolic (e.g. inverted U-shape) rather than monotonically decreasing. After each block, the AM estimates the parameters for each cognitive agent. Using the best fitting parameter estimates, the AM selects the sub-task allocation that maximizes the accuracy of the decision task.

Random

As a point of reference, we include a Random AM, which randomly allocates sub-tasks to agents after each block. Thus, an AM is minimally successful if it performs better than the Random AM.

Simulation Design

We performed a set of simulations to assess the ability of different AMs to improve accuracy in the ISR-MATB by dynamically reallocating tasks to agents of different types. Each team consisted of two cognitive agents of the same type. We crossed cognitive agent type with each AM type to create a total of (cognitive agent type: 5) X (AM type: 5) = 25 simulation conditions.

All simulation conditions have several design parameters in common: First, the duration of each simulation was 60 minutes in simulated time. Second, each AM made a decision to reallocate the tasks among the cognitive agents after every block of 10 trials. Third, each simulation condition was repeated 500 times in order to approximate the expected accuracy for each AM.

Performance Evaluation

Our analysis focused on decision accuracy because that is the most important criterion in ISR operations. We normalized accuracy as a percentage of maximum possible performance for each simulation using the following formula:

$$a_{\text{norm}} = \frac{a_{\text{AM}} - a_{\text{min}}}{a_{\text{max}} - a_{\text{min}}}$$

a_{min} and a_{max} are the minimum and maximum possible expected accuracy, and a_{AM} is the expected accuracy of the AM's allocation. Using a normalized accuracy metric has several benefits. First, it adjusts for differences in the range of possible performance, which varies according to the type of cognitive agent as well as the difference in performance between the cognitive agents in a team. Second, it allows one to identify whether further improvement is possible.

Results

The results of the simulation are summarized in Figure 2. In most cases, AMs performed better than chance (i.e. the Random AM). One exception to this finding is that AMs performed similar to chance for Fatigue-Dynamic agents. Another finding was that the greatest performance was achieved when the model of the AM matched the cognitive agent. For

example, the best AM for Yerkes-Dodson agents was the Yerkes-Dodson AM (see first sub-plot in Figure 2). However, the advantage of using a model that matches the agent was not consistently large. Furthermore, complex models tended to be less robust when their assumptions were violated. For example, the Capacity-Limited and Yerkes-Dodson AMs performed poorly for Constant and Random-Dynamic agents. By contrast, the Constant AM, which uses a simple model of cognitive agent performance, was more robust across different types of cognitive agents. Although the Constant AM did not always achieve the best performance, it performed moderately well even when its internal model did not match the cognitive agents.

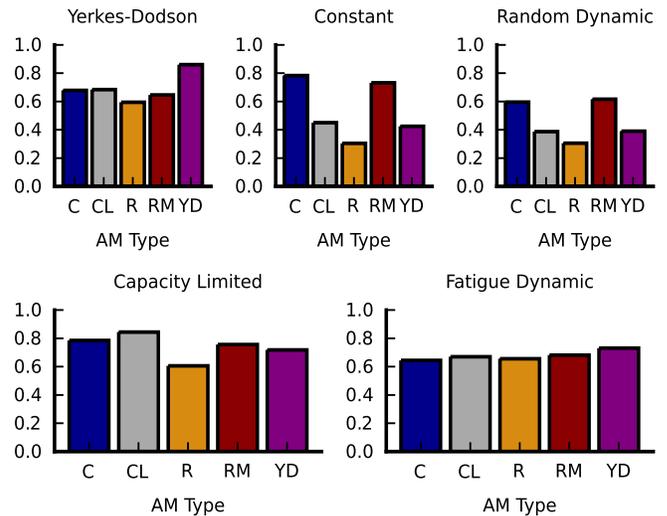


Figure 2: Sub-plots display % of maximum accuracy for each AM (colored bars) for a given agent type labeled in the sub-plot title. AM abbreviations: C: Constant, CL: Capacity-Limited, R: Random, RM: Recent-Maximum, YD: Yerkes-Dodson

Discussion

The goal of the present research was to explore how to integrate cognitive models into AMs to improve work productivity. AMs are designed to monitor performance of teams and dynamically allocate tasks to workers to improve performance. Cognitive models are an ideal candidate for augmenting the decision module of AMs because they can be used to predict the performance implications of alternative task distributions. Furthermore, the data requirements for most cognitive models are less onerous compared to deep neural networks and similar AI.

In our simulation study, we examined how AMs performed across a wide variety of conditions on a relatively complex ISR-themed task. We varied the type of cognitive agents that performed tasks and the internal model of the cognitive agent that the AM used to make task allocation decisions. One key finding is that AMs based on simple models were more robust

compared to those based on more complex models. The relationship between complexity and robustness is due, in part, to a well-known statistical phenomenon called variance-bias trade-off (Brighton & Gigerenzer, 2015). Models with more parameters—an indicator of complexity—produce more error variance due to over-fitting. In fact, an AM with a sufficiently complex internal model may perform at chance levels even if the internal model matches the cognitive agent performing the task. In addition, more complex models might be more brittle due to the increased number of assumptions that could be wrong.

One direction for future research is to investigate the use of cognitive architectures, such as ACT-R (Anderson et al., 2004). In the present research, we used simpler cognitive models because they are tractable and generate a wide variety of distinct performance profiles. However, cognitive architectures provide the opportunity to explore additional interventions, such as providing feedback to strengthen declarative memory, or prescribing more effective strategies for task completion.

Conclusion

Cognitive models have a wide range of applications. The present research demonstrates how cognitive models can be incorporated into technologies and the design process to improve task performance.

Acknowledgments

The opinions expressed herein are solely those of the authors and do not necessarily represent the opinions of the United States Government, the U.S. Department of Defense, the U.S. Air Force, or any of their subsidiaries or employees. The contents have been reviewed and deemed Distribution A. Approved for public release. Case number: AFRL-2021-1002.

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological review*, *111*(4), 1036.

Brighton, H., & Gigerenzer, G. (2015). The bias bias. *Journal of Business Research*, *68*(8), 1772–1784.

Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual rt task during sustained operations. *Behavior research methods, instruments, & computers*, *17*(6), 652–655.

Frame, M., Lopez, J., & Boydston, A. (2019a). *Dynamically managing task allocation between humans and machines in surveillance operations* (Tech. Rep.). Wright State Research Institute Beaver Creek United States.

Katidioti, I., Borst, J. P., van Vugt, M. K., & Taatgen, N. A. (2016). Interrupt me: External interruptions are less disruptive than self-interruptions. *Computers in Human Behavior*, *63*, 906–915.

McClelland, J. L. (2009). The place of modeling in cognitive science. *Topics in Cognitive Science*, *1*(1), 11–38.

McNeese, N. J., Demir, M., Cooke, N. J., & Myers, C. (2018). Teaming with a synthetic teammate: Insights into human-autonomy teaming. *Human factors*, *60*(2), 262–273.

Newell, A. (1990). *Unified theories of cognition*. Harvard University Press.

Nguyen, A., Yosinski, J., & Clune, J. (2015). Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 427–436).

Patterson, R. E., Lochtefeld, D., Larson, K. G., & Christensen-Salem, A. (2019). Computational modeling of the effects of sleep deprivation on the vigilance decrement. *Human factors*, *61*(7), 1099–1111.

Rackauckas, C., Ma, Y., Martensen, J., Warner, C., Zubov, K., Supekar, R., ... Edelman, A. (2020). Universal differential equations for scientific machine learning. *arXiv preprint arXiv:2001.04385*.

Riefer, D. M., Knapp, B. R., Batchelder, W. H., Bamber, D., & Manifold, V. (2002). Cognitive psychometrics: Assessing storage and retrieval deficits in special populations with multinomial processing tree models. *Psychological Assessment*, *14*(2), 184.

Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. *Human factors*, *48*(2), 362–380.

Santiago-Espada, Y., Myer, R. R., Latorella, K. A., & Comstock Jr, J. R. (2011). The multi-attribute task battery ii (matb-ii) software for human performance and workload research: A user's guide.

Sieck, W. R., & Yates, J. F. (2001). Overconfidence Effects in Category Learning: A Comparison of Connectionist and Exemplar Memory Models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(4), 1003–1021. doi: 10.1037//0278-7393.27.4.1003

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, *12*(1), 97–136.

Won, J. C., Condon, G. R., Landon, B. R., Wang, A. R., & Hannon, D. J. (2011). Assessing team workload and situational awareness in an intelligence, surveillance, and reconnaissance (ISR) simulation exercise. In *2011 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSima)* (pp. 163–167).

Yechiam, E., Busemeyer, J. R., Stout, J. C., & Bechara, A. (2005). Using cognitive models to map relations between neuropsychological disorders and human decision-making deficits. *Psychological science*, *16*(12), 973–978.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Punishment: Issues and experiments*, 27–41.

Zhou, D.-X. (2020). Universality of deep convolutional neural networks. *Applied and computational harmonic analysis*, *48*(2), 787–794.

Derivation of Metric Scales from Ordinal Data with Guttman-Goode’s Scaling

Víthor Rosa Franco (vithorfranco@gmail.com)

Postgraduate Program in Psychology, Department of Psychology, São Francisco University, Campinas Campus, Rua Waldemar César da Silveira, Jardim Cura D’Ars, 13045510 - Campinas, SP, BRA

Keywords: Measurement theory; Guttman scale; Goode scaling; Psychometrics

Introduction

Psychometric modeling usually assumes that the observed behavior is caused by a set of metric latent variables. For instance, the Rasch model, one of the most traditional models from the Item Response Theory (Embretson & Reise, 2013), assumes that the probability of getting an answer right (or saying yes, or agreeing to the statement, or simply that $X = 1$) is equal to a logistic transformation of an additive interaction between the respondent’s true score θ and the item’s difficulty b . Formally, the model is represented as:

$$P(X = 1) = \frac{1}{1 + \exp(\theta - b)}. \quad (1)$$

This type of model is used mainly to estimate and develop interval measures for θ and b . Perline et al. (1979) argued that this is possible because the Rasch model is a stochastic variant of the Additive Conjoint Measurement Theory (Luce & Tukey, 1964). The Additive Conjoint Measurement Theory is a formal theory of continuous quantities which allows for the derivation of interval scales from ordinal data, as long as some empirical relations are observed.

However, some authors have disputed this view that the Rasch model is a stochastic variant of the Additive Conjoint Measurement Theory (e.g., Michell, 2008). More specifically, it has been argued that if the Rasch model is a probabilistic version of the Guttman scale (Guttman, 1944), which allows only for θ and b to be measured in the same ordinal scale, then the Rasch model provides an interval measure only because it is modeling response error. This apparent inconsistency is called the Rasch paradox.

On the other hand, the Rasch paradox has also been disputed (e.g., Borsboom & Zand Scholten, 2008). Regardless of whether the Rasch paradox is real or not, it would be interesting for psychometric researchers if interval, or even ratio, (i.e., metric) scales could be derived from Guttman scales without reliance on response errors. The aim of the present study is to propose a procedure that combines the probabilistic Guttman scaling (Proctor, 1970) with Goode’s method (Coomb’s, 1964) to obtain metric scales from dichotomous psychometric data. We call this procedure the Guttman-Goode’s Scaling (GGS).

Guttman-Goode’s Scaling

The GGS procedure combines two methods for deriving interval and ratio scales from psychometric data. The first is the probabilistic Guttman scaling (Proctor, 1970). Guttman scales assume that the respondent will answer $X = 1$ if and

only if $\theta > b$. Otherwise, the respondent will answer $X = 0$. If this condition is exactly met, the matrix (or Guttman scalogram) of response patterns averaged by sum scores (for an instrument with 5 items) will be equal to the matrix represented in Table 1. It is possible to see that all cells are equal to 0 or 1, representing that all individuals with a specific ordinal θ level answered to the items in the same way (e.g., a correct answer, 1, or an incorrect answer, 0).

Table 1: Perfect Guttman scalogram of response patterns averaged by sum scores for an instrument with 5 items.

θ level	Item 1	Item 2	Item 3	Item 4	Item 5
0	0	0	0	0	0
1	1	0	0	0	0
2	1	1	0	0	0
3	1	1	1	0	0
4	1	1	1	1	0
5	1	1	1	1	1

However, real data seldomly result in a perfect Guttman scalogram, as represented in Table 2, which was calculated from a toy dataset with actual answers from respondents. Therefore, traditional Guttman scaling cannot be applied to this type of scenarios. The probabilistic Guttman scaling, then, estimates the probability of both the order of the items as well as the ordinal θ level by assuming that only the $v + 1$, where v is the number of items, levels of θ are distinguishable. This differs from the Rasch model, for instance, that allows for more than $v + 1$ values of θ to be estimated.

Table 2: Empirical Guttman scalogram of response patterns averaged by sum scores for an instrument with 5 items.

θ level	Item 1	Item 2	Item 3	Item 4	Item 5
0	0	.253	.126	.149	.092
1	1	0	.243	.216	.027
2	1	1	0	.285	.228
3	.727	1	1	0	0
4	.760	.880	1	1	0
5	.836	.873	.945	.727	1

After estimating the ordinal θ level and the order of the items, our procedure uses Goode’s method to analytically derive an interval scale from an ordered metric scale of respondents and items. The ordered metric scale is a scale derived from the data dependent on empirical relations

regarding the distance between a respondent ordinal θ level and two items' orders (i.e., b_1 and b_2), formally stated as:

$$\overline{\theta I_1} > \overline{\theta I_2}. \quad (2)$$

Equation 2 is simply an order relation of order relations (i.e., order relation of distances). Research in measurement theory has shown that ordered metric scales impose constraints on the uniqueness of numerical representations that can be derived from simple ordinal data (Coombs, 1964, p. 359). We propose that, for Guttman scales, the order relation of distances can be found by taking the average of the rows and the complement (i.e., 1 minus) of the average of the columns of the empirical Guttman scalogram. Because we know that the ordinal θ level represented with 0 is the smallest possible value and that the ordinal θ level represented with 5 is the largest possible value, we can use the aforementioned averages to create a dominance matrix, such as the one represented in Table 3. This table is created based on the distance between each point (i.e., a θ level or an item) and the ordinal θ level represented with 0.

Table 3: Dominance matrix of θ levels and item orders.

	I4	θ_1	I3	θ_2	I1	θ_3	I2	θ_4	I5	θ_5
I4	0	1	1	1	1	1	1	1	1	1
θ_1	0	0	1	1	1	1	1	1	1	1
I3	0	0	0	1	1	1	1	1	1	1
θ_2	0	0	0	0	0	1	1	1	1	1
I1	0	0	0	1	0	1	1	1	1	1
θ_3	0	0	0	0	0	0	1	1	1	1
I2	0	0	0	0	0	0	0	1	1	1
θ_4	0	0	0	0	0	0	0	0	1	1
I5	0	0	0	0	0	0	0	0	0	1
θ_5	0	0	0	0	0	0	0	0	0	0

Note. I is an acronym for "Item".

For the next step of Goode's method, one must choose the value for the smallest distance (represented as Δ_0) and then analytically derive the next distances, Δ_j , for each distance j . We adapt Goode's original equation to the current scenario and propose the following equation for calculating Δ_j :

$$\Delta_j = \Delta_0(CS(j) + 1) + CS(j), \quad (3)$$

where $CS(\cdot)$ is the sum of the column representing the distance j . For instance, $CS(I4)$ is equal to 0 and $CS(\theta_5)$ is equal to 9. Finally, the last step involves attributing values for each point. Arbitrarily, the smallest point θ_0 may be set to 0. The other points can simply be attributed their Δ_j values, as these were calculated based on the points distance in relation to θ_0 .

After analytically deriving all the scale values, which are measured in an interval level, one may wish to estimate how well this numeric approximation represents the data. One way of doing this is using a logistic or hyperbolic tangent function on the linearly transformed scale values and compare the results with the empirical Guttman scalogram. For the logistic function, we propose:

$$AM(\theta, I) = \frac{1}{1 + \exp(\psi_\theta - \psi_I)}, \quad (4)$$

where ψ_θ and ψ_I are, respectively, the normalized interval level measure for the ordinal θ level and for the item. For the hyperbolic tangent function, we propose:

$$DM(\theta, I) = 2 \frac{1}{1 + \exp[-2(\varphi_\theta \varphi_I)]} - 1, \quad (5)$$

where $\varphi_I = \exp(\psi_I)$, as the exponential transform of the normalized interval scale values results in a ratio scale (Fishburn, 1974). Applying this procedure to the data that generated Table 2 results in a RMSE of .015 to the logistic function approach and in a RMSE of .184 to the hyperbolic tangent function approach. This result suggests that an interval representation is better than a ratio representation of the points.

Final Considerations

The GGS procedure can be used with any data following a direct response design (such as attitude or performance psychometric scales). The main advantage of the GGS procedure is that, different from Item Response Theory models, the scales are derived from ordered metric information in the data and, therefore, should be less reliant on response error. However, it should be noted that this is an initial implementation of the GGS procedure and limitations abound. For instance, Table 3 presents an intransitivity for θ_2 which is not dealt with. We also do not estimate the distances between intermediary points such as $\overline{\theta_2 I_1} > \overline{\theta_1 I_2}$, which hides an implicit assumption that a unidimensional representation is the most appropriate (Coombs, 1964). Future studies should deal with these limitations to provide more robust metric scales for psychometric data.

References

Borsboom, D., & Zand Scholten, A. (2008). The Rasch model and conjoint measurement theory from the perspective of psychometrics. *Theory & Psychology, 18*, 111–117.

Coombs, C.H. (1964) *A theory of data*. Wiley.

Embretson, S. E., & Reise, S. P. (2013). *Item response theory for psychologists*. Lawrence Erlbaum.

Fishburn, P. C. (1974). von Neumann-Morgenstern Utility Functions on Two Attributes. *Operations Research, 22*(1), 35-45.

Guttman, L. (1944). A basis for scaling qualitative data. *American Sociological Review, 9*, 139–150.

Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology, 1*(1), 1-27.

Michell, J. (2008). Conjoint measurement and the Rasch paradox: A response to Kyngdon. *Theory & Psychology, 18*(1), 119-124.

Perline, R., Wright, B. D., & Wainer, H. (1979). The Rasch model as additive conjoint measurement. *Applied Psychological Measurement, 3*(2), 237-255.

Proctor, C. H. (1970). A probabilistic formulation and statistical analysis for Guttman scaling. *Psychometrika, 35*, 73-78.

Criticality Perception in Dynamic Traffic Scenarios: An ACT-R Model

Noémi Földes (noemi.foeldes@uni-ulm.de)

Department of Human Factors, Albert-Einstein-Allee 41
Ulm University, Germany

Moritz Held (moritz.held@uni-oldenburg.de)

Department of Psychology, Küpkersweg 74, 26129
University Oldenburg, Germany

Tanja Stoll (tanja.stoll@uni-ulm.de), Martin Baumann (martin.baumann@uni-ulm.de)

Department of Human Factors, Albert-Einstein-Allee 41
Ulm University, Germany

Keywords: criticality perception; dynamic driving scenario; ACT-R; situation awareness; cognitive modeling

Introduction

Determining the criticality of a traffic situation is a basic task that has to be accomplished in driving. Several theories assume that human drivers' evaluation of the criticality of a dynamic traffic situation is strongly determined by the time-to-collision (TTC) that is the time until two objects will collide if they both maintain speed and course (e.g., Heesen et al., 2012; Tamke et al., 2011). The evaluation of the situation's criticality strongly influences the drivers' action decisions in these situations. One of such dynamic situations where the evaluation of criticality is mainly based on the TTC to other vehicles is a lane change scenario.

Stoll et al. (2018) investigated in a video-based study the following lane change scenario where the criticality of the situation was systematically varied: Participants (Ego) drove on the left lane of a German 2-lane highway. They observed a passenger car (RU1) approaching a slow vehicle (RU2) on the right lane that might cut in to the participants' left lane. They were asked (1) whether they would accelerate, decelerate or maintain speed in this situation and (2) to rate the criticality of the situation on a scale from 1 (not critical) to 5 (very critical). Stoll et al. (2018) varied the criticality of the situation by the TTC between participants and RU1 (TTC_{Ego} , either 2, 4 or 6 s) and TTC between RU1 and RU2 (TTC_{RU1} , either 2, 4 or 6 s) at the time participants had to make their decision.

Even though findings suggest a relationship between perceived criticality and selecting the preferred action (maintaining speed was associated with rather low criticality ratings compared to decelerating and accelerating), the TTC values did not reliably trigger typically preferred actions, resulting in a large variance among participants. More importantly, this variance calls for more clarification on

exactly how critical vs. non-critical scenarios were perceived in the different TTC conditions.

We are developing a cognitive model using ACT-R (Anderson, 2007) to shed light on the complex cognitive processes of situation awareness (SA: perception, comprehension, projection; Endsley, 1995) in the highly dynamic traffic scenario of Stoll et al. (2018), in order to determine how participants evaluate the different conditions as critical or not.

Most importantly, we argue that not only the TTC, but the combination of perceived elements and the availability of memories containing them and which help build up a situation model (SM) are also part of the resulting perceptual decision (PD¹) participants make about criticality. We assume that these elements do not merely consist in the perception from the driver's own perspective, but the RU1's viewpoint and intention are taken into account as well. With other words, we suggest that the driver's SM includes the RU1's SM to a certain extent.

Method

ACT-R (Adaptive Control of Thought—Rational, Anderson, 2007) is a cognitive architecture with basic assumptions about human knowledge and about how information in the declarative memory (chunks) are used (production rules) to solve everyday tasks.

We are using ACT-R (Salvucci, 2006²) to recreate the driving scenario and to model participants' memory retrieval (MR) to build up SM and that leads to a PD about criticality.

General Assumptions of the Model

The ACT-R model assumes first of all that criticality perception does not mainly stem from perception, but from MR. Accordingly, even if the presented driving scenario was not familiar to participants, a "close enough" memory matching some of the perceived elements is retrieved to

¹ A PD can be considered as an intuitive decision (Thomson et al., 2015; or System 1 decision; Kahneman, 2011)

² Since its features are more suitable for driving than the original LISP version, we are using the Java version of ACT-R. (see <https://www.cs.drexel.edu/~salvucci/cog/act-r/>)

create SM. Further assumptions are listed below and depicted on Figure 1.

1. SM_{Ego} is created through MR.
2. MR takes place regardless, but it is approximative and resulting SM_{Ego} might be incorrect.
3. SM_{Ego} includes $Projection_{Ego}$ (i.e., how the situation is going to develop) and $Intention_{RU1}$ (i.e., RU1's action plan).

4. $Projection_{Ego}$ gets periodically confirmed by monitoring RUs. As long as $Projection_{Ego}$ is valid, no further MR takes place.
5. If $Projection_{Ego}$ is not valid:
 - (1) The model establishes new SM_{Ego} through MR and
 - (2) Makes PD (including a criticality decision [whether there interference with RU1 can be expected or not] and a certainty value [i.e., reliability of the SM_{Ego} based on how many times it needed updating]).

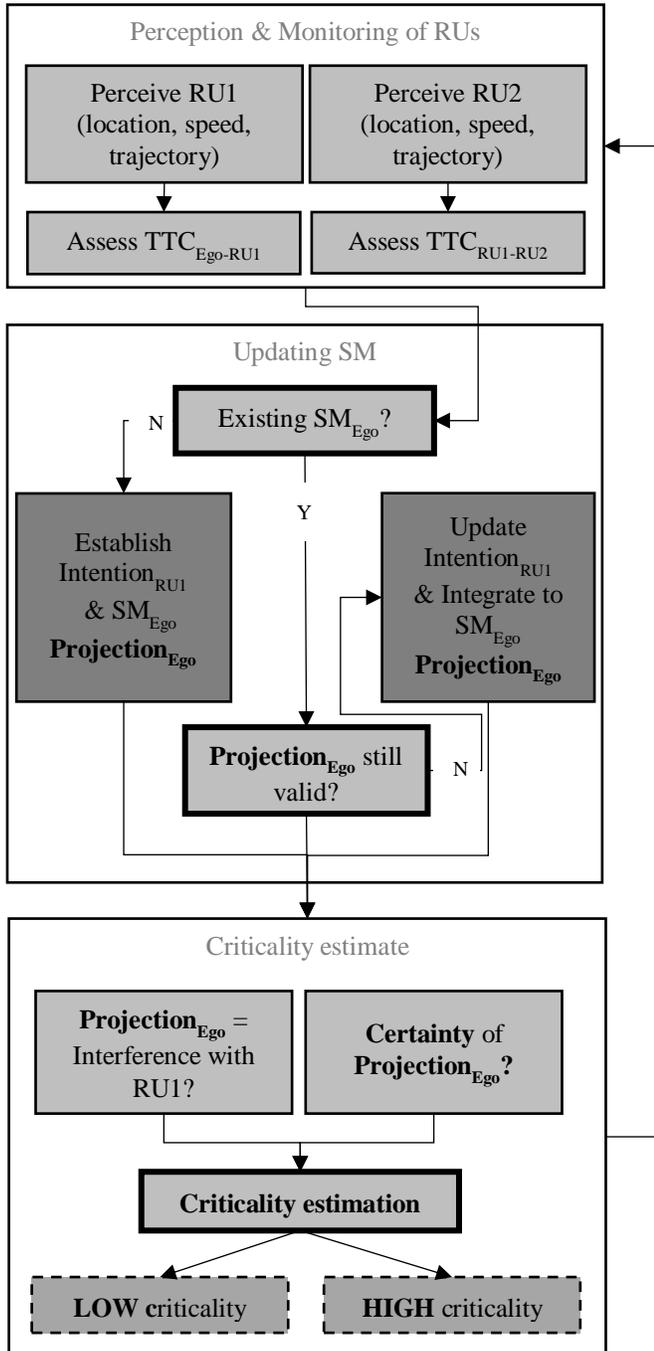


Figure 1: Flowchart representation of the model's main steps

We expect the model to reliably reproduce participants' subjective criticality ratings in the different TTC conditions in the study of Stoll et al. (2018).

Acknowledgments

This project was funded within the Priority Programme “CoInCar—Cooperatively Interacting Automobiles” of the German Science Foundation DFG.

References

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York, NY: Oxford University Press. ISBN 0-19-532425-0.

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), 32–64. <https://doi.org/10.1518/001872095779049543>.

Heesen, M., Baumann, M., Kelsch, J., Nause, D., Friedrich, M. (2012) *Investigation of cooperative driving behaviour during lane change in a multi-driver simulation environment*. In: Human Factors and Ergonomics Society (HFES) Europe Chapter Conference Toulouse

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Strausand Giroux.

Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors*, 48, 362-380.

Stoll, T., Müller, F. & Baumann, M. When cooperation is needed: the effect of spatial and time distance and criticality on willingness to cooperate. *Cogn Tech Work*, 21, 21–31 (2019). <https://doi.org/10.1007/s10111-018-0523-x>

Tamke, A., Dang, T., Breuel, G. (2011) *A flexible method for criticality assessment in driver assistance systems*. In: Intelligent vehicles symposium (IV). IEEE

Thomson, R., Lebiere, C., Anderson, J. R., & Staszewski, J. (2015). A General Instance-Based Learning Framework for Studying Intuitive Decision Making in a Cognitive Architecture. *Journal of Applied Research in Memory and Cognition*; Special Issue on Modeling and Aiding Intuitions in Organizational Decision Making

Sampling Heuristics for Active Function Learning

Rebekah A. Gelpi¹ (rebekah.gelpi@mail.utoronto.ca), Nayan Saxena² (nayan.saxena@mail.utoronto.ca),
George Lifchits³ (glifchits@cs.toronto.edu), Daphna Buchsbaum⁴ (daphna@brown.edu),
Christopher G. Lucas⁵ (clucas@ed.ac.uk)

¹Department of Psychology, University of Toronto

²Department of Statistical Sciences, University of Toronto

³Department of Computer Science, University of Toronto

⁴Department of Cognitive, Linguistic, and Psychological Sciences, Brown University

⁵School of Informatics, University of Edinburgh

Abstract

People are capable of learning diverse functional relationships from data; nevertheless, they are most accurate when learning linear relationships, and deviate further from estimating the true relationship when presented with non-linear functions. We investigate whether, when given the opportunity to learn actively, people choose samples in an efficient fashion, and whether better sampling policies improve their ability to learn linear and non-linear functions. We find that, across multiple different function families, people make informative sampling choices consistent with a simple, low-effort policy that minimizes uncertainty at extreme values without requiring adaptation to evidence. While participants were most accurate at learning linear functions, those who more closely adhered to the simple sampling strategy also made better predictions across all non-linear functions. We discuss how the use of this heuristic might reflect rational allocation of limited cognitive resources.

Keywords: Function learning; active learning; sampling

Introduction

People must often learn and generalize from relationships between continuous quantities, where these relationships can take diverse forms. Temperatures rise and fall with the changing of seasons, trees grow steadily from saplings until they are fully mature, crops have a “sweet spot” of climatic constraints such as humidity and cold, foraging for food rests on contextual variables such as animal populations and water supply, and diseases can exponentially increase in the absence of constraints. In order to represent these relationships and use them to make accurate predictions, we must learn the underlying function from sparse observations to be able to predict unseen outcomes in a variety of new scenarios.

In addition to the general challenge of learning a function faithfully, most investigations of explicit function learning have focused on the human ability to use observed data to interpolate between previously observed points and to extrapolate beyond the limits of their experience. Given sufficient evidence, people can learn a wide variety of functional relationships (Bott & Heit, 2004; Lucas, Sterling, & Kemp, 2012; Wilson, Dann, Lucas, & Xing, 2015), but their inductive biases strongly favor linear relationships: people tend to learn linear relationships better than non-linear ones (Brehmer, 1974) and often extrapolate linearly even having learned that a relationship is non-linear in the data they have observed (DeLosh, McDaniel, & Busemeyer, 1997; Kalish, 2013).

Given this systematic bias towards linearity, learning non-linear relationships may require more or better evidence to overwhelm this strong *a priori* belief. Standard function learning experiments—at least those that have revealed the human ability to learn non-linear relationships—tend to rely on providing overwhelming evidence, such as multiple presentations of the same data point, large numbers of training examples, or multiple blocks of training (e.g., DeLosh et al., 1997; Kalish, Lewandowsky, & Kruschke, 2004; Kalish, 2013). This is at odds with the view that people are efficient learners, able to make good use of sparse evidence.

We hypothesize that difficulties with learning non-linear relationships may be, at least in part, an artifact of using passive observational designs—that is, designs in which participants do not choose which points to learn about. The evidence participants are presented with in these experiments is often of limited utility, with the most informative observations presented alongside a large number of comparatively unhelpful ones, increasing participants’ attentional and memory burden, as well as fatigue. While active learning provides benefits beyond efficient selection of samples (e.g., Markant & Gureckis, 2014a; Markant, Ruggeri, Gureckis, & Xu, 2016), we focus here on testing whether people tend to choose useful samples, what policies may underlie the samples that people choose, and whether their sampling strategies facilitate better learning of linear as well as non-linear functions.

Previous work in active learning has suggested that people can effectively learn linear functions by focusing their sampling on regions of high uncertainty (Jones, Schulz, Meder, & Ruggeri, 2018). An uncertainty-based sampling policy could also be useful for non-linear functions, as maximizing information gain about the extrema of a function eliminates the need for extrapolation, which can otherwise be inaccurate in non-linear functions (DeLosh et al., 1997; Kalish et al., 2004). However, the computational demand of iteratively adjusting one’s sampling strategy could mean that, especially in more complex non-linear domains, people trade off optimal behaviour against the cognitive or temporal costs of doing so (Gershman, Horvitz, & Tenenbaum, 2015). In these circumstances, using less accurate or flexible heuristics could nonetheless be *resource-rational*, or optimal under constraint (Gigerenzer, 2008; Lieder & Griffiths, 2020).

In the following sections, we first present a Gaussian process-based framework for representing the task of func-

tion learning, describing how a learner can use their existing knowledge about a functional relationship to interpolate or extrapolate to unknown function values. We then introduce simple candidate strategies that an active learner might use to select samples and an experimental task to assess humans' sampling behaviour and predictions, and compare these to simulated learning under different sampling policies.

Gaussian Process Model

We use a Gaussian process (GP) model (e.g., Griffiths, Lucas, Williams, & Kalish, 2008; Lucas, Griffiths, Williams, & Kalish, 2015) to provide a general framework for understanding both rule-based and similarity-based function learning. This model uses samples $\mathbf{x}_n = (x_1 \dots x_n)$ to further approximate a learned function f by inducing a Gaussian distribution on the observed $y_i = f(x_i)$ values based on sampled x_i values (Rasmussen & Williams, 2006). For known function outputs f and new unknown points f_* the joint probability distribution is then defined as

$$\begin{pmatrix} f \\ f_* \end{pmatrix} = \mathcal{N}\left(\begin{bmatrix} \mu \\ \mu_* \end{bmatrix}, \begin{bmatrix} K & K_* \\ K_*^T & K_{**} \end{bmatrix}\right)$$

In the above equation, we have $K = k(x, x)$, $K_* = k(x, x_*)$ and $K_{**} = k(x_*, x_*)$, where k denotes the generalized class of Matérn kernel given by:

$$k(x_i, x_j) = \frac{1}{\Gamma(\nu)2^{\nu-1}} \left(\frac{\sqrt{2\nu}}{\ell} d(x_i, x_j)\right)^\nu K_\nu\left(\frac{\sqrt{2\nu}}{\ell} d(x_i, x_j)\right)$$

where d denotes Euclidean distance, K_ν is a modified Bessel function, and Γ is the standard Gamma function. Throughout this paper we use this model with smoothness $\nu = 1.5$ and length-scale $\ell = 0.1$ as these parameters provided consistent, natural and realistic function interpolations for varied distributions of known data and functional forms.

Sampling Strategies

Under this GP framework, as we get further from previously-observed points, our uncertainty increases, which could lead to more heterogeneous or inaccurate extrapolation if a function's minimum and maximum values are not known. Therefore, for a learner with limited opportunities to sample, we expect that the most effective and informative strategies will include sampling the extrema. We assess the informativeness of three policies: uniform random sampling, an equidistant sampling policy that selects the minimum and maximum values and interpolates equally between them, and uncertainty-based sampling.

Random sampling One simple strategy to learn about a function is to sample randomly from the possible domain of values. While random sampling is straightforward, decision-makers may display substantial sub-optimality in their sampling choices if they happen to sample multiple points in

close proximity that are unlikely to be maximally informative. Given human inductive biases towards linearity, this could result in difficulty extrapolating or interpolating non-linear functions where no samples have been drawn. We represent this policy as drawing samples from a distribution where each sample $x_i \sim \text{Uniform}(0, 1)$.

Equidistant sampling As mentioned previously, sampling the minimum and maximum feasible values within the confines of the problem to be solved minimizes the need for extrapolation, and partitioning the remaining space among the remaining samples likewise balances interpolation between the remaining points, making it suitable for relatively accurate prediction across many commonly-encountered functions. While somewhat inflexible, this may represent a highly tractable, efficient heuristic for sampling within a limited domain. We represent this sampling approach as taking the N samples to be drawn and allotting them in such a way that the points sampled roughly reflect $N - 1$ equal partitions of the space to be sampled:

$$x_i \sim \text{Beta}\left(1 + \frac{i\rho}{N-1}, 1 + \frac{(N-1-i)\rho}{N-1}\right) \quad (i \in \{0 \dots N-1\})$$

To reflect a moderate preference for sampling from the peaks of the beta distributions while allowing for some samples to be drawn from nearby values, we chose a free parameter value of $\rho = 10$ that denotes the size of the peaks of the sampling distribution.

Adaptive sampling The two above algorithms reflect sampling procedures that do not dynamically update based on new information. In order to choose the best samples, however, it may be effective to adapt one's strategy to account for already sampled points. Given that people preferentially draw samples in regions that resolve current uncertainty, even when this does not maximize information gain (Markant & Gureckis, 2014b), we designed a sampling algorithm motivated by a myopically ideal strategy that adapts to new information by identifying the point of highest uncertainty.

This sampling procedure uses a GP model that greedily chooses points from the domain-space by iteratively fitting the model on the previously sampled points, and choosing the next sample as the point on the posterior distribution of GP functions with the highest variance.

To test the effectiveness of these algorithms and compare their performance to human sampling strategies as well as the learned functions, we designed an experimental task in which participants must select a small number of samples to try to optimize their performance in a prediction task.

Experimental Design

Recruitment and Procedure 98 adult participants ($M_{\text{age}} = 31.8, SD_{\text{age}} = 11.1$) were recruited through Prolific and paid £1.00 for completing an online learning task presented via a web-based program.

Familiarization and Exposure In the task, participants were told that they would be playing the role of a scientist. The scientist’s job was to learn about a number of possible drugs, each being researched for their role in improving patients’ well-being. Ultimately, the goal of the participant was to learn the relationship between the length of time that a drug was provided to patients and the patient’s well-being scores.

Before participating in the experimental trials, participants were familiarized with a warm-up trial where they were presented with two empty horizontal bars. The first bar was labelled *orange juice consumed* and the second bar was labelled *hours of sleep*. Participants were told that they could drag the first bar to change the amount of orange juice consumed. After confirming their choice, they learned about the amount of sleep of an individual who consumed the specified amount of orange juice, which is displayed on the second bar. Next, participants were shown a specified amount of orange juice consumed, and asked to predict the amount of sleep on the second bar by dragging the bar to the predicted amount.

Experimental Task After completing the familiarization trial, participants completed the experimental task. Participants learned about four total drugs, presented in a randomized order of four blocks of trials. Each drug had a predefined relationship to the well-being of the patient: positive linear ($y = 0.8x + 0.1$), exponential ($y = 100^{(x-1)}$), quadratic ($y = 0.95 - \frac{(x-0.5)^2}{0.3}$), or periodic ($y = 0.5 \sin(7x) + 0.5$). For each drug, participants completed a block of trials with the same format as the warm-up. They first sampled time points since the patient began taking the drug, to learn about its effect on well-being at a given time point, by manipulating a blue horizontal bar (Figure 1). After sampling 5 points to learn about, participants were asked to predict the well-being of a patient at 8 randomly presented pre-selected time points (values of $x \in \{0.01, 0.15, 0.29, 0.43, 0.57, 0.71, 0.85, 0.99\}$), displayed on the blue bar. Participants were not given the exact values of data points, so the domains of x and y presented here are arbitrary.



Figure 1: Participants chose 5 points to learn about the drug’s effect on well-being by changing the value of the blue bar (top), and then made 8 predictions using the red bar (bottom).

Results

Simulations To analyze the effectiveness of the identified sampling strategies and establish comparison baselines, we generated synthetic data for 100 participants under each sampling strategy: random, equidistant, and adaptive sampling, and fitted Gaussian process models for each function (Figure 2) using the `sklearn` library in Python. The correlation coefficient and the pooled and average mean squared error (MSE) of the GP fits were then computed and compared (Table 1).

Sampling randomly (row 1) resulted in reasonably good approximations of the true function at the centre but demonstrated wider deviations at the minimum and maximum values. As a large number of the simulated participants failed to sample from the extrema of the function, this led to more inaccurate extrapolations in areas where no samples were drawn. The pooled and average MSE of simulations of this policy demonstrates a comparative disadvantage at approximating the functions when compared to equidistant and uncertainty minimization sampling.

In contrast, the process of sampling approximately equidistant x values (row 2) appeared to perform better at producing an estimate for the true functions. A considerable improvement in terms of the fit of the Gaussian process model can be visibly noticed as compared to random sampling. Notably, equidistant sampling generated the lowest pooled and average individual MSE of the three sampling strategies.

Finally, similar to the equidistant heuristic strategy, the adaptive sampling algorithm (row 3) frequently sampled from regions in the neighbourhood of zero and one. Overall, the adaptive sampling strategy yielded comparable, though slightly higher, MSE values to the equidistant strategy.

Given the cognitive costs of adaptive sampling, even when calculated approximately rather than through the representation of the full posterior distribution of candidate functions, we do not expect people to utilize this strategy to draw samples. Nevertheless, the similarity of this adaptive strategy to a simple heuristic strategy of sampling evenly spaced points in addition to the minimum and maximum values, suggests that if people use such a strategy, they will select comparably informative points to learn about, improving their subsequent predictions of the true function’s value.

Sampling Task Gaussian process models were also fitted on samples for each of the 98 human participants across all four functions. These models were then compared against previously established baselines produced by fitting GP models on generated data from different sampling strategies.

As observed in Figure 2 (row 4), the models fitted on human samples were able to predict the true function and were comparable to the equidistant or adaptive sampling strategies in terms of pooled approximation error (MSE), although people’s choices showed considerable variability and some individuals chose less informative points to learn about; for example, 15 participants did not sample a value above 0.5 for at least one of the functions. As a result, the average

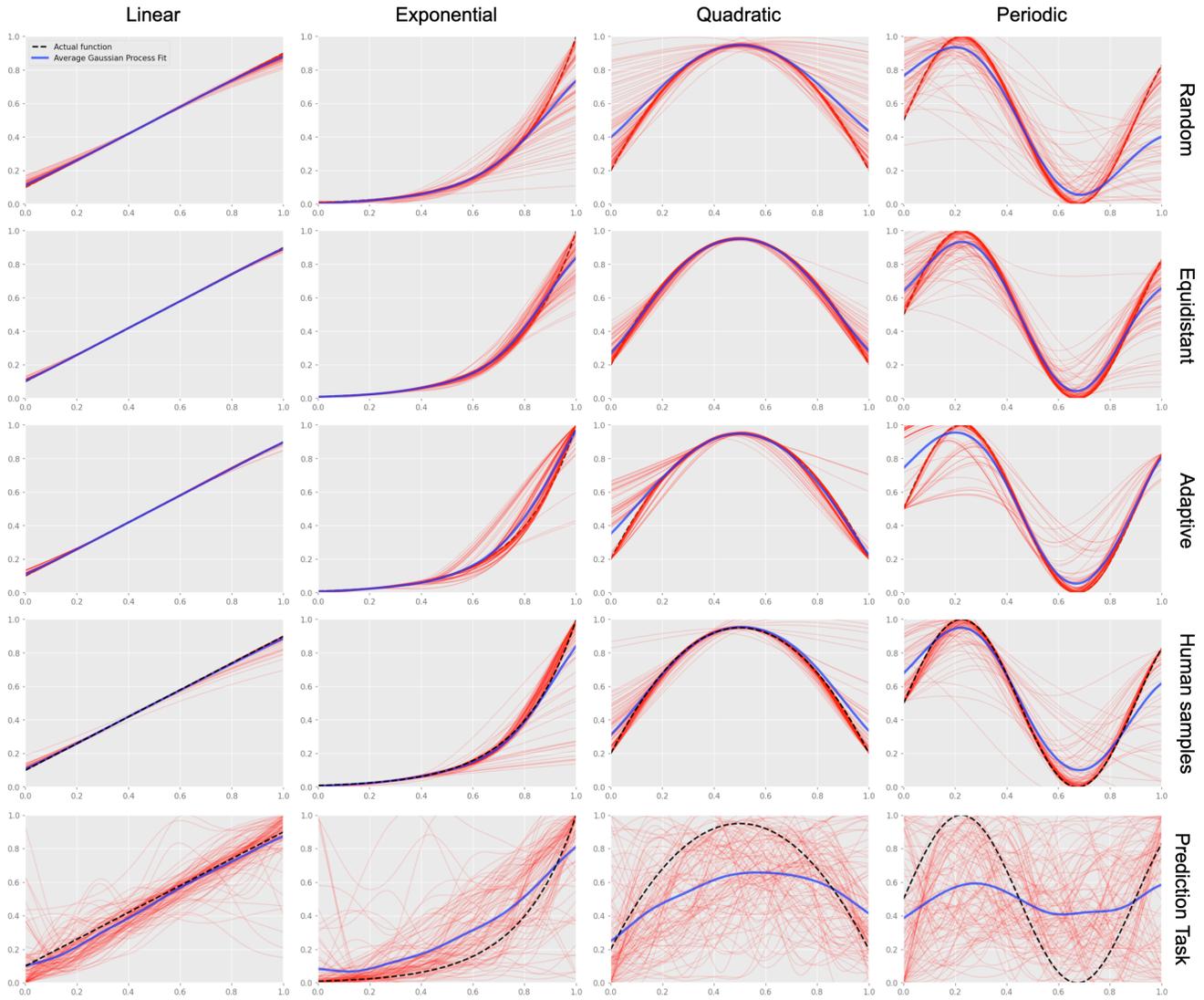


Figure 2: Gaussian process best fits for random sampling (row 1), equidistant sampling (row 2), adaptive sampling (row 3), human samples (row 4), and human predictions (row 5). Red lines represent individual GP fits for samples with a given sampling strategy or fits for individual learned functions on the prediction task. The blue line represents the mean of all GP fits.

Function	Sampling strategy	Pooled MSE	Average MSE	r
Linear	Random	$3.97 \cdot 10^{-5}$	0.0001	> 0.999
	Equidistant	$1.63 \cdot 10^{-7}$	$6.88 \cdot 10^{-6}$	> 0.999
	Adaptive	$4.87 \cdot 10^{-6}$	$2.23 \cdot 10^{-5}$	> 0.999
	Human	$2.18 \cdot 10^{-5}$	0.011	> 0.999
Exponential	Random	0.003	0.008	0.991
	Equidistant	0.001	0.002	0.995
	Adaptive	0.001	0.004	0.999
	Human	0.001	0.044	0.998
Quadratic	Random	0.044	0.015	0.995
	Equidistant	0.0004	0.002	0.997
	Adaptive	0.002	0.005	0.989
	Human	0.002	0.099	0.996
Periodic	Random	0.014	0.046	0.940
	Equidistant	0.002	0.009	0.995
	Adaptive	0.005	0.023	0.993
	Human	0.006	0.154	0.980

Table 1: Pooled MSE, average individual MSE, and correlation coefficients for sampling strategies and human samples.

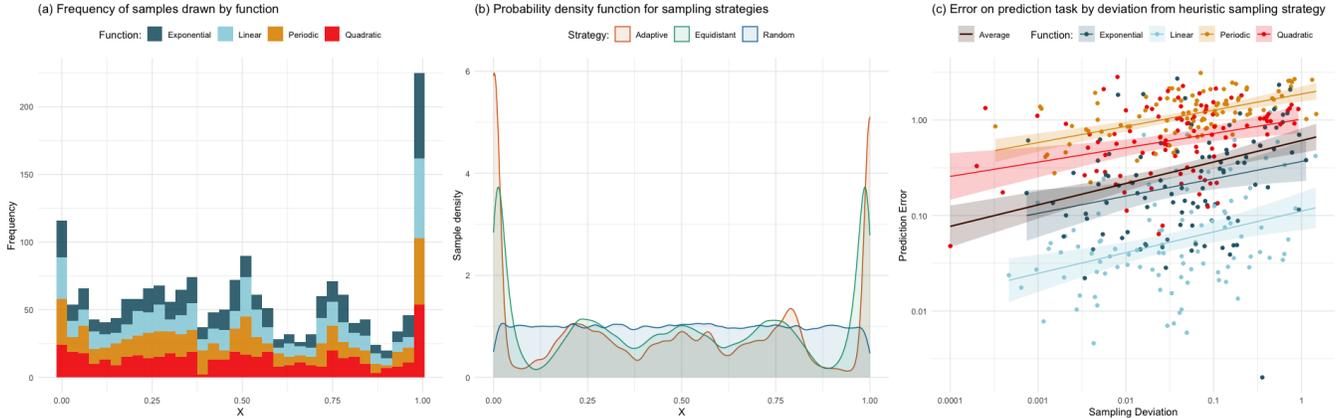


Figure 3: (a) Histogram showing frequency of human sampled points across the x axis by function; participants’ sampling behaviour was comparable across all four function families. (b) Estimated probability density functions for the random (blue), equidistant heuristic (green), and adaptive (orange) sampling strategies. (c) Participants were most accurate in predicting values for the linear function (light blue), followed by the exponential (dark blue), quadratic (red), and periodic (orange), in order, and were more accurate when their choices closely matched the equidistant strategy.

Human sampling (vs.)	Equidistant	Random	Adaptive
KL divergence (D_{KL})	0.052	0.077	0.103
Hellinger distance (H)	0.113	0.140	0.161

Table 2: Distance metrics for sampling strategies.

MSE for individual fits was relatively high. Nevertheless, in aggregate, the probability density of participants’ sampled choices showed the greatest similarity to the equidistant heuristic model predictions, compared to the random sampling and adaptive sampling strategies (Table 2), as measured by Kullback-Leibler divergence (a measure of difference between probability distributions) and Hellinger distance (a measure of similarity between distributions).

Prediction Task To measure participants’ prediction errors as well as their sampling behaviour, we calculated the sum of squared errors of participants’ predictions from the true function, as well as the deviation of their samples from the predictions of the equidistant heuristic strategy. As participants’ errors for both the prediction task and the samples were not normally distributed, sum of squared error (SSE) scores were scaled and log transformed before performing inference.

We first tested the hypothesis that prediction accuracy would be correlated with adherence to the equidistant sampling heuristic by running a Bayesian mixed-effects linear regression, with function and sampling deviation as predictors for prediction accuracy, and random intercepts for baseline accuracy, using the `brms` package in R, with uninformative priors of $\mathcal{N}(0, 3)$ placed on coefficients and a zero-truncated Cauchy(0, 2) prior placed on the standard deviation.

Confirming prior findings on the difficulty of learning non-linear relationships, participants were best at the linear, with increasing error, in order, on predictions for the exponential function ($\beta = 0.87$, 95% CI: [0.71, 1.04]), followed by the quadratic function ($\beta = 1.61$, 95% CI: [1.45, 1.77]) and the periodic function ($\beta = 1.96$, 95% CI: [1.81, 2.12]) relative to baseline performance on the linear function (Figure 3c).

Across all functions, there was strong evidence that greater deviation from sampling points consistent with the equidistant heuristic was associated with greater prediction errors, ($\beta = 0.19$, 95% CI: [0.12, 0.26]).

Further, while participants in the quadratic and periodic conditions had greater errors on average, this did not appear to be because participants were failing to infer high-level features of the functions such as their non-monotonicity. To investigate this phenomenon, we calculated the standard deviation of interpolated slopes between participants’ predicted points as a measure of the variability of the predicted function. Higher slope variability reflects larger deviation from a constant slope, such as a rising and falling function, while lower slope variability reflects a constant slope, closer to a flat or linear relationship. There was strong evidence that participants had more slope variability in predicting periodic ($\beta = 1.38$, 95% CI: [1.00, 1.75]) and quadratic ($\beta = 0.93$, 95% CI: [0.71, 1.15]) functions. Participants with lower variability made fewer errors in predicting the linear function ($\beta = 0.93$, 95% CI: [0.68, 1.16]), but no such relationship existed for the periodic ($\beta = -0.38$, 95% CI: [-0.73, 0.04]) or quadratic ($\beta = -0.10$, 95% CI: [-0.39, 0.19]) functions.

Participants making predictions for non-linear functions may have learned broader features of the functions, such as the fact that the function had both increases and decreases at various points, while not necessarily encoding exact values or appropriate parametrizations for the functions, confirming prior findings that participants readily extrapolate non-linear functions even in sparse environments (León-Villagrà, Preda, & Lucas, 2018). Failure to encode exact values while retaining some qualitative representation of the higher-order structure of the function might also explain why participants disproportionately estimated the minimum x value of the function to likewise have a y value of zero, as participants may have relied more heavily on inductive biases for an intercept value of zero (Kwantes & Neal, 2006).

General Discussion

Despite strong inductive biases towards positive linear relationships, people are able—at least in aggregate—to learn a variety of functional relationships, even when given very sparse evidence, and appear to be able to apply a relatively simple heuristic strategy, sampling the minimum and maximum x values and evenly spaced points in between, that requires little cognitive effort while providing a comparable outcome, and comparable information gain, to a more computationally expensive adaptive strategy. We also found that those who made choices more in line with this policy were more accurate in the prediction task.

Respondents showed vast variability in the informativeness of their samples and the accuracy of their predictions; however, aggregated results showed that a recognizable parametrization of the true function was learned, perhaps the product of a “wisdom of crowds” effect (Steyvers, Miller, Hemmer, & Lee, 2009) averaging out individual errors. Figure 2 shows the averages are compressed toward the center of the range relative to the true function, which is expected if some judgements are corrupted by additive zero-mean noise that is truncated at the limits of the values participants can select, or sometimes selected uniformly at random. Nevertheless, as Wilson et al. (2015) have pointed out, averaged responses can eliminate important statistical structure in human predictions of functional relationships, and our analysis revealed that even inaccurate individuals’ predictions suggested that they had learned, for example, when a function was non-monotonic.

While the proposed sampling strategy we introduced is relatively inflexible, this policy could reflect the use of rational metareasoning (Lieder & Griffiths, 2017), with participants deploying a heuristic with a favourable trade-off between its utility in giving relatively informative evidence for a variety of common functional relationships (including the most *a priori* plausible, positive linear), while requiring little cognitive effort to adapt to existing sampled points. This also coheres with previous findings in active function learning, where participants’ choices most closely fit a simpler linear regression policy rather than a generalized GP policy when learning in linear domains (Jones et al., 2018).

In this view, the equidistant heuristic may not be deterministically employed in all situations in which people must choose limited information to learn about a relationship, but could perhaps be used situationally in a rational way. Future research could place learners in a situation where use of a similar heuristic would lead to less informative evidence; if this heuristic is a rationally-adapted strategy trading off accuracy and efficiency, then we predict learners would adapt and deploy a different strategy for learning on such a task. Nevertheless, for limited domains, simple heuristics such as the one we have outlined may be a valuable element of the human cognitive toolkit to approximate optimal learning strategies.

Acknowledgements

We acknowledge the support of the Social Sciences and Humanities Research Council of Canada to RAG [CGS-M SSHRC] and DB [435-2018-0890].

References

- Bott, L., & Heit, E. (2004). Nonmonotonic Extrapolation in Function Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(1), 38–50.
- Brehmer, B. (1974). Hypotheses about relations between scaled variables in the learning of probabilistic inference tasks. *Organizational Behavior and Human Performance*, 11(1), 1–27.
- DeLosh, E. L., McDaniel, M. A., & Busemeyer, J. R. (1997). Extrapolation: The Sine Qua Non for Abstraction in Function Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(4), 968–986.
- Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, 349(6245), 273–278.
- Gigerenzer, G. (2008). *Rationality for mortals: How people cope with uncertainty*. Oxford University Press.
- Griffiths, T. L., Lucas, C. G., Williams, J. J., & Kalish, M. L. (2008). Modeling human function learning with Gaussian processes. In *Advances in Neural Information Processing Systems* (Vol. 21, p. 553–560).
- Jones, A., Schulz, E., Meder, B., & Ruggeri, A. (2018). Active function learning. In *Proceedings of the 40th Annual Meeting of the Cognitive Science Society* (pp. 578–583).
- Kalish, M. L. (2013). Learning and extrapolating a periodic function. *Memory & Cognition*, 41, 886–896.
- Kalish, M. L., Lewandowsky, S., & Kruschke, J. K. (2004). Population of Linear Experts: Knowledge Partitioning and Function Learning. *Psychological Review*, 111(4), 1072–1099.
- Kwantes, P. J., & Neal, A. (2006). Why people underestimate y when extrapolating in linear functions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(5), 1019–1030.
- León-Villagrà, P., Preda, I., & Lucas, C. G. (2018). Data Availability and Function Extrapolation. In *Proceedings of the 40th Annual Meeting of the Cognitive Science Society* (pp. 2017–2022).
- Lieder, F., & Griffiths, T. L. (2017). Strategy selection as rational metareasoning. *Psychological Review*, 124(6), 762–794.
- Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 43.
- Lucas, C. G., Griffiths, T. L., Williams, J. J., & Kalish, M. L. (2015). A rational model of function learning. *Psychonomic Bulletin & Review*, 22(5), 1193–1215.

- Lucas, C. G., Sterling, D., & Kemp, C. (2012). Superspace extrapolation reveals inductive biases in function learning. In *Proceedings of the 34th Annual Meeting of the Cognitive Science Society* (pp. 713–718).
- Markant, D. B., & Gureckis, T. M. (2014a). Is it better to select or to receive? Learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General*, *143*(1), 94–122.
- Markant, D. B., & Gureckis, T. M. (2014b). A preference for the unpredictable over the informative during self-directed learning. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (pp. 958–963).
- Markant, D. B., Ruggeri, A., Gureckis, T. M., & Xu, F. (2016). Enhanced Memory as a Common Effect of Active Learning. *Mind, Brain, and Education*, *10*(3), 142–152.
- Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. Cambridge, Mass: MIT Press.
- Steyvers, M., Miller, B., Hemmer, P., & Lee, M. D. (2009). The Wisdom of Crowds in the Recollection of Order Information. In *Advances in Neural Information Processing Systems* (pp. 1785–1793).
- Wilson, A. G., Dann, C., Lucas, C. G., & Xing, E. P. (2015). The Human Kernel. In *Advances in Neural Information Processing Systems* (Vol. 28).

Simulating Human Periodic Tapping and Implications for Cognitive Models

Pierre G. Gianferrara (pgianfer@andrew.cmu.edu)

Shawn Betts (sabetts@andrew.cmu.edu)

Dan Bothell (db30@andrew.cmu.edu)

John R. Anderson (ja0s@andrew.cmu.edu)

Department of Psychology, Carnegie Mellon University, 5000 Forbes Ave
Pittsburgh, PA 15213 USA

Abstract

This project's purpose was to simulate human periodic motor behavior in a simple self-paced tapping task that involved period error correction and feedback processing. When humans try to tap at a certain period, their inter-tap times are normally distributed with a standard deviation that is proportional to the period. When they try to change the period of their tapping, they do so in a single tap instead of a progressive correction taking place over multiple taps. We calibrated ACT-R's new periodic tapping motor extension based on human experimental results and showed that ACT-R can simulate human motor behavior. Future research can leverage these findings and ACT-R's periodic tapping motor extension to simulate fast-paced skilled motor behavior in complex perceptual-motor environments.

Keywords: ACT-R; modeling; motor; period; tapping; error; correction; skill; automaticity

Introduction

Cognitive scientists have recently shown a growing interest in video games and have started to uncover evidence supporting their potential cognitive benefits (Bediou *et al.*, 2018). From a psychological standpoint, video games can be useful as a way to investigate complex skill learning processes involving the integration of perceptual, cognitive, and motor information (Anderson *et al.*, 2019). In terms of skill acquisition, it is generally acknowledged that skill learning involves a shift from high-level processing of task-related declarative information to the progressive automatization of motor skills (Ackerman, 1988; Anderson, 1982). Acquiring skill in a motor task often involves progressively lower levels of motor variability, potentially due to improved feedback control (Shmuelof, Krakauer & Mazzoni, 2012). In a motor timing video game specifically, skilled behavior was found to be predicted by decreased motor timing variability and increased rhythmicity in motor behavior (Gianferrara, Betts & Anderson, 2020).

Another characteristic of video games is that they often require fast-paced actions which tend to be shorter than a second, and are characterized by more rhythmic motor actions than speeds slower than a second (Gianferrara, Betts & Anderson, 2020). In the brain, actions in the sub-second range are more likely to recruit sub-cortical structures implicated in the motor system such as the basal ganglia and the cerebellum whereas actions in the supra-second range are more likely to recruit cortical structures (Wiener, Turkeltaub

& Coslett, 2010). From a modeling perspective, cognitive architectures ought to include suitable motor mechanisms that may account for skilled motor behavior at fast speeds.

We augmented the motor module in the adaptive control of thought rational (ACT-R) architecture with a motor extension to enable ACT-R to engage in rhythmic motor behavior. The starting point for such motor extension is to model human behavior in a self-paced periodic tapping task. Most existing work on periodic tapping has commonly been presented in the context of sensorimotor synchronization studies. Such studies often investigate the process whereby participants first learn to align their taps to a periodic auditory stimulus (synchronization phase) and then continue to tap at that same period (continuation phase; Repp, 2005; Wing, 1980). Though synchronization-continuation paradigms are useful to model periodic tapping and error correction, they often do not provide an account of error correction based on external non-periodic feedback, in which case the adjustment of one's tapping period does not rely on sensorimotor synchronization with a periodic sensory cue.

The goal of this project was to calibrate ACT-R's periodic tapping motor extension based on the human experimental results from a novel game called *ChemLab*, which involves self-paced tapping. In this task, participants learn to adjust their tap frequency based on external feedback that they need to attend to. We first present our results and choice of ACT-R parameterization. We then conclude with some remarks and important implications for future cognitive models.

ChemLab Periodic Tapping Video Game

The goal of *ChemLab* is to fill as many rows of 8 beakers as possible by periodically pressing the space bar. Each beaker's total capacity was set to 100 pixels and each tap within the right tapping interval resulted in an incremental increase of 1/8 of the beaker capacity as well as a brief mid-pitched sine tone (625 Hz). Thus, 8 taps were required to completely fill a beaker. When participants did not press the space bar within the right tapping interval, one out of two possible outcomes could happen: 1) When taps were too fast, a panel with the message "too fast" immediately turned red and a brief high-pitched sine tone (890 Hz) was triggered. Each too-fast tap was penalized by a loss of 5% of the max beaker capacity. The "too fast" light only turned off when taps were at least as slow as the lower (fast) bound of the prescribed tapping interval. 2) When taps were too slow, a panel with the

message “too slow” immediately turned blue and a brief low-pitched sine tone (460 Hz) was triggered. Unlike “too fast” feedback, the beaker level progressively decreased at a constant rate of 0.125 % of the beaker’s max capacity every 1/60 s. The “too slow” light turned off and the beaker level stopped shrinking when taps became at least as fast as the upper (slow) bound of the tapping interval. An illustration of the *ChemLab* interface is depicted in Figure 1. One can play *ChemLab* by clicking on the following link: <http://andersonlab.net/demos/chemlab-v1/>

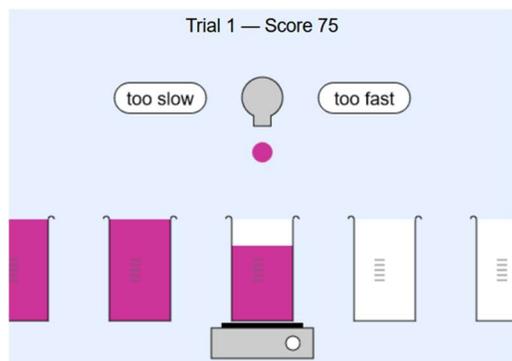


Figure 1: *ChemLab* video game interface.

Experimental Methods

Experimental Design

In this experiment, players completed 9 *ChemLab* sessions of 5 minutes each. In each session, players filled rows of beakers that were selected in a pseudorandom order. Rows of beakers, named trials, belonged to one out of 6 possible conditions that are introduced in Table 1. Each condition included two speeds with two consecutive tapping intervals. The four possible tapping intervals were [200-300 ms], [300-500 ms], [500-800 ms], and [800-1200 ms]. Tapping intervals were non-overlapping and had a range whose width increased at slower speeds. The serial order of trials within sessions was indicated at the top of the screen, along with the score.

Table 1: Description of the 6 *ChemLab* conditions

Condition	Speed 1	Points /beaker	Speed 2	Points /beaker
C1	200-300 ms	10	300-500 ms	20
C2	300-500 ms	20	200-300 ms	10
C3	300-500 ms	20	500-800 ms	30
C4	500-800 ms	30	300-500 ms	20
C5	500-800 ms	30	800-1200 ms	40
C6	800-1200 ms	40	500-800 ms	30

Each trial included 8 beakers that were divided into two phases: the pre-switch phase, and the post-switch phase. Beakers from the pre-switch phase and post-switch phase respectively shared the same tapping interval (“Speed 1” and “Speed 2” in Table 1). When the first post-switch beaker came up the subject would get feedback that they were too fast or too slow and they would have to adjust the period of

their tapping accordingly. The transition between the pre-switch and post-switch phases was scheduled pseudo-randomly and could either happen after the completion of 3, 4, or 5 beakers. For each condition, points were earned proportionally to the width and speed of the tapping interval such that slower intervals led to a higher reward than faster intervals. The total reward per trial was computed prior to the start of that trial by computing the sum of points per beaker within each phase (see Table 1) and then adding up the sums from each phase respectively. The total number of points for a trial was then divided by 8 (since there are 8 beakers in each trial), and 1/8 of the total was earned after the completion of each beaker within trials regardless of phase.

Measures

In this experiment, periodic tapping skills were measured in terms of performance within sessions, and in terms of tap variability. One critical *ChemLab* measure related to skill and period error correction was tap feedback.

Performance Score The main way of assessing subjects’ *ChemLab* performance was to compute each participant’s game score within 5-minute sessions using the scoring system described in Table 1.

Motor Behavior & Tapping Variability We assessed motor behavior by measuring the time between consecutive keypresses’ onsets within beakers. This time interval is often referred to as inter-press interval (IPI) in the literature (Diedrichsen & Kornysheva, 2015). Using this measure, it is possible to compute the coefficient of variation (CV), which is the standard deviation divided by the mean of the IPIs. Following previous work on video games, we assessed a logarithmic transformation of CV which measures motor variability and has been shown to be linearly related to performance in a motor timing video game (Gianferrara, Betts & Anderson, 2020).

Finally, we estimated participants’ tap regularity levels across speeds by computing the autocorrelation of vectorized tap holds and releases following the methodology from previous work (Gianferrara, Betts & Anderson, 2020). In this computation, keypress holds and releases had a temporal resolution of 1/60 s and we measured the autocorrelation of 100 lags of 1/60 s. We then extracted the correlation coefficient corresponding to the first non-zero positive peak of the autocorrelation function and used this as our tapping regularity estimate.

Tap Feedback In *ChemLab*, skill learning and period error correction mostly happened via the online processing of feedback that followed each individual tap. As mentioned earlier, taps could be categorized as “OK”, “too fast”, or “too slow”. Recording the feedback type that resulted from individual taps is useful because that helps the researcher understand how period error correction happens as a result of exposure to feedback.

Human Participants

A total of thirty-two human participants completed the *ChemLab* experiment. Out of these, one participant was excluded because of poor performance (less than 100 points per session in the last 7 *ChemLab* sessions). A second participant was excluded because their average performance was close to 3 SDs below the mean ($z = -2.9$; $M = 1034$ points; $SD = 151$ points), and their average tap variability level was 4 SDs above the mean in terms of the log CV of the IPIs ($z = 4.0$; $M = -1.35$; $SD = 0.55$).

The 30 remaining participants were aged 22 to 50 years-old ($M = 32.8$, $SD = 7.1$). Twenty were male and 10 were female. All participants were recruited on Amazon Mechanical Turk (mTurk). Subjects earned a base pay of \$4 for completing the experiment, in addition to a bonus which was proportional to their performance (in points) as specified in Table 1. On average, participants earned a bonus of \$5.50.

Procedure

To qualify for the experiment, participants needed to correctly answer at least 3 out of 4 multiple choice questions on an English comprehension quiz. The experiment then proceeded as follows: Participants first filled out short background questionnaires. They then read a quick description of *ChemLab* which included instructions on how to proceed. Once ready, participants completed 9 *ChemLab* sessions lasting 5 minutes each. Finally, they filled out some additional questionnaires where they provided feedback and wrote about strategies they found helpful.

Human Results

Behavioral Results

We first present some general results pertaining to human performance and human behavior in the *ChemLab* experiment. Figure 2 provides an illustration of human performance. Figure 2a shows that humans' average game score progressively increased in the 2 first sessions and eventually reached a learning plateau at game 3 onwards

when the average game score was consistently greater than 1000 points, which corresponded to more than 90 % of subjects' max performance score in *ChemLab*. Since this study is mostly concerned with skilled motor behavior and the modeling of periodic tapping, we elected to focus on the last 7 sessions at which most of the task-specific skills have been acquired. These included a total of 1014 trials across all subjects and speeds.

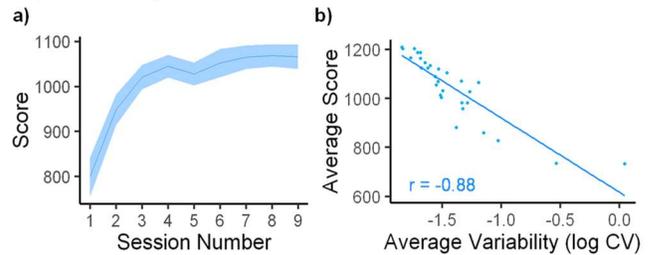


Figure 2: a) Mean game score (performance) over the 9 *ChemLab* sessions. The shaded area indicates the standard error of the means b) Correlation between subjects' average game score across sessions and tap variability as the logarithmic coefficient of variation.

Figure 2b compares individual subjects' performance and motor behavior during the learning plateau (last 7 sessions). Subjects' tap variability is measured in terms of the logarithmic CV and plotted against subjects' average game score across sessions. The main result is that game score is negatively correlated with tap variability ($r = -0.88$) meaning that the best performing subjects were also the ones with the lowest levels of tap variability. In terms of motor behavior, tap regularity levels defined with the autocorrelation ranged between $r = 0.43$ and $r = 0.49$ across the four different speeds.

With respect to subjects' adaptation to the new tapping period after the switch point, we found that human participants successfully transitioned from speed to speed as can be seen on Figure 3a. Table 2 displays the taps' categories

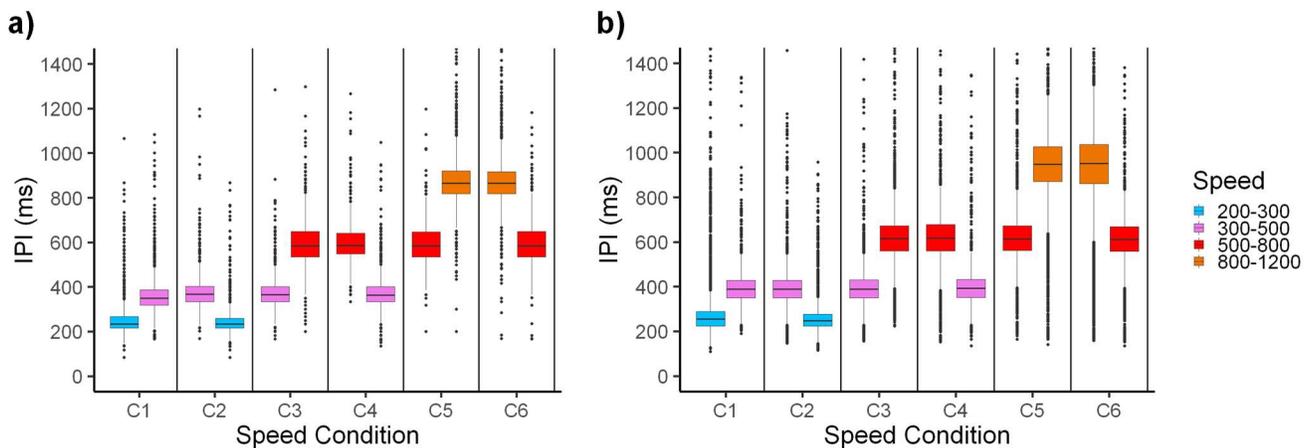


Figure 3: Inter-press interval (IPI) boxplot across the 6 conditions in the pre-switch and post-switch phases. Each speed corresponds to a different color. Human IPIs are depicted to the left (a) and ACT-R model IPIs are depicted to the right (b).

Table 2: Human and ACT-R model tap category proportions across speeds and feedback types.

Feedback	200-300 ms		300-500 ms		500-800 ms		800-1200 ms	
	Humans	ACT-R	Humans	ACT-R	Humans	ACT-R	Humans	ACT-R
OK	79.62 %	71.05 %	87.56 %	87.90 %	91.17 %	86.38 %	82.35 %	81.55 %
Fast	10.04 %	11.43 %	8.61 %	7.26 %	7.42 %	9.96 %	16.44 %	14.67 %
Slow	10.36 %	17.52 %	3.83 %	4.85 %	1.41 %	3.67 %	1.21 %	3.79 %

in the assigned tapping interval in the last 2 beakers (stable behavior) of either phase, sorted according to speed and agent (humans vs. ACT-R model). Overall, human subjects executed taps that were in the correct tapping interval ~80% of the time or more.

Feedback Processing

We then investigated participants’ response to feedback at the time of the period switch. To reiterate, the tapping interval switched to a consecutive speed bracket after the completion of 3, 4, or 5 beakers (this number was generated pseudorandomly), and players then learned to execute taps at the new speed for the remainder of the trial beakers until they completed the final (8th) beaker.

To explore the process of period error correction, we first computed the proportion of each tap category (“OK”, “too fast” and “too slow”) for the 8 first IPIs directly following the speed switch. Tap category proportions were computed across all trials from all subjects (see Figure 4). Figure 4’s top row illustrates cases in which the speed slowed down, thus resulting in “too fast” feedback, and Figure 4’s bottom row illustrates cases in which the speed sped up, thus resulting in “too slow” feedback. Overall, one can see that the majority of participants tended to persevere their taps at the old speed for 1 to 3 taps before adjusting their tap period, though most players needed at least 2 taps before initiating the correction.

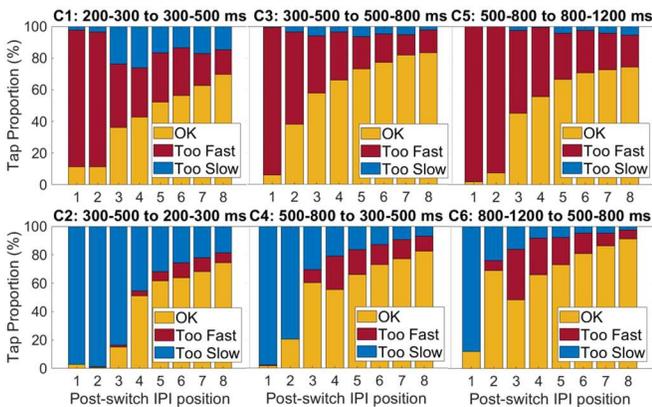


Figure 4: Evolution of tap category proportion as a function of post-switch IPI position following a period switch across the 6 ChemLab conditions.

Although Figure 4 suggests that participants may progressively correct their taps’ period, we found that this result was due to variation in when the period was corrected and was not indicative of continuous error correction with progressively smaller correcting steps. Instead, we found that period correction happened as a first-order process. Figure 5 shows the difference in IPI as a percentage of the previous IPI at the time of error correction (Pos 0) and at the tap position directly before (Pos -1) and directly after (Pos +1), regardless of the tap serial order in the post-switch beaker.

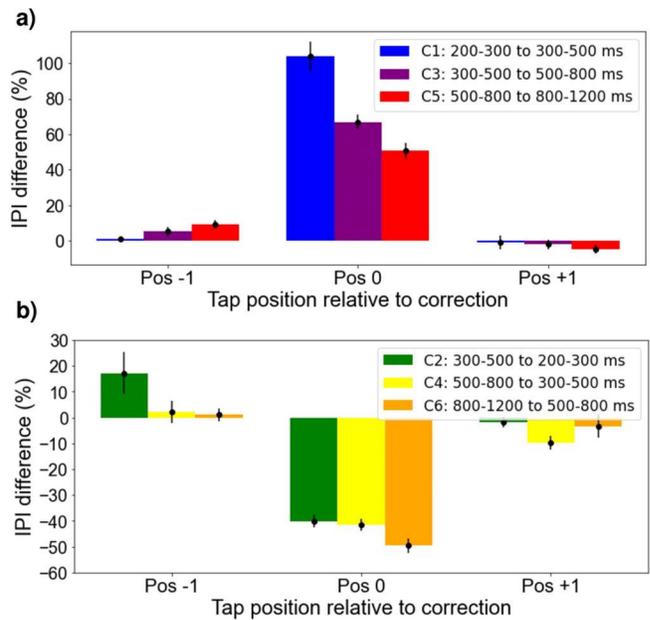


Figure 5: 1st order feedback processing in fast feedback conditions (a) and slow feedback conditions (b). Error bars correspond to the standard error of the means (SEM).

As can be seen on Figure 5, the IPI difference at Pos 0 was greater than at Pos -1 and Pos +1 in conditions in which the tapping interval got slower, but no significant difference relative to no difference (0 %) was found at Pos -1 and Pos +1 (standard deviations at these positions all included 0%). Conversely, the IPI difference at Pos 0 was smaller than at Pos -1 and Pos +1 in conditions in which the tapping interval got faster, but no significant difference relative to no difference (0 %) was found at Pos -1 and Pos +1.

ACT-R Model of Periodic Tapping

The next step was to integrate into ACT-R a model of tapping and period correction that was consistent with these results. To reiterate, a goal of the project was to use human experimental results in a simple tapping paradigm to calibrate the parameterization of motor parameters in ACT-R.

Periodic Tapping in ACT-R

A motor extension was added to ACT-R, which includes a few basic actions. First, taps can be initiated by making a request to the manual module with information pertaining to the hand, finger, and specific tapping period. Once periodic tapping has been initiated, the manual buffer corresponding to the tapping hand (“manual-right” or “manual-left”) continues the tapping action repeatedly. Note that periodic tapping does not require ACT-R to issue specific motor commands for each individual tap, which would not be feasible at the fastest tapping rates. This process is assumed to carry on automatically due to basal ganglia neural activity (Wu, Hallett & Chan, 2015). To stop the period, another request to the manual module can be made in a subsequent production, and ACT-R will then stop periodic tapping once ready. During periodic tapping, upcoming taps are automatically scheduled relative to the previous ones at the time of key release, unless a stop request has been initiated.

The periodic tapping motor extension also includes an additional “tap” buffer which can be accessed to determine the current tap period (in seconds), and a count of the number of taps made at that period. ACT-R can request that the motor module adjust the period at which it is tapping. The “periodic-tap” motor extension code has been created in the Lisp programming language and will be made available to users in an upcoming ACT-R release. We next review parameterization of the periodic tapping motor extension.

ACT-R Periodic Tapping Parameterization

In this paper, we are using the results from the *ChemLab* experiment to calibrate the ACT-R model of periodic tapping. This section is specifically focusing on the choice of noise parameter that governs the variability of taps across speeds. To address the variability in timing between individual taps, we investigated consecutive IPI % differences in an iterative fashion in the last 2 beakers of the pre-switch and post-switch phases. For each beaker, we recorded each tap’s IPI % difference relative to the IPI from the previous tap and sorted the IPI % tap differences according to speed. We thus obtained 4 IPI % tap difference frequency distributions which are displayed on Figure 6. As can be seen, the motor noise distribution is centered around 0 % and is normally distributed. One crucial finding was that variability in taps’ period across speeds can best be specified in terms of % IPI difference instead of a fixed IPI difference duration, which fits with past sensorimotor synchronization findings (Repp, 2005; Wing, 1980) and may partially be due to fingers’ biomechanical constraints (Loehr & Palmer, 2009).

The noise on the tap timing was generated using the same

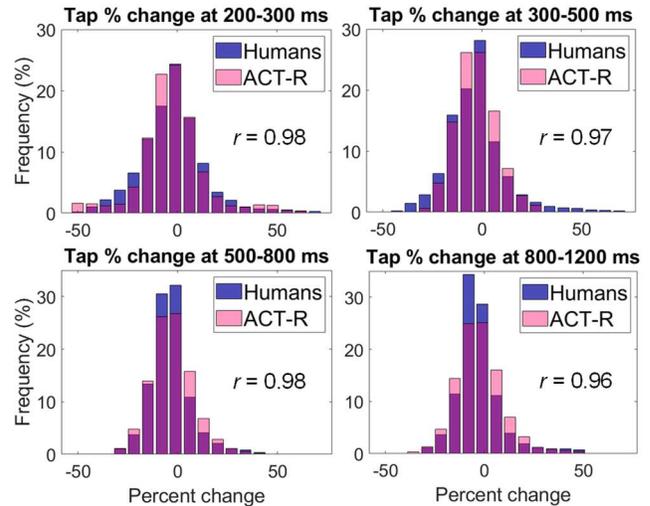


Figure 6: Overlap between human and ACT-R model percent change in tap IPIs. Bins have a width of 7%.

logistic distribution that is used for generating the noise in the ACT-R procedural and declarative systems¹. The s value of the distribution that best fit the human data was found to be 0.04 (see Figure 6). This corresponds to a standard deviation approximating 7% of the current tap period. The correlation between humans and ACT-R ranged between $r = 0.96$ and $r = 0.98$ across the four speeds.

ACT-R Model of *ChemLab*

Modeling performance in the *ChemLab* experiment not only required us to refine the parameterization of the periodic tapping motor extension in ACT-R, but it also necessitated identifying the key task-specific components of the experimental paradigm that were critical for learning. In this experiment, feedback was the most important experimental feature. Specifically, we needed to create a model that could simulate humans’ response to feedback, and error correction.

Responding to Feedback Humans’ response to feedback in *ChemLab* was not uniform within subjects as suggested the results displayed in Figure 4. While most period corrections happened shortly after feedback detection and processing, some other corrections happened after a few more taps. In ACT-R, we decomposed this process into three steps represented as separate ACT-R productions: 1) feedback detection, 2) feedback processing, 3) response to feedback.

The first step was to simulate perceptual feedback detection. Our data suggest that there may be perceptual delay and feedback processing differences, which have been hypothesized to be a function of skill level and past exposure to video games (Bediou *et al.*, 2018; Bejjanki *et al.*, 2014). We used ACT-R’s visual-search buffer to model humans’ visual detection of color changes that indicated an error, although auditory “too fast” and “too slow” feedbacks

¹Note that ACT-R uses logistic instead of an actual normal for computational efficiency (Anderson & Lebiere, 1998)

may have played a facilitatory role in feedback detection (Repp & Penel, 2002, 2004). Upon detecting a color change, ACT-R put the interpretation (“too fast” or “too slow”) into the imaginal module. To fit human performance, we set the mean time for this action to 50 ms, and the imaginal module adds noise to that from a uniform distribution of +/- 16ms (1/3 of the action duration).

Finally, the last step was to respond to feedback, which was implemented as a first-order process in accordance with the results from Figure 5. Based on our experimental investigation of feedback response, we found that the participants’ response to feedback was a probabilistic event which could be simulated with competing productions (“correct” vs. “do-not-correct”) and fixed utilities in ACT-R. Utilities were tuned using probabilistic estimates of error correction for fast and slow feedback respectively.

Error Correction Based on the subject data the model responded differently to “too-fast” and “too-slow” feedback. When exposed to “too-fast” feedback, the ACT-R model requested a period error correction from the manual module while maintaining the original tapping rate. When exposed to “too-slow” feedback, however, the model briefly stopped tapping to process the progressively decreasing beaker level caused by the slow taps (see *ChemLab* video game description), and it then made a request for a new tapping period.

The model attempted to correct errors and change the tapping period by adding or subtracting a correction from the period. This correction was expressed as a percentage of the original tapping period and was selected from a gamma distribution² generated with a shape parameter k and a scale parameter θ . To fit the subject data, we selected different gamma distributions for each speed.

One striking result was that the gamma distribution underlying “too fast” period corrections was closer to an exponential distribution than the gamma distribution underlying “too slow” corrections. Indeed, k estimates approximated 1 for “too fast” period corrections, regardless of speed (+/- 0.2). For “too slow” corrections, however, k estimates exceeded 2 across all speeds and increased as the tapping rate slowed down. These findings suggest that the shape of the period correction distribution may depend on task-specific feedback features, speed, and potentially feedback saliency.

ACT-R Model Results

We ran two hundred ACT-R model simulations of trials in each of the 6 conditions (1,200 model runs in total). All models were initialized with the same parameters. We then tested whether we could replicate human results from Figure 3a and Table 2. Figure 3b illustrates the model transition from

speed to speed in each of the 6 *ChemLab* conditions. To reiterate, IPIs were measured in the last 2 beakers of the pre-switch and post-switch phase, which reflect stable periodic tapping behavior.

We then computed the proportion of tap categories across speeds in either phase and reported these proportions in Table 2 (see ACT-R results). As can be seen, similar tap category proportions were found in ACT-R. A Chi-squared contingency test summarizing within-speeds tap proportion comparisons between ACT-R and humans ($df = 4*2 = 8$) revealed that both proportions were of a similar magnitude ($\chi^2(df = 8, N = 24) = 5.59, p = 0.69$).

Conclusions

The goal of this project was to simulate human motor behavior in a simple self-paced periodic tapping task in which period error correction was driven by visual and auditory feedback. By calibrating our novel periodic tapping motor extension in ACT-R, we showed that it is possible to replicate the general patterns of human behavior and periodic tapping. Some implications are worth noting.

In terms of motor behavior, we found two general mechanisms pertaining to human skill learning. First, we saw that the noise around periodic taps was proportional to the taps’ mean and could be simulated as a percentage of the period instead of a fixed time duration, which replicates results from the sensorimotor synchronization literature (Repp, 2005). Second, in the context of the *ChemLab* experiment, we also saw that feedback processing happened as a first-order process akin to reaction time processes.

Because the core periodic tapping code was built as a motor extension in the ACT-R architecture, it is possible for other modelers to use our code as a template of periodic tapping and build upon our work to model human behavior in fast-paced video games involving repetitive motor actions. We look forward to expanding our understanding of skilled motor behavior in complex perceptual-motor environments.

Acknowledgements

This research was supported by the Office of Naval Research Grant N00014-15-1-2151 and AFOSR/AFRL award FA9550-18-1-0251.

References

- Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of experimental psychology: General*, *117*, 288-318.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological review*, *89*, 369-406.
- Anderson, J. R., Betts, S., Bothell, D., Hope, R., & Lebiere, C. (2019). Learning rapid and precise skills. *Psychological review*, *126*, 727-760.
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.
- Bediou, B., Adams, D. M., Mayer, R. E., Tipton, E., Green,

² We utilized the “random-gamma-ml” function from the “cl-randist” Lisp package:
<http://github.com/lvaruzza/cl-randist/tree/master>

- C. S., & Bavelier, D. (2018). Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. *Psychological bulletin*, *144*, 77-110.
- Bejjanki, V. R., Zhang, R., Li, R., Pouget, A., Green, C. S., Lu, Z. L., & Bavelier, D. (2014). Action video game play facilitates the development of better perceptual templates. *Proceedings of the National Academy of Sciences*, *111*, 16961-16966.
- Diedrichsen, J., & Kornysheva, K. (2015). Motor skill learning between selection and execution. *Trends in cognitive sciences*, *19*, 227-233.
- Gianferrara, P. G., Betts, S., & Anderson, J. R. (2020). Time-related Effects of Speed on Motor Skill Acquisition. In Stewart, T. C. (Ed.). *Proceedings of the 18th International Conference on Cognitive Modelling* (pp. 64-70). University Park, PA: Applied Cognitive Science Lab, Penn State.
- Loehr, J. D., & Palmer, C. (2009). Sequential and biomechanical factors constrain timing and motion in tapping. *Journal of Motor Behavior*, *41*, 128-136.
- Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic bulletin & review*, *12*, 969-992.
- Repp, B. H., & Penel, A. (2002). Auditory dominance in temporal processing: New evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1085-1099.
- Repp, B. H., & Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological research*, *68*, 252-270.
- Shmuelof, L., Krakauer, J. W., & Mazzoni, P. (2012). How is a motor skill learned? Change and invariance at the levels of task success and trajectory control. *Journal of neurophysiology*, *108*, 578-594.
- Wiener, M., Turkeltaub, P., & Coslett, H. B. (2010). The image of time: a voxel-wise meta-analysis. *Neuroimage*, *49*, 1728-1740.
- Wing, A.M. (1980). The Long and Short of Timing in Response Sequences. In G. E. Stelmach, & J. Requin (Eds.), *Tutorials in motor behavior*. Amsterdam: North-Holland.
- Wu, T., Hallett, M., & Chan, P. (2015). Motor automaticity in Parkinson's disease. *Neurobiology of disease*, *82*, 226-234.

When Do You Buy? Predicting an Individual's Decision in Optimal Stopping Problems

Manuel Guth¹ and Marco Ragni^{1,2}

¹ Cognitive Computation Lab, Technical Faculty, Albert-Ludwigs-University Freiburg

² Danish Institute of Advanced Studies, Denmark
guthm@tf.uni-freiburg.de, ragni@sdu.dk

Abstract

Prices, e.g., for flight tickets can change almost daily. To minimize the costs, we have to decide when to take an action, i.e., when to buy. Suchs decision tasks are called *optimally stopping problems*. This paper reconsiders the strongest cognitive models that are able to predict the average decision maker, adapts them and investigate their predictive accuracy *on the individual level*, i.e., how good are models in predicting *when* a participant decides for an action. To perform this analyses, several steps are necessary: (i) Identify data sets that provide raw data for an individual, (ii) develop an individual testing framework to assess the models, (iii) implement and adapt existing models for the individual, and (iv) consider baseline models to assess the goodness-of-fit of the models for the individual. The best and second-best models achieved an overall prediction accuracy of 84.9% and 84.1% respectively. Five of the ten examined models managed to beat a strong baseline, proving that they did in fact managed to model the individual decision process.

Introduction

The *Optimal Stopping Problem* is implicitly present in many aspects of everyday life. When searching for the partner to spend life with, buying airplane tickets for the next holiday trip, or deciding when and with whom to fill an open job position. All tasks demand to decide whether to keep the current option (partner/ticket price/applicant) or to keep on searching for a better option. When declining an option, it is not known if a better option will eventually present itself. It is also often not possible to go back to one of the previous options, as possible partners might not be available anymore, ticket prices change from day to day, and a once rejected applicant might have started to work for another company.

Formally, the *Optimal Stopping Problem* is the task of finding in a sequence of timepoints $1 \leq i \leq n$ with a possibly unknown n for associated values (or options) $(y_i)_i$ the time i when to perform an action to maximize the desired potential outcome, i.e., increase profits or minimize costs. The options can change randomly and the quality of the future options cannot be estimated. In some cases, the number of options can be limited, e.g., one has to buy a plane ticket eventually if the vacation is planned on a certain date. If the last option is reached, it has to be chosen. The difficulty lies in evaluating if either the currently presented option is worth keeping, given the knowledge about the previously seen options and some domain knowledge (e.g., average plane ticket prices), or if a better option will occur in the future.

Most current models are assessed by a “fitting”-analyses of the response distribution and not on assessing the predictive accuracy of the next decision before an individual makes it (Guan & Lee, 2018; Lee & Wagenmakers, 2014; Seale & Rapoport, 2000; von Helversen & Mata, 2012; Zwick, Rapoport, Lo, & Muthukrishnan, 2003). The advantage of the latter method is that it allows even to falsify models, i.e., if they do not predict the right decisions and it can identify the underlying decision processes. Hence, we propose to assess models in the predictive setting using *CCOBRA* (Brand, Riesterer, & Ragni, 2020), a cognitive reasoning framework that allows predicting and adapting to an individual reasoner while evaluating a models performance. The data used stems from Baumann, Singmann, Gershman, and von Helversen (2020), which presents participants in an experiment with the task to buy an item for the cheapest possible price. The data includes artificially generated prices for fictional items with a normal, left- and right-skewed price distribution, as well as real prices for real items that can be used to evaluate the real-life performance of the predictive models. As part of this work, four models for human reasoning in *Optimal Stopping Problems* were implemented and adapted to individual human reasoning. The original models are presented in Baumann et al. (2020) and are based on a threshold heuristic. For comparison of the performance of the adapted models, two baseline models were implemented, one random one and one that will follow the optimal strategy (Gilbert & Mosteller, 2006) to find the best option. The models were fine-tuned on one of the available data sets (= training data) and then evaluated on the other data sets. The results were then compared to each other and the similarities and differences between them were examined. The findings were also compared to the findings of previous studies, which showed similar results. An insight into the prediction performance of each model on each individual reasoner is also given as the final step to analyze the models.

Related Work

Lee (2006) proposes a hierarchical *Bayesian Model* to predict human decisions. Participants in the experiment were confronted with the problem of choosing the maximum out of a sequence of few numbers. The participants knew the generation principle of the numbers. The results supported a threshold-based model to explain the decisions of the partici-

pants, choosing the first maximal number that exceeds a certain threshold for each index. Since the overall performance did not increase during the experiment, participants did not learn from the previous problems.

An optimal strategy (that has the highest expected value) to select the best (in this case highest) option out of a finite list of options is described in Gilbert and Mosteller (2006). The idea is to start with a high threshold and decrease it over time based on the distribution the options were sampled from.

Guan, Lee, and Vandekerckhove (2015) considers the *Optimal Stopping Problem* for a sequence length of both, 5 and 10. In both cases, the reasoner had to find the highest option of the sequence. The authors propose a threshold model for human reasoning that takes into account how far the individual reasoner deviates from the optimal threshold at the current step. The deviation is controlled by two parameters β and γ . β determines how far above or below the threshold is from the optimal one and γ controls how fast the bias increases/decreases as the sequence progresses. Their results show that reasoners that set their initial threshold higher than the optimal, tend to decrease it faster than optimal, and reasoners that set their initial threshold too low, decrease it slower than optimal. Furthermore, β and γ remains stable for participants in the sequence of length 5 and 10. That allows to transfer the observed values in one tasks to this individual for other tasks.

The data used for this research and the models that set the foundation for the adapted models are presented in Baumann et al. (2020). The authors describe four models for predicting human reasoning in an *Optimal Stopping Problem* where the goal is to find (and buy) the cheapest price for an item in a sequence of 10 prices. The four models are the *Independent Threshold Model*, the *Linear Threshold Model*, the *Biased Optimal Model*, and the *Cut-off Model*. All of the models are based on the calculation of an acceptance probability θ_i that implements a sigmoid choice function with sensitivity parameter β and the current item i with price x_i and the position-dependent threshold t_i :

$$\theta_i = \frac{1}{1 + \exp\{\beta(x_i - t_i)\}}$$

The goal of each model is to provide the threshold t_i which changes depending on the task and the position in the sequence.

The *Independent Threshold Model (ITM)* assumes no dependence between the thresholds, it samples N independent random thresholds (from a uniform distribution) t_1, \dots, t_N where at position $N + 1$ the price must be accepted.

In the *Linear Threshold Model (LTM)*, the thresholds are constrained by a linear relation to each other and are defined by the initial threshold t_0 and the linear scaling factor δ :

$$t_{i+1} = t_i + \delta \cdot i$$

The *Biased Optimal Model (BOM)* is based on the model presented in Guan et al. (2015). It uses the optimal threshold t_i^* , a systematic bias parameter γ that reflects the divergence

from the optimal threshold, and the parameter α which describes how much the threshold decreases or increases as the sequence progresses.

$$t_i = t_i^* + \gamma + \alpha \cdot i$$

The *Cut-off Model (CoM)* assumes that the reasoner has a fixed cutoff value k that determines how long the sequence is explored before the first value that is lower than the already seen minimum is accepted.

Benchmark Data

The data used in this project stems from (Baumann et al., 2020). It recreates a scenario in which the decision-maker is planning a vacation and wants to buy the flight tickets online. The prices vary randomly from day to day and the customer wants to find the cheapest ticket. Each day the decision-maker checks the price and can either buy the ticket or reject the offer and wait for the next day. Since the vacation will start in ten days, the decision maker has to accept the tenth offer, no matter the price. Once a price is accepted the search is also finished. A total of three experiments with different price distributions are reported.

For the *first experiment*, 129 participants were set to answer the described ticket-shopping task. The prices were sampled from a normal distribution with a mean of 180 and a standard deviation of 20. Each subject finished 200 trials of the ticket-shopping task. In each trial, the participants searched through a sequence of ten prices. The subjects were aware that they could see up to ten prices and were always informed about the number of remaining prices. However, they could only see the price of the current product. It was not possible to go back to an already rejected price. If the subjects reached the tenth price they were forced to accept it.

For the *second experiment*, 172 participants were in the ticket-shopping task but with changing distributions from which the prices were sampled. Three different sample methods were used: Exp 2a – prices were sampled from a left-skewed PERT distribution PERT(40, 195, 200) with a mean of 170. Exp 2b – a normal PERT distribution PERT(90, 140, 190) with a mean of 140 was used. Exp 2c – a right-skewed PERT distribution (PERT(120, 125, 400)) with a mean of 170 was used. Each participant was assigned to only one experiment and had to select the lowest price out of a sequence of 10 prices for 200 trials.

The *third experiment* simulates an online shopping experience where the goal is to buy a certain product for the lowest price possible with the prices being presented sequentially. A total of 60 commodity products were selected and the prices collected from an online shop. Only products with an approximately normal distributed price range were chosen. For the experiment, the prices were sampled from a normal distribution with the mean and standard deviation estimated from the real prices. All 100 participants performed 120 trials divided into two blocks containing the same sixty products. The subjects were always aware of the number of remaining prices

and were also informed about the mean price of the product. Once a price was rejected it could not be chosen again and the tenth price had to be chosen if no previous buy was performed.

All models were evaluated using *CCOBRA* (Brand et al., 2020). *CCOBRA* is a cognitive reasoning framework that sequentially presents per person data to the model that is currently evaluated. In each step of the evaluation sequence, the model is presented with the current task, in this case the task would be one price of the sequence of the ten prices out of a problem the participant had to face. With the presented task, the models have to predict the answer the current reasoner gave for this task, which is then used to evaluate the models performance. After predicting, *CCOBRA* provides an adaption function in which not only the task had to be predicted just now, but also the given answer is presented to the model. This information can then be used to adapt the model to the current reasoner. For the evaluation in this paper, the data was prepared in a way that the available information for the task is the price for the current ticket/item and the sequence of the current task (how often a price was rejected in this iteration). The reaction time (how long the individual took to make the decision) and the mean price (180 for experiments 1, 170 for the left- and right-skewed, 140 for the normal task in experiment 2, and variable for the third experiment) are also given as further information.

Adapting Models to Predict the Individual

The previously presented models from Baumann et al. (2020) were adapted to work with variable price means by scaling the individual parameters with the mean of the task. The models were also able to adapt to the individual reasoner by updating the parameters during the prediction process which will be presented in the following sections. A genetic algorithm was used to search for the optimal parameters for every 30 questions asked. The current parameters were then updated with new optimal values by setting the new parameters to be *70% current parameter + 30% searched parameter*. For all models, the previously presented β parameter was initialized to 0.21 and the genetic algorithm searched in the interval of $[0, 2]$.

Random Model (RM)

The *Random Model* represents the most simple decision maker by randomly selecting one of the options. A model that can't beat the random baseline would probably be better off just guessing the answer.

Independent Threshold Model (ITM)

The *Independent Threshold Model* samples its ten independent thresholds from a *uniform distribution* between 60% and 120% of the mean value for the current task. During the adaption to the individual reasoner, only β is searched with the genetic algorithm.

Linear Threshold Model (LTM)

The *Linear Threshold Model* starts with an initial threshold t_0 , which in this case is a percentage of the mean value of the current task. It is defined as: $t_0 = \text{meanvalue} \cdot t\%$. The linear increase δ is also represented as a percentage of the mean value of the current task. It is defined as: $\delta = \text{meanvalue} \cdot \delta\%$. The threshold calculation is then done via: $t_{i+1} = t_i + \text{meanvalue} \cdot \delta\% \cdot i$. For the basic version and the adaption, the parameters are initialized with $\delta\% = 0.005$ and $t\% = 0.7$. During the search for a better fitting value in the adaption phase, $\delta\%$ was limited to the interval of $[0, 0.1]$ and $t\%$ searched between 0% and 100%.

Optimal Threshold Model (OTM)

The *Optimal Threshold Model* uses the mathematical optimal threshold to determine whether to buy for the current price or to wait for the next opportunity. A way to calculate the optimal thresholds to find the highest number in a sequence is described in Gilbert and Mosteller (2006). In order to calculate the thresholds for the lowest number in the sequence, the threshold generation process was inverted. This results in a list of optimal thresholds (percentage of the mean price):

Table 1: Optimal thresholds for each value in the sequence. At 10 there is a 'must buy'. Any option below the threshold is accepted.

Pos.	1	2	3	4	5	6	7	8	9
Opt.	72	78	84	91	99	109	121	137	160

At each option that is to be predicted, the *Optimal Threshold Model* checks whether the current price is below the optimal threshold and predicts the *buy* option. If the current price is above the optimal threshold, it decides not to buy. Since there are no free parameters that can be optimized for this model, there is also no adaption variant for it.

Biased Optimal Model (BOM)

The *Biased Optimal Model* takes the optimal threshold for the current position in the sequence and adds two parameters to it (γ and α). Since those are also dependent on the magnitude of the current prices, they are also represented by a percentage of the optimal threshold t^* . The calculation for the *Biased Optimal Model* is therefore done like the following:

$$t_i = t_i^* + \gamma \cdot t_i^* + \alpha \cdot i \cdot t_i^*$$

Both parameters γ and α are both initialized to 0, which means that without any adaption, the *BOM* is equal to the *OTM*. During the adaption to the individual reasoners, the genetic algorithm searches for the optimal α value in the range of $[-0.2, 0.1]$ and for the optimal γ value in the range $[-0.5, 0.3]$.

Cut-off Model (CoM)

The *Cut-off Model* explores the sequence a fixed number of steps (k) and then takes the next option that is lower than the

previously seen lowest value. During the initialization of the *CoM* k is set to 5. The genetic algorithm for the adaption part of the model searches for k in the range of [1, 10].

Results

All models were tested and tuned on the data of Exp 1. These results were used to improve the models by adapting the parameters for better performance. The data of all other experiments was not used in the training process, and only used in the final evaluation, to avoid overfitting on the data. Experiment 2 and its variant for skewness demonstrates the models' power to adapt, given different price distributions. A model that performs well on a left-skewed distribution (more cheap prices in the sequence) might in term perform worse on a right-skewed distribution (more costly prices). The dataset of Exp 3 (real prices) gives insight into the models' ability to adapt to real-life situations. During the evaluation, each model was run five times on each dataset, to account for the randomness of the genetic algorithm (Ritter, Schoelles, Quigley, & Klein, 2011; Byrne, 2013). All later presented results are therefore the mean of five evaluation runs. Overall the results show a good performance for most models on the datasets of the first and second experiment. The best performing models were the *Biased Optimal Model* and the *Linear Threshold Model* both with adaption. The best performing model was the *LTM* with adaption, reaching an accuracy of 88.74%. On the third experiment, however, all models had a significant drop in performance, with the *Cut-off Model* being the best performing one. Overall the *Independent Threshold Model* showed the worst performance.

Experiment 1

The evaluation results for this experiments can be found in Table 2. The clear prediction performance winner is the *Biased Optimal Model* with adaption. It achieved an 86.7% mean accuracy on the prediction. Next up is the *Linear Threshold Model* with adaption which scored an 85.0% accuracy. The *Cut-off Model* with and without adaption as well as the *Linear Threshold Model* without adaption scored at around 81% accuracy. Due to the nature of how the data is presented in the datasets, simply predicting that the current reasoner will not buy for the current price will lead to a high prediction performance (in this case 79.1%). This is because once a reasoner accepted a certain price, the remaining prices for this sequence were skipped. This leads to an over-representation of *don't buy* answers in the dataset. This prediction performance can therefore be seen as the barrier that shows if the model truly learned the reasoning process. As to be expected, the *Biased Optimal Model* without adaption and the *Optimal Threshold Model*, with 73.5% prediction accuracy, share the same performance since the *BOM* without adaption represents the *OTM*. The *Independent Threshold Model* with and without adaption achieved around 68% accuracy with the adaption model even performing slightly worse. The random model scored around 50% accuracy as it is to be expected in an two possible outcomes random choice.

Experiment 2a: left skewed prices

The results for the second experiment's first condition, with a left-skewed distribution (more cheap prices), are presented in Table 2. Interestingly, all models managed to improve their performance in comparison to the first experiment. This time, the *Linear Threshold Model* with adaption with 88.7% accuracy performed slightly better than the *Biased Optimal Model* with adaption that reached 88.7% prediction performance. With 87.2% accuracy, the *LTM* without adaption managed to improve its performance drastically compared to the first experiment.

Experiment 2b: normally distributed prices

In Table 2 the results for the second condition of the second experiment (normal-distributed prices) are shown. Once again, the *BOM* and *LTM*, both with adaption, are the best performing models. This time, like in the first experiment where the prices were also normal-distributed, the *Biased Optimal Model* with adaption performed slightly better, with 88.0% accuracy, than the *Linear Threshold Model* with adaption which reached 86.6% accuracy.

Experiment 2c: right skewed prices

The results for the second experiment's third condition (left-skewed distribution, more expensive prices) are presented in Table 2. The *Biased Optimal Model* with adaption continued with the trend of being one of the strongest models and showed the best performance of all models with 87.3%. However, compared to the other two conditions of the second experiment, the *BOM* with adaption showed the worst performance in this condition. With an accuracy of 86.1%, the *Linear Threshold Model* with adaption also showed slightly worse performance than under the previous two conditions of the second experiment. Nevertheless, the *LTM* with adaption proved to be a solid predictor of the decision-makers in the second experiment. Both, the *Cut-off Model* with and without prediction, showed constantly good results in the entirety of the second experiment and always managed to beat the *never buy* threshold. In this case, they achieved an accuracy of 81.2% and 80.5% respectively. The *BOM* without adaption and the *Optimal Threshold Model* consistently achieved a performance of around 74% accuracy. The trend of both *Independent Threshold Models* being the worst in the portfolio also continued under this condition and with an accuracy of 68.8% both reached a performance as low as ever.

Experiment 3: real prices

In the third experiment, the performance of all models dropped significantly in comparison to the first and second experiment (cp. Table 2). None of the models managed to beat the *don't buy* threshold. Most notably the *Cut-off Model* with and without adaption is now leading the scoreboard with an accuracy of 77.8% and 76.1% respectively. This represents a loss of roughly three percentage points compared to the first and second experiments. The *Linear Threshold Model* with adaption lost roughly thirteen percentage points compared to the previous experiments and dropped

Table 2: The median predictive accuracy of the cognitive models for each experiment and all experiments, and the random and don't buy baseline models. In bold is the highest predictive performance. The median is calculated from five evaluation runs.

	<i>Random</i>	<i>dontbuy</i>	<i>ITM_A</i>	<i>ITM</i>	<i>OTM</i>	<i>BOM</i>	<i>LTM</i>	<i>CoM</i>	<i>CoM_A</i>	<i>LTM_A</i>	<i>BOM_A</i>
Exp. 1	49.8	79.1	67.9	68.1	73.5	73.5	81.0	81.1	81.4	85.0	86.7
Exp. 2 (left skew.)	50.0	81.1	69.7	69.8	73.9	73.9	87.2	81.5	82.9	88.7	88.5
Exp. 2 (normal)	50.1	79.2	70.0	70.0	74.7	74.7	82.1	81.2	81.7	86.6	88.0
Exp. 2 (right skew.)	50.0	78.7	68.8	68.8	74.1	74.1	83.7	80.5	81.2	86.1	87.2
Exp. 3	50.0	78.2	64.2	64.1	71.2	71.2	73.1	76.1	77.8	74.0	73.8
All	50.0	80.5	68.1	68.2	73.5	73.5	81.4	80.1	81.0	84.1	84.9

to an accuracy of 73.9%. Surprisingly, the *LTM* without adaption scored only slightly lower with 73.1% accuracy, showing that in this case, the adaption did not bring any meaningful advantage. Overall the *LTM* without adaption dropped about ten percentage points in performance compared to the previous experiments. A similar observation can be done with the *Biased Optimal Model*. The adaption-based model scored 73.8% accuracy while the non-adapting one scored only slightly lower with 71.2% prediction accuracy. The adaption-based *BOM* loses roughly fourteen percentage points of prediction performance compared to experiments one and two, while the non-adaption model loses only around three percentage points, which is of course the same development as the *Optimal Threshold Model*. Once again, the *Independent Threshold Model*, with and without adaption, showed the worse performance of all models with only around 64% accuracy. Compared to the other experiments this represents a loss of around four to five percentage points.

Individual Prediction Performance

Fig. 1 shows the individual participant performance for all models on the third experiment. The *COM* model and its adaption show no meaningful difference in the overall distribution of the individual prediction performance other than that the adaption manages to score slightly higher on almost all participants showing a better fit to the individual. The adaption of the *LTM* clearly reduces the overall spread of the individuals, while also the outliers both at the bottom and at the top are clearly reduced by the adaption.

For the *Biased Optimal Models*, the non-adapting model shows a wider spread of the individual performances. However, there are no real outliers to the bottom of the performance so the basic model manages to fit all participants at least to a certain amount. The model with adaption clearly shows how the performance is improving by shrinking the spread between the individuals and bringing them all to higher accuracy. Nevertheless, there are some participants the adaption model does not quite manage to fit and therefore there are also more outliers to the bottom. The *ITM* and its adaption did not show any improvements. It is therefore most likely, that the β value does not yield any performance increase in a setting with random thresholds.

Discussion

The overall bad performance of the *Independent Threshold Model* with and without adaption reinforces the previous findings (Baumann et al., 2020; Guan & Lee, 2018; Guan et al., 2015) that human reasoners change their thresholds following a certain strategy (e.g. linear). A random (bounded) threshold for each step in the sequence would mean that a decision-maker could refuse a cheap price in the current step only to accept a way to high price in the next step. This behavior is not typical for a human reasoner, which is shown by the bad performance of the models. The below-average performance of the *Optimal Threshold Model* and the *Biased Optimal Model* without adaption, which are the same, confirms the findings of Guan et al. (2015) that human reasoners in *Optimal Stopping Problems* tend to set their initial threshold too low or too high, and then either decrease them to slow or increase them to fast. The *Cut-off Model* showed a solid performance and managed to beat the *don't buy threshold* consistently. This shows that the reasoners tend to at least somewhat try to explore the sequence rather than making a hasty decision. This effect has already been mentioned in Baumann et al. (2020). The outstanding performance of the *Linear Threshold Model* with and without adaption once again strengthens the assumption that human reasoners tend to use linear threshold in *Optimal Stopping Problems*. The good performance of the *LTM* shows that most reasoners, at least to a certain extend, choose a linear threshold to guide their decision. However, some outliers might apply a different technique and an in-depth evaluation of those outliers could help to further increase the ability to predict the individual reasoner. The *Biased Optimal Model* with adaption shows a similar prediction accuracy to the *LTM* with adaption. This behavior is to some extent understandable since the optimal thresholds for the first ten options in the sequence are roughly linear as well. Additionally, the parameters for γ and α found by the genetic algorithm during the adaption were mostly negative meaning they also counteracted the nonlinear effect of the optimal threshold.

Both the *BOM* and *LTM* show a strong performance over all experiments (cp. Table 2). Only in the third experiment, both did not outperform the *don't buy* baseline. There might

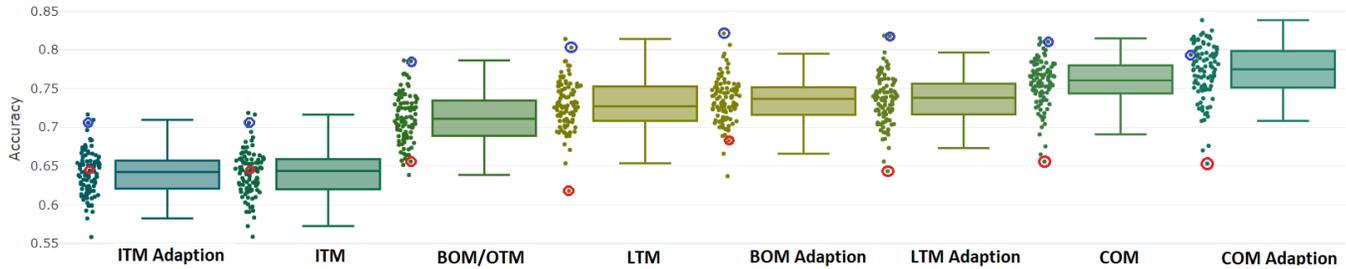


Figure 1: Prediction results of the models on each participant in the third experiment. Random model omitted for space reasons, BOM/OTM are the same with no adaption. Red and blue circle mark the same participants respectively.

be several reasons for this drop in performance: One reason for the change in performance could lie in the participants themselves. The fact that they had to deal with real items that they know and have an understanding of the price developments, could have nudged a behavior change compared to the fictional price setting of the first and second experiment. Another reason might be that the reasoners just did not behave as predictable as in the other tasks. Since the prices varied quite a lot in their magnitude, a reasoner that previously accepted a costly product at a 20% discount could not accept the same 20% discount for a cheaper item since the absolute discount of the cheap item does not appear as high.

Consider two types of participants in Fig. 1 and 2, depicted by a red and blue circle accordingly. The blue participant is predicted adequately by almost all models in contrast to the low predictive accuracy for the red participant across models. This indicates that either the red-marked participant gave more random answers and is less predictable by any meaningful models or the reasoner developed a strategy that is beyond what the implemented models can cover. The well-predicted blue reasoner, however, seems to use a strategy that is covered by almost all models.

In conclusion, this paper adapted and evaluated core decision making models for *Optimal Stopping Problems* to predict decisions performed by individuals. Using genetic algorithms allowed the models to find the optimal parameters for each individual. The findings support previous research that showed that human reasoners tend to use a linear threshold in *Optimal Stopping Problems* to rate the current option. The analysis shows that thresholds are variable among decision makers and that adapting to the individual can bring a vast improvement in the prediction capabilities of the models.

Acknowledgements

Support to MR by the Danish Institute of Advanced Studies and the DFG (RA 1934/4-1, RA1934/8-1) is gratefully acknowledged.

References

Baumann, C., Singmann, H., Gershman, S. J., & von Helversen, B. (2020). A linear threshold model for optimal

stopping behavior. *Proceedings of the National Academy of Sciences*, 117(23), 12750–12755.

Brand, D., Riefterer, N., & Ragni, M. (2020). *Ccobra-framework*. <https://github.com/CognitiveComputationLab/ccobra>. GitHub.

Byrne, M. D. (2013). How many times should a stochastic model be run? an approach based on confidence intervals. In *Proceedings of the 12th International Conference on Cognitive Modeling, Ottawa*.

Gilbert, J. P., & Mosteller, F. (2006). Recognizing the maximum of a sequence. In *Selected Papers of Frederick Mosteller* (pp. 355–398). Springer.

Guan, M., & Lee, M. D. (2018). The effect of goals and environments on human performance in optimal stopping problems. *Decision*, 5(4), 339.

Guan, M., Lee, M. D., & Vandekerckhove, J. (2015). A hierarchical cognitive threshold model of human decision making on different length optimal stopping problems. In *CogSci*.

Lee, M. D. (2006). A hierarchical bayesian model of human decision-making on an optimal stopping problem. *Cognitive Science*, 30(3), 1–26.

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

Ritter, F. E., Schoelles, M. J., Quigley, K. S., & Klein, L. C. (2011). Determining the number of simulation runs: Treating simulations as theories by not sampling their behavior. In *Human-in-the-loop simulations* (pp. 97–116). Springer.

Seale, D. A., & Rapoport, A. (2000). Optimal stopping behavior with relative ranks: The secretary problem with unknown population size. *Journal of Behavioral Decision Making*, 13(4), 391–411.

von Helversen, B., & Mata, R. (2012). Losing a dime with a satisfied mind: Positive affect predicts less search in sequential decision making. *Psychology and Aging*, 27(4), 825.

Zwick, R., Rapoport, A., Lo, A. K. C., & Muthukrishnan, A. (2003). Consumer sequential search: Not enough or too much? *Marketing Science*, 22(4), 503–519.

Inferring a Cognitive Architecture from Multi-Task Neuroimaging Data: A Data-Driven Test of the Common Model of Cognition Using Granger Causality

Holly S. Hake (hakehs@uw.edu)

Department of Psychology, University of Washington, Seattle, WA 98195 USA

Catherine Sibert (sibert@uw.edu)

Department of Psychology, University of Washington, Seattle, WA 98195 USA

Andrea Stocco (stocco@uw.edu)

Department of Psychology, University of Washington, Seattle, WA 98195 USA

Abstract

A common complaint levied at analyses based on cognitive architectures is their lack of connection to observed functional neuroimaging data, particularly for architectural models that rely on high level, theoretical components of cognition. Previous work has connected task-based functional MRI data to the Common Model of Cognition (CMC), using a top-down modeling approach. Here, a bottom-up method, Granger Causality Modeling (GCM), is applied to the same task-based data to infer a network of causal connections between brain regions based on their functional connectivity. The resulting network shares many connections with those proposed by the Common Model, and also suggests important additions to the Common Model, likely related to the role of episodic memory in control.

Keywords: Cognitive Modeling, Cognitive Architecture, Granger Causality, Functional Connectivity

Introduction

In the field of cognitive architectures, an important topic of discussion is the relationship between the components of an architecture and their relationship to the brain. Some architectures, like SPAUN (Eliasmith et al., 2012), LISA (Hummel & Holyoak, 2005), and Leabra (O'Reilly et al., 2016), are designed to mimic the brain's biological circuits and rely on artificial neurons as their building blocks. These systems take a circuit-level approach to cognitive modeling, based on the notion that function arises from form. An alternate, functional approach forms the basis of another class of architectures, such as Soar (Laird, 2019), or ACT-R (Anderson, 2007), whose building blocks are more abstract and high-level cognitive components such as perceptual systems and memory that have been then mapped *post-hoc* to particular brain regions (e.g., Anderson, Fincham, Qin, & Stocco, 2008).

Ultimately, the success of both bottom-up and top-down approaches depends on one fundamental aspect, that is, the exact nature of the functional connections between the assumed components, or the underlying brain's architecture. Surprisingly, the fields of systems-level neuroscience and the fields of cognitive architectures have rarely interacted in this domain. In this paper, we attempt to resolve some of the tensions between the competing methodologies by using Granger Causality Modeling

(GCM) of low-level functional brain activity to find causal connections between brain regions associated with high-level cognitive components. The networks produced by these connections are then compared to existing frameworks of theoretical architectures.

Functional Connectivity

Most research aimed at understanding brain architecture has been done through the analysis of *functional connectivity*, a data-driven and bottom-up method of determining the degree of connection between brain regions through statistical dependencies--typically, the Pearson correlation between times series in different brain regions (e.g., Fox et al., 2005). Through this method, network neuroscientists have identified several distinct networks of brain regions, such as the Default Mode Network (Raichle et al., 2001). However, while functional connectivity analysis can detect the presence of such networks, it can be difficult to characterize the specific function or role that they play in higher level cognition. Furthermore, correlation coefficients have no directionality attached to them, which makes it impossible to draw causal conclusions about the role of different regions and the flow of information along a network.

The Common Model of Cognition

A number of recent studies have tried to connect architecture frameworks to functional brain activity in a top-down fashion, by imposing architectural constraints on a network of connected brain regions. In particular, these studies have capitalized on the Common Model of Cognition (CMC), an abstract description of the principles common to multiple architectures (Laird, Lebiere, & Rosenbloom, 2017). The CMC proposes that, at the highest level, cognition arises from the interaction of five cognitive components, corresponding to Perception, Action, Long-Term Memory, Procedural Memory, and Working Memory. These components can be associated with five corresponding large-scale brain circuits, and a network of directional connections can be drawn between them. Most recently, Stocco et al. (2021) showed that the CMC

outperforms a selection of six other architectures in fitting data across six different task paradigms spanning seven different domains, suggesting that it provides a reasonably accurate system-level description of the brain’s architecture.

Unfortunately, all of the previous tests of the CMC (Stocco et al., 2018; Steine-Hanson et al., 2018; Stocco et al., 2021, have employed a top-down approach, comparing the relative fit of different possible architectures. This approach was partially constrained by the choice of one particular method of the analysis of effective connectivity, Dynamic Causal Modeling (DCM: Friston, Harrison, & Penny, 2003). The authors justified the choice of DCM because it allows for the distinguishing of the directionality of connections, while the most commonly used functional connectivity measures are based on partial correlations and are non-directional. While DCM allows for directional estimates, it relies on the top-down implementation of a plausible architecture, and it also limits the use of a bottom-up, data-driven approach. Because the space of possible architectures, even when only five components are considered, is extremely large, it is possible that a better candidate architecture exists, but was simply not included among those examined by Stocco et al. (2021).

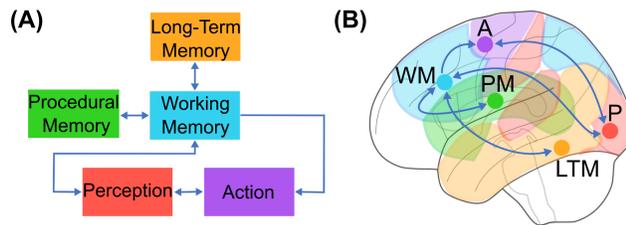


Figure 1: (A) The Common Model of Cognition (CMC); (B) Proposed associations between components and anatomical brain regions.

Granger Causality Modeling

In this paper, we re-analyze the data from the Stocco et al (2021) paper using Granger Causality Modeling. In GCM, the existence of a causal effect between two time series x and y is established by comparing two models (Granger, 1969), one auto-regressive linear model in which the value of y at times t depends only on its past value at time $t-1$:

$$y(t) = \beta_0 + \beta_1 y(t-1)$$

and an alternative model that includes the effect of the past state of x :

$$y(t) = \beta_0 + \beta_1 y(t-1) + \beta_2 x(t-1)$$

If the second model is significantly better than the first, then x is said to Granger-cause y . Although it was originally developed and applied in the field of economics, Granger causality has been successfully applied to

neuroimaging data (Roebreck et al., 2005; Deshpande et al., 2008) and offers similar advantages and comparable performance to DCM (Friston et al., 2012). In this paper, we apply this method to test the existence of all possible causal connections between the five components proposed by the CMC.

Materials and Methods

Participants

The study presented herein consists of an extensive analysis of a large sample ($N = 200$) of neuroimaging data from the Human Connectome Project (HCP), the largest existing repository of healthy young adult neuroimaging data.

Task fMRI Data

The HCP task-fMRI data encompasses seven different paradigms designed to capture a wide range of cognitive capabilities. Of these paradigms, six were included in our analysis (the seventh was a motor localization task). A full description of these tasks and the rationale for their selection can be found in the original HCP papers (Barch et al., 2013; Van Essen et al., 2013).

Data Processing and Analysis

Image Acquisition and Preprocessing. MRI images were acquired and minimally preprocessed according to HCP guidelines (Barch et al., 2013; Van Essen et al., 2013). Scans were taken on a 3T Siemens Skyra using a 32-channel head coil with acquisition parameters set at TR = 720 ms, TE = 33.1 ms, FA = 52°, FOV = 208 × 180 mm. Each image contained 72 2.0mm oblique slices with an in-plane 2.0 x 2.0 mm resolution. Images were acquired with a multi-band acceleration factor of 8X. These raw images then underwent minimal preprocessing including unwarping, motion realignment, and normalization to the standard MNI template. From there, the images were then smoothed with an isotropic 8.0 mm full-width half maximum Gaussian kernel.

Regions of Interest Definition. Regions of Interest (ROIs) for each task and participant were defined using the method described in Stocco et al. (2021) and available on the paper’s online repository. For each CMC component, a group-level centroid was first identified by running a canonical GLM analysis that compared the stimuli against their task-specific baseline and then locating the peak of a statistical parametric map within the general areas associated with that CMC component (Figure 1). Because all tasks show stronger activation in the left hemisphere than in the right, all the group-level centroids were located in the left hemisphere.

To account for individual-level variability in functional neuroanatomy, the group-level coordinates were then used as the starting point to search in 3D space for the closest activation peak within each individual statistical parameter map. Figure 2 illustrates the distribution of the individual

coordinates of each region for each task, overlaid over a corresponding group-level statistical map of task-related activity (as in Stocco et al., 2021). Each individual coordinate is represented by a point; the ≈ 200 points for each region form a cloud that captures the spatial variability in the distribution of the individual coordinates for that region. Next, the individualized ROI coordinates were used as the center of a spherical ROI with an 8mm radius. All voxels within the sphere whose response was significant at a minimal threshold of $p < .50$ (that is, a 50% probability of showing a response) were included as part of the ROI.

Finally, for each ROI of every participant in every task, a representative time course of the BOLD signal was extracted as the first principal component of the time series of all of the voxels within the sphere. The resulting time series, one per component, were then entered into a Granger causality model.

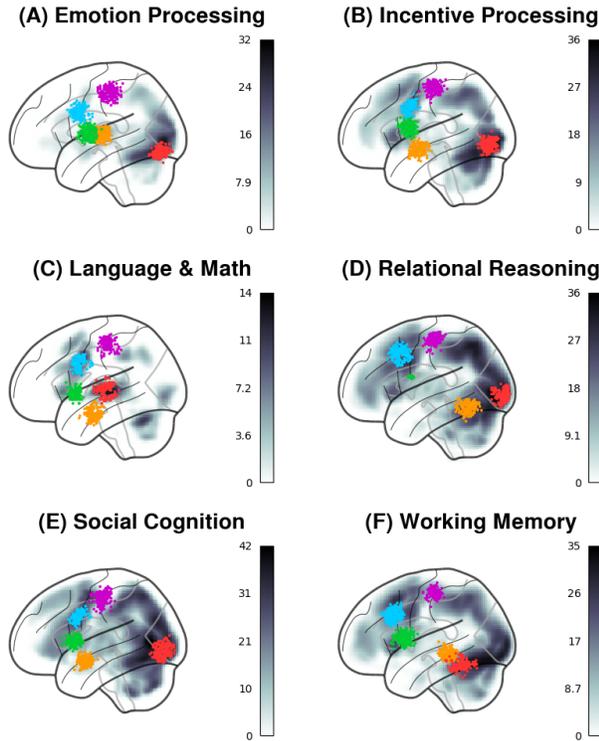


Figure 2: Location of ROI centroids across the six tasks of the Human Connectome Project; variations account for individual differences in functional anatomy.

Granger Causality Model

A multivariate Granger causality model was then set up, in which the BOLD response at time t across all regions, $\mathbf{x}(t)$, was modeled as the contribution of all of the regions (including itself) at lags 1, 2, ... k :

$$\mathbf{x}(t) = \beta_0 + \beta_1 \mathbf{x}(t-1) + \dots + \beta_k \mathbf{x}(t-k) \tag{1}$$

To determine the optimal lag value, ten models were created by varying k from $k = 1$ to $k = 10$, and the value of k that gave rise to the model with the lowest Bayesian Information Criterion was selected. Across all participants and tasks, the maximum lag that was observed was $k = 6$, and the modal was $k = 2$. Note that, when $k > 1$, there are multiple different parameter estimates that quantify the directional effect of a region on another region, one for each lag. To reduce the dimensionality of these estimates, only the most significant lag (i.e., the one with the smallest p -value) was selected.

For each participant, a subject-level inferred architecture was then created by discretizing the matrix of connections and maintaining only directed links with $p < .05$. To infer a group-level architecture from the individual-level architectures, the most likely directed links between regions need to be inferred from the frequency of their distribution in the sample of participants. To determine the probability that each directed connection c is part of the group-level architecture, we modeled the probability of it appearing across all participants as a binomial distribution, with a prior probability of $P = .50$.

Results

Group level connection maps for each of the six tasks used in the HCP dataset are shown in Figure 3. The figures show a connectivity matrix representation of the inferred architectures for each task, where the brightness of each matrix cell reflects the probability that the corresponding directional connection should be included in the architecture.

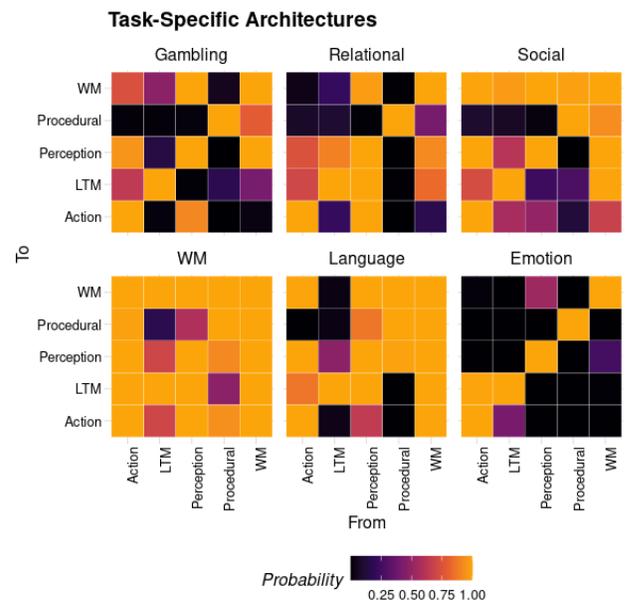


Figure 3: Task-specific connection grids for each of the six HCP task paradigms. Each grid square represents a potential causal connection between two regions, and the brightness of the square reflects the probability of that link at a group level.

The figure highlights the different connections utilized in each task domain, as well as some commonalities shared by all tasks (i.e., the preponderance of connections to and from the WM component, corresponding to the PFC).

As pointed out by Stocco et al. (2021), however, an efficient architecture should be *stable* and maintain its functional characteristics across different tasks. Therefore, to derive a general architecture from these six task-specific ones, we considered each task as an independent experiment to test these connections, and we used Fisher’s (1932) method to combine the p -values from each task. According to this method, the distribution of the log of p -values from independent tests follows a χ^2 distribution with $2N$ degrees of freedom, and the p -value of each connection can be calculated from the χ^2 cumulative distribution function as follows:

$$p_{\text{global}} = p(\chi^2_{2N} > \sum_{\text{task}} \log p_{\text{task}})$$

The results are shown in Figure 4, which represents the connectivity matrix of an architecture inferred across participants *and* domains.

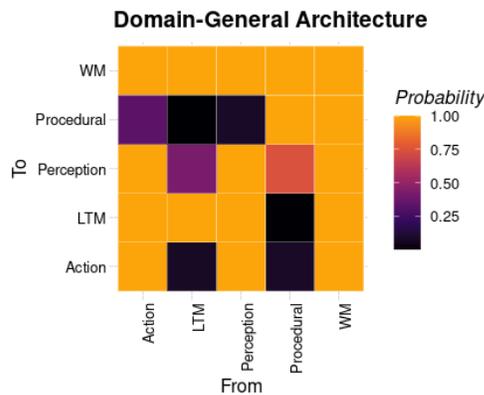


Figure 4: The connection grid for a general architecture incorporating all six HCP tasks. Each grid square represents a potential causal connection between regions, and the brightness of the square reflects the probability of that connection being present.

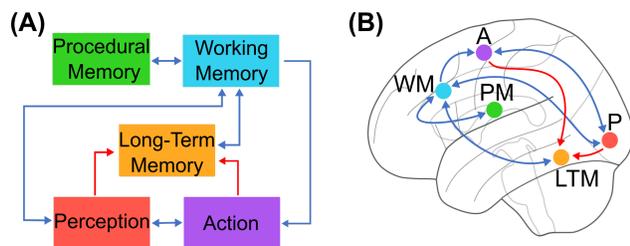


Figure 5: (A) A visual representation of the architecture inferred from the Granger causality model; (B) Proposed associations between components and anatomical brain regions. Arrows: dark blue, connections present in both CMC and GM; red, connections unique to Granger model.

The results support an architecture that is similar, but not exactly identical, to the CMC. If a strict 95% threshold is applied to the map of connections inferred from the GCM analysis, 22 of the possible 25 connections are shared between the human-derived network and the CMC (Fig. 5).

Comparing the Other Architectures

In the previous DCM based analysis of architecture structures, Stocco et al. (2021) were not able to incorporate data-driven inferences about connections. Instead, the plausibility of the CMC was evaluated by comparing its predictions against a set of representative alternative architectures across tasks (Fig. 6). These architectures, divided into two categories, or “families”, represent the possible organizational structures of general purpose architectures. All consist of the same five regions or components present in the Common Model, but provide differing accounts of the connections between them.

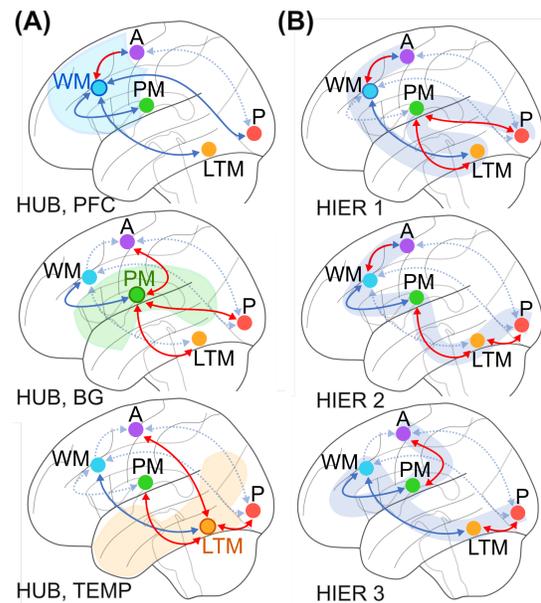


Figure 6: (A) Three variations of Hub-and-Spoke (HUB) models, and of (B) Hierarchical (HIER) models. Arrows: dark blue, connections present in both CMC and alternate models; red, connections unique to alternate models; and dotted, connections present in CMC and absent in alternate models.

The “Hub and Spoke” model family designates a single region as the “Hub” of model activity, with bidirectional connections between it and all other regions. These “Spoke” regions, however, do not connect to one another, and activity passing from one spoke region to another must also pass through the hub. Of the five CMC components, arguments can be made for each of the three memory modules serving in the capacity of a hub: working memory could drive activity from the prefrontal cortex (Hub PFC), long term memory could drive activity from temporal

regions (Hub Temporal), or procedural memory could drive activity from the basal ganglia (Hub Procedural).

An alternate account of model structure is posed by members of the “Hierarchical” family. In this account, the architecture serves as a feed-forward system where activity originates in the perception region, travels through the successive memory regions, and culminates in the action region. With the limited number of regions and the fixed position of the perception and action regions, the potential models in this family vary only in the order of the three memory modules. An additional constraint, the assumption that long term memory (LTM) will proceed working memory (WM), leaves the position of procedural memory as the only degree of freedom. It is either the first of the three memory modules (Hierarchical 1), the middle module (Hierarchical 2), or the final module before action (Hierarchical 3).

To test whether the results of our Granger causality model converge with those previously reported with DCM (a test of convergent validity), we performed the same comparison of architectures done by Stocco et. al (2021). To do so, we first derived the theoretical network architectures of the six alternate architectures examined in that study for comparison against the network architecture derived using GCM. These networks are represented in the form of connectivity matrices in Figure 7.

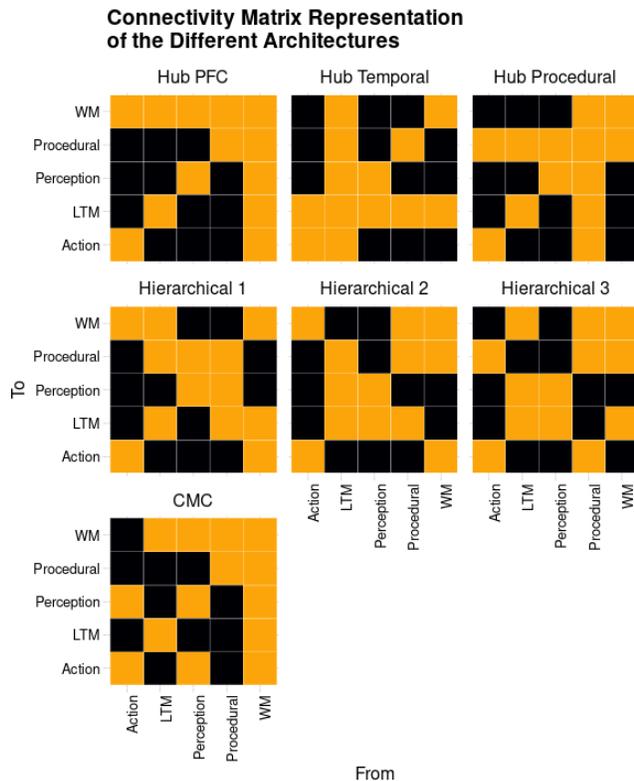


Figure 7: Connectivity matrix representations of the six alternative architectures (Figure 6) and the CMC.

For each alternate architecture, as well as the CMC, we examined the degree of similarity between the network of connections suggested by the GCM analysis and the connections theorized by the architecture. We considered three metrics. The first is the *correlation* between the predicted and observed directed connections in the vector of 25 possible edges in the networks. The second is the proportion of *overlap* between the two vectors, defined as the proportion of exactly matched connectivity predictions or, equivalently, the complement of the proportional Hamming distance between the two vectors of connections. The third and final metric is *likelihood*, defined as the Z-scores of predicted vs. expected number of successes in a binomial distribution of 25 connections. The results of each of the three metrics are compared in Figure 8. For all criteria, the CMC reflects the greatest similarity to the network architecture uncovered by the GCM analysis.

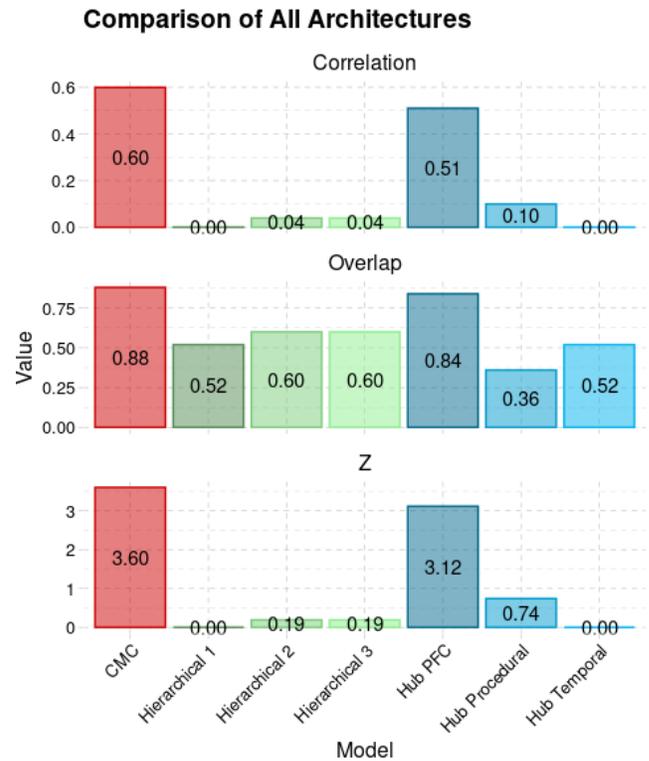


Figure 8: Comparison of the CMC (in red) and six alternate architectures in terms of three measures of similarity to the GCM network.

Discussion

In this paper, we have presented an analysis of Human Connectome Data using the same ROIs as in Stocco et al (2021), but employing Granger causality instead of DCM, to analyze fMRI data. A replication of the original comparison (Stocco et al., 2021) between the Common Model of Cognition and six alternative architectures largely confirmed the previous study’s findings, namely, that the CMC performs better than the alternative

architectures at explaining effective connectivity within and across all tasks.

Granger causality modeling, however, offers the unique opportunity of deriving a new architecture from data. The new architecture, represented in Figure 5, is a modified version of the CMC with the addition of projections from Perception and Action regions to the LTM component.

It is interesting to note that both connections are unidirectional, i.e., Perception and Action feed to LTM but do not receive projections back. Instead, bidirectional connections exist between LTM and WM. This particular connectivity structure seems to be best adapted to implement a form of instance-based learning, whereby successful episodic memories are formed by encoding previous stimuli, and actions and outcomes are stored to be later retrieved and guide behavior. In this case, direct connections from Perception and Action to LTM would support the encoding of stimuli and actions, respectively, while the connectivity from WM to LTM could support the encoding of evaluation of the outcome (performed by the WM component). Previous episodes could be later retrieved through the directed connection between LTM and WM. Moreover, the existence of additional functional links to LTM suggests that the large-scale brain organization seems to contain multiple hubs of different importance.

Limitations

However, these findings should be considered in light of a number of potential limitations. The first is that, while we ultimately aggregated the results into a single task-independent network, a significant amount of variability exists between the network architectures that can be inferred from the specific tasks. While DCM is intrinsically top-down and limited to examining the fit of specific network models, GCM does not suffer such limitation. Thus, the degree to which an task-independent architecture could be derived from individual tasks is debatable, and reflects the underlying assumption that, at a very high level, brain activity showcases a common invariant architecture. This hypothesis, of course, is not universally accepted and should be examined independently in future studies.

A second limitation is that the estimates of connectivity obtained through Granger causality might change when a larger set of component regions are included; thus, these results cannot be considered stable until the exact number of ROIs is considered canonical. It should be noted that, however, this limitation is also common to DCM and thus similarly affects previous work in this area (e.g., Steine-Hanson et al., 2019; Stocco et al., 2021).

Finally, it should be noted that, although these results do suggest that a better architecture (depicted in Figure 5) might outperform the original Common Model architecture and the others tested by Stocco et al. (2021), they do not necessarily imply so. This is because GCM is a different method than DCM, it is entirely possible that the

architecture of Figure 5 would not perform as well when its effective connectivity is measured within the DCM framework. Therefore, possible future studies should re-investigate the superiority of this new architecture using the DCM-based comparison, as done in Stocco et al (2021).

Implications for the Common Model of Cognition

These limitations notwithstanding, these results do support credibility to the principles of the CMC. The architecture that was identified through GCM differs only minimally from the CMC, and the CMC remains the architecture that most closely matches our results across the set of potential architectures tested by Stocco et al. (2021). Our new findings, however, suggest important modifications to its structure. We consider these results as an exciting starting point for the future examination of large scale-connectivity of the brain.

Acknowledgement

This research was supported by grant FA9550-19-1-0299 from the Air Force Office of Scientific Research to AS.

References

- Anderson, J. R. (2007). How can the human mind occur in the physical universe? *Oxford University Press*.
- Anderson, J. R., Fincham, J. M., Qin, Y., & Stocco, A. (2008). A central circuit of the mind. *Trends in cognitive sciences*, 12(4), 136–143.
- Barch, D. M., Burgess, G. C., Harms, M. P., Petersen, S. E., Schlaggar, B. L., Corbetta, M., & WU-Minn HCP Consortium. (2013). Function in the human connectome: Task-fMRI and individual differences in behavior. *NeuroImage*, 80, 169–189.
- Deshpande, G., LaConte, S., James, G. A., Peltier, S., & Hu, X. (2009). Multivariate Granger causality analysis of fMRI data. *Human brain mapping*, 30(4), 1361–1373.
- Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., & Rasmussen, D. (2012). A large-scale model of the functioning brain. *science*, 338(6111), 1202–1205.
- Fisher, R. A. (1932). *Statistical methods for research workers*. Edinburgh: Oliver and Boyd.
- Friston, K., Moran, R., & Seth, A. K. (2013). Analysing connectivity with Granger causality and dynamic causal modelling. *Current opinion in neurobiology*, 23(2), 172–178.
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, 102(27), 9673–9678.
- Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, 19, 1273–1302.
- Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37, 424–438.

- Hummel, J. E., & Holyoak, K. J. (2005). Relational reasoning in a neurally plausible cognitive architecture: An overview of the LISA project. *Current Directions in Psychological Science*, 14(3), 153-157.
- Laird, J. E. (2019). *The Soar cognitive architecture*. MIT press.
- Laird, J. E., Lebiere, C., & Rosenbloom, P. S. (2017). A Standard Model of the Mind: Toward a Common Computational Framework across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics. *AI Magazine*, 38, 13–26.
- O'Reilly, R. C., Hazy, T. E., & Herd, S. A. (2016). The leabra cognitive architecture: How to play 20 principles with nature. *The Oxford handbook of cognitive science*, 91, 91-116.
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, 98, 676–682.
- Roebroeck, A., Formisano, E., & Goebel, R. (2005). Mapping directed influence over the brain using Granger causality and fMRI. *Neuroimage*, 25(1), 230-242.
- Steine-Hanson, Z., Koh, N., & Stocco, A. (2018). Refining the Common Model of Cognition Through Large Neuroscience Data. *Procedia Computer Science*, 145, 813–820.
- Stocco, A., Laird, J. Lebiere, C., & Rosenbloom, P. (2018). Empirical evidence from neuroimaging data for a Standard Model of the Mind. In C. Kalish, M. Rau, J. Zhou, and T. T. Rogers (Eds.), *Proceedings of the 40th Annual Meeting of the Cognitive Science Society*. Madison, WI, pp. 1094-1099.
- Stocco, A., Sibert, C., Steine-Hanson, Z., Koh, N., Laird, J. E., Lebiere, C. J., & Rosenbloom, P. (2021). Analysis of the human connectome data supports the notion of a “Common Model of Cognition” for human and human-like intelligence across domains. *NeuroImage*, 235, 118035.
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., & WU-Minn HCP Consortium. (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage*, 80, 62–79.

Physio-Cognitive Modeling: Explaining the Effects of Caffeine on Fatigue

Tim Halverson (thalverson@aptima.com)

Aptima, Inc., 2555 University Blvd, Suite 300
Fairborn, OH 45324 USA

Christopher W. Myers (christopher.myers.29@us.af.mil)

Jeffery Gearhart (jeffery.gearhart.ctr@us.af.mil)

Matthew W. Linakis (matthew.linakis.1@us.af.mil)

Glenn Gunzelmann (glenn.gunzelmann@us.af.mil)

711th Human Performance Wing, Air Force Research Laboratory
Wright-Patterson AFB, OH 45433 USA

Abstract

Most computational theories of cognition lack a representation of physiology. Understanding the effects of compounds present in the environment on cognition is important for explaining and predicting changes in cognition and behavior given exposure to toxins, pharmaceuticals, or the deprivation of critical compounds like oxygen. This research integrates physiologically-based pharmacokinetic (PBPK) model predictions with ACT-R's fatigue module to predict the effects of caffeine on fatigue. Parameter mapping between PBPK model parameters and ACT-R are informed by neurophysiological literature and established mappings between ACT-R modules and brain regions. Predicted caffeine concentrations in the brain are used to modulate a parameter in the fatigue module to explain caffeine's effects on multiple performance metrics.

Keywords: caffeine; fatigue; ACT-R; physiologically based pharmacokinetic modeling; computational modeling

Introduction

Human cognition is intimately tied to the environment. Indeed, there have been decades of research and discovery on how subtle differences in interactive tasks impact cognitive performance (Anderson, 1990; Gray & Boehm-Davis, 2000). Cognitive performance can also be altered through the deliberate or accidental exposure to compounds in the environment, such as pharmaceuticals, nutraceuticals, and toxins. For example, most countries limit alcohol consumption to avoid accidents that stem from alcohol-induced impairments to cognitive processing (Japan's blood alcohol concentration is 0.03; Canada's is 0.08). In the current paper, an approach toward integrating models of physiology with models of cognition to explain and predict the impacts of chemical compounds on cognitive performance is described and evaluated.

Caffeine is one of the most widely used chemical compounds (Barone & Roberts, 1996), and its effects on performance (Aidman et al., 2021) and fatigue mitigation (Lorist & Tops, 2003) are well documented. Fatigue negatively affects many cognitive functions, including attention, memory, learning, and executive function (Jackson & Van Dongen, 2011). A moderate use of caffeine seemingly reverses some of fatigue's negative effects (Bonnet & Arand, 2012), but too much caffeine decreases those benefits and increases negative subjective experiences (Kaplan et al., 1997).

Integrating the effects of caffeine and fatigue into cognitive architectures through the integration of models of physiology will provide a broader and more detailed understanding of cognition. Even with the wide variety of studies related to

caffeine, it is difficult to accurately account for the effects of caffeine on cognition (Lorist & Tops, 2003). The accrual and integration of theories into a single framework to better understand cognition is precisely the promise of cognitive architectures (Newell, 1973).

The objective of the presented research was to develop a task-independent mechanism within a cognitive architecture to account for the fatigue mitigating effects of caffeine. In the following sections we review the literature on caffeine, physiological and cognitive modeling, and an earlier approach to integrating the modeling approaches. We then present observed data of fatigue mitigation through caffeine and present a model that accounts for the data.

Related Literature

Research has been conducted across constituent facets of physio-cognitive modeling. In the following sections we first provide background on our target compound, caffeine. Next, models of physiology are described. Finally, prior attempts to integrate computational models of physiology and cognition are provided.

Caffeine

Caffeine is a widely used stimulant known to provide benefits to cognitive performance (Kamimori et al., 2015). Caffeine and its metabolites (e.g., paraxanthine) act as adenosine antagonists (inhibitors) on two types of receptors: A₁ and A_{2A} (Kaplan et al., 1997). A₁ receptors are distributed throughout the brain, but are most concentrated in the thalamus, cerebral cortex, and hippocampus. A_{2A} receptors are less widely distributed, existing in dopamine rich regions like the striatum, but are more central to caffeine's stimulatory effects on cognition (Lorist & Tops, 2003).

Physiologically-based Pharmacokinetic Models

A physiologically-based pharmacokinetic (PBPK) model is an *in silico* representation of the movement of chemicals in the arterial blood, flowing to each major organ or lumped tissue compartment(s), including the brain. They provide the time-course of compounds via ordinary differential equations to account for absorption, distribution, metabolism, and excretion processes (Pearce, Setzer, Strobe, Sipes, & Wambaugh, 2017). Thus, PBPK models enable predictions

of the amount and time course of a compound in the brain and enable dose-response predictions.

There are three components to PBPK models: 1) species-specific physiological parameters, 2) chemical-specific parameters, and 3) experiment-specific details for the studies to be simulated. Species-specific physiological parameters are the organ weights and blood flow rates for the defined organs in the PBPK model and are derived from the closest like species when not available. Chemical-specific parameters that are unique for each chemical are the tissue solubility (partition coefficient), metabolism of the parent compound, and plasma and tissue binding characteristics.

Atomic Components of Thought–Rational (ACT-R)

ACT-R is a cognitive architecture that specifies how cognitive processes interact to produce cognition and overt behavior (Anderson, 2009). Models developed within ACT-R posit a common set of processes and mechanisms, which are instantiated as a computer simulation. Independent modules operate in parallel and include declarative memory, vision, attention, and motor modules. Procedural memory coordinates the behavior of the architecture through a set of production rules. Production rules follow an “if-then” structure that encodes the conditions under which specific actions are taken.

Prior research in ACT-R has related the striatum and the thalamus to the architecture’s action-selection system (Anderson, 2009). Both of these regions are associated with adenosine receptors, which would suggest ACT-R’s action-selection system is likely to be affected by caffeine.

Including Physiology within Cognitive Modeling

A few previous research efforts have integrated physiological mechanisms into computational cognitive modeling. Some cognitive architectures include physiological constraints from the brain (e.g., spiking neural networks in Spaun; Eliasmith et al., 2012), but the vast majority of architectures tend not to include physiological constraints. A few efforts have integrated simplified aspects of non-brain physiology into computational cognitive models, like Ritter, Kase, Klein, Bennett, and Schoelles (2009) that explored how ACT-R parameters could be varied to explain effects of stress and caffeine. Work by Dancy, Ritter, Berry, and Klein (2015) used a more complete model of human physiology to affect behavior within ACT-R (i.e., HumMod; Hester et al. 2011).

The research of Dancy and colleagues inspired the development of a similar, yet novel, approach to integrating models of physiology with models of cognition. This novel approach provided compound blood concentrations to ACT-R mechanisms through PBPK models. The result was a cognitive model capable of predicting cognitive performance effects of a common volatile organic compound, toluene (Fisher et al., 2017). The present research extends the research by Fisher et al. in three ways: (1) another compound, caffeine, is explored, (2) the mapping of PBPK predictions to ACT-R parameters is informed by neurophysiological literature, and (3) the research is focused on how caffeine mitigates fatigue, and so

ACT-R’s fatigue module is used (Walsh, Gunzelmann, & Van Dongen, 2017).

Observed Data

Sleep deprivation data were collected and analyzed by McIntire, McKinley, Goodyear, and Nelson (2014). All participants were kept awake for 30 hours, and some were given caffeine. A summary of the study and data are provided here (additional details can be found in the original paper).

Thirty active-duty military personnel (22 male) participated in the study and were compensated for their time. Participants were randomly assigned to one of three conditions: transcranial direct current stimulation (tDCS) active stimulation with placebo caffeine, caffeine with sham tDCS, and sham tDCS and placebo caffeine. Data from the active stimulation condition was omitted from the present study. Two participants’ data were excluded from data presented here as those two were non-compliant.

The psychomotor vigilance task (PVT) was used to assess alertness. On each trial, digits were presented that show the number of milliseconds since the stimulus was presented. Each trial lasts for a minute, or until the participant responds by pressing a button. The interstimulus interval varied randomly from 2 to 12 s. The total task duration was 10 minutes.

Participants were instructed to sleep for at least 7 hours for the two nights prior to the study. Participants awoke at 6 a.m. and were awake for 30 continuous hours. One session of PVT was administered every two hours starting at 6 p.m. Participants in the caffeine condition received 200mg of caffeine chewing gum at 3:15 a.m. Participants in the control group received gum without caffeine.

All data were normalized to 2 a.m. values, just prior to caffeine administration. McIntire et al. (2014) found a significant difference in mean response times, and a marginal difference in lapses, between caffeinated and control participants (see Figure 1, solid lines). No mention is made of false starts, but Figure 1 shows little to no difference in false starts. Lapses are responses that occur 500 ms after stimulus presentation or later. False starts are responses that occur 150 ms after stimulus presentation or earlier. Both are common PVT metrics used in the sleep literature to understand the effects of fatigue (Lim & Dinges, 2008).

Model

In this section, the constituent parts of the model are discussed. The ACT-R model is described first, followed by the PBPK model. Finally, variants of the model are discussed along with the strengths and weaknesses of each.

ACT-R Model

This modeling builds on previous research that integrates ACT-R with biomathematical models (BMM) of fatigue (Walsh et al., 2017), and PBPK models (Fisher et al., 2017). The initial ACT-R model was identical to that in previous research investigating sleep loss and vigilance with the PVT (Veksler & Gunzelmann, 2018). Initial parameters of the

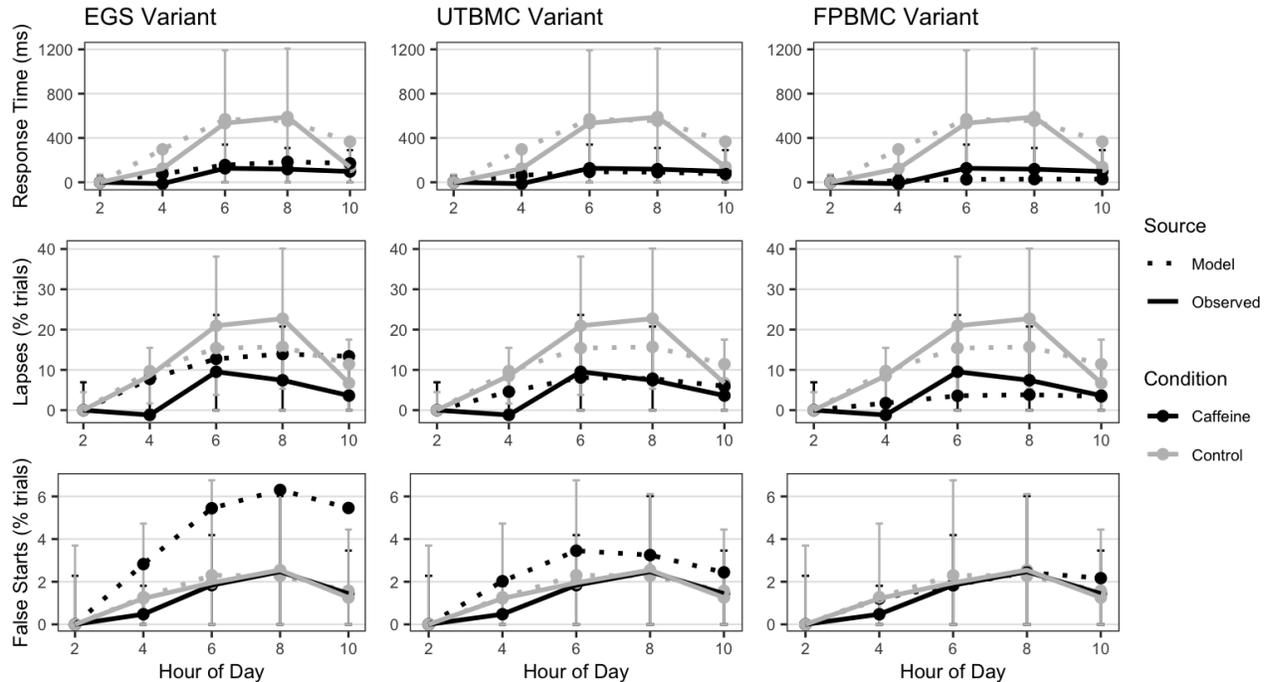


Figure 1: Predictions of the best fitting models for the EGS, UTBMC, and FPBMC variants. Error bars indicate ± 1 standard deviation of participant means.

model were set to the mean of individual participants' parameter estimates from Walsh et al. (2017; Table 5).

The PVT model contains only three productions: *wait*, *attend*, and *respond*. False starts, which are responses before or within 150ms of stimulus onset, can occur due to partial matching between the wait and attend goals. Additional details on the model can be found in Veksler and Gunzelmann (2018).

The fatigue module accounts for the effects of sleep homeostasis and circadian rhythms. The module consists of a theory of *microlapses* and a BMM of fatigue. The BMM predicts alertness levels based on sleep schedule and the time of day. Lower levels of alertness increase the likelihood of microlapses, a brief interruption of cognitive processing. Microlapses affect ACT-R's production utility mechanism by reducing the utility of all productions. The production utility decrement caused by microlapses is determined by the fatigue module's *FPBMC* parameter. Microlapses also impact a fatigue compensatory mechanism that decreases ACT-R's utility threshold. The degree of compensation by this mechanism is determined by the fatigue module's *UTBMC* parameter. Additional details on the fatigue module can be found in Walsh et al. (2017).

The model was initially fit to the control data. Solid lines in Figures 1 and 3, and Table 1, show the fit of this baseline model. The same parameters were varied as in Walsh et al. (2017), and the best fitting values were very near the mean parameters found in that study:

- Initial utility (IU) = 5.1
- Utility threshold (UT) = 4.62
- Production utility noise (EGS) = 0.43
- Default action time (DAT) = 0.04
- Fatigue production utility BMM constant (FPBMC) = 0.025
- Utility threshold BMM constant (UTBMC) = 0.0155
- Fatigue procedural decrement (FPDEC) = 0.99

PBPK Model

The blood pharmacokinetics of caffeine after oral consumption of a 200 mg caffeine gum by a 70 kg individual (i.e. 2.86 mg/kg), was simulated using the R package "high throughput toxicokinetics" (httk; Pearce et al., 2017). In addition to the default tissue compartments in the PBPK model structure selection of the httk platform (lung, G.I. tract, liver, kidney, rest of body), we added a brain and an adipose tissue compartment (fat), in order to address the main pharmacodynamic target tissue for caffeine, that of the central nervous system (CNS) and other peripheral tissue concentrations (fat). This allows mapping of the pharmacokinetics of caffeine in the CNS in addition to the plasma compartment.

The blood plasma concentration time-course from the controlled human pharmacokinetic study of Syed, Kamimori, Kelly, and Eddington (2005) after an acute oral chewing gum dosage is plotted in Figure 2, with the PBPK model results overlaid on the data. An excellent match to the data confirms the ability to predict accurately the concentration of caffeine in plasma after this unique dosing route.

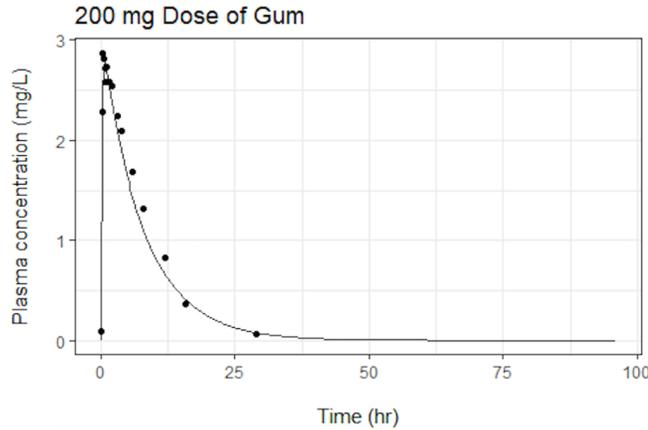


Figure 2: PBPK model results (solid line) of human caffeine plasma (dots) concentration time-course after an acute caffeine dose (200 mg) in chewing gum (n=1 human subject).

Model Variants

Table 1 shows a list of the model variants and their fit to the observed data. Each model variant explored the use of a single parameter affecting production selection. Production selection was the focus of this research for two reasons. First, the effects of alertness in previous modeling of the PVT have been explained with procedural effects (Walsh et al., 2017). Second, pharmacological research has noted that a primary mechanism for stimulation by caffeine is as an antagonist of A_{2A} receptors in regions of the *basal ganglia*, most notably in the *striatopallidal* and *striatonigral* pathways (Fredholm, Bättig, Holmén, Nehlig, & Zvartau, 1999). The striatopallidal pathway in the basal ganglia has been mapped to ACT-R procedural processor, with the striatum more directly linked to production matching and the pallidum more directly linked to production selection (Anderson, 2009).

Production Noise Parameter (EGS). In the fatigue module, as alertness decreases, noise plays a larger role in production selection. Our initial hypothesis was that caffeine may offset some of the effects of noise as alertness decreases.

The left plot in Figure 1 shows the best fitting predictions of the model with EGS varied as a function of caffeine presence. Noise was increased to 0.5 (from the baseline model’s 0.43) in the caffeinated condition. As shown in the first row of Table 1, the fit is not good. An increase in noise increases false starts, just as a decrease in noise (not shown) decreases false starts. In the observed data, the presence of caffeine has no effect on false starts.

Utility threshold compensatory mechanism (UTBMC). The next parameter explored was UTBMC. This parameter determines how utility threshold is affected by the fatigue module. Alertness predictions are scaled by UTBMC and summed with the utility threshold. Changes to UTBMC affects the complex interaction between activation of the

Table 1: Model fits for the baseline model and four variants.

Variant	Mean RT		Lapses		False Starts	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
Baseline	146	.95	5.1	.89	0.3	.80
EGS	68	.88	7.6	.65	3.5	.91
UTBMC	44	.82	3.2	.98	1.3	.76
FPBMC	77	.91	3.8	.79	0.5	.80
with PBPK	55	.93	2.7	.73	0.5	.99

model’s response production, “misfiring” of the model’s response production (due to partial matching), and microlapses.

Previous research has associated ACT-R action selection, of which production utility threshold is a part, with the *pallidum* in the brain (Anderson, 2009). Other research has identified regions rich in dopamine receptors, especially striatopallidal regions, as playing pivotal roles in caffeine’s effect on behavior (Lorist & Tops, 2003). Therefore, caffeine could modulate production utility thresholds, with greater caffeine concentrations making it more likely that a production will fire and therefore less likely that a micro-lapse will occur.

The center plot in Figure 1 shows the best fitting predictions for the model with a UTBMC value of 0.018 when caffeine is present. These predictions are a substantial improvement over the previous mechanisms. Mean response time predictions remain good, and a differentiation of lapses as a function of caffeine presence is predicted. However, there is a slight, but substantial, increase in the number of false starts, which is not present in the observed data.

Fatigue production utility decrement (FPBMC). The final parameter explored was FPBMC. This parameter determines how production utilities are affected by the fatigue module. Alertness predictions are scaled by FPBMC, and then production utilities are scaled by one minus the scaled alertness predictions. A decrease in FPBMC results in higher utilities, and an increase results in lower utilities.

Just as with UTBMC, the literature suggests a link between FPBMC and caffeine effects. Production utilities are as much a part of action selection in ACT-R as utility threshold, and so are also associated with the *pallidum* (Anderson, 2009) and could also be modulated by caffeine (Lorist & Tops, 2003).

The right plot in Figure 1 shows the best fitting predictions for the model with a FPBMC value of 0.02 with caffeine and 0.025 without caffeine (control). As with the UTBMC model, this variant predicts a differentiation of lapses as a function of caffeine. As shown in Table 1, the fits to the false starts is better than with the UTBMC, and the fits to the response times and lapses are comparable to those with the UTBMC model.

Scaled by caffeine predictions. Once we had a good candidate parameter that could account for changes in performance due to caffeine, the fatigue module (Walsh et al., 2017) was modified to allow caffeine concentration predictions to modify the FPBMC parameter similar to the method used by Fisher et al. (2017). The modified FPBMC parameter, $FPBMC_p$, is:

$$FPBMC_p(t) = FPBMC - \beta PBPK(t)$$

where $PBPK$ is the predicted concentration of caffeine in the brain tissue at time t , $FPBMC$ is the fatigue module’s production utility decrement parameter, and β determines the degree to which the $PBPK$ predictions modulate $FPBMC$. The $PBPK$ values used were the mean caffeine concentrations during each task presentation in the McIntire et al. (2014) protocol. The concentrations varied little during the ten minutes of task presentation. The scaling parameter was varied a few times until the ACT-R predictions approximated the observed performance.

Figure 3 and Table 1 show the results for the best fitting β parameter, which was 0.0045. The model continues to do a good job of predicting most of the trends that the static $FPBMC$ model did, with most metrics improving slightly and the R^2 for false starts improving substantially.

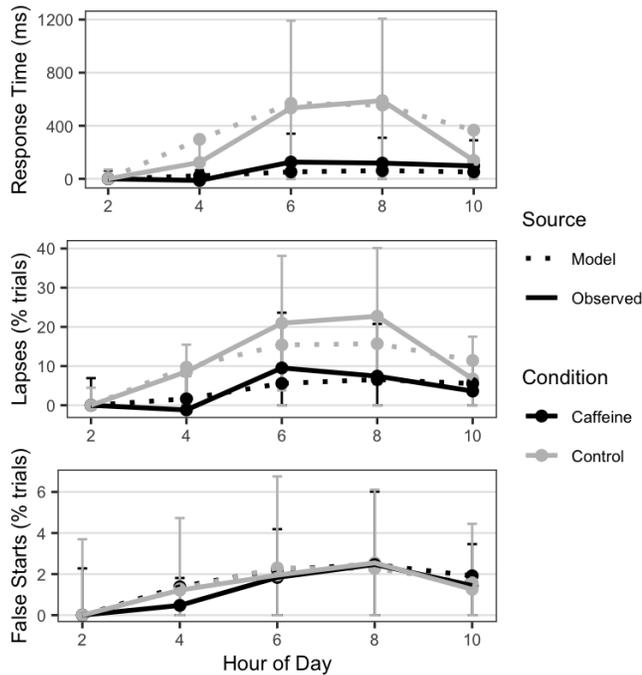


Figure 3: Predictions of the best fitting model varying $FPBMC$ modulated by $PBPK$ predictions. Error bars indicate ± 1 standard deviation of participant means.

Discussion

This work investigated the effects of caffeine on fatigued people. While the effects of caffeine have been studied exten-

sively in psychology and physiology, few formal models have been used to study these effects; exceptions include Ritter et al. (2009) and Ramakrishnan et al. (2016). Ritter et al. investigated the effects of 200 or 400 mg of caffeine on the serial subtraction task without sleep restriction on three ACT-R parameters related to vocalization and memory retrieval (SYL, BLC, and ANS). Ramakrishnan et al. present a mathematical model that predicts human performance on the PVT from a large number of protocols, with different sleep restrictions and caffeine administration. While both models explain the data well, neither model seems to be informed by the underlying physiological processes.

The mapping of physiological to cognitive processes is not trivial. In both formal physiological and cognitive models, there are many variables that could potentially interact to produce behavior (Dancy et al., 2015). In this work, we limited our parameter space to those parameters associated with ACT-R’s action selection mechanism, as the PVT is largely procedural. In addition, the caffeine literature suggests a critical connection between caffeine and action selection, with caffeine affecting A_{2A} receptors concentrated in dopamine rich areas like the basal ganglia (Lorist & Tops, 2003). Still, future research will need to employ other tasks that recruit other cognitive processes, as caffeine has also been shown to affect memory (Loke, 1988) and motor processes (Loke, 1988).

Walsh et al. (2017) integrated a mathematical model of alertness with a theory of microlapses to create the ACT-R fatigue module. The work presented here builds on that to explain how caffeine mitigates the effects of fatigue. The modeling revealed that caffeine may effectively “reverse” some of the decrement in production utility that result from fatigue. This reversal is supported by the physiology literature. Caffeine is an adenosine inhibitor and adenosine plays a role in sleep homeostasis (Landolt, 2008). This inhibition was implemented by scaling the fatigue production utility decrement ($FPBMC$) parameter as a function of predicted caffeine concentration in the brain. This one parameter captured the three substantial trends in the observed data without the need to vary other parameters explored in this research, namely the fatigue module’s compensatory mechanism (UTBMC) and production utility noise (EGS).

This research requires validation of the link function between the $PBPK$ caffeine level predictions and the fatigue module’s $FPBMC$ parameter. While the use of the $PBPK$ model gives us some confidence that our mechanism will account for variations in caffeine, data from additional studies that include multiple administrations and dosages of caffeine will be required. Correspondingly, the mechanism does not currently account for potential negative effects of too much caffeine or individual differences (Kaplan et al., 1997). Future research will include extending the $PBPK$ -to- $FPBMC$ link function to account for known, physiological processes like paraxanthine (a metabolite of caffeine) and adenosine pharmacodynamics.

Conclusion

This research explains the effects of caffeine as a moderation of fatigue's effects on procedural utility. This is done by extending previous research that integrated biomathematical models of alertness (Walsh et al., 2017) and PBPK models (Fisher et al., 2017). Utilizing physiologically-valid predictions of compound levels in the brain, such as caffeine, to vary parameters of cognitive modules mapped to relevant neural mechanisms has the potential to increase the fidelity and accuracy of cognitive models of human performance.

Acknowledgments

This research was supported by the Air Force Research Laboratory Line of Effort on Physiocognitive Modeling.

References

- Aidman, E., Balin, M., Johnson, K., Jackson, S., Paech, G. M., Pajcin, M., ... Banks, S. (2021). Caffeine may disrupt the impact of real-time drowsiness on cognitive performance: a double-blind, placebo-controlled small-sample study. *Scientific Reports*, *11*(1), 1–10.
- Anderson, J. R. (1990). *The adaptive character of thought*. Lawrence Erlbaum Associates.
- Anderson, J. R. (2009). *How can the human mind occur in the physical universe?* Oxford University Press.
- Barone, J. J., & Roberts, H. R. (1996). Caffeine consumption. *Food and Chemical Toxicology*, *34*(1), 119–129.
- Bonnet, M. H., & Arand, D. L. (2012). Utility of caffeine: evidence from the laboratory. In N. J. Wesensten (Ed.), *Sleep deprivation, stimulant medications, and cognition* (pp. 82–92). Cambridge University Press.
- Dancy, C. L., Ritter, F. E., Berry, K. A., & Klein, L. C. (2015). Using a cognitive architecture with a physiological substrate to represent effects of a psychological stressor on cognition. *Computational and Mathematical Organization Theory*, *21*(1), 90–114.
- Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, C., & Rasmussen, D. (2012). A large-scale model of the functioning brain. *Science*, *338*(6111), 1202–1205.
- Fisher, C. R., Myers, C., Hassan, R., Stevens, C., Hack, C. E., Gearhart, J. M., & Gunzelmann, G. (2017). A Cognitive-Pharmacokinetic Computational Model of the Effect of Toluene on Performance. In *Proceedings of the annual meeting of the cognitive science society* (pp. 355–360).
- Fredholm, B. B., Bättig, K., Holmén, J., Nehlig, A., & Zvartau, E. E. (1999). Actions of caffeine in the brain with special reference to factors that contribute to its widespread use. *Pharmacological Reviews*, *51*(1), 83–133.
- Gray, W. D., & Boehm-Davis, D. a. (2000). Milliseconds matter: an introduction to microstrategies and to their use in describing and predicting interactive behavior. *Journal of experimental psychology. Applied*, *6*(4), 322–35.
- Hester, R. L., Brown, A. J., Husband, L., Iliescu, R., Pruett, D., Summers, R., & Coleman, T. G. (2011). HumMod: A modeling environment for the simulation of integrative human physiology. *Frontiers in Physiology*, *2*, 12.
- Jackson, M. L., & Van Dongen, H. P. A. (2011). Cognitive effects of sleepiness. In M. J. Thorpy & M. Billiard (Eds.), *Sleepiness: Causes, consequences and treatment* (pp. 72–81). Cambridge University Press.
- Kamimori, G. H., McLellan, T. M., Tate, C. M., Voss, D. M., Niro, P., & Lieberman, H. R. (2015). Caffeine improves reaction time, vigilance and logical reasoning during extended periods with restricted opportunities for sleep. *Psychopharmacology*, *232*(12), 2031–2042.
- Kaplan, G. B., Greenblatt, D. J., Ehrenberg, B. L., Goddard, J. E., Cotreau, M. M., Harmatz, J. S., & Shader, R. I. (1997). Dose-dependent pharmacokinetics and psychomotor effects of caffeine in humans. *Journal of Clinical Pharmacology*, *37*(8), 693–703.
- Landolt, H.-P. (2008). Sleep homeostasis: A role for adenosine in humans? *Biochemical Pharmacology*, *75*(11), 2070–2079.
- Lim, J., & Dinges, D. F. (2008). Sleep deprivation and vigilant attention. *Annals of the New York Academy of Sciences*, *1129*, 305–322. doi: 10.1196/annals.1417.002
- Loke, W. H. (1988). Effects of caffeine on mood and memory. *Physiology & Behavior*, *44*(3), 367–372.
- Lorist, M. M., & Tops, M. (2003). Caffeine, fatigue, and cognition. *Brain and Cognition*, *53*(1), 82–94.
- McIntire, L. K., McKinley, R. A., Goodyear, C., & Nelson, J. (2014). A Comparison of the Effects of Transcranial Direct Current Stimulation and Caffeine on Vigilance and Cognitive Performance During Extended Wakefulness. *Brain Stimulation*, *7*(4), 499–507.
- Newell, A. (1973). You Can't Play 20 Questions with Nature and Win: Projective Comments on the Papers of this Symposium. In W. G. Chase (Ed.), *Visual information processing* (pp. 283–308). Academic Press.
- Pearce, R. G., Setzer, R. W., Strope, C. L., Sipes, N. S., & Wambaugh, J. F. (2017). htk : R Package for High-Throughput Toxicokinetics. *Journal of Statistical Software*, *79*(4).
- Ramakrishnan, S., Wesensten, N. J., Kamimori, G. H., Moon, J. E., Balkin, T. J., & Reifman, J. (2016). A unified model of performance for predicting the effects of sleep and caffeine. *Sleep*, *39*(10), 1827–1841.
- Ritter, F. E., Kase, S. E., Klein, L. C., Bennett, J., & Schoelles, M. J. (2009). Fitting a model to behavior tells us what changes cognitively when under stress and with caffeine. In *2009 AAAI fall symposium series* (pp. 109–115).
- Syed, S. A., Kamimori, G. H., Kelly, W., & Eddington, N. D. (2005). Multiple dose pharmacokinetics of caffeine administered in chewing gum to normal healthy volunteers. *Biopharmaceutics & Drug Disposition*, *26*(9), 403–409.
- Veksler, B. Z., & Gunzelmann, G. (2018). Functional Equivalence of Sleep Loss and Time on Task Effects in Sustained Attention. *Cognitive Science*, *42*(2), 600–632.
- Walsh, M. M., Gunzelmann, G., & Van Dongen, H. P. A.

(2017). Computational cognitive modeling of the temporal dynamics of fatigue from sleep loss. *Psychonomic Bulletin & Review*, 24(6), 1785–1807.

On Disjunctions and the Weak Completion Semantics

Islam Hamada (islamhamada222@gmail.com)

Technische Universität Dresden, Fakultät Informatik, 01062 Dresden, Germany

Steffen Hölldobler (sh43@posteo.de)

Technische Universität Dresden, Fakultät Informatik, 01062 Dresden, Germany
and North Caucasus Federal University, Stavropol, Russian Federation ¹

Abstract

The weak completion semantics is a three-valued, non-monotonic theory which has been shown to adequately model various cognitive reasoning tasks. In this paper we extend the weak completion semantics to model disjunctions and exclusive disjunctions. Such disjunctions are encoded by integrity constraints and skeptical abduction is applied to compute logical consequences. We discuss various examples and relate the approach to the elimination of disjunctions in the calculus of natural deduction.

Keywords: logic programming; human reasoning; disjunctions; weak completion semantics; natural deduction.

Introduction

Logic programs are universal in that they can compute any computable function (Sebelik & Stepanek, 1982). Hence, if high-level cognitive processes like reasoning are computable, then – from a functional point of view – they can be computed by logic programs. Although originally developed in the context of classical binary logic, modern logic programming approaches are richer: they cover non-monotonicity (Clark, 1978), they can be learned (Muggleton, 1992), they can be interpreted over multi-valued logics (Fitting, 1985), and they can be mapped onto artificial neural or connectionist networks (Hölldobler & Kalinke, 1994).

The *weak completion semantics* (WCS) is a logic programming approach to model human reasoning. Based on ideas originally developed by Stenning and van Lambalgen (2008), it is a three-valued, non-monotonic theory which is knowledge-rich, can handle inconsistent background knowledge, and has been shown to adequately model the average case in various human reasoning tasks like the suppression task (Dietz, Hölldobler, & Ragni, 2012), human syllogistic reasoning (Oliviera da Costa, Dietz Saldanha, Hölldobler, & Ragni, 2017), and human conditional reasoning (Cramer, Hölldobler, & Ragni, 2021). Thus, the WCS offers solutions for the five fundamental problems attributed to the classical binary logic approach in the psychology of reasoning by Oaksford and Chater (2020).

The WCS differs significantly from *mental logic* (Rips, 1994) and the *mental model theory* (Johnson-Laird, 1983). Mental logic is based on syntactic rules which are valid in classical binary logic. However, as pointed out by López-Astorga (2015) they are not a complete system like, for ex-

ample, the calculus of natural deduction of Gentzen (1935); some problematic rules like the *introduction of disjunction* are omitted and other rules like the *introduction of implication* have certain restrictions. The WCS generates models much like the mental model theory. But, whereas the models in the mental model theory are classical binary models, the models in WCS are ternary. Moreover, the computation of WCS's models is rigorously defined by means of a fixed point construction and skeptical abduction.

The semantics of logic programs is usually defined model-theoretically, fixpoint-theoretically, and operationally. This applies to WCS as well. It has an operational semantics given by implementations in PROLOG, PYTHON, and ASP. Hölldobler and Kencana Ramli (2009) have shown that each program and its weak completion² admits a least model under the three-valued Łukasiewicz (1920) logic, which can be computed as the least fixed point of a semantic operator specified by Stenning and van Lambalgen (2008). In other words, in the WCS a least model is constructed and reasoning is with respect to this model. Moreover, the least model can be computed by a connectionist network (Dietz Saldanha, Hölldobler, Kencana Ramli, & Palacios Medinacelli, 2018).

As an example consider the conditional sentence *if Ella is studying, then she will be hungry*. Following Stenning and van Lambalgen (2008) this sentence is represented by the logic program $\mathcal{P} = \{h \leftarrow s \wedge \neg ab_s, ab_s \leftarrow \perp\}$, where h and s denote that *Ella will be hungry* and *Ella is studying*, respectively, and ab_s is an abnormality predicate. ab_s encapsulates everything that could prevent the conditional from holding and is assumed to be false initially.³ If the conditional sentence is given as first premise, then a three-valued model is constructed, where ab_s is mapped to false and h and s are mapped to unknown. This initial model is represented by $\langle \emptyset, \{ab_s\} \rangle$. If the sentence *Ella is studying* is given as second premise, then the mapping of s is updated to true. Consequently, $s \wedge \neg ab_s$ is true and, hence, h must be true as well. This second model is represented by $\langle \{s, h\}, \{ab_s\} \rangle$. In fact, it is the least model of the weak completion of \mathcal{P} . Reasoning is performed with respect to this model and, hence, we conclude *Ella will be hungry*.

²See next section for a formal definition.

³In this paper the abnormalities will always be false. However, in other applications like the suppression task they are important to model exceptional cases and enabling relations (Dietz et al., 2012).

¹The authors are mentioned in alphabetical order.

The WCS is a rigorously defined formal theory. But is it also a cognitive or psychological theory? Are the models constructed by the WCS mental models in the sense of Craik (1945) or Johnson-Laird (1983)? Is it plausible from a cognitive or psychological point of view that humans construct models in a similar way as the WCS? Is it plausible from a biological point of view that a human brain constructs models in a similar way as the connectionist encoding of WCS?

These are challenging questions. So far, we have approached them by considering human reasoning tasks and experimental data from the literature as well as conducting own experiments. But, the WCS cannot deal with disjunctions like *Ella is studying or Ella is running, or both*, or *Ella is studying or Ella is running, but not both*. It is the aim of this paper to overcome this limitation. To this end, we propose to represent disjunctions by sets of integrity constraints like $\{\perp \leftarrow \neg s \wedge \neg r\}$ or $\{\perp \leftarrow \neg s \wedge \neg r, \perp \leftarrow s \wedge r\}$, where the former represents the disjunction $s \vee r$ and the latter the exclusive disjunction $s \oplus r$. Furthermore, we require that models satisfy the integrity constraints. If this does not hold, then abduction (e.g. (Kakas, Kowalski, & Toni, 1992)) is applied and if several (minimal) explanations can be found, then their consequences are combined skeptically. We will demonstrate that this not only solves various human reasoning tasks involving (exclusive) disjunctions but that there is also a striking similarity to the elimination of disjunctions in the *calculus of natural deduction* as defined by Gentzen (1935). But, the WCS avoids some of the drawbacks of natural deduction like the usage of the logical law *ex falso quodlibet* (or *falsum*). The paper demonstrates that logic programming and the WCS can also model human disjunctive reasoning.

The Weak Completion Semantics

We assume the reader to be familiar with classical binary logic (e.g. (van Dalen, 1997)). Let \top , \perp , and \cup be truth constants denoting *true*, *false*, and *unknown*, respectively. A *literal* is an atom or the negation of an atom. A (*logic*) *program* is a finite set of clauses of the form $B \leftarrow \text{Body}$, where B is an atom and Body is either \top , or \perp , or a finite, non-empty conjunction of literals. Clauses of the form $B \leftarrow \top$, $B \leftarrow \perp$, and $B \leftarrow L_1 \wedge \dots \wedge L_n$ are called *facts*, *assumptions*, and *rules*, respectively, where L_i , $1 \leq i \leq n$, are literals.

In this paper, \mathcal{P} denotes a program. An atom B is *defined* in \mathcal{P} if and only if \mathcal{P} contains a clause of the form $B \leftarrow \text{Body}$. In the program $\mathcal{P} = \{h \leftarrow s \wedge \neg ab_s, ab_s \leftarrow \perp\}$ presented in the introduction the atoms h and ab_s are defined, whereas s is undefined. We restrict our attention to propositional programs although the WCS extends to first-order programs as well (see e.g. (Hölldobler, 2015) and the conclusion).

Consider the following transformation: (1) For all defined atoms B occurring in \mathcal{P} , replace all clauses of the form $B \leftarrow \text{Body}_1, B \leftarrow \text{Body}_2, \dots$ by $B \leftarrow \text{Body}_1 \vee \text{Body}_2 \vee \dots$. (2) Replace all occurrences of \leftarrow by \leftrightarrow . The resulting set of equivalences is called the *weak completion* of \mathcal{P} . It differs from the completion defined by Clark (1978) in that unde-

defined atoms are not mapped to false, but to unknown instead.

As shown by Hölldobler and Kencana Ramli (2009), each weakly completed program admits a least model under the three-valued Łukasiewicz (1920) logic (see Table 1). This model will be denoted by $\mathcal{M}_{\mathcal{P}}$. It can be computed as the least fixed point of a semantic operator introduced by Stenning and van Lambalgen (2008). Let \mathcal{P} be a program and I a three-valued interpretation represented by the pair $\langle I^{\top}, I^{\perp} \rangle$, where I^{\top} and I^{\perp} are the sets of atoms mapped to true and false by I , respectively, and atoms which are not listed in either set are mapped to unknown by I . We define $\Phi_{\mathcal{P}} I = \langle J^{\top}, J^{\perp} \rangle$,⁴ where

$$\begin{aligned} J^{\top} &= \{B \mid \text{there is } B \leftarrow \text{Body} \in \mathcal{P} \text{ and } I \text{Body} = \top\}, \\ J^{\perp} &= \{B \mid \text{there is } B \leftarrow \text{Body} \in \mathcal{P} \text{ and} \\ &\quad \text{for all } B \leftarrow \text{Body} \in \mathcal{P} \text{ we find } I \text{Body} = \perp\}. \end{aligned}$$

Following Kakas et al. (1992) we consider *abductive frameworks* $\langle \mathcal{P}, \mathcal{A}_{\mathcal{P}}, IC, \models_{wcs} \rangle$, where \mathcal{P} is a program,

$$\begin{aligned} \mathcal{A}_{\mathcal{P}} &= \{B \leftarrow \top \mid B \text{ is undefined in } \mathcal{P}\} \\ &\cup \{B \leftarrow \perp \mid B \text{ is undefined in } \mathcal{P}\} \end{aligned}$$

is the *set of abducibles*, IC is a finite set of *integrity constraints* of the form $\perp \leftarrow \text{Body}$, where Body is a non-empty and finite conjunction of literals, and $\mathcal{M}_{\mathcal{P}} \models_{wcs} L$ if and only if $\mathcal{M}_{\mathcal{P}}$ maps the literal L to true. Let O be an *observation*, i.e., a finite set of literals. O is *explainable* in the abductive framework $\langle \mathcal{P}, \mathcal{A}_{\mathcal{P}}, IC, \models_{wcs} \rangle$ if and only if there exists a non-empty $X \subseteq \mathcal{A}_{\mathcal{P}}$ called an *explanation* such that (1) $\mathcal{M}_{\mathcal{P} \cup X} \models_{wcs} L$ for all $L \in O$ and (2) $\mathcal{M}_{\mathcal{P} \cup X}$ satisfies IC . The literal L *follows credulously* from \mathcal{P} and O if and only if there exists an explanation X for O such that $\mathcal{M}_{\mathcal{P} \cup X} \models_{wcs} L$. L *follows skeptically* from \mathcal{P} and O if and only if O can be explained and for all explanations X for O we find $\mathcal{M}_{\mathcal{P} \cup X} \models_{wcs} L$. One should observe that if an observation O cannot be explained, then *nothing follows* credulously as well as skeptically. In case of skeptical consequences this is an application of the *Gricean implicature of an existential statement from a universal one* (Grice, 1975): humans normally do not quantify over things which do not exist.

Given premises, general knowledge, and observations, *reasoning in the WCS* is modeled in five steps:

1. Reasoning towards a program \mathcal{P} following Stenning and van Lambalgen (2008).
2. Weakly completing the program.
3. Computing the least model $\mathcal{M}_{\mathcal{P}}$ of the weak completion of \mathcal{P} under the three-valued Łukasiewicz logic.
4. Reasoning with respect to $\mathcal{M}_{\mathcal{P}}$.
5. If observations cannot be explained or integrity constraints are violated, then applying skeptical abduction.

⁴Whenever we apply a unary operator like $\Phi_{\mathcal{P}}$ to an argument like I , then we omit parenthesis and write $\Phi_{\mathcal{P}} I$ instead.

F	$\neg F$	\wedge				\vee				\leftarrow				\leftrightarrow			
\top	\perp	\top	\top	\perp	\perp	\top	\top	\perp	\perp	\top	\top	\perp	\perp	\top	\top	\perp	\perp
\top	\perp	\top	\top	\perp	\perp	\top	\top	\perp	\perp	\top	\top	\perp	\perp	\top	\top	\perp	\perp
\perp	\top	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp
\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp	\perp

Table 1: The truth tables for the Łukasiewicz logic. One should observe that $\perp \leftarrow \perp = \perp \leftrightarrow \perp = \top$ as shown in the grey cells.

Models and Integrity Constraints

Consider the following scenario: *John is playing basketball. If John is playing basketball, then his shirt is not clean anymore. What follows?* Let b and c denote that *John is playing basketball* and *John's shirt is clean*, respectively. This scenario can be represented by

$$\{b \leftarrow \top, \neg c \leftarrow b \wedge \neg ab_b, ab_b \leftarrow \perp\},$$

where ab_b is an abnormality predicate which is assumed to be false. Unfortunately, this set of formulas is not a program because $\neg c$ is not an atom. However, by introducing the antonym *dirty* and specifying that *if a shirt is not dirty, then it is clean* we can represent the scenario by the program

$$\{b \leftarrow \top, d \leftarrow b \wedge \neg ab_b, ab_b \leftarrow \perp, c \leftarrow \neg d \wedge \neg ab_d, ab_d \leftarrow \perp\},$$

where d denotes that *John's shirt is dirty* and ab_d is another abnormality predicate which is also assumed to be false. Its weak completion is

$$\{b \leftrightarrow \top, d \leftrightarrow b \wedge \neg ab_b, ab_b \leftrightarrow \perp, c \leftrightarrow \neg d \wedge \neg ab_d, ab_d \leftrightarrow \perp\}$$

and admits the least model $\langle \{b, d\}, \{c, ab_b, ab_d\} \rangle$ under Łukasiewicz logic. We conclude that *John's shirt is not clean*. But in order to complete the specification we must add the integrity constraint $\perp \leftarrow d \wedge c$ because *a shirt cannot be clean and dirty at the same time*. Luckily, the least model satisfies the integrity constraint. This idea was applied several times by Oliviera da Costa et al. (2017) in modeling human syllogistic reasoning using the WCS.

Although each weakly completed program admits a least model under Łukasiewicz logic, this model does not have to satisfy integrity constraints. As an example consider a scenario where *a judge has ordered Peter not to come near location ℓ* , which can be expressed by the integrity constraints $\perp \leftarrow \ell$. This integrity constraint is satisfied if and only if ℓ is mapped to \perp . Now suppose that *somebody has seen Peter at the location ℓ* , which can be expressed by the program $\mathcal{P} = \{\ell \leftarrow \top\}$. Weakly completing \mathcal{P} we obtain $\{\ell \leftrightarrow \top\}$ whose least model $\mathcal{M}_{\mathcal{P}} = \langle \{\ell\}, \emptyset \rangle$ maps ℓ to \top . In this case, the conflict can not be resolved. One should observe that interpretations can be partially ordered with respect to the subset relation \subseteq , where for interpretations I and J we define $I = \langle I^{\top}, I^{\perp} \rangle \subseteq \langle J^{\top}, J^{\perp} \rangle = J$ if and only if $I^{\top} \subseteq J^{\top}$ and $I^{\perp} \subseteq J^{\perp}$. The partially ordered set of interpretations for the example discussed in this paragraph is shown in Figure 1.

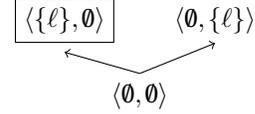


Figure 1: The partially ordered set of interpretations for the weak completion $\{\ell \leftrightarrow \top\}$ of the program $\mathcal{P} = \{\ell \leftarrow \top\}$ and the set $IC = \{\perp \leftarrow \ell\}$, where ℓ is the only atom and each arrow denotes \subseteq . The least model of the weak completion of \mathcal{P} is shown with a black border. All interpretations map the union of $\{\ell \leftrightarrow \top\}$ and IC to either false or unknown; this set of formulas has no model.

In classical binary logic anything follows from a set of formulas for which there is no model. For example, in the *calculus of natural deduction* there is the *falsum* rule which allows to deduce any formula from a contradiction (Gentzen, 1935). It is quite unlikely that humans do this as well and in the *mental model theory* (Johnson-Laird & Byrne, 1991) nothing follows from an empty set of models. Likewise, in the WCS, programs always have a least model and reasoning is with respect to this model. Furthermore, if skeptical abduction is applied, then there must be at least one explanation to conclude a formula from a program and an observation.

Consider a scenario where it is known that *if Linda is in Amsterdam, then she visits her most favorite club*. Ignoring abnormalities for the moment this can be encoded by the program $\mathcal{P} = \{c \leftarrow a\}$, where c and a denote that *Linda visits her most favorite club* and *Linda is in Amsterdam*, respectively. Weakly completing the program yields $\{c \leftrightarrow a\}$, whose least model $\mathcal{M}_{\mathcal{P}}$ is $\langle \emptyset, \emptyset \rangle$. This model does not satisfy the integrity constraint $\perp \leftarrow c$ as c is mapped to unknown. But $\perp \leftarrow c$ is satisfied by the non-least model $\langle \emptyset, \{a, c\} \rangle$ of $\{c \leftrightarrow a\}$. This model can be computed. One should observe that a is undefined in \mathcal{P} and, hence, the set $\mathcal{A}_{\mathcal{P}}$ of abducibles is $\{a \leftarrow \top, a \leftarrow \perp\}$. The empty observation $O = \emptyset$ can be explained by the minimal explanation $X = \{a \leftarrow \perp\}$. Adding X to the program \mathcal{P} and weakly completing the extended program yields $\{c \leftrightarrow a, a \leftrightarrow \perp\}$ whose least model $\mathcal{M}_{\mathcal{P} \cup X}$ is $\langle \emptyset, \{a, c\} \rangle$ (see Figure 2).

As another example consider a scenario discussed by Khemlani, Byrne, and Johnson-Laird (2018), where it is known that *Lisa is in Cambridge or Ben is in Dublin, or both*. This can be encoded by the integrity constraint $\perp \leftarrow \neg c \wedge \neg d$,

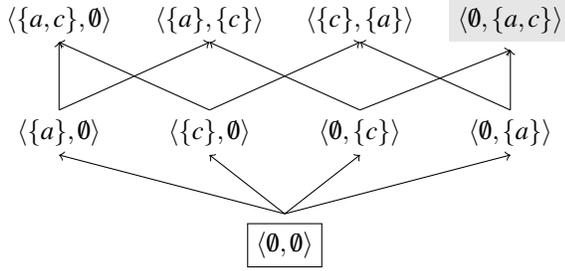


Figure 2: The partially ordered set of interpretations for the weak completion $\{c \leftrightarrow a\}$ of the program $\mathcal{P} = \{c \leftarrow a\}$ and the set $IC = \{\perp \leftarrow c\}$, where a and c are the only atoms. The least model of the weak completion of \mathcal{P} is shown with a black border. The model for the union of the weak completion of $\mathcal{P} \cup \{a \leftarrow \perp\}$ and IC is shown in grey.

where c and d denote that *Lisa is in Cambridge* and *Ben is in Dublin*, respectively. This integrity constraint is satisfied if either c is mapped to true, or d is mapped to true, or both. Now suppose that the program \mathcal{P} is empty, i.e., we know nothing about Lisa and Ben except the above-mentioned disjunction. The empty interpretation $\langle \emptyset, \emptyset \rangle$ is the least model of the weak completion of the empty program and maps c and d to unknown. Thus, the least model violates the integrity constraint. In this case, we find two minimal models $\langle \{c\}, \emptyset \rangle$ and $\langle \{d\}, \emptyset \rangle$ of the weak completion of \mathcal{P} which satisfy the integrity constraint. There are more models satisfying the integrity constraint like $\langle \{c, d\}, \emptyset \rangle$ or $\langle \{c\}, \{d\} \rangle$, but they are larger than at least one of the two minimal models. The minimal models can be computed by considering the empty observation $O = \emptyset$. As neither c nor d are defined in the empty program \mathcal{P} we find $\mathcal{A}_{\mathcal{P}} = \{c \leftarrow \top, c \leftarrow \perp, d \leftarrow \top, d \leftarrow \perp\}$. The sets $\mathcal{X}_c = \{c \leftarrow \top\}$ and $\mathcal{X}_d = \{d \leftarrow \top\}$ are the minimal explanations for O . We obtain $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}_c} = \langle \{c\}, \emptyset \rangle$ and $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}_d} = \langle \{d\}, \emptyset \rangle$. They are the minimal models for the weak completion of the program and the integrity constraint (see Figure 3). One should observe that the non-minimal models can be computed by non-minimal explanations for the empty observation.

The models shown in Figure 3 are the three-valued models for the disjunction $c \vee d$. Moreover, if we want to compute logical consequences then it suffices to consider the minimal models and to compute their skeptical consequences.

Finally, if we modify the last example and consider the exclusive disjunction $c \oplus d$, then we obtain Figure 4. The two models $\{c, \neg d\}$ and $\{d, \neg c\}$ can be computed by explaining the empty observation. They are the three-valued models for the exclusive disjunction $c \oplus d$.

Short Summary

If integrity constraints are not satisfied by the least model of the weak completion of a program, then we try to explain the empty observation. In doing so, we may find no expla-

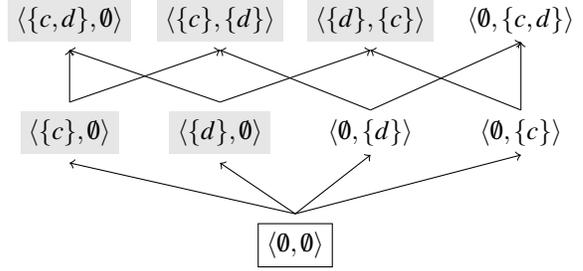


Figure 3: The partially ordered set of interpretations for the weakly completed program $\mathcal{P} = \emptyset$ and the set of integrity constraints $IC = \{\perp \leftarrow \neg c \wedge \neg d\}$, where c and d are the only atoms. The least model of the weak completion of \mathcal{P} is shown with a black border. Models for the union of the weak completion of \mathcal{P} and IC are shown in grey.

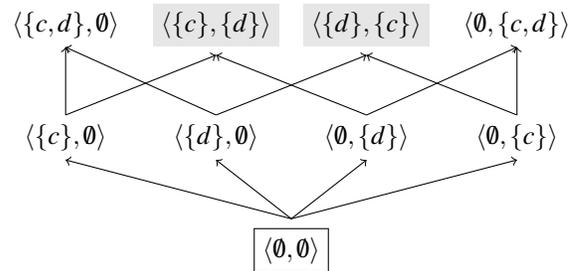


Figure 4: The partially ordered set of interpretations for the weakly completed program $\mathcal{P} = \emptyset$ and the set of integrity constraints $IC = \{\perp \leftarrow \neg c \wedge \neg d, \perp \leftarrow c \wedge d\}$, where c and d are the only atoms. The least model of the weak completion of \mathcal{P} is shown with a black border. Models for the union of the weak completion of \mathcal{P} and IC are shown in grey.

nation, or a single minimal explanation, or several minimal explanations. Each explanation leads to a model of the weak completion of the program and the integrity constraints.

Disjunctions

In this paper we want to extend the WCS by disjunctions. Let \mathcal{P} be a program and \mathcal{D} a set containing disjunctions and exclusive disjunctions of literals. If \mathcal{D} contains the disjunction $L_1 \vee L_2$, where L_1 and L_2 are literals, then this is encoded by the integrity constraint $\perp \leftarrow \neg L_1 \wedge \neg L_2$, where we assume that double negations are eliminated, i.e. $\neg \neg A$ is replaced by the semantically equivalent A for each atom A . If \mathcal{D} contains the exclusive disjunction $L_1 \oplus L_2$, then this is encoded by the integrity constraints $\perp \leftarrow \neg L_1 \wedge \neg L_2$ and $\perp \leftarrow L_1 \wedge L_2$.

Example 1

Consider the following scenario discussed by Johnson-Laird, Byrne, and Schaeken (1992): *Lisa is in Cambridge or Ben is in Dublin, or both. Lisa is not in Cambridge. What follows?* Let c and d denote that *Lisa is in Cambridge* and *Ben is in*

Dublin, respectively. The disjunction is encoded by the set $\mathcal{D} = \{c \vee d\}$ and, hence, by $IC = \{\perp \leftarrow \neg c \wedge \neg d\}$. The negative sentence is represented by the program $\mathcal{P} = \{c \leftarrow \perp\}$. As d is undefined in \mathcal{P} we obtain the set of abducibles $\mathcal{A}_{\mathcal{P}} = \{d \leftarrow \top, d \leftarrow \perp\}$. \mathcal{P} is weakly completed to $\{c \leftrightarrow \perp\}$, whose least model $\mathcal{M}_{\mathcal{P}}$ is $\langle \emptyset, \{c\} \rangle$. This model does not satisfy IC as it maps c to false, d to unknown, and $\neg c \wedge \neg d$ to unknown. But the empty observation can be explained by the minimal explanation $\mathcal{X} = \{d \leftarrow \top\}$. We obtain $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}} = \langle \{d\}, \{c\} \rangle$. We do not need to consider any other explanation as \mathcal{X} is the only minimal one. $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}}$ satisfies IC . We conclude that *Ben is in Dublin and Lisa is not in Cambridge*.

Example 2

Consider a variant of the previous example: *Lisa is in Cambridge or Ben is in Dublin, or both. Lisa is in Cambridge. What follows?* This scenario is modeled as Example 1 except that the program \mathcal{P} is $\{c \leftarrow \top\}$. The program is weakly completed to $\{c \leftrightarrow \top\}$, whose least model $\mathcal{M}_{\mathcal{P}}$ is $\langle \{c\}, \emptyset \rangle$. This model satisfies $IC = \{\perp \leftarrow \neg c \wedge \neg d\}$ as it maps c to true and $\neg c$ as well as $\neg c \wedge \neg d$ to false. $\mathcal{M}_{\mathcal{P}}$ satisfies IC and we conclude that *Lisa is in Cambridge*. But we have no idea where *Ben* is.

Example 3

Consider the following scenario which has also been discussed by Johnson-Laird et al. (1992): *Linda is in Amsterdam or Cathy is in Majorca, but not both. Cathy is not in Majorca. What follows?* Let a and m denote that *Linda is in Amsterdam* and *Cathy is in Majorca*, respectively. The exclusive disjunction is encoded by the set $\mathcal{D} = \{a \oplus m\}$ and, hence, $IC = \{\perp \leftarrow \neg a \wedge \neg m, \perp \leftarrow a \wedge m\}$. The negated sentence is represented by the program $\mathcal{P} = \{m \leftarrow \perp\}$. As a is undefined in \mathcal{P} we obtain the set of abducibles $\mathcal{A}_{\mathcal{P}} = \{a \leftarrow \top, a \leftarrow \perp\}$. Program \mathcal{P} is weakly completed to $\{m \leftrightarrow \perp\}$, whose least model $\mathcal{M}_{\mathcal{P}}$ is $\langle \emptyset, \{m\} \rangle$. This model does not satisfy the first element of IC as it maps a , $\neg a$, and $\neg a \wedge \neg m$ to unknown. But the empty observation can be explained by the minimal explanation $\mathcal{X} = \{a \leftarrow \top\}$ and we obtain $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}} = \langle \{a\}, \{m\} \rangle$. $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}}$ satisfies IC and we conclude *Linda is in Amsterdam and Cathy is not in Majorca*.

Example 4

Consider a variant of the previous example: *Linda is in Amsterdam or Cathy is in Majorca, but not both. Cathy is in Majorca. What follows?* The scenario is modeled as Example 3 except that the program \mathcal{P} is $\{m \leftarrow \top\}$. The program is weakly completed to $\{m \leftrightarrow \top\}$, whose least model $\mathcal{M}_{\mathcal{P}}$ is $\langle \{m\}, \emptyset \rangle$. This model satisfies the first integrity constraint but violates the second one as it maps a and $a \wedge m$ to unknown. But the empty observation can be explained by the minimal explanation $\mathcal{X} = \{a \leftarrow \perp\}$ and we obtain $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}} = \langle \{m\}, \{a\} \rangle$. $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}}$ satisfies IC and we conclude *Cathy is in Majorca and Linda is not in Amsterdam*.

Example 5

Consider the following scenario: *It is late in the afternoon. Ella wants to study or to go running, or both. If she is studying, she will be hungry. If she is running, she will be hungry. What follows?* Let s , r , and h denote that *Ella is studying*, *Ella is running*, and *Ella is hungry*, respectively. The disjunction is encoded by $\mathcal{D} = \{s \vee r\}$ and, hence, $IC = \{\perp \leftarrow \neg s \wedge \neg r\}$. The two conditionals are represented by the program

$$\mathcal{P} = \{h \leftarrow s \wedge \neg ab_s, h \leftarrow r \wedge \neg ab_r, ab_s \leftarrow \perp, ab_r \leftarrow \perp\},$$

where ab_s and ab_r are two abnormality predicates which are assumed to be false. As s and r are undefined in \mathcal{P} we obtain $\mathcal{A}_{\mathcal{P}} = \{s \leftarrow \top, s \leftarrow \perp, r \leftarrow \top, r \leftarrow \perp\}$. The weak completion of \mathcal{P} is $\{h \leftrightarrow (s \wedge \neg ab_s) \vee (r \wedge \neg ab_r), ab_s \leftrightarrow \perp, ab_r \leftrightarrow \perp\}$. Its least model $\mathcal{M}_{\mathcal{P}}$ is $\langle \emptyset, \{ab_s, ab_r\} \rangle$. $\mathcal{M}_{\mathcal{P}}$ does not satisfy IC as it maps s and r to unknown. The empty observation can be explained by $\mathcal{X}_s = \{s \leftarrow \top\}$ and $\mathcal{X}_r = \{r \leftarrow \top\}$. We find $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}_s} = \langle \{s, h\}, \{ab_s, ab_r\} \rangle$ and $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}_r} = \langle \{r, h\}, \{ab_s, ab_r\} \rangle$. Both models satisfy IC . We do not need to consider any other explanation as \mathcal{X}_s and \mathcal{X}_r are the only minimal ones. Comparing $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}_s}$ and $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}_r}$ we skeptically conclude that *Ella will be hungry*, but we can neither skeptically conclude that *Ella is studying* nor that *Ella is running*.

Example 6

We can modify the previous example by assuming that the disjunction is exclusive. Hence, the integrity constraints are $\{\perp \leftarrow \neg s \wedge \neg r, \perp \leftarrow s \wedge r\}$. This will lead to the explanations $\mathcal{X}_1 = \{s \leftarrow \top, r \leftarrow \perp\}$ and $\mathcal{X}_2 = \{r \leftarrow \top, s \leftarrow \perp\}$ and the models $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}_1} = \langle \{s, h\}, \{ab_s, ab_r, r\} \rangle$ and $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}_2} = \langle \{r, h\}, \{ab_s, ab_r, s\} \rangle$. Again, we skeptically conclude that *Ella will be hungry*, but we have no idea whether *she was studying* or *she was running*.

Natural Deduction

Considering Example 5, there is a striking similarity to the elimination of a disjunction in the *calculus of natural deduction*. A disjunction like $s \vee r$ can be eliminated and a formula like h can be derived if (i) h can be derived assuming s and (ii) h can be derived assuming r . (i) and (ii) hold in the given scenario: Assuming s and knowing that $\neg ab_s$ holds, we can derive $s \wedge \neg ab_s$ by the *introduction rule for conjunction*; knowing that $h \leftarrow s \wedge \neg ab_s$ holds we can derive h by the *elimination rule for implication* (see the subproof shown in light gray in Figure 5). Likewise, we can derive h assuming r by utilizing that $\neg ab_r$ and $h \leftarrow r \wedge \neg ab_r$ hold (see the subproof shown in dark gray in Figure 5). The subproof shown in light gray corresponds to the computation of $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}_s}$ by iterating the $\Phi_{\mathcal{P}}$ operator as detailed in the introduction, whereas the subproof shown in dark grey corresponds to the computation of $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}_r}$. The final application of the *elimination rule for implication* corresponds to the computation of skeptical conclusions given $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}_s}$ and $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}_r}$.

$$\frac{s \vee r \quad \frac{\frac{[s] \quad \neg ab_s}{s \wedge \neg ab_s} (\wedge I) \quad h \leftarrow s \wedge \neg ab_s}{h} (\leftarrow E) \quad \frac{\frac{[r] \quad \neg ab_r}{r \wedge \neg ab_r} (\wedge I) \quad h \leftarrow r \wedge \neg ab_r}{h} (\leftarrow E)}{h} (\vee E)$$

Figure 5: A natural deduction proof of $\{s \vee r, h \leftarrow (s \wedge \neg ab_s), h \leftarrow (r \wedge \neg ab_r), \neg ab_s, \neg ab_r\} \vdash h$ in the notation of Hölldobler (2009), where $(\wedge I)$, $(\leftarrow E)$, and $(\vee E)$ denote the rules *introduction of conjunction*, *elimination of implication*, and *elimination of disjunction*, respectively. $[]$ denotes that the enclosed hypothesis has been cancelled.

$$\frac{c \vee d \quad \frac{\frac{[c] \quad \neg c}{\perp} (\neg E) \quad \frac{\perp}{d} (f)}{d} (\vee E) \quad [d]}{d} (\vee E)$$

Figure 6: A natural deduction proof of $\{c \vee d, \neg c\} \vdash d$, where $(\neg E)$ and (f) denote the rules *elimination of negation* and *falsum*, respectively.

A similar observation can be made for the other examples discussed previously. E.g., Figure 6 depicts a natural deduction proof of $\{c \vee d, \neg c\} \vdash d$ corresponding to Example 1. In this proof the *falsum* rule is applied. This rule is problematic in human reasoning as well as in proof search irrespective of whether it is applied top-down or bottom-up. Such an application – and, in fact, the whole subproof shown in light grey – is avoided in the search for explanations for the empty observation in the WCS, where we simply need to consider the possible explanations $\{d \leftarrow \top\}$ and $\{d \leftarrow \perp\}$. This is because we know that c is defined in the program $\mathcal{P} = \{c \leftarrow \perp\}$ and is mapped to \perp by the $\Phi_{\mathcal{P}}$ operator.

Conclusion

In this paper we have extended the WCS by disjunctions. Although we have only discussed examples with a single disjunction, the approach can handle finite sets of disjunctions. Each disjunction is represented by integrity constraints and skeptical abduction is applied to satisfy them. The extension is implemented in PYTHON and ASP. We are unaware of any benchmark sets for disjunctions, but have tested the discussed and related examples.

In a three-valued logic, exclusive disjunctions can also be used to assign true or false to an unknown atom. As an example consider the following puzzle which was published in the German weekly newspaper *DIE ZEIT* on July 26, 2020: *Antonia is looking at Berta while Berta is looking at Cleopatra. Antonia is wearing a red hat, Cleopatra is not wearing a hat, and it is unknown whether Berta is wearing a red hat. Is a person with a red hat looking at a person without a red hat?*

This puzzle can be solved by the techniques presented in this paper. Let a , b , and c denote *Antonia*, *Berta*, and *Cleopatra*, respectively, rX that X is wearing a red hat, $\ell(X, Y)$ that X is looking at Y and $goal$ that somebody with a red hat is looking at a person without a red hat. The scenario can be formalized as a program \mathcal{P} consisting of the following clauses:

$$\begin{aligned}
 \ell(a, b) &\leftarrow \top, \\
 \ell(b, c) &\leftarrow \top, \\
 ra &\leftarrow \top, \\
 rc &\leftarrow \perp, \\
 goal &\leftarrow \ell(X, Y) \wedge rX \wedge \neg rY.
 \end{aligned}$$

This is a first-order program. It can be turned into a propositional one by replacing the variables X and Y consistently by (all combinations of) the constants a , b , and c . The weak completion of this program admits the least model

$$\mathcal{M}_{\mathcal{P}} = \langle \{\ell(a, b), \ell(b, c), ra\}, \{rc\} \rangle.$$

The fact that it is *unknown whether Berta is wearing a red hat* can be represented by the exclusive disjunction $rb \oplus \neg rb$ and, hence, by the set $IC = \{\perp \leftarrow rb \wedge \neg rb\}$ of integrity constraints. This set is not satisfied by $\mathcal{M}_{\mathcal{P}}$ because $\mathcal{M}_{\mathcal{P}}$ maps rb to unknown. As rb is undefined in \mathcal{P} , the set $\mathcal{A}_{\mathcal{P}}$ of abducibles contains $rb \leftarrow \top$ and $rb \leftarrow \perp$. The empty observation can be explained by either $X_1 = \{rb \leftarrow \top\}$ or $X_2 = \{rb \leftarrow \perp\}$. We obtain

$$\begin{aligned}
 \mathcal{M}_{\mathcal{P} \cup X_1} &= \langle \{\ell(a, b), \ell(b, c), ra, rb, goal\}, \{rc\} \rangle, \\
 \mathcal{M}_{\mathcal{P} \cup X_2} &= \langle \{\ell(a, b), \ell(b, c), ra, goal\}, \{rc, rb\} \rangle.
 \end{aligned}$$

Reasoning skeptically we conclude that *goal* is true: *there is a person with a red hat looking at a person without a red hat*. If rb holds, then *Berta* is the person in question; if rb does not hold, then it is *Antonia*. One should observe that reasoning skeptically we can neither conclude that *Berta is wearing a red hat* nor that *Berta is not wearing a red hat*.

We have not yet considered nested disjunctions and disjunctive illusory inferences as discussed by Khemlani and Johnson-Laird (2009). This is one of the next goals in the development of the WCS.

Acknowledgement We like to thank Marco Ragni for many valuable comments and suggestions.

References

- Clark, K. L. (1978). Negation as failure. In H. Gallaire & J. Minker (Eds.), *Logic and databases* (p. 293-322). New York: Plenum.
- Craik, K. J. W. (1945). *The nature of explanation*. Cambridge: Cambridge University Press.
- Cramer, M., Hölldobler, S., & Ragni, M. (2021). When are humans reasoning with modus tollens? In *Proceedings of the 43rd annual conference of the cognitive science society*. (to appear)
- Dietz, E.-A., Hölldobler, S., & Ragni, M. (2012). A computational logic approach to the suppression task. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (p. 1500-1505). Cognitive Science Society.
- Dietz Saldanha, E.-A., Hölldobler, S., Kencana Ramli, C. D. P., & Palacios Medinacelli, L. (2018). A Core method for the weak completion semantics with skeptical abduction. *Journal of Artificial Intelligence Research*, 63, 51-86.
- Fitting, M. (1985). A Kripke–Kleene semantics for logic programs. *Journal of Logic Programming*, 2(4), 295–312.
- Gentzen, G. (1935). Untersuchungen über das logische Schließen. *Mathematische Zeitschrift*, 39, 176-210 und 405-431.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, p. 41-58). Academic Press, New York.
- Hölldobler, S. (2009). *Logik und Logikprogrammierung* (Vol. 1: Grundlagen). Heidelberg: Synchron Publishers GmbH.
- Hölldobler, S. (2015). Weak completion semantics and its applications in human reasoning. In U. Furbach & C. Schon (Eds.), *Bridging 2015 – bridging the gap between human and automated reasoning* (Vol. 1412, p. 2-16). CEUR-WS.org. (<http://ceur-ws.org/Vol-1412/>)
- Hölldobler, S., & Kalinke, Y. (1994). Towards a new massively parallel computational model for logic programming. In *Proceedings of the ECAI94 workshop on combining symbolic and connectionist processing* (p. 68-77).
- Hölldobler, S., & Kencana Ramli, C. D. P. (2009). Logic programs under three-valued Łukasiewicz’s semantics. In P. M. Hill & D. S. Warren (Eds.), *Logic programming* (Vol. 5649, p. 464-478). Springer-Verlag Berlin Heidelberg.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge: Cambridge University Press.
- Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hove and London (UK): Lawrence Erlbaum Associates.
- Johnson-Laird, P. N., Byrne, R. M. J., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review*, 99(3), 418-439.
- Kakas, A. C., Kowalski, R. A., & Toni, F. (1992). Abductive Logic Programming. *Journal of Logic and Computation*, 2(6), 719-770.
- Khemlani, S. S., Byrne, R. M. J., & Johnson-Laird, P. N. (2018). Facts and possibilities: A model-based theory of sentential reasoning. *Cognitive Science*, 1-38.
- Khemlani, S. S., & Johnson-Laird, P. N. (2009). Disjunctive illusory inferences and how to eliminate them. *Memory & Cognition*, 37, 615-623.
- López-Astorga, M. (2015). The formal discipline theory and mental logic. *Praxis Filosófica*, 41, 11-25.
- Łukasiewicz, J. (1920). O logice trójwartościowej. *Ruch Filozoficzny*, 5, 169-171. (English translation: On Three-Valued Logic. In: *Jan Łukasiewicz Selected Works*. (L. Borkowski, ed.), North Holland, 87-88, 1990.)
- Muggleton, S. (1992). *Inductive logic programming*. London: Academic Press.
- Oaksford, M., & Chater, N. (2020). New paradigms in the psychology of reasoning. *Annual Review of Psychology*, 71, 12.1-12.26.
- Oliviera da Costa, A., Dietz Saldanha, E.-A., Hölldobler, S., & Ragni, M. (2017). A computational logic approach to human syllogistic reasoning. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th annual conference of the cognitive science society* (p. 883-888). Austin, TX: Cognitive Science Society.
- Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. MIT Press.
- Sebelik, J., & Stepanek, P. (1982). Horn clause programs for recursive functions. In K. Clark & S.-Å. Tärnlund (Eds.), *Logic programming* (p. 324-340). New York: Academic Press.
- Stenning, K., & van Lambalgen, M. (2008). *Human reasoning and cognitive science*. MIT Press.
- van Dalen, D. (1997). *Logic and structure* (3rd ed.). Berlin: Springer-Verlag.

Structure Learning as a Mechanism of Overharvesting

Nora C. Harhen (nharhen@uci.edu)

Aaron M. Bornstein (aaron.bornstein@uci.edu)

Department of Cognitive Sciences, University of California, Irvine, Irvine, CA 92697
USA

Abstract

In patch leaving problems, foragers must decide between engaging with a currently available, but depleting, patch of resources or foregoing it to search for another, potentially better patch. Overharvesting, or staying in the patch longer than what is optimally prescribed, is widely observed in these problems. Most previous explanations for this phenomenon focus on how foragers' mis-estimations of the environment could produce overharvesting. They suggest that if the forager correctly learned the environment's quality, then they would behave according to Marginal Value Theorem (MVT). However, this proposal rests on the assumption that the forager has full knowledge of the environment's structure. Rarely does this occur in the real world. Instead, foragers must learn the structure of their environment. Here, we model foragers as pairing an optimal decision rule with an optimal learning procedure that allows for the possibility of heterogeneously-structured (i.e. multimodal) reward distributions. We then show that this model can appear to produce overharvesting, as measured by the common optimality criterion, when applied to the usual tasks, which employ homogeneous reward distributions. This model accounts for behavior in a previous serial stay/leave task, and generates novel predictions regarding sequential effects that agree with participant behavior. Taken together, these results are consistent with overharvesting reflecting optimality with respect to a different set of conditions than MVT and suggests that MVT's definition of optimality may need to be adjusted to account for behavior in more naturalistic contexts.

Keywords: foraging; structure learning; reinforcement learning; decision-making;

Introduction

Marginal Value Theorem (MVT; Charnov, 1976) provides the optimal decision rule for maximizing reward in patch-foraging tasks: leave the current patch of resources when the estimated reward rate drops below the average reward rate of the global environment. The rule sets aside the question of how the environment is learned - it is assumed that the forager has full knowledge of the environment: overall quality and any structure (states) to the distribution of rewards.

Foragers, including rodents, primates, and humans, however, demonstrate a consistent bias towards staying too long in the current patch, or "overharvesting" it. Several explanations have been proposed. Some accounts accept that the forager has full knowledge of the environment and attribute overharvesting to different biases and goals irrespective of the environment. These include sunk costs (Wikenheiser et al, 2013), impatient time preferences (Kane et al, 2019), and maximizing marginal utility over reward (Constantino & Daw, 2015). However, rarely is a forager fully certain of their

environment. Given this broken assumption, previous work has focused on adapting MVT to include learning of the environment's quality (e.g. average richness) and explored errors in this learning as a potential mechanism of overharvesting. For instance, biased updating of beliefs can explain overharvesting in non-stationary environments (Garrett & Daw, 2020). Uncertainty over environment quality from insufficient experience can also explain patterns of over and under harvesting (Kilpatrick, Davidson, & El Hady, 2021). However, this work suggests that with sufficient experience deviations from MVT optimality should diminish.

In previously proposed mechanisms of overharvesting, less focus has been placed on how the environment's structure (i.e. distribution of rewards) is learned relative to how its quality has been learned. Most prior work assumes that the forager knows the different patch types within the environment (e.g. richer or poorer). However, in naturalistic settings, foragers are not given this information, they must infer it from experience alone. Here, we question this assumption and propose that the structure learning process can itself explain the appearance of overharvesting. We developed a computational model of how foragers might learn the structure of the environment's reward distribution (the number of modes). First, we show in simulation that the model can generate overharvesting in a single patch-type environment. Then, we examine if the model can explain behavior in a previous serial stay/leave decision task. Finally, we test a novel prediction of the model — that harvesting behavior depends on the order of shifts in volatility — and show that human behavior agrees with the model's predictions. Taken together, these results demonstrate a novel mechanism for overharvesting and, more broadly, brings into question whether MVT is the right optimum to compare behavior to as its assumptions fail to meet the conditions of natural environments.

Methods

Model

Structure learning model We apply rational models of categorization (Anderson, 1991; Sanborn, Griffiths, & Navarro, 2006) to capture how foragers learn the latent structure of the environment. The model (adapted from Harhen, Hartley, & Bornstein, 2021) allows for the possibility that patches belong to different categories varying in richness. The num-

ber of patch categories is not pre-specified and is instead inferred from experience. The forager begins with assumptions of how their observations were generated. They assume that: 1. Rewards exponentially deplete with each harvest; 2. Each patch belongs to a category; 3. Each category is characterized by a unique distribution over depletion rates; 4. A new patch is more likely to belong to a popular category (i.e. many category members); 5. There is some probability that a new patch will belong to a new, previously unobserved category.

Foragers combine these prior beliefs with observed data (depletion rates) to generate new beliefs. The forager then compares the expected reward from harvesting the current patch, v_{stay} , to a reference point, v_{leave} . v_{stay} is estimated as the last received reward multiplied by the expected depletion rate. Having categorized the patch, the forager can better predict the upcoming depletion rate.

MVT's reference point averages overall all previous patches as it assumes homogeneous reward distribution. Our model allows for the possibility that the environment is heterogeneous (e.g. has multiple patch types or multiple modes), so the reward rate of one patch may not be predictive of all other patches' reward rate. Consequently, our model's reference point uses patch experiences integrated over a much shorter time-scale. The reference point for the current patch depends only on the reward rate of the last encountered patch of a different type/category.

Generative model. At trial t , c_{t-1} reflects the assignment of all patches up until the current trial. Each new patch can be assigned to an existing category or a new category. The prior probability of it belonging to an existing category, k , is proportional to the number of patches already assigned to that category, N_k , at trial t . The prior probability of it belonging to a new category is proportional to the parameter α which reflects how densely distributed patches are across categories.

$$P(c_t = k | c_{t-1}) = \begin{cases} \frac{N_k}{t-1+\alpha} & \text{if } k \text{ is an old cluster} \\ \frac{\alpha}{t-1+\alpha} & \text{if } k \text{ is a new cluster} \end{cases} \quad (1)$$

Each category is associated with its own normal distribution over depletion rates, parameterized by μ_c and σ_c^2 . When a patch is assigned to an existing category, depletion rates observed in that patch update the category-specific distribution. *Inference model.* Given a set of observed depletion rates up to trial t , D_t , the forager's posterior beliefs over patch assignments to categories, c_t , are described by:

$$P(c_t | D_t) = \frac{P(D_t | c_t) P(c_t)}{p(D_t)} \quad (2)$$

Exact computation of the posterior is computationally intractable, so we use particle filtering as an approximate inference algorithm (Gershman, Niv, & Blei, 2010; Sanborn, Griffiths, & Navarro, 2006). The posterior is represented with a set of m particles. Each particle represents a possible assignment of patches to categories. Some particles will have the same category assignments. The posterior probability of a

category assignment is proportional to the number of particles within the set which contain that assignment. To approximate the posterior distribution, we can average over the particles:

$$P(c_t | D_t) \approx \frac{1}{M} \sum_{m=1}^M \delta(c_t - c_t^{(m)}) \quad (3)$$

where $\delta(\cdot)$ is 1 when its input is 0, and 0 otherwise.

We can then approximate the prior distribution over category assignments for $t+1$ trials with

$$\begin{aligned} P(c_{t+1} | D_t) &= \sum_{c_t} P(c_{t+1} | c_t) P(c_t | D_t) \\ &\approx \sum_{c_t} P(c_{t+1} | c_t) \frac{1}{M} \sum_{m=1}^M \delta(c_t - c_t^{(m)}) \\ &= \frac{1}{M} \sum_{m=1}^M P(c_{t+1} | c_{t+1}^{(m)}) \end{aligned} \quad (4)$$

We can then approximate the posterior for trial $t+1$ with:

$$\begin{aligned} P(c_{t+1} | D_{t+1}) &\propto \sum_{c_t} P(d_{t+1} | c_{t+1}, D_t) P(c_{t+1} | D_t) \\ &\approx \frac{1}{M} \sum_{m=1}^M P(d_{t+1} | c_{t+1}, D_t) P(c_{t+1} | c_{t+1}^{(m)}) \end{aligned} \quad (5)$$

M samples are drawn from this distribution to create a new particle set. 50 particles were used for all simulations. An intermediate number of particles allows for fairly accurate prediction while being psychologically plausible and capable of capturing the variability and order sensitivity people display (Sanborn, Griffiths, & Navarro, 2006).

Prediction To predict how much the harvest will deplete next, possible depletion rates are sampled from the forager's inferred generative model of the environment, and these samples are averaged over. Depletion rates are sampled in the following way: first, a category is drawn with probability proportional to its posterior probability, and then, a depletion rate is drawn from the category-specific normal distribution over depletion rates. In our simulations, we used 1000 samples.

Single Patch Type Learning model Patches are assumed to all belong to the same category. This is equivalent to setting alpha to 0. This should generate behavior similar to what MVT would produce, with the additional ability to account for the variance of observed rewards.

Making a choice

To make a decision, the forager compares the value of staying with the value of leaving. The value of staying, v_{stay} , is the reward received from the last harvest multiplied by the predicted depletion rate. The value of leaving, v_{leave} , is calculated as the average reward rate in the last visited patch of a different category multiplied by the time that would be spent harvesting it. This serves as a more dynamic, shorter time scale reference point than MVT's.

Experiment 1: Simulating the structure learning model in single patch type environments

We propose that overharvesting may emerge from inferring more structure in the environment than what is actually present. In particular, inferring that the environment has multiple patch types when it is, in fact, a single highly variable patch type. Simulated agents visited patches to harvest for resources. They decided between harvesting the current exponentially depleting patch or spending more time to travel to a new, unharvested patch (harvest time = 3.5, travel time = 15.5 sec). Depletion rates were drawn from a Beta distribution parameterized by $a = 1.5$ and $b = 1.5$. The mean depletion rate was 0.5 with a SD of 0.25.

Experiment 2: Reanalysis of Constantino & Daw (2015)

We fit our model to data from Constantino & Daw (2015). Participants harvested trees for apples. After each harvest, they could decide between harvesting again or traveling to a new tree and incurring a time delay. The number of apples gained per harvest depleted exponentially. Participants foraged in four different environments that varied in their mean depletion rate and travel time delay. These two features controlled the overall richness of the environment (i.e. higher depletion rate \rightarrow richer, shorter travel time \rightarrow richer). Critically, in this experiment, participants were told when (though not how) the environment changed.

Experiment 3: Novel task

Participants We recruited 82 participants from Amazon Mechanical Turk (ages 23 - 63, Mean= 38, SD=10). Participation was restricted to workers who had completed at least 100 prior studies with at least 99% approval rate. Participants were paid \$6 as a base payment and a bonus contingent on performance (\$0-\$4). We excluded 7 participants for failing at least one catch trial or having average patch residence time 2 standard deviations above or below the group mean.

Procedure We adapted from Harhen et al (2021) a novel variant of the Constantino & Daw (2015). The task was framed as a space mining game where participants were told to collect as many space gems as possible. On each trial, participants had to decide if they wanted to continue digging for gems on the current planet or travel to a new planet. If they stayed and harvested, they watched a short animation of an astronaut digging and then the reward would be displayed. If they chose to travel to a new planet, they watched an animation of a flying rocket ship and then an image of a trial-unique alien was displayed for 5 seconds. Participants had 2 seconds to make their choice. If they did not make a decision, they had to wait another 2 seconds before making another choice. To ensure participants' reaction times did not affect their reward rate, the reaction time (RT) was subtracted from the ensuing dig or travel animation display time.

Participants completed 6 blocks lasting 5 minutes each. Blocks varied in the spread of depletion rates experienced.

Depletion rates in highly volatile blocks were sampled from a Beta distribution with parameters $a = 1$ and $b = 1$. The mean depletion rate was 0.5 with a SD of 0.29. Depletion rates in the least volatile blocks were sampled from a Beta distribution with $a = 20$ and $b = 20$ (Mean = 0.5, SD = 0.078). In the medium volatility block, depletion rates were sampled from a Beta distribution parameterized by $a = 4$ and $b = 4$ (Mean = 0.5, SD = 0.17). Participants were told when a new block began, but were not told if and how it changed. Participants were placed in one of two conditions that differed in the order of blocks encountered. In the high early volatility condition the first two blocks were the most volatile, and the third and fourth blocks were the least volatile. In the other condition (low early volatility), the order of the blocks first four blocks were reversed. In both conditions, the last two blocks were of medium volatility. By matching the last two blocks on volatility, we were able to directly compare behavior between the conditions.

Model fitting procedure

The MVT learning model's free parameters were beta (softmax temperature), c (stay/leave bias), alpha (learning rate), and ρ_0 (initial global reward rate). For both the tasks, the free parameters for the structure learning model were the prior over cluster dispersion (alpha), and prior over environment richness. For the Constantino & Daw (2015) task, participant data was characterized by the mean patch residence time (PRT) in each of the blocks. For the novel variant of the task, participant data was characterized by the mean patch residence time (PRT) relative to MVT optimal in each of the blocks. We compared this to the same measures predicted by the model. The loss for a parameter set was calculated as the sum of squared error between the participant's data and the model's simulated data averaged across 10 simulations. 500 sets of parameters were sampled from a Sobol Sequence, and the set of parameters that produced the lowest sum of squared error was chosen. Generating candidate parameter sets from a Sobol Sequence rather than a grid, can provide superior fits, particularly, when there are more than two parameters (Bergstra & Bengio, 2012).

Model comparison

To compare models, we used cross validation. We held out one test block and then fit the model using the PRTs for the remaining blocks. The model error was then measured by taking the absolute difference between the model prediction for the held-out block and the participant's measure for that block. We repeated this procedure for every possible combination of fit blocks and test block and then averaged over the errors to compute the cross validation score.

Results

Experiment 1: Simulating the structure learning model in single patch type environments

We first simulated variants of the model which differed in whether they allowed for the possibility of multiple patch

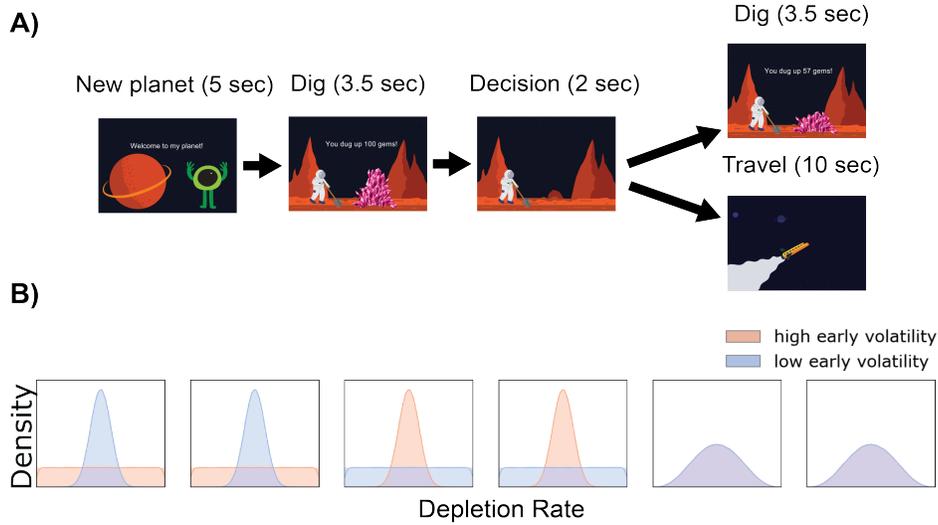


Figure 1: **Task Designs.** **A.** Participants sequentially decide whether to dig or travel to a new planet. **B. Novel task volatility structure** The experiment is broken up into six blocks. Blocks differ in the distribution from which depletion rates are sampled. Some have high variance, others have low variance, and some fall in between. The two conditions, high early volatility and low early volatility, differ only in the order of blocks.

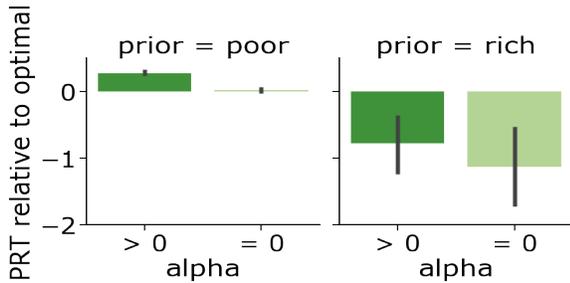


Figure 2: **Results from Experiment 1** Overharvesting and under harvesting behavior depends on both the prior over environment complexity, alpha, and the prior over environment richness. Error bars are 95% CI.

types in highly variable a single patch environment. When allowing for multiple patch types to be inferred ($\alpha > 0$), simulated agents did so. Consistent with our prediction, rational behavior that began with this mismatch between environment type and assumptions resulted in overharvesting when the environment was initially believed to be poor (Figure 2). However, when the true environment structure was assumed (single patch type, $\alpha = 0$), behavior was MVT-optimal. Underharvesting behavior emerged from an initial belief that the environment was rich regardless of assumptions about the environment’s structure.

Experiment 2: Reanalysis of Constantino & Daw (2015)

Constantino & Daw found that Marginal Value Theorem (MVT) with an error-driven learning rule better explained

participants’ data than a temporal-difference learning model. The MVT learning model had four free parameters: learning rate (α), softmax temperature (β), initial global reward rate (ρ_0), and stay-leave bias (c). c captured an individual’s bias to stay in the current patch. We reasoned that this bias parameter would be instrumental in capturing behavior that deviated from MVT optimality. To test this hypothesis, we fit the data with the MVT model with and without c . We found that c was indeed critical to capturing participant’s overharvesting behavior (Figure 3, $t(24) = -6.04$, $p < 0.0001$). Given the importance of this parameter, how does this bias emerge?

When comparing both the MVT and the structure learning model with a stay/leave bias, neither was superior to the other ($t(24) = -1.23$, $p = 0.23$). However, when comparing the MVT and structure learning model without the stay/leave bias, the structure learning model was superior (Figure 3, Table 1, $t(24) = 3.63$, $p = 0.001$). Taken together, these results suggest that the (nonstandard) stay/leave bias added to the MVT model in Constantino & Daw (2015) — added to place it on par with the temporal-difference learning model used as an alternative hypothesis in that study — was a primary factor in the fit quality of that model, perhaps due to the fact that the long blocks in that experiment allowed learning to reach a steady state. Here, we show that the optimal structure learning procedure can account for much if not all of the variance that this parameter adds, while rooting the behavior in a principled, rational learning approach.

Experiment 3: Novel Task

We next tested whether human behavior reflected a novel prediction of the structure learning model, namely sensitivity to the order in which patch volatility is experienced (Figure 1b).

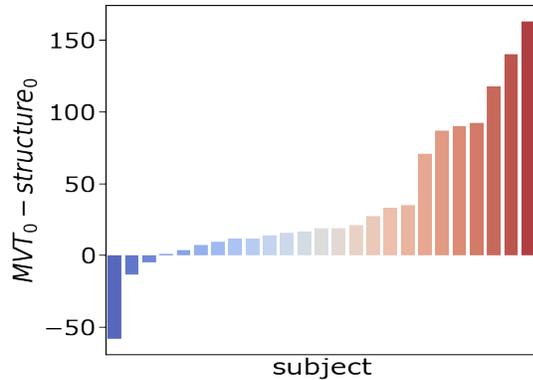


Figure 3: **Results from Experiment 2.** Each bar reflects the difference in cross-validation scores between the MVT learning model without c , the stay-leave bias parameter, and the structure learning model, also without a c , for an individual participant. Positive values indicate the structure learning model provides a better fit to the participant’s data. Overall, 22 out of 25 participants were better fit by the structure learning model than by MVT.

Model predictions & participant behavior Our model predicts that the order of volatility shifts in the environment will affect how patch categories are inferred and consequently, stay/leave decisions. When prior beliefs about environment structure and/or richness do not align with experience in the environment, the model infers more patches than there really are, leading to overharvesting. The pattern of experience in an initially predictable environment discourages inferring multiple patch types such that there is less of an influence of a prior bias towards complexity on foraging behavior.

Across the population, participants in both conditions overharvested roughly equally ($t(73)=0.21$ $p = 0.83$). We next examined the fit parameters to identify heterogeneity the population. Most participants were better fit with alpha as a free parameter (Figure 4a, $t(74) = 2.90$, $p = 0.004$). Participants in both conditions had a similar range of fit alpha parameters (Figure 4b, $t(73) = -0.73$, $p = 0.47$). However, matching the simulation results in Experiment 1, the inferred prior over environment richness differed between conditions. Participants in the high early volatility condition had lower prior estimates of environment richness or quality (Figure 4c, $t(73) = -3.30$, $p = 0.001$). Participants were split into high and low parameter groups (alpha and prior over environment richness) based the median value of the parameter. Behavior was simulated with the structure learning model using the best fitting parameters for each participant. We then the compared model generated behavior between the two groups. There were no differences in behavior between the high and low alpha group in either condition (Figure 4d, high early volatility - $t(36) = 0.32$, $p = 0.75$; low early volatility - $t(34) = 0.54$, $p = 0.59$). However, when splitting by prior over environment richness,

those in the low group overharvested more than those in the high group in both conditions (Figure 4e, high early volatility - $t(35) = -3.82$, $p < 0.001$; low early volatility - $t(35) = -4.105$, $p < 0.001$).

Discussion

We asked if the process of learning the environment’s structure could explain overharvesting behavior in certain contexts. To address this question, we developed a computational model of how foragers could learn environment structure and leverage it during decision making. First, in simulation, we showed that allowing the possibility of inferring multiple patch types results in overharvesting in highly variable single patch type environments. Next, we showed that our structure learning model could capture behavior in a previously collected stay/leave task. In this prior work, a model with error-driven learning of environment quality and a MVT decision rule was found to replicate participant’s behavior. However, its success in fitting the data critically depended on a stay-leave bias parameter to account for overharvesting. Our model, on the other hand, provided a superior fit to participants’ overharvesting relative to the MVT model without a stay-leave bias parameter. A possibility is that some of the variance explained by the stay-leave bias parameter emerged from the learning process formalized in our model. Finally, we tested a novel prediction of the structure learning model. Namely, that participant responses should be sensitive to the order of shifts in volatility. Participant behavior was consistent with this prediction, providing further evidence in favor of the model.

Taken together, these results suggest that seemingly sub-optimal behavior like overharvesting can be explained with statistically optimal learning of environment structure and a prior expectation of heterogeneous environments. This is consistent with previous work demonstrating that people will infer structure or observe non-existent patterns even when there is no incentive to do so (Yu & Cohen, 2009) and even when it’s disadvantageous (Collins & Frank, 2013; Gaissmaier & Schooler, 2008). This prior bias towards structure possibly emerges from it being frequently incentivized in the real world.

Potentially, MVT’s definition of optimality may need to be expanded. In particular, foraging has been suggested to provide a decision context that we were evolutionarily adapted to and consequently, likely to yield normative behavior. However, MVT assumes an environment that does not concord with naturalistic environments which tend to be heterogeneous, non-stationary, and exhibit multiple scales of spatio-temporal regularities. Prior work demonstrates that foragers do consider this multi-scale information in adapting their search strategies in naturalistic settings (Fagan et al., 2013). Future work could explore extending the model to include multiple scales of reference points — one integrating over a longer time scale like MVT and another integrating over a shorter time scale as presented here. The present work and

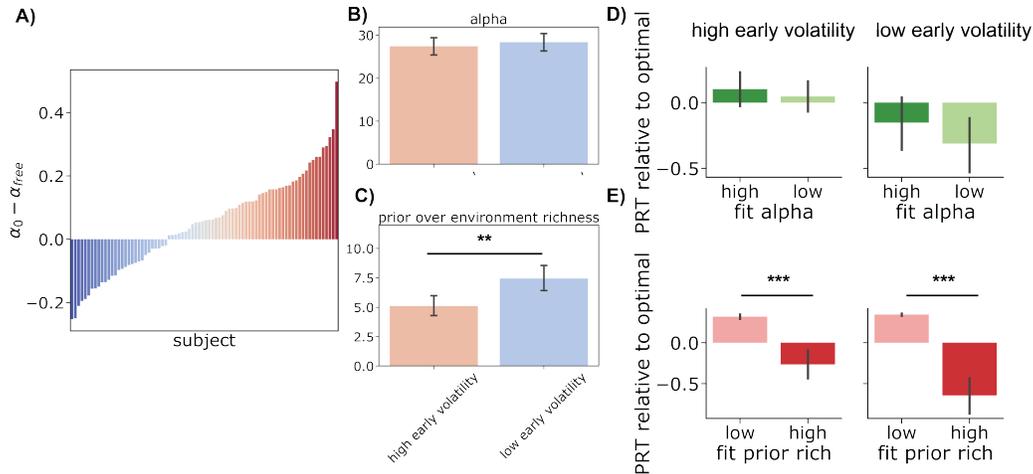


Figure 4: **Results from Experiment 3.** **A.** Each bar reflects the difference in cross-validation scores between the structure learning model with alpha fixed at 0 and the same model when alpha is a free parameter. Positive values indicate the structure learning model with free alpha provides a better fit to the participant’s data. Overall, 49 out of 75 participants were better fit by the structure learning model with free alpha than the alpha fixed at 0 model. **B-C.** Participants’ fit parameters for the structure learning model. **D-E.** Model simulated overharvesting/underharvesting behavior separated by a median split on fit parameters from the structure learning model. Error bars are 95% CI.

potential future work could suggest optimality in foraging may need to be redefined to incorporate dealing with the multiple scales of uncertainty that natural environments present foragers with.

Acknowledgements

This work was supported by NIMH P50MH096889 to AMB. NCH was supported by a National Defense Science and Engineering Graduate fellowship. The authors thank Sara M Constantino for providing data for Experiment 2.

References

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychol. Rev.*, *98*(3), 409–429.

Bergstra, J., & Bengio, Y. (2012). *Random search for hyperparameter optimization.*

Charnov, E. L. (1976, April). Optimal foraging, the marginal value theorem. *Theor. Popul. Biol.*, *9*(2), 129–136.

Collins, A. G. E., & Frank, M. J. (2013, January). Cognitive control over learning: creating, clustering, and generalizing task-set structure. *Psychol. Rev.*, *120*(1), 190–229.

Constantino, S. M., & Daw, N. D. (2015, December). Learning the opportunity cost of time in a patch-foraging task. *Cogn. Affect. Behav. Neurosci.*, *15*(4), 837–853.

Fagan, W. F., Lewis, M. A., Auger-Méthé, M., Avgar, T., Benhamou, S., Breed, G., ... Mueller, T. (2013). Spatial memory and animal movement. *Ecol. Lett.*, *16*(10), 1316–1329.

Gaissmaier, W., & Schooler, L. J. (2008, December). The smart potential behind probability matching. *Cognition*, *109*(3), 416–422.

Garrett, N., & Daw, N. D. (2020, July). Biased belief updating and suboptimal choice in foraging decisions. *Nat. Commun.*, *11*(1), 3417.

Gershman, S. J., Blei, D. M., & Niv, Y. (2010, January). Context, learning, and extinction. *Psychol. Rev.*, *117*(1), 197–209.

Harhen, N. C., Hartley, C. A., & Bornstein, A. M. (2021). Model-based foraging using latent-cause inference. *Proceedings of the 43rd Annual Conference of the Cognitive Science Society*, to appear.

Kane, G. A., Bornstein, A. M., Shenhav, A., Wilson, R. C., Daw, N. D., & Cohen, J. D. (2019, September). Rats exhibit similar biases in foraging and intertemporal choice tasks. *Elife*, *8*.

Kilpatrick, Z. P., Davidson, J. D., & El Hady, A. (2021, April). *Uncertainty drives deviations in normative foraging decision strategies.*

Sanborn, A. N., Griffiths, T. L., & Navarro, D. J. (2006, January). A more rational model of categorization.

Wikenheiser, A. M., Stephens, D. W., & Redish, A. D. (2013, May). Subjective costs drive overly patient foraging strategies in rats on an intertemporal foraging task. *Proc. Natl. Acad. Sci. U. S. A.*, *110*(20), 8308–8313.

Yu, A. J., & Cohen, J. D. (2008). Sequential effects: Superstition or rational behavior? *Adv. Neural Inf. Process. Syst.*, *21*, 1873–1880.

Measuring and Modelling How People Learn How to Plan and How People Adapt Their Planning Strategies to the Structure of the Environment

Ruiqi He (ruiqi.he@tuebingen.mpg.de)

Max Planck Institute for Intelligent Systems, Tübingen, Germany

Yash Raj Jain (yasshain@gmail.com)

Max Planck Institute for Intelligent Systems, Tübingen, Germany

Falk Lieder (falk.lieder@tuebingen.mpg.de)

Max Planck Institute for Intelligent Systems, Tübingen, Germany

Abstract

Often we find ourselves in unknown situations where we have to make a decision based on reasoning upon experiences. However, it is still unclear how people choose which pieces of information to take into account to achieve well-informed decisions. Answering this question requires an understanding of human metacognitive learning, that is how do people learn how to think. In this study, we focus on a special kind of metacognitive learning, namely how people learn how to plan and how their mechanisms of metacognitive learning adapt the planning strategies to the structures of the environment. We first measured people's adaptation to different environments via a process-tracing paradigm that externalises planning. Then we introduced and fitted novel metacognitive reinforcement learning algorithms to model the underlying learning mechanisms, which enabled us insights into the learning behaviour. Model-based analysis suggested two sources of maladaptation: no learning and reluctance to explore new alternatives.

Keywords: decision-making; planning; metacognitive learning; reinforcement learning; cognitive modelling

Introduction

In real life, we often have to make decisions in new situations. Often our decisions and actions result from learned experiences and reasoning upon them. However, it is still unknown how we learn which pieces of information we should take into account to efficiently make a well-informed decision. Answering this question requires understanding how people learn how to think (metacognitive learning). While direct decision-making has been studied extensively from the perspective of cognitive science (Wang & Ruhe, 2007) and machine learning (Niv, 2009), our contemporary understanding of how people learn how to decide remains shallow. There is some work on modelling how people learn to select between the decision-making strategies they already know (Lieder & Griffiths, 2017; Rieskamp & Otto, 2006; Erev & Barron, 2005) but there is little work on how people discover those decision strategies in the first place. In this study, we focus on a special kind of metacognitive learning, namely how people learn how to plan.

Our work is structured in two parts - measuring and then modelling metacognitive learning in terms of reinforcement learning algorithms. For this, we set up an experiment that utilises a process-tracing paradigm that makes planning observable. The resulting process-tracing data is then analysed by a recently developed computational method for inferring people's planning strategies and their changes over time. To

model how people learn how to plan, we formalised and tested three competing hypotheses about how people learn how to plan using three novel computational models. We tested our models against each other. The resulting best model was used to draw conclusions for different groups of participants.

By advancing our understanding of human metacognitive learning, this line of work may contribute to laying the foundations for improving metacognitive learning and helping people overcome maladaptive ways of decision-making.

Background

Mouselab MDP paradigm

A major obstacle to studying metacognitive learning is that we cannot directly observe people's cognitive strategies and how they change over time. To overcome this hurdle, we utilise a process-tracing paradigm that renders people's behaviour highly diagnostic of their planning strategies, namely the Mouselab Markov Decision Process (MDP) paradigm (Callaway, Lieder, Krueger, & Griffiths, 2017). In this paradigm, participants plan the route of a spider through a maze with the goal to maximise their score (see Figure 1) with the given number of trials. The score is the sum of the values of the nodes (the gray circle) on the path they choose to traverse. Each node harbours a gain or a loss, which are initially hidden but can be revealed by clicking on it. This explicit clicking action corresponds to evaluating the quality of a potential future state, which is a fundamental cognitive operation in planning. The cognitive cost of this operation is externalised by charging a fee of -1 for each node they reveal. Participants are thus encouraged to not immediately click every location, but instead, reveal information as necessary. In this way, the paradigm externalises the mental representation that people use for planning in terms of which nodes have been clicked and what their revealed values are.

Measuring metacognitive learning

The Mouselab-MDP paradigm can be used to measure the changes in people's strategy sequence. For this, we have previously developed a computational method that infers which planning strategy a participant used on each trial based on their clicks (Jain, Callaway, & Lieder, 2019; Jain et al., 2021). This method returns which of 79 predefined planning strategies a participant is most likely to have

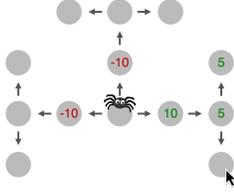


Figure 1: Example of the Mouselab paradigm for the constant variance condition with five nodes revealed

used. For detailed documentation of all 79 strategies please see <https://osf.io/zgshx/>. We can therefore measure metacognitive learning in terms of how the inferred strategy changed from each trial to the next.

Modelling metacognitive learning

To model metacognitive learning we will apply reinforcement learning algorithms to the problem of deciding how to decide (*meta-decision-making*). We will briefly introduce these two frameworks and how they can be combined.

Reinforcement learning Research suggests that human learning is partly driven by rewards and punishments, which they receive through trial and error (Niv, 2009). This learning mechanism has inspired reinforcement learning algorithms, which learn to estimate how much reward can be expected to receive from a certain action (a) in a given state (s). This estimate is updated according to the differences between received and predicted rewards δ :

$$Q(s, a) \leftarrow Q(s, a) - \alpha \cdot \delta \quad (1)$$

where α is the learning rate. To balance exploitation and exploration, the agent can choose its actions *probabilistically*, maximising the predicted action value, for example using the softmax rule (see for example Equation 3).

Meta-decision-making Previous work suggests that the brain is equipped with multiple decision systems that interact in numerous ways (Dolan & Dayan, 2013; Daw, 2018). In contrast to the Pavlovian and model-free systems, the model-based system supports flexible reasoning about which action might be best given available information, goals and preferences. To efficiently balance decision quality and decision time given enormous amount of information, the model-based system's flexibility necessitates a mechanism for selecting only relevant information, that is deciding how to decide, which is formally known as *meta-decision-making* (Boureau, Sokol-Hessner, & Daw, 2015). Recent work has formalised the problem of meta-decision-making as a meta-level MDP (Krueger, Lieder, & Griffiths, 2017; Griffiths et al., 2019):

$$M_{\text{meta}} = (\mathcal{B}, \mathcal{C} \cup \{\perp\}, T_{\text{meta}}, r_{\text{meta}}), \quad (2)$$

where belief states $b_t \in \mathcal{B}$ encode the model-based decision system's beliefs about the values of alternative courses

of actions. The temporal evolution of those belief states (b_1, b_2, \dots) is driven by the decision system's computations c_1, c_2, \dots according to the meta-level transition probabilities $T(b_t, c_t, c_{t+1})$. Finally, the meta-level reward function $r_{\text{meta}}(b_t, c_t)$ encodes the cost of performing the planning operation $c_t \in \mathcal{C}$ and the expected return of terminating planning ($c_t = \perp$) and acting based on the current belief state b_t .

Metacognitive reinforcement learning Planning strategies can be thought of as policies for solving metalevel MDPs. We can therefore formalise the problem of discovering effective planning strategies as solving a metalevel MDP for the optimal metalevel policy (Griffiths et al., 2019). Solving meta-decision-making problems optimally is often computationally intractable but the optimal solution can be approximated through reinforcement learning (Russell & Wefald, 1991; Callaway, Gul, Krueger, Griffiths, & Lieder, 2018). Hence, we assume that the brain approximates optimal meta-decision-making via reinforcement learning mechanisms that seek to approximate the optimal solution of the meta-level MDP defined in Equation 2 by either learning to approximate the optimal policy directly or by learning an approximation to its value function. Previous work has applied this idea to model how people learn to select between alternative cognitive strategies (Erev & Barron, 2005; Rieskamp & Otto, 2006; Lieder & Griffiths, 2017), learn how many steps to plan ahead (Krueger et al., 2017), and learn when to exert how much cognitive control (Lieder, Shenhav, Musslick, & Griffiths, 2018). However, this approach has yet to be applied to investigate how people discover and refine their cognitive strategies.

Experiment

To investigate metacognitive learning, we designed an experiment with three conditions using the Mouselab-MDP paradigm to measure how people adapt their planning strategies to different environments.

Methods

Each participant was randomly allocated to one of three conditions. Each condition presented the participants with a different environment. In the increasing variance environment, the range of possible rewards is larger at locations that are further away from the starting point at the centre of the maze. In the decreasing variance environment, the variance between possible node values decreases the further away from the starting point, that is the nodes that are closest to the centre have the largest range of possible values. In the constant variance environment, the variance between possible node values remains the same. The possible value of each node at any given step can be seen in table 1. Step 1 corresponds to the three nodes that are closest to the starting point in the middle, step 2 is the next node, step 3 is the last set of nodes that are furthest away from the starting point.

Environment	Step 1	Step 2	Step 3
Increasing	-4, -2, 2, 4	-8, -4, 4, 8	-48, -24, 24, 48
Decreasing	-48, -24, 24, 48	-8, -4, 4, 8	-4, -2, 2, 4
Constant	-10, -5, 5, 10	-10, -5, 5, 10	-10, -5, 5, 10

Table 1: Possible reward values for the three environments

Participants We recruited 174 participants, 58 for each condition, on CloudResearch. The recruitment was limited to participants who had completed 100+ HITS, had a score > 90 , and were located in the United States. Each participants received a base-pay of \$1.50 and a bonus up to \$5 based on their performance. They received minimal instructions and had to pass a quiz to demonstrate correct comprehension of the setup before starting the first trial.

Procedure Each participant was assigned to one condition was asked to complete 35 trials. The scores are displayed on the screen and are updated after each action (click, move). Planning is encouraged by a performance-depend bonus, which is 0.2 cents for each point of their final score after completion of all trials.

Results

To investigate whether people learn to adapt their planning strategies to the structure of the environment, the strategy sequences were analysed. To classify our participants’ planning strategies into adaptive and maladaptive ones, we first created a list of planning strategies that were used by the participants and then determined the expected score of the strategies in the list using computer simulations. For each environment, the five strategies with the highest score are labelled as adaptive, while the five low scoring strategies are labelled as maladaptive strategies. We illustrated (see Figure 2) and tested the aggregated proportion of the five adaptive and five maladaptive strategies for trends using Mann Kendall tests. The tests confirmed an increasing trend for the aggregated proportion of adaptive strategies in all environments ($S > 367, p < .001$ in all environments). In addition, the tests suggest a decreasing trend for the maladaptive strategies in the increasing ($S = -429, p < .001$) and decreasing variance environments ($S = -295, p < .001$) and no trend in the constant variance environment ($S = -83, p = .176$). This means that in all three environments the proportion of people who adopted using adaptive strategies gradually increased while the proportion using maladaptive strategies gradually decreased in all but one environment. Furthermore, for each of those five adaptive and five maladaptive strategies, we tested whether the proportion of people using that strategy increased or decreased across trials using a series of Mann Kendall tests (see <https://osf.io/zgshx/> for detailed results of the tests). Overall, the tests suggested an increasing trend or no trend for the adaptive strategies, while the data indicated decreasing trend or no trend for the maladaptive strategies (see Figure 3). For instance, for the increasing variance condition, we found that the frequency of the adaptive strategy to search the

final destinations for the best possible outcome (Strategy 21) steadily increased (Mann Kendall test: $S = 535, p < .001$), while the frequency of the maladaptive strategy to act without planning (Strategy 30) steadily decreased (Mann Kendall test: $S = -414, p < .001$).

These results suggest that people discover and learn to use adaptive strategies in all three environments. The effect is most prominent in the increasing variance condition and least prominent in the constant variance condition. This might be because discovering adaptive strategies is easiest when the environment has a clear structure that adaptive strategies can exploit.

Modelling metacognitive learning

Having empirically demonstrated that people discover and learn to use adaptive planning strategies, we now model the underlying computational mechanisms in terms of metacognitive reinforcement learning using two novel models: Learned value of computation (LVOC), direct adjustment of decision-making policy via gradient ascent (REINFORCE) and its non-learning variant, which postulates that people do not update their planning strategy. Each of these three models hypothesise a different learning mechanism.

Models of metacognitive reinforcement learning

Representations of the strategies The strategies people use in the Mouselab-MDP can be described in terms of a weighted combination of neuroscience-informed features. One example of a group of features are pruning features, which are related to assigning a negative value to thinking about a path whose expected value is below a certain threshold. Therefore, the learning trajectory can be expressed by how the weights of those features evolve over time. We have defined 52 different features (see <https://osf.io/zgshx/> for a detailed description).

The REINFORCE model The REINFORCE model assumes that people adjust their planning strategy directly by following its performance gradient ascent through the strategy space using a softmax policy (Williams, 1992):

$$\pi_{\theta}(c|b) = \frac{\exp\left(\frac{1}{\tau} \cdot \sum_{k=1}^{52} \theta_k \cdot f_k(b, c)\right)}{\sum_{c \in C_b} \exp\left(\frac{1}{\tau} \cdot \sum_{k=1}^{52} \theta_k \cdot f_k(b, c)\right)} \quad (3)$$

where b is the belief state, c is the click being considered and C_b is the set of clicks available in the belief state b . τ is the inverse temperature parameter and f_k are the neuroscience-informed features values described above. The larger the value of τ is, the more deterministically the highest value action is chosen. The parameters of the policy (θ) are updated once after each trial according to the learning rule:

$$\theta \leftarrow \theta + \alpha \cdot \sum_{t=1}^O \gamma^{t-1} \cdot r_{\text{meta}}(b_t, c_t) \cdot \nabla_{\theta} \ln \pi_{\theta}(c_t|b_t) \quad (4)$$

where α is the learning rate, γ is the discount factor and O is the number of planning operations the model performed on

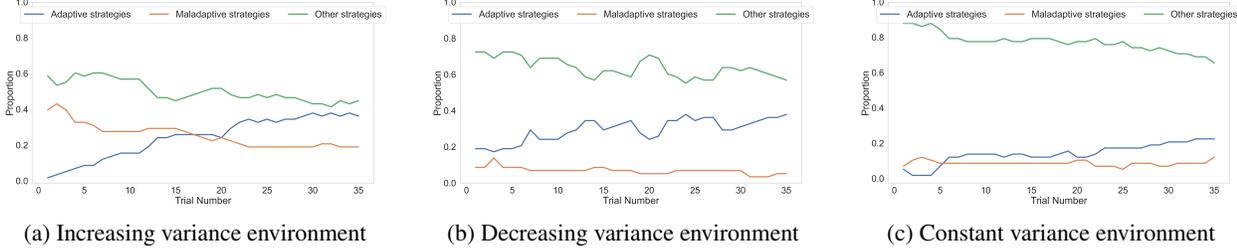


Figure 2: Proportion of aggregated strategy development throughout the trials for each environment.

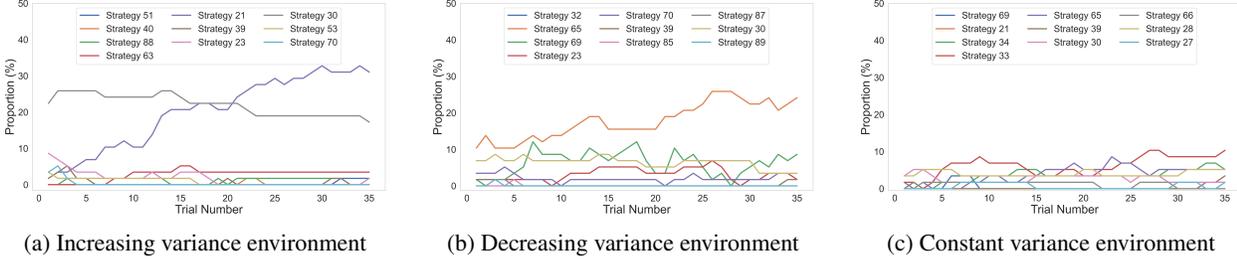


Figure 3: Trial-wise proportion of adaptive and maladaptive strategy for each environment. The first 5 strategies are adaptive strategies, the last 5 strategies are maladaptive strategies

that trial, that is the number of clicks plus one, which represents the termination of planning operation. The learning rate α was optimised using ADAM (Kingma & Ba, 2014). Both α and γ are treated as free parameter that are fit separately for each participant. In addition to the vanilla REINFORCE, a pseudo-reward (Ng, Harada, & Russell, 1999) is used to speed up learning. The value of the pseudo-reward on performing computation c_t in belief state b_t and transitioning to belief state b_{t+1} is given by

$$\text{PR}(b_t, c, b_{t+1}) = \mathbb{E}[R\pi_{b_{t+1}} | b_{t+1}] - \mathbb{E}[R\pi_{b_t} | b_{t+1}] \quad (5)$$

which is the difference between the expected value of the best path in belief state b_{t+1} according to the policy $\pi_{b_{t+1}}$ and the expected value of the best path in belief state b_{t+1} according to the policy π_{b_t} .

The LVOC model According to the LVOC model, people discover and change their strategy continuously by learning to predict the values of alternative planning operations (Krueger et al., 2017). The model assumes that people learn a linear approximation to the meta-level Q-function:

$$Q_{\text{meta}}(b_t, c_t) \approx \sum_{k=1}^{52} w_k \cdot f_k(b_t, c_t), \quad (6)$$

using the f_k and corresponding weights w_k . The LVOC model learns the weights w_k of those features by Bayesian linear regression of the bootstrap estimate $\hat{Q}(b_t, c_t)$ of the meta-level value function onto the features f_k . The bootstrap estimate

$$\hat{Q}(b_t, c_t) = r_{\text{meta}}(b_t, c_t) + \sum_{k=1}^{52} \mu_{k,h} \cdot f_k(b_{t+1}, c_{t+1}) \quad (7)$$

is the sum of the immediate meta-level reward and the predicted value of the next belief state b_{t+1} under the current meta-level policy. The predicted value of b_{t+1} is the scalar product of the posterior mean $\mu_{k,h}$ of the weights w_k , given the observations from all h preceding planning operations and the features $f_k(b_{t+1}, c_{t+1})$ of b_{t+1} and the cognitive operation c_{t+1} that the current policy selects given state. Given the posterior on the feature weights $\mathbf{w} = (w_1, \dots, w_{52})$, the next planning operation c_{t+1} is selected by a generalised version of Thompson sampling. That means, to make the k^{th} meta-decision, n weight vectors $\tilde{w}^{(1)}, \dots, \tilde{w}^{(n)}$ are sampled from the posterior distribution of the weights given the series of meta-level states, selected planning operations, and resulting value estimates experienced so far. That is,

$$\tilde{w}_t^{(1)}, \dots, \tilde{w}_t^{(n)} \sim P(\mathbf{w} | \mathcal{E}_t), \quad (8)$$

where the set $\mathcal{E}_k = \{e_1, \dots, e_t\}$ contains the meta-decision-maker's experience from the first t meta-decisions. To be precise, each meta-level experience $e_j \in \mathcal{E}_k$ is a tuple $(b_j, h_j, \hat{Q}(b_j, c_j; \mu_j))$ containing a meta-level state, the computation selected in it, and the bootstrap estimates of its Q-value. The arithmetic mean of the sampled weight vectors $\tilde{w}^{(1)}, \dots, \tilde{w}^{(n)}$ is then used to predict the Q-values of each possible planning operation $c \in \mathcal{C}$ according to Equation 6. The planning operation with the highest predicted Q-value is used for decision-making. For a fair comparison, the LVOC model also utilises the metacognitive pseudo rewards defined in Equation 5. The LVOC model has three free parameters: the mean vector $\boldsymbol{\mu}_{\text{prior}}$ and variance σ_{prior}^2 of its prior $\mathcal{N}(\mathbf{w}; \boldsymbol{\mu}_{\text{prior}}, \sigma^2 \cdot \mathbf{I})$ on the weights \mathbf{w} and the number of samples n .

Model fitting methods

To assess how well each model can capture how people learn how to plan, we fitted each model’s free parameters and priors on feature weights to the participant’s data and applied each model to the series of problems the participant had to solve.

The parameters of the models were fit by maximising a Multivariate-Normal pseudo-likelihood function defined in terms of the probability that the model would generate the participant’s trial wise scores as a function of its parameters. For a given participant i , the pseudo-likelihood function under model m is given by:

$$\mathcal{L}((\theta_{i,m}, \sigma_{i,m}) | \mathbf{r}_i) = \phi(\mathbf{r}_i; \hat{\mathbf{r}}_{i,m}(\theta), \sigma_{i,m} I) \quad (9)$$

where $\theta_{i,m}$ is the parameter vector used to fit the data from participant i with model m , \mathbf{r}_i is the vector of scores that the i^{th} participant obtained on trials 1 through 35, σ is the standard deviation of the errors between the observed scores and the model’s predictions $\hat{\mathbf{r}}_{i,m}(\theta_{i,m})$, and $\phi(\mathbf{x}; \mu, \Sigma)$ is the density function of the multivariate normal distribution. We estimate the parameters $\theta_{i,m}$ and $\sigma_{i,m}$ by maximising the pseudo-likelihood function in Equation 9 using Bayesian Optimisation (Bergstra, Yamins, & Cox, 2013). All selected models are then fit to the participant data using 400 iterations. In each iteration, the model’s prediction is estimated by averaging the model’s scores across 30 simulations.

Model selection

After the model-fitting, we performed individual-level and group-level model selection using the Akaike Information Criterion (AIC) (Akaike, 1998). On the level of individual participants, both learning models, LVOC and REINFORCE, seem to explain the learning behaviour reasonably better than the non-learning model (see Table 2). The number of participants whose data was best explained was the same for both learning models (71). To investigate the differences in

Environment	Model	Count
Increasing variance	non-learning	11
	REINFORCE	24
	LVOC	23
Decreasing variance	non-learning	10
	REINFORCE	28
	LVOC	20
Constant variance	non-learning	11
	REINFORCE	19
	LVOC	28

Table 2: Count of individual participants’ best fitted model.

which model explains a participant’s data best, we divided the participants into three groups: participants who were not using adaptive strategies in the beginning but learned to do so were classified as *highly adaptive learners*, participants using maladaptive strategies in the last trial were classified as *maladaptive participants*, and the other participants are labelled as *moderately adaptive participants*. The group-level model comparison provided strong evidence in favour of the REINFORCE model (average AIC = 308.31) over the LVOC

model (average AIC = 315.94) and over its non-learning variant (average AIC = 341.43). As shown in Figure 4, the REINFORCE model was able to capture how the participants’ performance throughout the experiment in all three conditions. Most importantly, the REINFORCE model was able to capture the improvement in people’s performance as they adapt their planning strategies to the structure of the increasing variance environment (Figure 4a).

Increasing (n=58)	non-learning	REINFORCE	LVOC
Highly adaptive (n=21)	387.44	343.54	346.97
Maladaptive (n=11)	184.42	174.68	205.36
Mod. adaptive (n=26)	375.56	341.55	351.25
Decreasing (n=58)	non-learning	REINFORCE	LVOC
Highly adaptive (n=16)	369.84	326.66	320.53
Maladaptive (n=3)	202.95	198.94	197.86
Mod. adaptive (n=39)	370.27	306.39	324.88
Constant (n=58)	non-learning	REINFORCE	LVOC
Highly adaptive (n=11)	349.30	330.33	334.64
Maladaptive (n=7)	326.66	316.72	309.72
Mod. adaptive (n=40)	307.08	290.15	294.28

Table 3: Averaged AIC for each model grouped by participants. Best performance is marked in bold.

Model-based analysis

Due to its superior performance, REINFORCE was chosen to perform model-based analysis to gain insights into the learning behaviour and more specifically how they differ among groups of participants.

We hypothesised that maladaptive participants would have lower learning rates than the other two groups and tested our hypothesis using Wilcoxon rank-sum tests on the fitted learning rates. In addition, exploratory Wilcoxon rank-sum tests were conducted on the other parameters γ , which quantifies the influence of immediate meta-level rewards as opposed to the reward received later during the trial, and τ , which describes to which extent the participant explores different strategies (see Table 4). For the increasing variance environment, the tests imply that the distribution of inverse temperature parameters differs significantly between maladaptive ($M = 233.94, SD = 380.68$) and moderately adaptive participants ($M = 37.89, SD = 88.61$). This suggests that maladaptive participants might choose their planning operations more deterministic and thereby perform less cognitive exploration of alternative planning strategies. Participant-level analyses confirmed that 9 out of the 11 maladaptive participants started with a maladaptive strategy and either did not change their strategy or only changed it once. This suggests that the reason why some people find it difficult to steer away from maladaptive decision strategies is that they fail to explore alternative decision strategies. In the decreasing variance environment, the learning rate also differed significantly between maladaptive participants ($M < 0.0001, SD < 0.0001$) and the other two groups (highly adaptive: $M = 0.007, SD = 0.018$; moderately adaptive: $M = 0.009, SD = 0.026$). The small learning rate suggests that maladaptive participants did not learn at all. In the constant variance environment, the significant

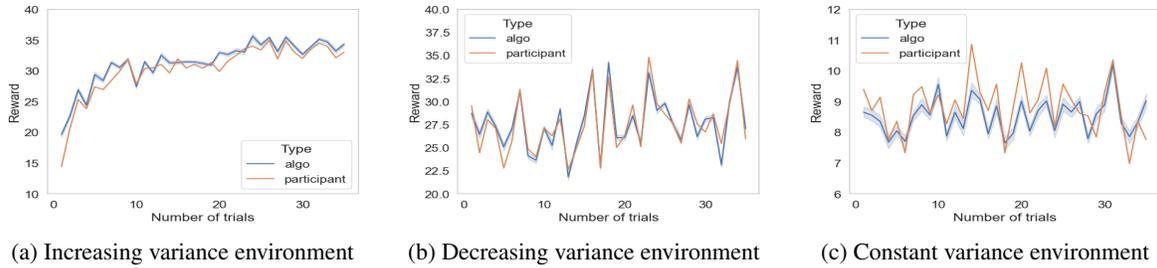


Figure 4: Averaged REINFORCE model performance visualised against participants' performance

difference in inverse temperature implies that highly adaptive learners ($M = 9.12, SD = 26.51$) explore more than the maladaptive ones ($M = 32.46, SD = 54$), which aligns with the adaptive strategy for this environment.

Parameter	Comparison	T	p
Inverse temperature (increasing variance)	Malad. vs. Mod. ad.	2.79	.005
Inverse temperature (constant variance)	Malad. vs. Highly. ad.	2.22	.026
Learning rate (decreasing variance)	Malad. vs. Highly. ad.	-2.01	.044
	Malad. vs. Mod. ad.	-1.98	.048

Table 4: Results of Wilcoxon rank sum test on the fitted parameters

Conclusion and further work

We first measured how people adapt their planning strategies to different environments and then modelled the underlying learning mechanisms in terms of metacognitive reinforcement learning. Using a process-tracing method, we found that participants discovered different types of planning strategies depending on what was adaptive for the environment they were in. Concretely, the proportion of adaptive strategies significantly increased in all environments, while the proportion of maladaptive strategies significantly decreased in almost all environments. After having confirmed that people adapt to all three environments, we proceeded to develop and test two new models of metacognitive reinforcement learning. Our models extend previous models of metacognitive learning (Lieder & Griffiths, 2017; Krueger et al., 2017; Lieder et al., 2018) to the problem of strategy discovery. They achieve this by learning a policy for selecting individual planning operations. In addition, innovation of our models is that they learn not only from external rewards but also from intrinsically generated pseudo-rewards for gaining valuable information. Model selection suggested that the REINFORCE model best describes how people learn how to plan. Our model-based analysis of individual differences in metacognitive learning highlighted two potential causes of maladaptive planning: no learning and reluctance to explore. Further work could look into how to motivate learning and exploration - for example by gamification.

Acknowledgement

We thank Valkyrie Felso for code-related help and Frederic Becker for his support in demonstrating how to run experiments.

References

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected papers of Hirotugu Akaike* (pp. 199–213). Springer.

Bergstra, J., Yamins, D., & Cox, D. (2013). Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In *International conference on machine learning* (pp. 115–123).

Boureau, Y.-L., Sokol-Hessner, P., & Daw, N. D. (2015). Deciding how to decide: Self-control and meta-decision making. *Trends in cognitive sciences*, 19(11), 700–710.

Callaway, F., Gul, S., Krueger, P., Griffiths, T. L., & Lieder, F. (2018). Learning to select computations. In *Uncertainty in artificial intelligence: Proceedings of the thirty-fourth conference*.

Callaway, F., Lieder, F., Krueger, P. M., & Griffiths, T. L. (2017). Mouselab-mdp: A new paradigm for tracing how people plan. In *The 3rd multidisciplinary conference on reinforcement learning and decision making*.

Daw, N. D. (2018). Are we of two minds? *Nature Neuroscience*, 21(11), 1497.

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, 80(2), 312–325.

Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological review*, 112(4), 912.

Griffiths, T. L., Callaway, F., Chang, M. B., Grant, E., Krueger, P. M., & Lieder, F. (2019). Doing more with less: meta-reasoning and meta-learning in humans and machines. *Current Opinion in Behavioral Sciences*, 29, 24–30.

Jain, Y. R., Callaway, F., Griffiths, T. L., Dayan, P., Krueger, P. M., & Lieder, F. (2021). A computational process-tracing method for measuring people's planning strategies and how they change over time.

Jain, Y. R., Callaway, F., & Lieder, F. (2019). Measuring how people learn how to plan. In *Cogsci* (pp. 1956–1962).

- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Krueger, P. M., Lieder, F., & Griffiths, T. (2017). Enhancing metacognitive reinforcement learning using reward structures and feedback. In *Cogsci*.
- Lieder, F., & Griffiths, T. L. (2017). Strategy selection as rational metareasoning. *Psychological Review*, *124*(6), 762–794.
- Lieder, F., Shenhav, A., Musslick, S., & Griffiths, T. L. (2018). Rational metareasoning and the plasticity of cognitive control. *PLoS computational biology*, *14*(4), e1006043.
- Ng, A. Y., Harada, D., & Russell, S. (1999). Policy invariance under reward transformations: Theory and application to reward shaping. In *Icml* (Vol. 99, pp. 278–287).
- Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154.
- Rieskamp, J., & Otto, P. E. (2006). Ssl: a theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*(2), 207.
- Russell, S., & Wefald, E. (1991). Principles of metareasoning. *Artificial intelligence*, *49*(1-3), 361–395.
- Wang, Y., & Ruhe, G. (2007). The cognitive process of decision making. *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, *1*(2), 73–85.
- Williams, R. J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, *8*(3), 229–256.

Utilizing ACT-R to Investigate Interactions between Working Memory and Visuospatial Attention while Driving

Moritz Held^{1,2} (Moritz.held@uol.de), Jelmer Borst² (j.p.borst@rug.nl), Anirudh Unni¹ (anirudh.unni@uol.de), Jochem W. Rieger¹ (jochem.rieger@uol.de)

¹Dept. of Psychology, University of Oldenburg, Germany
²Bernoulli Institute, University of Groningen, the Netherlands

Keywords: ACT-R; driving; working memory

Introduction

Autonomous driving is an area which has seen rapid growth in recent years. A long-held belief in this field is that more automation will equate more safety. However, some researchers continue to challenge this conviction in an argument for adaptive automation (Hancock et al., 2013).

In the context of driving man-machine systems implementing adaptive automation are envisioned to continuously engage the driver in the driving task and at the same time dynamically adapt the task-load depending on the driver’s momentary cognitive ability. A key step towards this approach is to continuously monitor the driver’s mental state and predict when the automation system should take more responsibilities and when to give them back to prevent drivers from mentally disengaging in the driving task.

Predicting mental workload has been done in recent studies using neuroimaging data (e.g., fNIRS; Unni et al., 2017; Scheunemann et al., 2019) but has come with limitations as different types of cognitive workload were interacting instead of adding at the brain level, which led to a decrease in prediction accuracy for two cognitive concepts relevant to driving: working memory load and visuospatial attention. In this study, we developed a cognitive model in the cognitive architecture ACT-R that integrates these two cognitive concepts to provide insights into how, when and where these concepts interact.

Methods

The model used in this study was a modification of the Java ACT-R driving model¹, which itself was a re-implementation of the Lisp ACT-R model (Salvucci, 2006). The model performed two tasks simultaneously using threaded cognition (Salvucci & Taatgen, 2011): a driving task and an n-back task based on Unni et al. (2017).

In the driving task, the model must maintain a safe position on the road while driving along a three-lane highway with some concurring traffic.

To manipulate visuospatial demands, the road alternates between a regular highway with standard lane width (3.5m) and a construction site with narrower lanes (2.5m), where the leftmost lane is blocked by red-white pylons.

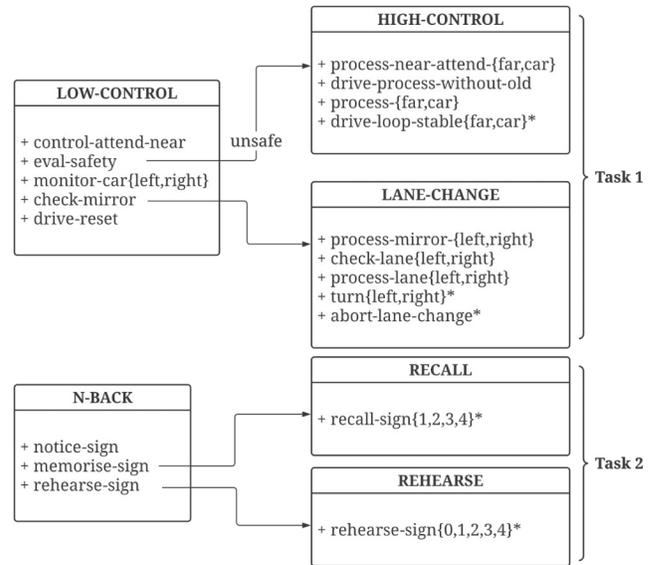


Figure 1: Model production system. Box titles indicate goal types and names below indicate production rules. Asterisks indicate production rules that allow the model to return to parent goal.

The second task consists of a modified n-back task. As the model is driving along the highway, speed signs appear on the right side of the road every 20s, indicating a speed. The model must always drive according to the speed *n* signs back. Thus, upon encounter of a speed sign, the model has to memorize the sign, update a mental list of task-relevant signs, recall the appropriate sign depending on the n-back level and drive according to its speed. The difficulty in the n-back task ranges from 0-back to 4-back.

Driving model

The driving model is an adaptation of the model presented by Salvucci (2006). As the control loop of this driving model is independent of lane-width and it can thus not account for the effect of narrower lanes, we added a ‘low-control loop’ to the model. When the car is at least 0.7m away from either lane edge, the model enters this low-control loop. During this loop, the model can only fire productions from the low-control loop, as can be observed in Figure 1. This loop does not involve steering-control. When the position of the car becomes too close to the lane-edges, the model re-enters the high-control loop to steer back to the center. After re-entering the high-control loop,

¹ <https://www.cs.drexel.edu/~salvucci/cog/act-r/>

the model cannot switch back to the low-control loop for a short period to avoid drifting outside the safe-zone immediately after entering.

As the construction site has a narrower lane-width, the car does not enter the low-control loop in the construction site as it will never be sufficiently far away from the lane edge.

If other cars need to be overtaken, lane-changes are initiated after the model checks the appropriate mirror and lane. If the model is not in the right-most lane, it attempts to change lanes after making checks in a similar manner.

N-back model

The n-back model works by a sequential memorizing mechanism. Each sign is stored in declarative memory with a unique ID when encountered. To successfully recall a sign, the model sequentially goes through the speed signs back in time to remember the desired speed. Importantly, the number of backward steps is dependent on the n-back level, e.g., in a 3-back, the model goes back three times. Errors are modeled by partial matching.

The model rehearses the task-relevant sequence of signs up to three times or until the rehearsal process is interrupted by the encounter of the next sign (cf. Salvucci & Taatgen, 2011).

Results

As can be seen in Figure 2, n-back performance decreased with increasing n-back level, showing a similar effect as human participants (Scheunemann et al., 2019). There was no difference in n-back performance between lane-widths. This effect can be explained by the fact that the model must perform a higher number of retrievals in higher n-back levels. This leads to a higher chance of a mismatch when compared to lower n-back levels.

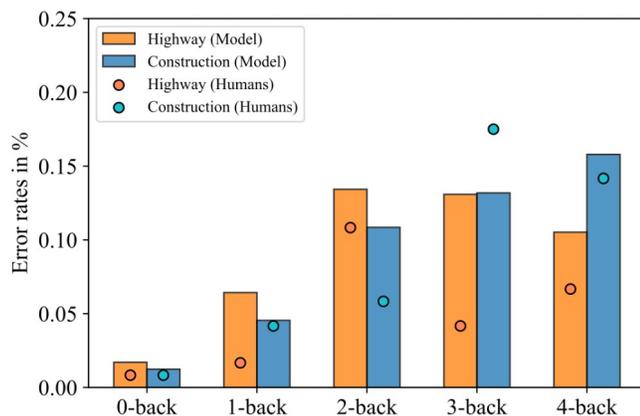


Figure 2: Average error rates in the n-back task. Human Data is taken from Scheunemann et al. (2019)

Analysis of steering reversal rates revealed increased steering reversals in the construction condition across all n-back levels, indicating an increase in driving difficulty (Fig.

3). Additionally, steering reversal rates decrease with increased n-back difficulty.

As the difficulty in the driving task increases, the model spends more time in the high-control loop, which leads to increased steering movements and thus reversals. Moreover, the increase in n-back difficulty requires a higher number of productions to successfully accomplish this task, which has the opposite effect on steering reversals. As more time is spent on the n-back task, less time is available for driving.

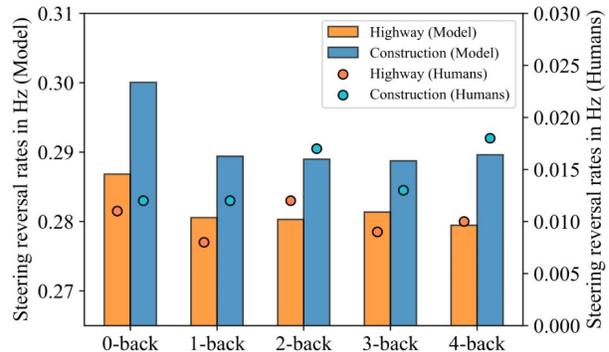


Figure 3: Average steering reversal rates.

Discussion

The ACT-R model is able to show how both tasks compete for available resources: driving behavior is influenced by n-back level because of a competition for access to procedural and declarative memory. These results indicate an interaction at common task-unspecific level.

Because there is limited behavioral data available regarding driving behavior with varying lane-widths, some model parameters had to be estimated when developing the model (e.g., overtaking distance). We are currently conducting a behavioral study with human participants to remedy these factors and further validate the model.

References

Hancock, P. A., Jagacinski, R. J., Parasuraman, R., Wickens, C. D., Wilson, G. F., & Kaber, D. B. (2013). Human-automation interaction research: Past, present, and future. *Ergonomics in Design*, 21(2), 9–14.

Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors*, 48(2), 362–380.

Salvucci, D., & Taatgen, N. A. (2011). *The multitasking mind*. Palgrave MacMillan.

Scheunemann, J., Unni, A., Ihme, K., Jipp, M., & Rieger, J. W. (2019). Demonstrating Brain-Level Interactions Between Visuospatial Attentional Demands and Working Memory Load While Driving Using Functional Near-Infrared Spectroscopy. *Front. in Hum. Neurosci.*, 12, 542.

Unni, A., Ihme, K., Jipp, M., & Rieger, J. W. (2017). Assessing the driver’s current level of working memory load with high density functional near-infrared spectroscopy: A realistic driving simulator study. *Frontiers in Human Neuroscience*, 11, 167.

Validating and Refining Cognitive Process Models Using Probabilistic Graphical Models

Laura M. Hiatt
US Naval Research Laboratory
Washington, DC USA
laura.hiatt@nrl.navy.mil

Connor Brooks
University of Colorado Boulder
Boulder, CO USA
connor.brooks@colorado.edu

J. Gregory Trafton
US Naval Research Laboratory
Washington, DC USA
greg.trafton@nrl.navy.mil

Abstract

We describe a new approach for developing and validating cognitive process models. In our methodology, graphical models (specifically, hidden Markov models) are developed both from human empirical data on a task, as well as from synthetic data traces generated by a cognitive process model of human behavior on the task. Differences between the two graphical models can then be used to drive model refinement. We show that iteratively using this methodology can unveil substantive and nuanced imperfections of cognitive process models that can then be addressed to increase their fidelity to empirical data.

Introduction

Building cognitive process models of human behavior is a challenging task that has rich rewards. Such models can be used for many purposes, including to better understand where people make errors (Hiatt & Trafton, 2015), to better understand how to effectively teach people new skills (Lee, Anderson, Betts, & Anderson, 2011), and to design effective computer interfaces (Cao, Ho, & He, 2018).

Because human thought processes can be only observed indirectly, such as by recording external behavior (e.g., reaction times, responses to questions, etc.), each cognitive process model serves as a de facto theory of human behavior on the task being modeled. To validate those theories, the simulated behavior of the models on the task can be compared to the behavior of human participants on the task using measures such as statistical goodness-of-fit measures (e.g., R^2 , RMSE, etc.). A strong fit indicates that the model is a good theoretical candidate for how humans complete the task.

However, there are two related issues with this development methodology that we address in this paper, both of which stem from the fact that a strong or weak statistical fit does not necessarily indicate which parts of the model may fit human behavior better than others. First, a weak statistical fit does not necessarily give any actionable information for which part of the model to modify to improve it. Second, a strong statistical fit for part of the model's task may mask issues with another area.

Here, we use hidden Markov models (HMMs), a type of probabilistic graphical model, as a new analysis tool for validating the efficacy of a cognitive model on a human behavioral task. First, we use the cognitive model to generate synthetic data of human behavior on a task. Then, we train two HMMs using two datasets: (1) human empirical behavior on

the task, and (2) the generated synthetic data. By comparing these HMMs both qualitatively and quantitatively, we can not only measure how similar the datasets (and thus the underlying behaviors, or models of behavior) are, but also, because HMMs are a form of graphical model, visually see ways in which the datasets differ. We show that using this process can improve cognitive model fidelity by both of these qualitative and quantitative metrics, as well as improve the predictive accuracy of the cognitive model when predicting a human's next action on a task.

In the following sections, we describe: the task; the empirical data that was collected; the initial model of the task; the graphical model approach; and the revision of the cognitive model. Finally, we discuss the resulting improvement in predictive performance.

Task Description

In order to study how cognitive models of complex behavior can leverage machine learning as a tool for improvement, we turn to a supervisory control task. Specifically, we considered how people performed while interacting with the Research Environment for Supervisory Control of Heterogeneous Unmanned Vehicles (RESCHU) (Boussemart & Cummings, 2008) simulator. RESCHU is an interactive supervisory control task that requires complex decision making, problem solving, and reasoning.

Figure 1 shows the simulation, which has three panels: a map panel, a status panel, and a payload panel. The map panel (Figure 1, right) displays unmanned aerial vehicles (UAVs) (blue/red half ovals), targets (orange/green diamonds) towards which UAVs are moving, and hazard areas (yellow circles) which should be avoided by UAVs and can change location over time. The status panel (Figure 1, bottom left) shows the status of the UAVs and includes information on vehicle damage, time until the vehicle reaches a waypoint or target, and time remaining in the simulation. The payload panel (Figure 1, top left) is used by the operator to perform a manual visual acquisition task once the UAV has reached the target. It is not critical to this work so we largely omit its consideration; it is more fully described in (Breslow, Gartenberg, McCurry, & Trafton, 2014).

The goal for an operator's session in RESCHU was to monitor and guide the five UAVs to reach as many targets as possible, and complete the corresponding payload tasks, while

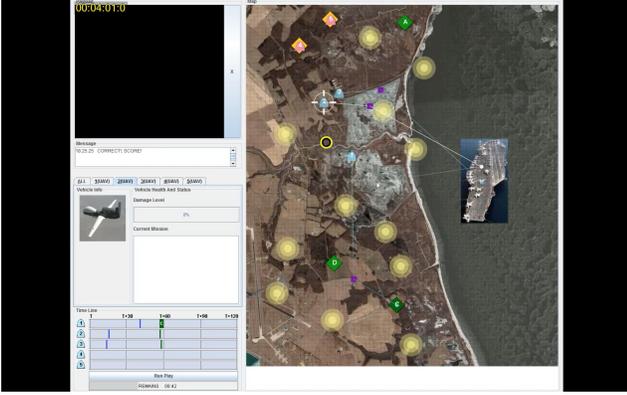


Figure 1: A screenshot of the RESCHU environment simulator used in this experiment.

avoiding damaging the UAVs in the hazard areas.

At the start of the simulation, the UAVs were randomly assigned to targets; thus, the UAVs might not be directed towards the optimal target. After a target was reached and the visual acquisition payload task was complete, the target “reset”, and the UAV was randomly assigned to a new currently-unassigned target, which again might not be optimal.

Because of this suboptimal automation, as well as the changing location of the hazard areas, a critical subtask of RESCHU was changing the target of a UAV. Operators could do this at any time by using the mouse to click on the UAV and then clicking on the UAV’s new target. We focus on this subtask here because of its import as well as because, as we will see, even in this “simple” subtask there is variation in how operators perform it, making it an ideal subtask to test our proposed methodology. The end goal of the modeling effort is to be able to predict an operator’s final selected target, so the interface can better assist with the selection process (e.g., pre-selecting the target, highlighting the target, etc.).

Empirical data

The empirical data we consider was based on ten participants using RESCHU. Participants were provided extensive training on the RESCHU system, through an online tutorial, in-person instruction, and walk-throughs. Participants also had as much time as they wanted to use the entire system until they were well-versed in the intricacies of RESCHU. Participants were all volunteers (no incentives), healthy, with age less than 30 years. Details on the methodology of the study are available in (Breslow et al., 2014).

After a participant was fully trained on RESCHU, they were seated approximately 66 cm from the computer monitor and were calibrated on an SMI RED eye tracker operating at 250 Hz, which collected eye tracking data once the experiment began. A fixation was defined using the dispersion method based on a minimum of 15 eye samples within 60 ms and within 50 pixels (approximately 3° of visual angle) of each other, calculated in Euclidian distance. Fixations on specific objects were automatically identified after all data

collection was completed. The simulation also logged all operator *actions*, i.e. mouse clicks, indicating what object was clicked on (i.e., selected) at different times.

All instances of changing a UAV’s target were manually extracted from the simulation. A total of 200 sequential process traces, with eye fixations and mouse clicks listed in chronological sequence, were created by these participants. For this subtask, a process trace contained sequences of the set of possible observations of the operator’s behavior, $\{uav1, uav2, \dots, haz1, haz2, \dots, tar1, tar2, \dots, Action-SelectUAV1, Action-SelectUAV2, \dots, Action-SelectTar1, Action-SelectTar2, \dots\}$. Note that if the observation corresponds to a mouse click, it is designated specifically as an “Action-Select”; otherwise, it is an eye gaze fixation.

We were able to create a simple coding scheme for the empirical data to categorize different strategies people used (e.g., a planning strategy needed to fixate on both the UAV and the target before selecting the UAV; an opportunistic strategy picked and selected a UAV before looking for a target). The coding scheme was implemented computationally and run on all the empirical data to provide preliminary evidence of different strategies.

Hidden Markov Models for Comparing Datasets

A hidden Markov model (HMM) is a graphical model that stochastically transitions between states using the Markov assumption (i.e., transitions depend on only the current state). The *hidden* term refers to how states are not directly observable; instead, states stochastically emit observations that give clues to what state the model is currently in. Figure 3 shows example HMMs that we will discuss later in our analysis.

A typical hidden Markov model (HMM) is formally defined by the tuple $\langle S, Z, A, B, \pi \rangle$:

- S is the set of states.
- Z is the set of observations.
- A is an $|S| \times |S|$ matrix defining transition probabilities between states $a_{i,j} = p(x^t = s_j | x^{t-1} = s_i)$.
- B is an $|S| \times |Z|$ matrix defining observation probabilities of the states $b_{i,k} = p(o^t = z_k | x^t = s_i)$.
- π is a vector with initial state probabilities $\pi_i = p(x^0 = s_i)$.

In this definition, x_t and o_t represent the true state and emitted observation at time t , respectively.

Learning HMMs

In order to create an HMM that models a dataset, existing techniques for learning HMMs can be used. Learning for HMMs can refer to both learning the topology of the HMM (i.e., learning what states connect to others; Singer & Ostendorf, 1996; Siddiqi, Gordon, & Moore, 2007), as well as learning the parameters of it (i.e., learning the values of A and B ; Rabiner, 1989). Here, we adapt the basic strategy from (Singer & Ostendorf, 1996) for learning the HMM topology using repeated “state splitting,” followed by standard Baum-Welch parameter learning (Rabiner, 1989). The Baum-Welch algorithm uses successive iterations of forward

and backward passes through all available data traces to iterate toward the transition and observation probabilities that maximize the probability of the observed data. Repeating these two steps of state splitting and Baum-Welch parameter learning results in joint learning of both the topology and parameters of the HMM.

The training procedure begins by reading in a set of observation sequences. We prepend each sequence with a *START* observation and append each with an *END* observation to denote the fixed beginning and end of the sequence.

Then, an HMM is initialized with a start state that emits only *START*, a middle state that emits all observations other than *START* and *END*, and an end state that emits only *END*. The start state transitions only to the mid state, the middle state transitions to itself and the end state, and the end state terminates the sequence.

From this point, state-splitting is used to expand the HMM topology, one new state at a time. At every expansion step, a new possible HMM is created and trained for each candidate split, with every state (other than the specialized start and end states) considered a candidate for splitting.

For each new candidate HMM, both of the newly split states are initialized with the same transitions as the original state. The new HMM's parameters are then trained using Baum-Welch learning. After all possible HMMs for the current expansion step have been created and trained, the candidate HMM that maximizes the likelihood of the dataset is selected for use or as the basis of the next expansion step; the rest are discarded. This procedure can end once either a fixed number of states is reached, or some measure of fit (such as the likelihood of a validation set) stops improving.

HMMs as Dataset Representations

Learning the HMM directly from a dataset produces a graphical model that represents some underlying structure in the data. In particular, if two HMMs are created from datasets generated by two different sources, the learned HMMs can provide a graphical representation of the differences in the sources underlying the two datasets. This provides a novel way to analyze and validate cognitive process models: comparing HMMs learned from synthetic data from a cognitive model with those learned from human empirical data.

Intuitively, we should be able to learn HMMs that do not depend on the specific label of current action (i.e., looking at *tar1* vs. *tar2*), but instead consider its meaning (i.e., looking at a target of interest vs. looking at a target not of interest). There is an additional consideration, therefore, to using HMMs as an analysis tool for cognitive process models on tasks where a “template”-style process is applied to various versions of a task. As an example, in the UAV re-routing subtask, the same processes are followed no matter which UAV is being rerouted to which target. But because the specific targets of interest differ across observation sequences, learning a single standard HMM on a dataset for this subtask would essentially be meaningless for understanding the process generating the data: the HMM will not be able to meaningfully

differentiate between targets of and not of interest.

To address this need, we next describe how we collapse observations into composite observation sets in order to train effective HMMs on these types of “template” tasks.

Composite Observations for HMMs

The first step in collapsing observations is to divide all possible raw observations into classes (e.g., group together all hazard observations into a *haz* class, etc.). For each class, we represent it using either one or two composite observations. For classes like *hazard*, all observations can generally be collapsed into one composite observation *haz*. For classes like *target*, however, they cannot: that would meaningfully impact an HMM's ability to understand the operator's observations and actions. Such classes have two composite observations: *tar+* which indicates the target of interest for a given sequence, and *tar-* which indicates targets not of interest.

It is straightforward to convert raw observation traces into composite observation traces. Each raw observation is replaced with either its single composite symbol, or with the appropriate dual composite (i.e., if the raw observation is *tar2* and *tar2* is the selected target, it is replaced with *tar+*; otherwise, it is replaced with *tar-*). The resulting composite dataset can then be used to train an HMM as described above¹.

HMM Comparison Measures

With two HMMs in hand, we can compare them to find structural or other similarities or differences in the underlying data. We consider two ways to compare HMMs: one quantitative, and one qualitative. Quantitatively, comparing HMM topology and parameters involves calculating the similarity of the expected outputs produced by two HMMs. A classic approach to do so, described by Juang and Rabiner (1985), is to estimate the Kullback-Leibler divergence between the probability distributions of observation sequences generated by the two HMMs. This estimation is done through a Monte Carlo approach by repeatedly generating observation sequences from one HMM, calculating the probability of each of these sequences being emitted from *both* HMMs, and comparing the two probability values. Although this measure is asymmetric, we transform it to a symmetric measure by calculating it in both directions and averaging them. As the number of observations goes to infinity, the estimate approaches the true Kullback-Leibler divergence of the two HMMs.

Qualitatively, a comparison of HMMs can be done by visually viewing them, and comparing transitions between states as well as observation probabilities. For example, if one HMM always begins by entering a state where it looks at a UAV, while another always begins by entering the state where it looks at a hazard, that can be viewed as a meaningful qualitative difference between the models.

¹An additional, subtle, benefit of an HMM that reasons over composite UAVs, hazards, targets, etc., is that it allows for adding or removing additional items (such as more targets) at run time without going through training again.

Cognitive Model Development Cycle

Our cognitive model development cycle began by one of the authors of this paper hand-writing a cognitive model in ACT-R/E (Trafton et al., 2013) to capture human performance on the UAV re-routing subtask. The original models are identical to that described in (Trafton, Hiatt, Brumback, & McCurry, 2020); we describe them generally here, but interested readers can refer to that work for their specifics.

Original Model Description

While performing the re-routing subtask, people could use a variety of cognitive strategies; here, we focused on modeling two strategies introduced when discussing the empirical data: a *planning* strategy and an *opportunistic* strategy.

The planning strategy captures the insight that people sometimes plan a few actions ahead, or search for the best action to do, when performing a task. Here, the strategy first searches for a UAV whose target needs to be changed (because it is on a collision course with a threat or is far away from its target, etc.), by using its perception to see the interface, and its memory to interpret what it sees. It then holds the UAV in working memory while searching for a better target (e.g., one that does not intersect a hazard or that is closer). After identifying the UAV and the new target of interest, the model executes the actions to change the UAV's target (i.e., clicking on the UAV, then clicking on the new target).

The opportunistic strategy occurs when people may not have time, resources or inclination to plan ahead. The model of this strategy sequences its actions differently. It first searches for a UAV whose target needs to be changed, and then immediately clicks to select the UAV without a specific target yet in mind. Next, the model searches for a target where the UAV could be sent. After an appropriate target is found, it clicks on it to change the UAV to go to that target.

Note that while the differences in strategies are subtle, they are different in their actions (i.e., mouse clicks) as well as patterns of eye-movements. Also, the planning strategy clearly requires greater utilization of working memory (e.g., needing to hold the UAV in mind while searching for an appropriate target); however, we note that even though the RESCHU task is dynamic, there seems to be enough time to execute both strategies within the constraints of the subtask.

Analysis Set-Up

There are two steps to setting up the analysis of the cognitive models: acquiring the right data, and training the HMMs.

Data Evaluating the cognitive process models using machine learning techniques requires both empirical as well as synthetic data stemming from the cognitive models. The empirical data was described when introducing the RESCHU task. For the synthetic data, we used the developed cognitive models to generate observation traces indicating the model's theory of how people perform the task. Critically, these models generate traces of observational data that were identical in form to the traces that were generated from the human par-

ticipants, including eye fixations and mouse clicks listed in chronological order. All together, both the planning model and the opportunistic model were each run 20,000 times to generate 20,000 individual, distinct traces of synthetic human performance for each strategy.

HMM Training HMMs can then be trained using each data source: empirical planning, empirical opportunistic, synthetic planning, and synthetic opportunistic. Before training, each dataset was converted into its composite dataset as described above. Each HMM was limited to 5 splits while learning the topography, resulting in a total of 8 states per HMM (six "split" states plus the start and end states). For each candidate split, parameters were learned using 500 Baum-Welch iterations. Figure 3a shows one such trained HMM.

Analysis

The trained HMMs can next be compared both quantitatively and qualitatively. Quantitatively, the HMM distance measure, described above, calculates how related the different HMMs are. The first columns of Table 1, show these values for the original cognitive process models (PlanSynOrig, OppSynOrig) compared to the empirical data (PlanEmp, OppEmp). While there is not an exact "ideal" target for these values, lower values indicate that the data generated by the two models overlap more and, as such, lower values are better. These values show that the cognitive model of the planning strategy is closer to that of the empirical data than the model for the opportunistic strategy. These values do not, however, offer any insight into why this is the case.

In contrast, Figures 3a and 3b show a qualitative comparison of the empirical planning HMM and the synthetic planning HMM. As they show, there are several qualitative ways in which the HMMs differently characterize the behavior of their respective data. Notably, as we will discuss in the following section, there are several differences in what observations occur directly before and after actions (mouse clicks).

Revision

The HMM analysis highlighted several differences between the cognitive models and the empirical data. These differences suggested changes in the models that potentially could lead to not only a greater theoretical understanding of dynamic tasks but also improved fits and better prediction. Two changes were made to both strategy models based on the revealed differences between them and the empirical data.

Looking at the target of interest before selecting it Before clicking the target of interest, the original cognitive models looked directly at it and then selected it very consistently. However, the empirical data did not show such a strong relationship. One possible explanation for this is that people used a sort of embodied cognition by focusing on a target, moving their mouse to that target, looking around more, and then simply clicking the mouse to finalize the selection. We implemented this aspect in the models by providing a 50% chance that people would use a form of embodied cognition.

	PlanSynOrig	PlanSynRev		OppSynOrig	OppSynRev
PlanEmp	0.209	0.181	OppEmp	0.312	0.181

Table 1: Distances between the empirical-based HMMs and original and revised process model-based HMMs for the planning and opportunistic strategies. Lower scores indicate greater similarity between the HMMs.

Looking at the UAV immediately after selecting it The empirical data suggested that people frequently (50%) looked at the target of interest immediately after selecting the UAV. In contrast, the original cognitive model used a GOMS-style approach, generally looking at the UAV immediately after selecting it to confirm that it was selected. To address this, instead of increasing the probability that the target was focused on immediately after selection, we instead decreased the probability that the model confirmed the UAV was selected. By verifying that the UAV had been selected less often, we expected that the target would be examined sooner.

Note that for both of these changes, we did not perform any sort of parameter-space search; we simply changed the aforementioned probabilities to 50%. We assume that if we had performed additional parameter-space search, the model fit would be better, perhaps at the cost of over-generalization.

Results

We first look at the HMM measures to evaluate whether the model fit to the empirical data has improved after revision. Table 1 shows that the quantitative fit did, in fact, improve. The planning strategy showed a moderate improvement, while the opportunistic strategy showed a large improvement. Qualitatively, as Figure 3 shows, the new planning cognitive model shows a stronger topological and parameter similarity to the empirical HMM; the opportunistic shows a similar improvement.

As an additional measure, we have suggested that an effective way to evaluate a cognitive model is by predicting specific steps during task execution (Breslow et al., 2014; Trafton et al., 2020; Ratwani & Trafton, 2011). In (Trafton et al., 2020), we showed that we could predict what target a UAV would be directed to by using synthetic data generated by a cognitive model to train a CNN. If the HMM analysis added value by improving the cognitive model, we should see an improvement to the predictive capabilities of the CNN: the cognitive model should capture the patterns in the data better, allowing better prediction of the specific target selection.

Thus, we trained CNNs using synthetic data from the original and revised cognitive models to see whether the revision resulted in better predictive performance. For each CNN, 10-fold cross validation was used to divide the empirical data into training and testing data. All conditions used the same folds for training and testing, and all models were evaluated on the empirical data. In addition to training CNNs for the planning and opportunistic strategies, we also trained a combined strategy CNN that used half of each. As Figure 2 shows, in all cases, the revision improved the predictive performance, showing that not only did our proposed process

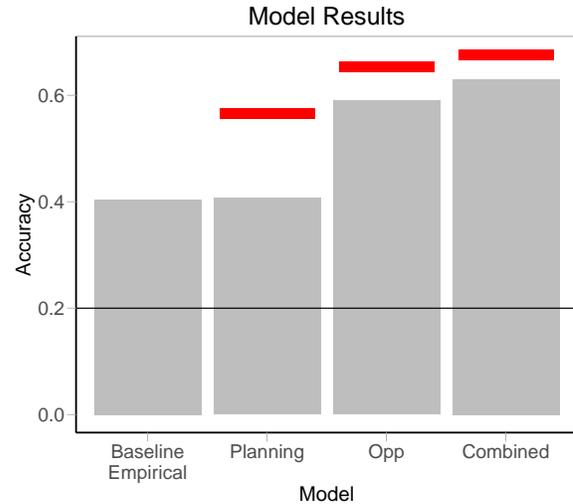


Figure 2: Original model predictive accuracy (bar) and updated model predictive accuracy (red rectangle). The black horizontal line is chance.

improve HMM-based metrics, but it improved the models' eventual predictive power as well.

Discussion

In this paper, we have described a new methodology that cognitive modelers can use to develop, analyze and verify cognitive process models. In it, we learn hidden Markov models from data from empirical experimentation, as well as from synthetic data generated by candidate cognitive models. By comparing those HMMs both qualitatively and quantitatively, we can see the cognitive models' goodness of fit, as well as determine concrete ways in which the cognitive model can be improved. This is different than validating models with statistical goodness-of-fit measures, which do not offer concrete pointers for improving cognitive models. We also show that this process can lead to increased predictive performance by the cognitive model. In future work, we plan to expand this methodology to demonstrate it for models in additional cognitive architectures and of additional tasks.

Interestingly, this methodology revealed an issue with how visual-based actions are typically done in cognitive models; namely, look at, encode, prep to act, then act upon it. The HMM built for the empirical data suggests a different paradigm for dynamic tasks: look at, encode, prep to act, and then *possibly look elsewhere* before acting. This type of insight is not possible to glean from summary statistics or other typical measures of cognitive process models; a graph-

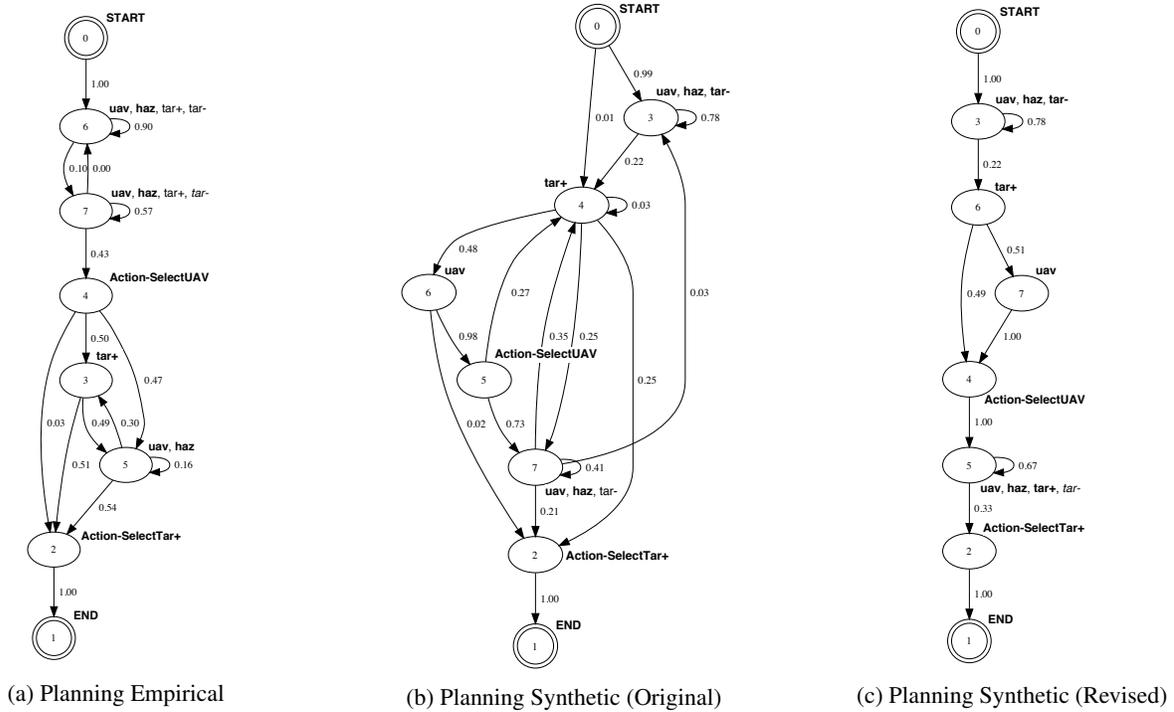


Figure 3: HMMs of the planning strategy built from different datasets. The numbers of the states indicate the order in which they were created. The meaning of the states can be derived from the observations they emit, which are shown in the figures. Bolded observations occur with greater than 20% likelihood; italicized observations occur with less than 10% likelihood.

ical model allowed us to find this approach.

Another key insight is that, perhaps surprisingly, the improvements in HMM similarity and in predictive performance did not mirror one another. For the HMM measure, the opportunistic model saw the greater improvement after its revision. Considering predictive performance, the planning model saw the greater improvement. This is because the different representations of the HMM and the CNN lead them to capture different aspects of the data: the HMM focuses on underlying structure; the CNN focuses on non-linear patterns. This subtlety only highlights the need for more diverse tools for cognitive model analysis like the one that is proposed here.

Acknowledgements

We thank the Office of Naval Research for funding for this work (LH and GT). We also thank Tony Harrison for useful discussions. The views expressed in this paper are those of the authors and do not reflect the official policy or position of the US Navy, Department of Defense or the US Government.

References

Boussemart, Y., & Cummings, M. (2008). Behavioral recognition and prediction of an operator supervising multiple heterogeneous unmanned vehicles. *Humans operating unmanned systems*.

Breslow, L. A., Gartenberg, D., McCurry, J. M., & Trafton, J. G. (2014). Dynamic operator overload: A model for

predicting workload during supervisory control. *IEEE Transactions on Human-Machine Systems*, 44(1).

Cao, S., Ho, A., & He, J. (2018). Modeling and predicting mobile phone touchscreen transcription typing using an integrated cognitive architecture. *International Journal of Human-Computer Interaction*, 34(6), 544-556.

Hiatt, L. M., & Trafton, J. G. (2015). An activation-based model of routine sequence errors. In *Proceedings of the International Conference on Cognitive Modeling*.

Juang, B.-H., & Rabiner, L. R. (1985). A probabilistic distance measure for hidden markov models. *AT&T Technical Journal*, 64(2), 391-408.

Lee, H. S., Anderson, A., Betts, S., & Anderson, J. R. (2011). When does provision of instruction promote learning? In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*.

Rabiner, L. R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257-286.

Ratwani, R. M., & Trafton, J. G. (2011). A real-time eye tracking system for predicting and preventing postcompletion errors. *Human-Computer Interaction*, 26(3).

Siddiqi, S. M., Gordon, G. J., & Moore, A. W. (2007). Fast state discovery for HMM model selection and learning. In *Artificial Intelligence and Statistics* (pp. 492-499).

Singer, H., & Ostendorf, M. (1996). Maximum likelihood successive state splitting. In *IEEE International Con-*

ference on Acoustics, Speech, and Signal Processing.

Trafton, J. G., Hiatt, L. M., Brumback, B., & McCurry, J. M. (2020). Using cognitive models to train big data models with small data. In *International Conference on Autonomous Agents and Multi-Agent Systems*.

Trafton, J. G., Hiatt, L. M., Harrison, A. M., Tamborello, II, F. P., Khemlani, S. S., & Schultz, A. C. (2013). ACT-R/E: An embodied cognitive architecture for human-robot interaction. *Journal of Human-Robot Interaction*, 2(1), 30-55.

Competence Assessment by Stimulus Matching: An Application of GOMS to Assess Chunks in Memory

Hadeel Ismail (hi71@sussex.ac.uk)

Department of Informatics, University of Sussex
Brighton, BN1 9QJ, UK

Peter C-H. Cheng (p.c.h.cheng@sussex.ac.uk)

Department of Informatics, University of Sussex
Brighton, BN1 9QJ, UK

Abstract

It has been shown that in hand-written transcription tasks temporal micro-behavioral chunk signals hold promise as measures of competence in various domains (e.g., Cheng, 2014). But data capture under that an approach requires the use of graphics tablets which are relatively uncommon. In this paper we propose and explore an alternative method – Competence Assessment by Stimulus Matching (CASM). This new method uses simple mouse-driven interfaces to produce temporal chunk signals as measures of learner’s ability. However, it is not obvious what features of CASM will produce effective competence measures and the design space of CASM tasks is large. Thus, this paper uses GOMS modelling in order to explore the design space to find factors that will maximize the discrimination of chunk measures of competence. Results of a pilot experiment show that CASM has potential in using chunk signals to measure competence in the domain of English language.

Keywords: chunking; GOMS; language competence; pause analysis; stimulus matching

Introduction

This paper concerns the assessment of learners’ competence in knowledge rich domains, using the analysis of computer logs of micro-behaviors in task activities. Moss, Kotovsky, and Cagan (2006), in the domain of engineering, and Arslan, Keehner, Gong, Katz, & Yan (2020), in the domain of mathematics, used drag and drop tasks to examine the underlying cognitive processes in either replicating subject-related diagrams or solving mathematical problems, respectively. Another study analyzed pauses during text composition by means of key-stroke logging (Schilperoord, 2002). These methods were successful in extracting and associating behavioral signals with cognitive processes, by logging actions at a time scale of ≈ 10 seconds.

An alternative approach that holds some promise is to log and analyze micro-behaviors at a time scale of 1 second and less. Machine learning was used to analyze large amounts of data logged during freehand writing (Stahovich & Lin, 2016) and drawing (Oviatt, Hang, Zhou, Yu, & Chen, 2018) during problem-solving tasks. Their findings revealed significant correlations between pause durations and proficiency levels.

In contrast, Cheng and colleagues have used cognitive chunking theory to develop methods that require less data

using short transcription tasks. According to Cowan (2001) and Miller (1956), “chunking” is a process by which perceived information are grouped and stored in working memory (WM), and since information is presented as units, people tend to group these units into “chunks” of meaningful information. The number of “chunks” stored is constrained by one’s mental capacity, however Cowan (2001) also points out that the capacity is also affected by the amount of prior knowledge one holds in long term memory in the expert domain. So, in the experiments carried out by Cheng and colleagues, they examined differences in pause behavior of novices and experts whilst engaging in transcription tasks to probe chunk structures in memory. Cheng and Rojas-Anaya (2007) observed individuals copying mathematical equations freehand and could distinguish level of experience. However, their sample size was small and participants had large differences in their mathematical expertise. Extending the approach Cheng (2014) showed strong correlations between competence and the third quartile (Q3) pauses. Similarly, Zulkifli (2013) asked learners of English as a second language to transcribe English sentences freehand and found Q3 to be an effective measure of competence. Albehajjan and Cheng (2019) show the possibility of measuring programming competency using the same method. Overall, it seems that pause based measures in transcription tasks have some potential for assessing competence in various domains.

Despite the promise of freehand transcription, one limitation is the need for a graphics tablet, an uncommon IT equipment. Thus, it would be desirable to combine mouse driven tasks (Arslan et al., 2020; Moss et al., 2006) with the benefits of capturing micro-behaviors. Cheng (2015) used a mouse and a response grid on a screen to measure temporal chunk signals related to mathematical competency. Participants copied the stimuli by clicking on the matching symbols that appeared on the grid. Results showed that clicking to select symbols has potential as a means to measure mathematical competence.

In this paper we propose an alternative approach to the assessment of competence administered on a standard computer by means of mouse clicking: *Competence Assessment by Stimulus Matching* (CASM). A preliminary CASM task design has been created (Fig. 1), that takes into consideration the different factors that would encourage the

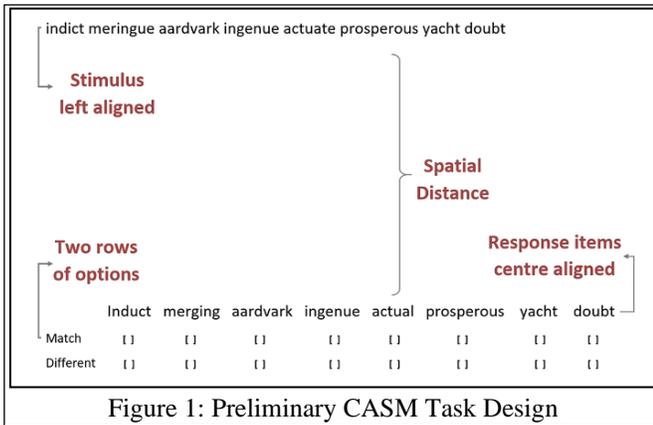


Figure 1: Preliminary CASM Task Design

use of chunking. The task is presented as a split screen with the stimulus at the top and the response area at the bottom. The response area includes words that either match or differ from the stimulus. Participants are expected to verify the match or mismatch and use the mouse to mark their responses as quickly and as accurately as possible. The time course of clicks in the check boxes will reflect certain aspects of the test takers language competence, so measures of competence may be devised for the task.

A key problem is how to design CASM tasks to produce behaviors that maximally differentiate high and low competence. Will the micro-behaviors of experts and novices differ substantially on the task in Fig.1, and so potentially provide effective temporal chunks measures of competence? This paper considers the possibilities, but the possible space of design is large. Some of the factors influencing this task include: the large spatial distance, the deliberate misalignment of words, the use of low frequency words and multi-syllabic words. So, how can we effectively yet efficiently explore the space? A task analysis approach is adopted, in particular a somewhat novel approach to the application of GOMS modeling is used to assess chunks in memory in order to further determine how the different design factors impact the task environment.

Task Design Space

The aim is to develop chunk-based Competence Assessment by Stimulus Matching (CASM) tasks that rely on mouse clicking, in contrast to Cheng and colleagues pen-on-paper transcription approach. The key issue is the design space, where many variables provide us with a plethora of design choices, from which we must choose those that impact the distribution of pauses that maximally differentiate experts from novices.

Screen Layout and Stimulus Positioning: The layout may encourage the use of chunking to provide experts with an advantage over novices. Firstly, the spatial distance between the stimulus and the response areas may be made deliberately large to impose a task load on individuals, who must shift their gaze vertically. In turn this may encourage them to chunk as much as possible. Cheng (2014, 2015) used distant positioning to improve the Q3 pause measures of competence. Secondly, the misalignment of the stimulus

and the response is assumed to encourage experts to use chunks to save the effort of switch gaze, and place some difficulty on the novices who, because of their limited language knowledge, might take longer to locate the point where they last left as they shift their gaze.

Presentation Mode: In presenting the stimuli, one approach is to have it visible throughout the duration of the task; “constant display” (Cheng & Rojas-Anaya, 2007; Cheng, 2014; Zulkifli, 2013). The second is “voluntary view”, where the appearance of the stimuli requires an action by the individual (Albehajjan and Cheng, 2019).

Stimulus and Response Composition: The general approach here is to play with effects of stimulus and response composition or decomposition. This applies at the whole stimulus (sentence), word (compound words) and part word (syllable) levels. If working at the word level, one option is to present stimuli words in a way that, if two were combined, they would make up a compound word which may differentially benefit the expert by increasing their chunk size by treating the two words as one unit rather than two for a novice (e.g., “counter measure”). We would expect the benefit to be reflected in the pauses in the task and hence in measures of competence.

Stimulus Content: Content manipulations include word frequency (high and low), word length, sentence structure (simple, complex, incorrect), semantic meaning, etc. Zulkifli (2013) shows that such manipulations can be applied in ways that benefit experts to use their knowledge which may be revealed in chunking measures.

Method

The steps taken to carry out the task analysis are: (1) Design a number of task variations. (2) Use GOMS to develop flow charts that predict the processes employed by experts and novices. (3) Calculate the durations for each process, to predict differences in pause distributions and lengths. (4) Run a pilot study to evaluate the modeling results.

GOMS, is a well-established systematic approach to cognitive task analysis that is usually applied during system design to test for usability aspects, choose between candidate designs and understand user behavior (Card, Moran & Newell, 1983). However, our motivation is not to understand user performance, per se, but rather to find designs that constrain their behavior so that micro-behavioral signals of competence are as robust as possible.

While the GOMS models are usually applied to understand how the external task environment affects the individual’s behavior, we on the other hand apply the analysis in a way to understand the internal processing of chunks, leading to how that impacts the design of the task. So, within the framework of GOMS, in our approach, *goals* are related to the size of the chunk an individual can hold in memory. Not only this is affected by the layout of the interface (*externally*) but its largely constrained by their level of familiarity with the words presented (*internally*). Among the *operators* of particular interest to us are those classified as cognitive operators. Those that deal with the

decomposition of a chunk are decisional processes concerned with whether certain elements makeup a chunk or not. Others are related to retrieving chunks from memory, comparing and verifying. The *methods* are the internal loop processing by which the sequence of operators to achieve a certain sub-goal. *Selection rules* are choices that test takers will make to choose between alternative methods based on the chunks they possess, which will be manifest as different micro-behaviors and that chunk measures will attempt to measure.

Allocating Time Durations

All operators are allocated specific time durations that were mostly extracted from past GOMS studies.

1. **Word/syllable recognition:** The time for recognizing a six-letter word, a syllable or a letter is 340ms (John & Newell, 1989).
2. **Cognitive operators:** Cognitive operators include those processes that involve holding a chunk in memory, decision making, verifying, and comparing. According to the literature, the average duration for mental processes is between 50 and 70ms (Gray & Boehm-Davis, 2000; Olson & Olson, 1990; John &

Newell, 1989). The proposed tasks involve low-level cognitive processing, so 50ms is chosen.

3. **Chunk retrieval:** This process was allocated a duration of 50ms, following similar studies involving immediate copying (John, 1988, as cited in Olson & Olson, 1990).
4. **Mouse move:** A quick pilot experiment was conducted on the author and an additional participant. The average time for moving between response items was 500ms, and 700ms for moving from the top screen to the bottom. The second was used as the duration of the action to reveal stimuli in voluntary display tasks.
5. **Eye movement:** The time for a saccade is 30ms (Russo, 1978, cited in Card et al.,1983).

Analysis

Task Analysis: Flowcharts

Since the design space is large it is impossible to examine all combinations of variables here, so we focus on the design in Fig. 1 as an exemplar. The main features of the design are the layout, use of low frequency words, inclusion of disyllabic and trisyllabic words, and presenting the stimulus in constant display mode. The flowcharts in Fig. 2

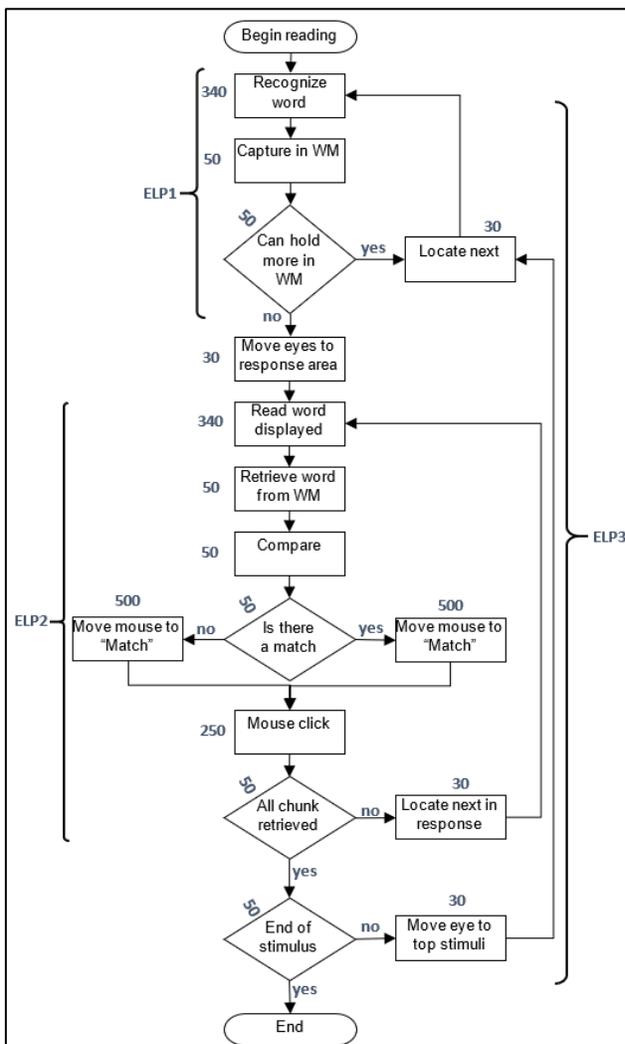


Figure 2: Expert Flowchart

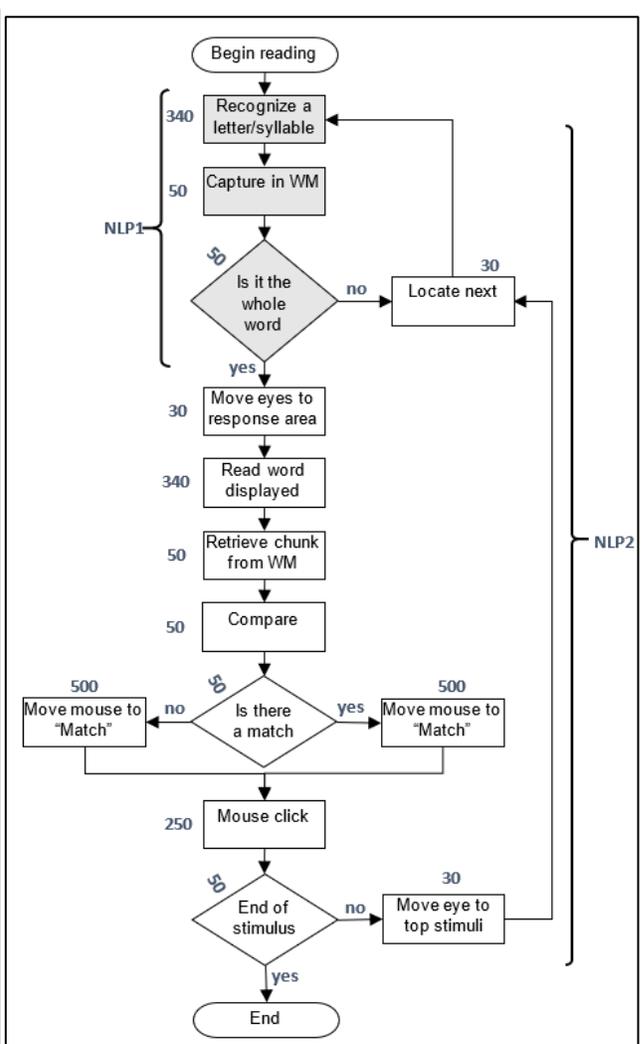


Figure 3: Novice Flowchart

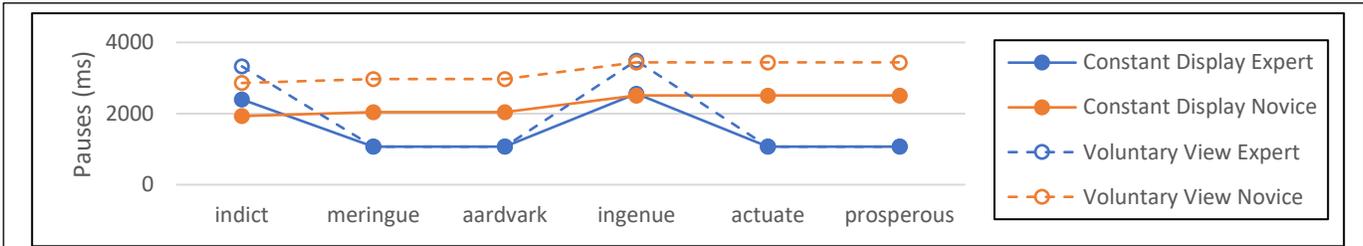


Figure 4: Predicted pattern of pauses for experts and novices in the constant display and voluntary view conditions

and Fig. 3 are high-level representations that conceptualize how an ideal expert and an ideal novice, in the English language, would approach the indicated task. For the purpose of this study, our definition of an expert is someone who possess a vocabulary that enables them to fluently read a piece of text while simultaneously processing its meaning. A novice, on the other hand, is someone with a small vocabulary size, and therefore their reading is much slower as they exert much of their cognitive effort in phonetically processing presented words.

In general, the processing of chunks suggested in both flowcharts act in nested loops. This is similar to Crump and Logan’s (2010) inner-outer loop theory of typing, where the outer loop receives words from reading that are then individually passed to an inner loop that translates the word into letters for keystrokes. In our case, there are different loops that work together in a nested fashion for grouping bits of a chunk, decomposing them, transferring them individually to be compared, and then back again to process the next chunk.

Expert Flowchart, Fig. 2: For the sake of analysis the expert is assumed to chunk three words at a time, so they are predicted to have the following pattern of steps:

1. Begin by viewing stimulus, looping three times around ELP1 to create a chunk of three words. By the third loop, the WM is assumed to have reached its capacity and therefore a decision is made to end WM loading. Time elapsed to this point totals 1380ms ($3 \times (340 + 50 + 50) + (2 \times 30)$).
2. The eyes shift to the response area (time duration 30ms). With this movement, the second loop of processes (ELP2) is triggered, which includes reading the word displayed, selecting target word from WM, comparing the words, deciding and finally moving the mouse to click. Accordingly, the step duration is 990ms. The total time, from the start to the first mouse click, the initial pause, is 2400ms.
3. The clicking action of the first word takes 250ms.
4. The expert would then continue to loop through ELP2 to make their second and third response for the words “meringue” and “aardvark” respectively (Fig 4). Pauses for these two responses are both 1070ms each.
5. Once the first three-word chunk is complete, they loop up to the stimulus to gather the next chunk of three words (ELP3). The process of deciding to do this and looking up takes 160ms. This duration is the first part of the pause that precedes the first click in the next group of words.

This analysis is depicted on the solid blue line in Fig 4, which shows pause duration for successive words. The first

point is the pause before “indict”, comprised of steps 1 and 2. The second and third points are the result of step 4. The fourth point, the pause prior to “ingenue”, is comprised of step 5 and 1 again. Hence, experts are expected to exhibit long pauses for grouping words into chunks, with short pauses between responses from within the chunks.

Novice Flowchart, Fig. 3: A novice is assumed to process unfamiliar words by breaking them into parts and then regrouping them to form a chunk. Therefore, for modeling purposes a novice would process a word by the number of syllables it contains. In Fig. 4, the first half of the words are disyllabic while the others are trisyllabic. Hence, a novice’s steps for processing are assumed as follows:

1. Begin by looping through NLP1 twice taking 910ms ($2 \times (340 + 50 + 50) + 30$). They then move their eyes to the response area (30ms) to process the presented word and make a move to click (990ms). So, prior to making their first click their total pause would be 1930ms.
2. Next, they click to make a response (250ms).
3. Finally, they would loop up for the next word, NLP2, with the duration for deciding, gazing up and locating the next item is 110ms. This will be calculated as part of the pause that precedes the next response click. These pause durations are represented on the solid orange line in Fig 4. While the first point is comprised of process 1, the rest are composed of processes 1 and 3. The small rise in the duration of the final three points to 3440ms is the result of processing trisyllabic words, where the number of times they loop through NLP1 (in step 1) would increase to three. Accordingly, a novice is predicted to experience long pauses between all clicks, and slightly longer pauses when the number of syllables in a word increases. Overall, the predicted profiles of the expert and novice are substantially different.

Effects of Various Factors

Other factors and their potential effects were analyzed in the same manner. By changing the display of the stimuli from constant display to voluntary view, the stimulus is now concealed and may only be revealed by hovering the mouse over it in the top area. As a result, extra processes are added to the expert’s and novice’s models for the hover actions. This increases the lengths of long pauses, so further increases the difference in profiles between experts and novices in Fig. 4 for the voluntary view condition, with two of the expert’s pauses increasing, whereas all the novice’s pauses are higher. The first half of rows in Table 1 summarizes all of the separate pieces of analysis for the

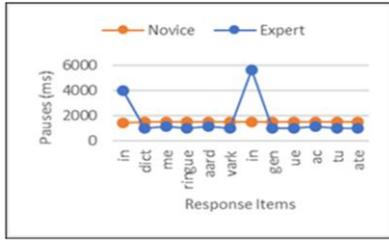


Figure 5: Pause pattern in matching parts of words with parts of words

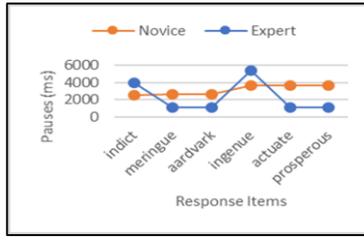


Figure 6: Pause pattern in matching parts of words with words

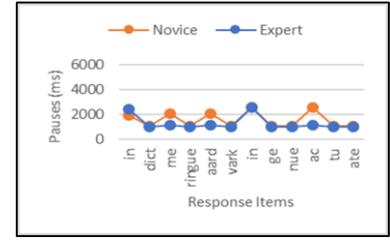


Figure 7: Pause pattern in matching words with parts of words

presentation factor, showing higher discrimination among individuals under the voluntary view mode. The median was chosen to represent the data, however in calculating the mean, a similar pattern of data existed; showing no difference in the overall results.

Models were created to analyze the effect of pairing different types of stimuli with responses, the range of data between the first row and last row of the first half of Table 1 summarizes these modelling results. In addition to matching words with words, we looked at the possibility of pairing parts of word in the stimuli with parts of words in the response (i.e., syllables with syllables). Such presentation alters the expert’s model to include two additional loops, one at the start to group syllables into words, and one at the end to decompose the chunked words back to their syllables. This in turn affects the shape of their pause pattern (Fig. 5). A novice on the other hand, is predicted to treat each syllable as a separate chunk, processing each syllable in one large loop causing them to shift their gaze frequently between syllables. Accordingly, their pause pattern is a straight line (Fig. 5).

The other possibility is to pair parts of words in the stimuli with words in the response, for example matching the syllables “in” “dict” with the word “indict”. As with the previous task, experts are expected to chunk syllables and form words in their WM and then matching them directly with whole words in the response. The graph, Fig. 6, for this model predicts that an expert’s pause pattern would be similar to that found in Fig. 4, however with an increase in the long pauses, in particular, prior to chunking trisyllabic words. If novices were assumed to treat each syllable as a separate chunk, the model predicts that they would be shifting their gaze many times prior to clicking a response causing their overall pause durations to be higher than previously seen (Fig. 6). The difference in pause measures is the highest for this task design (Table 1).

Finally matching words in the stimuli with parts of words (opposed to the above task) was tested. The expert’s pattern of pauses is similar to those found in Fig.5 however, with a decrease in the overall duration (Fig.7). On the other hand, a novice’s pause pattern differs from those depicted in Figs. 4, 5 and 6 with long pauses prior to matching the first part of that particular word (Fig. 7). The reason behind the change in pattern is due to the number of loops experienced by the novice. While their processing was always composed of either one or two loops, in this task a third loop appears at the bottom of the model for decomposing the chunk, and comparing parts. This design has the least effect on the pause measures (Table 1).

Evaluating Model Results

To test the model, a pilot study was conducted with two participants. The participants were picked and classified after assessing their vocabulary size using a standard vocabulary size test (Nation & Beglar, 2007), with the high competent (HC) individual scoring at the 16,800-word level and the less competent (LC) at the 8,100-word level.

Based on the predictions in pause measures, the pilot was developed to include four blocks of twelve trials under the conditions of matching word for word and part to word in both constant display and voluntary view. Although, the number of participants was limited, the amount of data was substantial; 48 pause measures were extracted from 384 mouse clicks per individual. The mean of median pauses was calculated for each block separately (Table 1). Overall, findings reveal that patterns between the model and observations are consistent, with the LC experiencing higher pause durations than the HC across all types of tasks. Specifically, out of the 48 trials, only two of the LC trials scored better, i.e., having shorter pauses. It is worth noting however that the value of those measures were small

Table 1: The effects of design variables on pause durations

Model vs. Pilot	Type of Display	S-R Composition	Median		
			Novice	Expert	Differ.
Modelling Results	Constant Display (CD)	Word to word	2275	1070	1205
		Part to word	3175	1120	2055
		Part to part	1520	1020	500
		Word to part	1070	1020	50
	Voluntary View (VV)	Word to word	3205	1070	2135
	Type of Display	S-R composition	Mean of Medians		
			Novice	Expert	Differ.
Pilot Experiment Results	Constant Display (CD)	Word to word	2269	1287	982
		Part to word	3856	2502	1354
	Voluntary View (VV)	Word to word	2116	942	1174
		Part to word	4235	1569	2666

(≈ 150 ms), occurring on items that contained low frequency words and would not be expected to distinguish participants well. Furthermore, confirming our predictions, higher discriminations were observed under voluntary view, especially when combined with part to word tasks.

Discussion

The aim of the present study was twofold. Firstly, to introduce the method of Competence Assessment by Stimulus Matching. CASM attempts to combine the benefits of mouse driven tasks for assessing chunking behavior (c.f., Arslan et al., 2020; Moss et al., 2006) with the benefits of temporal chunk measures for micro-behavior analysis (c.f., Albehajjan & Cheng, 2019; Cheng, 2014; Zulkifli, 2013). In other words, CASM aims to obtain measures of competence based on rich chunk level data at a time scale of 1s with the convenience of standard IT interface devices. From the preliminary results it appears that CASM has potential to compete with the freehand transcription approach and also Cheng's (2015) method that used a mouse and a selection grid. The magnitudes of predicted differences of pauses between the expert and novice are comparable to the magnitudes observed in our pilot as well as the empirical evaluation of those previous approaches.

The second aim was to explore some of the large design space of CASM tasks by using GOMS models to examine the effects of different factors on the processes of chunks. A reason for using GOMS and not a sophisticated cognitive model such as ACT-R (Anderson, 1998), is that we were looking at an efficient method for finding effective designs without all of the detail and effort required to build a full cognitive model. The aim is not to explain in precise detail all of the cognitive steps associated with doing the task, therefore what we needed was an engineering tool and not a scientific one. The produced models provided us with useful guides for designing CASM tasks, as they represent general differences in the processes of an ideal expert and an ideal novice. In between these two models would exist various intermediate levels. Someone who is gradually learning the language may behave according to a mixture of the models. Their decomposition of words may vary depending on their level of familiarity with the words presented, so their looping structure would differ. Variations at the level of individual loop structures would not affect the overall results as these differences would be reflected on the expert's and novice's models, however the number of each type of loop that exist within a model determines the difference.

In using GOMS to analyze the tasks, it was possible to assess chunks in memory and predict pause behaviors. The modelling results show how different patterns of nested loops affect the shape of pause distributions. In the task of matching words with words (Fig 4), an expert's pattern included few long pauses separated by successive short pauses, while novices were shown to have long pauses between clicks. This is explained by how their language knowledge affects the process of chunking. Experts are

expected to recognize words in a fluent manner, providing them with the advantage of loading into their memory as many words as possible (see ELP1 in Fig. 2), explaining the few long pauses. The short pauses however, are due to the transfer of words in memory from ELP1 to ELP2. Novices, on the other hand, spend time in processing a word, by breaking it apart into syllables and then regrouping them (see NLP1 in Fig. 3). This lengthy process is expected to load their WM, limiting their ability to hold one word in a chunk and causing frequent gaze shifting between responses. This indicates that behaviors are very much determined by the chunking structure of the participants.

In terms of the design space what task factors are predicted to mostly distinguish between different competence levels? First, the spatial distance between the stimulus and response plays a role in encouraging the use of chunks (Cheng, 2014). If they were close, then experts and novices might rely on quick gazes rather than chunking, causing both to exhibit similar patterns.

Second, for the presentation mode, the analysis showed no effect on the pattern of pauses but a greater difference between pause measures was identified under voluntary view (Table 1). Confirmed by the pilot study, this mode seems potentially more effective than constant display.

Third, with respect to stimulus and response composition, pairing syllables in the stimuli with words in the response seems to be the most effective option. According to GOMS, constructing the stimulus in this way makes it easier for novices to recognize a syllable and move to the response for comparison. However, the complexity of having multi-syllabic words in the response forces novices to shift their gaze as many times as required to have all parts of the word matched. Predictions were confirmed by the results of the pilot study showing longer pauses for novices in these types of tasks, making it seem most effective in exploiting the difference between experts and novices (Table 1).

Fourth, the difference between the model and pilot results are reasonably close, which drives us to conclude that there is potential for such approach. However, one explanation for the absolute difference between the model and each participant being relatively large may be due to variations in strategies within each participant. To control for that, task instructions are being tightened.

GOMS has helped in visualizing the kind of designs most suitable for developing CASM tasks that use temporal chunk measures to assess competency in natural language. We are planning on carrying out further empirical studies.

References

- Anderson, J. R. (1998). *The atomic components of thought*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Albehajjan, N., & Cheng, P. C.-H. (2019). Measuring programming competence by assessing chunk structures in a code transcription task. In *Proc. 41st Ann. Conf. of the Cog Sci Soc* (pp. 76-82).
- Arslan, B., Jiang, Y., Keehner, M., Gong, T., Katz, I. R., & Yan, F. (2020). The Effect of Drag-and-Drop Item

- Features on Test-Taker Performance and Response Strategies. *Educational Measurement: Issues and Practice*, 39(2), 96-106.
- Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc
- Cheng, P. C-H., & Rojas-Anaya, H. (2007). Measuring Mathematical Formula Writing Competence: An Application of Graphical Protocol Analysis. In *Proc. of the 29th Ann. Conf. of the Cog Sci Soc* (pp.869-874).
- Cheng, P. C.-H. (2014). Copying equations to assess mathematical competence: An evaluation of pause measures using graphical protocol analysis. In *Proc. of the 36th Ann. Conf. of the Cog Sci Soc* (pp.319-324).
- Cheng, P. C. H. (2015). Analyzing chunk pauses to measure mathematical competence: Copying equations using 'centre-click' interaction. In *Proc. of the 37th Ann. Conf. of the Cog Sci Soc* (pp.345-350).
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav Brain Sci*, 24(1), 87-114. doi:10.1017/S0140525X01003922
- Crump, M. J. C., & Logan, G. D. (2010). Hierarchical Control and Skilled Typing: Evidence for Word-Level Control over the Execution of Individual Keystrokes. *J Exp Psychol Learn Mem Cogn*, 36(6), 1369-1380.
- Gray, W. D., & Boehm-Davis, D. A. (2000). Milliseconds matter: an introduction to microstrategies and to their use in describing and predicting interactive behavior. *Journal of Experimental Psychology: Applied*, 6(4), 322-335.
- John, B. E., & Newell, A. (1989). Cumulating the science of HCI: from s-R compatibility to transcription typing. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems* (pp.109-114).
- Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological review*, 101(2), 343-352. doi:10.1037/0033-295X.101.2.343
- Moss, J., Kotovsky, K., & Cagan, J. (2006). The Role of Functionality in the Mental Representations of Engineering Students: Some Differences in the Early Stages of Expertise. *Cognitive Science*, 30(1), 65-93.
- Nation, I.S.P. & Beglar, D. (2007) A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Olson, J. R., & Olson, G. M. (1990). The Growth of Cognitive Modeling in Human-Computer Interaction Since GOMS. *Human-comp. interaction*, 5(2-3), 221-265.
- Oviatt, S., Hang, K., Zhou, J., Yu, K., & Chen, F. (2018). Dynamic Handwriting Signal Features Predict Domain Expertise. *ACM Trans. Interact. Intell. Syst.*, 8(1), 1-21
- Schilperoord, J. (2003). On the cognitive status of pauses in discourse production. In T. Olive & C. M. Levy (Eds.), *Contemporary tools and techniques for studying writing* (pp.61-88). Dordrecht: Kluwer.
- Stahovich, T. F., & Lin, H. (2016). Enabling data mining of handwritten coursework. *Computers&graphics*, 57, 31-45.
- Zulkifli, M. (2013). *Applying Pause Analysis to Explore Cognitive Processes in the Copying of Sentences by Second Language Users*. (PhD), University of Sussex, Brighton, UK

Simulating Proficiency and Exposure Effects on Cross-language Structural Priming in Simultaneous Bilinguals

Yung Han Khoe (yunhan.khoe@ru.nl)

Centre for Language Studies, Radboud University
Nijmegen, The Netherlands

Gerrit Jan Kootstra (g.kootstra@let.ru.nl)

Centre for Language Studies, Radboud University
Nijmegen, The Netherlands

Rob Schoonen (rob.schoonen@ru.nl)

Centre for Language Studies, Radboud University
Nijmegen, The Netherlands

Stefan Frank (stefan.frank@ru.nl)

Centre for Language Studies, Radboud University
Nijmegen, The Netherlands

Abstract

Bilingual speakers are more likely to use a syntactic structure in one language if they have recently encountered that same structure in another language. This cross-language structural priming effect is predicted to be positively modulated by second language proficiency according to a developmental account by Hartsuiker and Bernolet (2017). We propose to extend this account from *sequential* bilinguals to *simultaneous* bilinguals. In this latter group, syntactic structures develop in parallel and can integrate from the onset. Therefore, we do not expect proficiency or other measures of development, such as exposure, to modulate cross-language structural priming in these bilinguals. In simulated cross-language structural priming experiments, we explored how proficiency affects priming of transitives. We use an implicit learning model of sentence production to model the simultaneous English-Spanish bilinguals in these simulations. Furthermore, we investigated whether the priming effect is modulated by exposure to the non-dominant language, which only Kutasi et al. (2018) also analyzed. We found no evidence for any modulating effects for either proficiency or exposure, which is in line with the previously reported behavioral result of Kutasi et al. (2018). Together, our modeling results and Kutasi et al.'s (2018) behavioral results support an extended version of the developmental account of cross-language structural priming that predicts a modulating effect of proficiency in *sequential* bilinguals, but not in *simultaneous* bilinguals.

Keywords: cross-language structural priming; multilingualism; proficiency; syntax; error-driven implicit learning

Introduction

Structural priming is the tendency of speakers to reuse syntactic structures that they have recently encountered. It occurs in real life discourse and it is a phenomenon that can give insight in how syntax is represented in the human mind. Structural priming has been demonstrated to occur between different languages. In a study on priming of transitives in Spanish-English bilinguals (Hartsuiker, Pickering, & Veltkamp, 2004), for example, participants were more likely to use a passive target sentence in English (e.g., “*The bottle is hit by the bullet*”) after hearing a passive Spanish sentence (“*El camión es perseguido por el taxi*”) than after hearing an

active Spanish sentence (“*El taxi persigue el camión*”). This shows that syntactic representations can be shared between languages. Cross-language structural priming has been investigated in pairs of relatively similar languages such as English and Spanish, but also in languages from different families such as English and Korean (Shin & Christianson, 2009). Cross-language priming has been demonstrated for different syntactic structures such as transitives, datives (Loebell & Bock, 2003) and genitives (Bernolet, Hartsuiker, & Pickering, 2013). It has been shown to occur not only in adults but also in children (Vasilyeva et al., 2010).

Different accounts of structural priming have been proposed. One account explains it as the result of residual activation of syntactic representations and combinatorial nodes (Pickering & Branigan, 1998). Another account explains it as the result of error-driven implicit learning (Chang, Dell, & Bock, 2006; Chang, Dell, Bock, & Griffin, 2000). In this account, prediction error leads to strengthening of connections between representations that support the use of a syntactic structure, which in turn leads to increased production of that structure, which is measurable in behavioral experiments as a priming effect.

Different models of within-language structural priming have been implemented. Specifically, the Dual-path model (Chang, 2002) was used to simulate monolingual priming of transitives in English (Chang et al., 2006) and of datives in German (Chang, Baumann, Pappert, & Fitz, 2015). It has also been extended to a bilingual model, which was used to study cross-linguistic transfer (Tsoukala, Frank, Van Den Bosch, Kroff, & Broersma, 2021) and code-switching (Tsoukala, Broersma, Van Den Bosch, & Frank, 2021), and it is the only model in which cross-language structural priming has been demonstrated (Khoe, Tsoukala, Kootstra, & Frank, 2020). A hybrid model by Reitter, Keller, and Moore (2011), in which priming is primarily activation-based, has been used to simulate priming in one language but not between different lan-

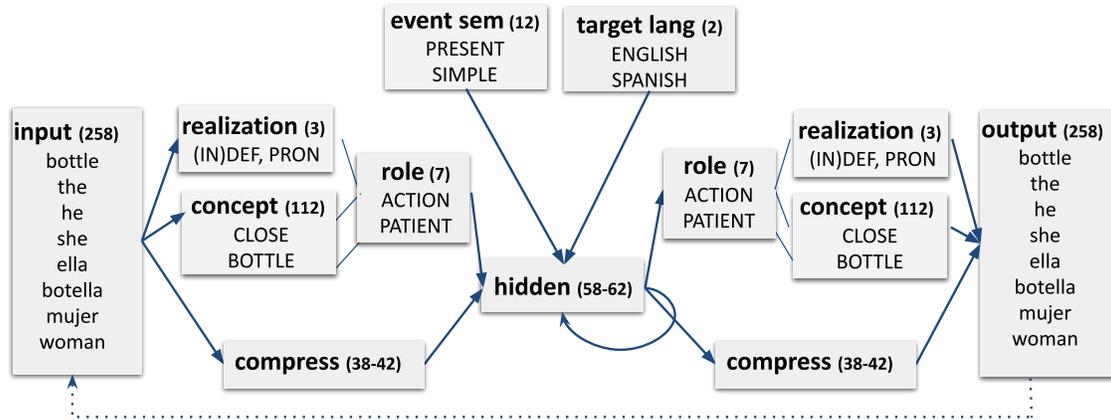


Figure 1: **Bilingual Dual-path.** The model is a next-word prediction network that converts messages into sentences. It is a simple recurrent network-based model (the lower path, via the ‘compress’ layers) that is augmented with a semantic stream (upper path) that contains information about concepts, thematic roles, event semantics, and the target language. The number of units per layer is shown in parentheses. The numbers of units for the hidden and compress layers vary across simulations. (Figure adapted from Tsoukala et al., 2017)

guages.

It is still an open question how second language (L2) proficiency affects cross-language structural priming. Hartsuiker and Bernolet (2017) have hypothesized that as L2 learners become more proficient, their L2 syntactic representations become more integrated with the representations that they already have for their native language (L1). In this developmental account, the increased integration will then result in increased cross-language structural priming.

In a number of cross-language structural priming studies, proficiency or amount of exposure to the L2 were investigated as predictors of the strength of the priming effect. In four cases, increased cross-language structural priming was found for more proficient participants. The results presented by Bernolet et al. (2013) revealed a positive effect of proficiency on the strength of priming between Dutch and English genitives. A reanalysis by Hartsuiker and Bernolet (2017) of an experiment performed by Schoonbaert, Hartsuiker, and Pickering (2007) also revealed that cross-language priming of datives in Dutch-English bilinguals was stronger for participants who were more proficient in their L2. Similarly, Favier, Wright, Meyer, and Huettig (2019) found that proficiency positively modulated priming of datives and passives from Irish to English. In their investigation of priming between Korean and English transitives, Hwang, Shin, and Hartsuiker (2018) found a priming effect that increased in magnitude as participants were more proficient in their L2, English. In contrast, three other studies did not yield evidence that proficiency modulates priming. The results reported by Hartsuiker, Beerts, Loncke, Desmet, and Bernolet (2016) for priming of relative clause attachment and of datives in multilingual speakers of Dutch (L1), French (L2), English (L2), and German (L2) did not reveal such an effect. Similarly, the results reported by Kutasi et al. (2018) for English and Gaelic transitives did not reveal any effect of either proficiency or exposure. Huang et al. (2019) also found no correlation between self-rated proficiency and the priming effect of datives in trilingual speakers of Mandarin, Cantonese, and English.

These conflicting results might be partly explained by considering the type of bilinguals that were involved in the studies. Whereas all but one of the participants in Kutasi et al.’s (2018) study were *simultaneous* bilinguals, the participants in the other studies were all or in majority *sequential* bilinguals. In our interpretation of the developmental account by Hartsuiker and Bernolet (2017), proficiency is expected to affect cross-language structural priming in *sequential* bilinguals, who start learning a second language after they have acquired their L1, but not in *simultaneous* bilinguals, who acquire their two languages at the same time. These simultaneous bilinguals would develop syntactic representations for both languages at the same time, which could integrate from the onset. The results of the study by Kutasi et al. (2018) is in line with this extended account, as they did not reveal an effect of either proficiency or exposure in the non-dominant language on cross-language structural priming.

While the number of behavioral studies on the effect of proficiency on cross-language structural priming is growing, proficiency differences have not been studied using implemented models of cross-language structural priming. In the above mentioned study by Khoe et al. (2020), the aim was to model balanced simultaneous bilinguals, and the models were therefore trained using approximately equal numbers of sentences in the two languages, that varied only minimally.

In the present work we explore the effect of proficiency and exposure in the non-dominant language on cross-language structural priming in simultaneous bilinguals, whom we model using an implicit learning model of sentence production. We do this by varying the amount of input in the two different languages that the model receives during training. We then perform cross-language structural priming experiments with these model instances as participants. We analyse the results of these experiments to determine whether proficiency or exposure in the non-dominant language modulate cross-language structural priming in the model.

Method

Model

We trained instances of the Bilingual Dual-path model¹ of sentence production (Figure 1) on miniature versions of English and Spanish to serve as simulated participants in cross-language priming experiments. The Bilingual Dual-path model extends the meaning path of the original Dual-path model (Chang, 2002) with a target language layer that indicates the intended output language.

The training input to the model consists of sentences in two artificial languages that are paired with messages that encode their meaning (see examples below, under: Artificial languages). The model instances receive input in both languages from the start of training to simulate simultaneous English-Spanish bilinguals, who start acquiring both English and Spanish from infancy. The model learns to convert a message into a sentence by predicting sentences word by word and adjusting its connection weights based on prediction error using back-propagation. A difference between the Dual-path architecture and other Recurrent Neural Networks is that the network has connections with fixed weights between concepts and roles of the message to be expressed.

Artificial languages The artificial versions of English and Spanish² that we used include the same nine sentence types for each language: Animate intransitive, Animate with-intransitive, Inanimate intransitive, Locative, Transitive (in active or passive form), Cause-motion, Benefactive transitive, State-change, and Locative alternation³. The two languages together have 275 unique lexical items. In addition to nouns, verbs, adjectives, determiners, and prepositions, these lexical items include inflectional morphemes such as a past tense marker (Spanish: ‘-pas’; English: ‘-pst’) and a past participle marker (Spanish: ‘-prf’; English: ‘-par’). The message semantics contain 121 concepts and 7 thematic roles. Only singular verbs, pronouns, nouns, and adjectives were used. Verbs and pronouns were always in third person form.

Of the transitives in our artificial languages, 75% were actives and 25% were passives. This skew in favor of actives is more in line with the frequencies of these constructions in natural language than the balanced frequencies of actives and passives that was used by Khoe et al. (2020).

In the training and test input, any message that can be expressed using two different syntactic structures has a strong bias towards one of those structures. This was implemented by creating differences in activation of thematic roles based on how each structure emphasizes those roles in the sentence. Biasing towards an active sentence (1, 2), for example, was

¹The Bilingual Dual-path model can be downloaded from: <https://gitlab.com/ykhoe/bilingual-dual-path/-/tree/ICCM2021>

²The files that the model requires to generate the artificial language input, and the input for the priming experiment can be found here: <https://github.com/khoe-yh/cross-lang-struct-priming>

³Examples for these sentence types can be found in Chang et al. (2006)

done by giving the agent a higher activation (X:1) than the patient (Y:0.5 or Y:0.75). In the same way, a bias towards a passive sentence (3, 4) was achieved with a higher activation for the patient (Y:1), than for the agent (X:0.5 or X:0.75). In the priming experiment, we gave the de-emphasized roles in target messages an activation of 0.75.

1. Spanish Active: el padre romper -pas la botella .
X = def, FATHER, M;
ACTION-LINKING = BREAK;
Y = def, BOTTLE;
EVENT-SEM = X:1, Y:0.5, PAST,
SIMPLE, ACTION-LINKING;
TARGET-LANG = es
2. English Active: the father break -pst the bottle .
[...];
EVENT-SEM = X:1, Y:0.5, [...];
TARGET-LANG = en
3. Spanish Passive: la botella es romper -prf por el padre .
[...];
EVENT-SEM = X:0.5, Y:1, [...];
TARGET-LANG = es
4. English Passive: the bottle is break -par by the father .
[...];
EVENT-SEM = X:0.5, Y:1, [...];
TARGET-LANG = en

Model training and testing We trained 120 model instances that function as simulated participants in our experiments. To simulate proficiency differences in the English-Spanish models, we trained the models with a percentage of sentences in Spanish, the non-dominant language, sampled from a truncated normal distribution (lower bound: 0%, upper bound: 50%) with a mean of 35%, and a standard deviation of 15, and the rest was in English. A set of 8,000 unique message-sentence pairs was generated for each model participant. 80% of these sentences were used for training, while 20% were set aside for testing the accuracy of the model. Following Chang et al. (2006), the message was excluded from 25% of training pairs. The models iterated over their training sets 16 times. After each of these 16 epochs, model accuracy was tested using the test set. The training set was shuffled at the beginning of each epoch.

Model configuration Differences between individual simulated participants were also created through small variation in model parameters. The number of hidden-layer units was sampled from a uniform distribution between 58 and 62, while the number of compress layer units was sampled from a uniform distribution between 38 and 42. The fixed weight value for concept–role connections was sampled from a uniform distribution between 13 and 17.

Priming experiment

Simulated participants Table 1 gives an overview of measures of proficiency and exposure for the non-dominant lan-

Table 1: Meaning accuracy, syntactic accuracy, and input in the non-dominant language (Spanish) for the 120 simulated participants in our experiment.

	Mean	Standard Deviation
Meaning accuracy	59.8%	20.0
Syntactic accuracy	95.1%	8.7
Input	29.8%	11.3

guage (Spanish) of the 120 simulated participants in our experiment. We operationalized proficiency in the non-dominant language as either syntactic accuracy or meaning accuracy in that language. Syntactic accuracy was measured as the percentage of sentences out of all test sentences for which all the words had the correct part of speech. Meaning accuracy was measured as the percentage of syntactically accurate sentences that convey the target message without any additions. Exposure to the non-dominant language was operationalized as the percentage of sentences in the training input in that language.

The standard deviations of these measures suggest that the heterogeneity in our sample of simulated participants is comparable to that in the participant samples of Kutasi et al. (2018) and Favier et al. (2019). Both studies report self-rated proficiency measures on a 7-point scale. The standard deviations for these measures ranged from 0.51 to 1.00 in the study by Kutasi et al. (2018), and from 0.61 to 1.12 in the study by Favier et al. (2019).

Experimental trials In addition to the training and test sets, we generated a single set of experimental trials that was used to perform the priming experiment on all of the model participants. Each trial consisted of a combination of a unique prime sentence and a unique target message that did not have any semantic overlap in terms of their verb, agent, and patient. Following Kutasi et al. (2018), we only use prime sentences in the non-dominant language, which in our case is Spanish. We had equal numbers of trials with active and passive primes, and equal numbers of trials with active- and passive-bias target messages. We had 50 prime-target combinations that all occurred as each of the 4 different trial types. Each experiment thus consisted of 200 trials.

Procedure The priming experiment was performed on the models after 16 training epochs. As was done by Chang et al. (2006) and Chang et al. (2015), we presented the models with prime sentences without a message, while learning was turned on in the model. After each prime, a response was elicited from the model by presenting it with a target message.

We aimed to simulate a cross-language structural priming effect that is similar in strength to what is found experimentally. Since the strength of the effect is largely determined by the learning rate, we used a range of different learning rates. In Khoe et al. (2020), a learning rate of 0.2 was used during the experiment. This resulted in priming effects that were

Table 2: Percentage of included responses, and percentage of passive sentences produced after a passive prime or after an active prime, at learning rates of 0.02, 0.04 or 0.06.

	Learning rate		
	0.02	0.04	0.06
Responses included	61.0%	60.8%	60.4%
Passives after passive prime	37.6%	38.0%	38.5%
Passives after active prime	37.4%	37.5%	37.4%

stronger than such effects found in behavioral experiments. For the present study, we therefore used learning rates between 0.02 (the learning rate at the end of training) and 0.06 (the average of the learning rates at the start and the end of training).

After each trial, the connection weights were reset to the values they had before starting the priming experiment. The state in which the model encounters each trial was thus the same for all of the trials, hence, there was no between-trial priming or any other learning effect during the experiment. This means that we did not need to (pseudo-)randomize the order of the trials across model participants.

Results

Descriptive results

Our analyses only included responses that correctly expressed the target message, with either an active or a passive structure. However, we disregarded errors involving definiteness of articles or missing periods. Table 2 shows the percentage of responses that was included on the basis of these criteria for each of the three learning rates at which the experiment was run. The table also shows the percentage of these responses that were passives after a passive prime or after an active prime.

Bayes Factor analyses

We analyzed the data from our experiment with Bayesian logistic mixed-effects models, with a logit link function, using the function `brm` from the package `brms` (Bürkner et al., 2017; Bürkner, 2018, version 2.12.0) in R (R Core Team, 2013, version 3.5.1). These analyses were not pre-registered and should therefore be considered exploratory.

The models predicted a binary dependent variable, IS PASSIVE, that indicated whether the sentence that the model produced was passive (1), or not (0). The null model included three centered continuous predictors: MEANING ACCURACY, SYNTACTIC ACCURACY, and INPUT, and two contrast-coded predictors PRIME STRUCTURE (Active = -0.5, Passive = 0.5), and TARGET-MESSAGE BIAS (Active = -0.5, Passive = 0.5). We fit random intercepts for model participants and items, as well as by-participant random slopes for PRIME STRUCTURE. The alternative models only differed from the null model in including an interaction between PRIME STRUCTURE and either MEANING ACCU-

Table 3: Bayes Factors that compare models including interactions between each of the three predictors of interest and Prime Type with a null model without any such interaction, for priming experiments with a learning rate of 0.02, 0.04, or 0.06, where the prior for the interaction had a standard deviation of either 0.5 or 1. A Bayes Factor smaller than 1 favors the null model whereas a Bayes Factor larger than 1 favors the alternative model that includes an interaction.

Learning Rate	Standard Deviation			
	0.25	0.5	0.75	1
Meaning accuracy				
0.02	0.111	0.052	0.035	0.025
0.04	0.091	0.046	0.032	0.027
0.06	0.124	0.051	0.035	0.025
Syntactic accuracy				
0.02	0.377	0.175	0.136	0.118
0.04	0.273	0.160	0.105	0.077
0.06	0.329	0.179	0.103	0.077
Input				
0.02	0.237	0.105	0.075	0.058
0.04	0.169	0.094	0.056	0.044
0.06	0.212	0.079	0.057	0.044

RACY, SYNTACTIC ACCURACY, or INPUT. We computed Bayes Factors that compare the null model to these alternative models.

We calculated Bayes Factors using bridge sampling (Bennett, 1976; Meng & Wong, 1996; Gronau et al., 2017), with four chains and 8000 iterations, including a warm-up phase of 2000 iterations. Because an uninformative prior for the predictor of interest can make a Bayes Factor biased towards the null model (Lee & Wagenmakers, 2014), we report Bayes Factors across four different values of the standard deviation (σ) for the prior of the interaction of interest (Normal(0, σ)), ranging from a value appropriate for an informative prior (i.e., $\sigma = 0.25$) to a value appropriate for a regularizing prior (i.e., $\sigma = 1$). Regularizing priors (Normal(0,1)) were used for all other predictors in our models. These priors give a minimal amount of information with the objective of yielding stable inferences. Prior means were 0, and did thus not bias towards specific effects. The only exception to this was the TARGET-MESSAGE BIAS predictor for which we excluded negative values by using a prior with a Gamma distribution (Gamma(1, 0.5)).

Table 3 shows that the Bayes Factors are all smaller than 1, and thus provide evidence in favor of the null model. Based on the scale proposed by Jeffreys (1998), we interpret this evidence as ranging from anecdotal to very strong. As expected, when a smaller standard deviation is used for the prior, the Bayes Factors are mostly closer to 1, and thus provide less

conclusive evidence for the null model. The Bayes Factors do not suggest a clear effect of learning rate on the strength of the evidence for the null model.

Null model estimates

Because our exploratory analysis does not yield any evidence for modulating effects of proficiency or exposure on priming, we do not report estimates from the analyses that included interactions between PRIME STRUCTURE and any of our three predictors of interest. Instead, we provide a summary of the results from the null models for priming experiments with three different learning rates in Table 4. In line with our expectations, the estimates for the PRIME STRUCTURE predictor are higher for higher learning rates. The credible intervals for the PRIME STRUCTURE predictor contain only positive values at learning rates of 0.04 and 0.06, which indicates strong evidence for a priming effect. At a learning rate of 0.02, the credible interval that includes some negative values indicates weaker evidence for a priming effect.

Discussion

In the present work, we explored whether proficiency or exposure modulate cross-language structural priming in simultaneous bilinguals, simulated using an implicit learning model of sentence production. Our results indicate anecdotal to strong evidence against such modulating effects in the model. This is in line with the results reported by Kutasi et al. (2018). Taken together, those behavioral results and our modeling results provide support for an extended version of the developmental account of cross-language structural priming (Hartsuiker & Bernolet, 2017) that not only predicts a modulating effect of proficiency in *sequential* bilinguals, but that also explicitly predicts the absence of such an effect in *simultaneous* bilinguals.

Limitations and further work

One limitation of our simulations lies in a difference between the languages and syntactic structures involved in our simulated experiments and those in the experiments that Kutasi et al. (2018) conducted. The main question that Kutasi et al. (2018) addressed in their study, was whether cross-language priming can occur for structures with different word order between languages. For this reason, they studied bilinguals who spoke English and Scottish Gaelic, for which active as well as passive word order is different. In contrast, the English and Spanish transitives in our experiments have the same word order between the two languages for both actives and passives. We could therefore come closer to simulating the results from Kutasi et al. (2018) by using the English-Dutch model reported on by Khoe et al. (2020) in which English passives are verb-medial, while Dutch passives are verb-final.

The participants that were involved in the other studies that investigated the possible modulating effect of proficiency on cross-language structural priming were sequential bilinguals. An obvious follow up to the present study is to simulate

Table 4: Summary of the fixed effects in the Bayesian logistic mixed-effects null models with different learning rates ($N = 14, 633, 14,594, \text{ and } 14,491$, for experiments with learning rates of 0.02, 0.04, and 0.06 respectively).

Predictor	Learning rate	Estimate			95% CrI			P(Est. > 0)		
		0.02	0.04	0.06	0.02	0.04	0.06	0.02	0.04	0.06
INTERCEPT		1.04	1.02	0.83	[-0.29, 2.62]	[-0.14, 2.39]	[-0.16, 1.99]	0.93	0.96	0.95
PRIME STRUCTURE		0.52	1.08	1.30	[-0.38, 1.45]	[0.29, 1.91]	[0.49, 2.12]	0.87	1.00	1.00
TARGET-MESSAGE BIAS		27.59	25.16	22.98	[21.91, 34.83]	[20.36, 31.41]	[18.98, 28.03]	1.00	1.00	1.00
MEANING ACCURACY		0.05	0.09	0.08	[-0.06, 0.17]	[-0.00, 0.20]	[0.00, 0.17]	0.81	0.97	0.97
SYNTACTIC ACCURACY		0.01	-0.08	-0.08	[-0.26, 0.29]	[-0.31, 0.15]	[-0.29, 0.13]	0.54	0.26	0.23
INPUT		-0.12	-0.09	-0.07	[-0.27, 0.02]	[-0.22, 0.03]	[-0.18, 0.03]	0.05	0.07	0.09

cross-language structural priming in these sequential bilinguals, and to determine whether proficiency or exposure does modulate priming in these simulations, as predicted by the developmental account of Hartsuiker and Bernolet (2017).

References

- Bennett, C. H. (1976). Efficient estimation of free energy differences from Monte Carlo data. *Journal of Computational Physics*, 22(2), 245–268.
- Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2013). From language-specific to shared syntactic representations: The influence of second language proficiency on syntactic sharing in bilinguals. *Cognition*, 127(3), 287–306.
- Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, 10(1), 395–411.
- Bürkner, P.-C., et al. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), 1–28.
- Chang, F. (2002). Symbolically speaking: A connectionist model of sentence production. *Cognitive Science*, 26(5), 609–651.
- Chang, F., Baumann, M., Pappert, S., & Fitz, H. (2015). Do lemmas speak German? a verb position effect in German structural priming. *Cognitive Science*, 39(5), 1113–1130.
- Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113(2), 234.
- Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, 29(2), 217–230.
- Favier, S., Wright, A., Meyer, A., & Huettig, F. (2019). Proficiency modulates between-but not within-language structural priming. *Journal of Cultural Cognitive Science*, 3(1), 105–124.
- Gronau, Q. F., Sarafoglou, A., Matzke, D., Ly, A., Boehm, U., Marsman, M., ... Steingroever, H. (2017). A tutorial on bridge sampling. *Journal of Mathematical Psychology*, 81, 80–97.
- Hartsuiker, R. J., Beerts, S., Loncke, M., Desmet, T., & Bernolet, S. (2016). Cross-linguistic structural priming in multilinguals: Further evidence for shared syntax. *Journal of Memory and Language*, 90, 14–30.
- Hartsuiker, R. J., & Bernolet, S. (2017). The development of shared syntax in second language learning. *Bilingualism: Language and Cognition*, 20(2), 219.
- Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science*, 15(6), 409–414.
- Huang, J., Pickering, M. J., Chen, X., Cai, Z., Wang, S., & Branigan, H. P. (2019). Does language similarity affect representational integration? *Cognition*, 185, 83–90.
- Hwang, H., Shin, J.-A., & Hartsuiker, R. J. (2018). Late bilinguals share syntax unsparingly between L1 and L2: Evidence from crosslinguistically similar and different constructions. *Language Learning*, 68(1), 177–205.
- Jeffreys, H. (1998). *The theory of probability*. Oxford, Oxford University Press.
- Khoe, Y. H., Tsoukala, C., Kootstra, G. J., & Frank, S. L. (2020). Modeling cross-language structural priming in sentence production. In *Proceedings of the 18th International Conference on Cognitive Modeling (ICCM 2020)*.
- Kutasi, T., Suffill, E., Gibb, C. L., Sorace, A., Pickering, M. J., & Branigan, H. P. (2018). Shared representation of passives across Scottish Gaelic and English: Evidence from structural priming. *Journal of Cultural Cognitive Science*, 2(1-2), 1–8.
- Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.
- Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics*, 41(5; ISSU 387), 791–824.
- Meng, X.-L., & Wong, W. H. (1996). Simulating ratios of normalizing constants via a simple identity: a theoretical exploration. *Statistica Sinica*, 831–860.
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, 39(4), 633–651.
- R Core Team. (2013). R: A language and environment for statistical computing.
- Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science*, 35(4), 587–637.
- Schoonbaert, S., Hartsuiker, R. J., & Pickering, M. J. (2007).

- The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, 56(2), 153–171.
- Shin, J.-A., & Christianson, K. (2009). Syntactic processing in Korean–English bilingual production: Evidence from cross-linguistic structural priming. *Cognition*, 112(1), 175–180.
- Tsoukala, C., Broersma, M., Van Den Bosch, A., & Frank, S. L. (2021). Simulating code-switching using a neural network model of bilingual sentence production. *Computational Brain & Behavior*, 4, 87–100.
- Tsoukala, C., Frank, S. L., & Broersma, M. (2017). “He’s pregnant”: Simulating the confusing case of gender pronoun errors in L2 English. In *the 39th Annual Meeting of the Cognitive Science Society (CogSci 2017)* (pp. 3392–3397).
- Tsoukala, C., Frank, S. L., Van Den Bosch, A., Kroff, J. V., & Broersma, M. (2021). Modeling the auxiliary phrase asymmetry in code-switched Spanish–English. *Bilingualism: Language and Cognition*, 24, 271–280.
- Vasilyeva, M., Waterfall, H., Gámez, P. B., Gómez, L. E., Bowers, E., & Shimpi, P. (2010). Cross-linguistic syntactic priming in bilingual children. *Journal of Child Language*, 37(5), 1047–1064.

Toward Undifferentiated Cognitive Models

Colin Kupitz¹ (colin.kupitz.ctr@us.af.mil), Aaron Eberhart² (aaron.eberhart@gmail.com)
Daniel Schmidt³ (danielschmidt98@gmail.com), Christopher Stevens³ (christopher.stevens.28@us.af.mil)
Cogan Shimizu² (coganshimizu@ksu.edu), Pascal Hitzler² (hitzler@ksu.edu)
Dario D. Salvucci⁴ (salvucci@drexel.edu), Benji Maruyama³ (benji.maruyama@us.af.mil),
& Christopher W. Myers³ (christopher.myers.29@us.af.mil)

¹Oak Ridge Institute for Science & Education at AFRL

²Kansas State University

³Air Force Research Laboratory

⁴Drexel University

Abstract

Autonomous systems are a new frontier for pushing sociotechnical advancement. Such systems will eventually become pervasive, involved in everything from manufacturing, healthcare, defense, and even research itself. However, proliferation is stifled by the high development costs and the resulting inflexibility of the produced systems. The current time needed to create and integrate state of the art autonomous systems that operate as team members in complex situations is a 3-15 year development period, often requiring humans to adapt to limitations in the resulting systems. A new research thrust in interactive task learning (ITL: Laird et al., 2017) has begun, calling for natural human-autonomy interaction to facilitate system flexibility and minimize users' complexity in providing autonomous systems with new tasks. We discuss the development of an undifferentiated agent with a modular framework as a method of approaching that goal.

Keywords: cognitive model; cognitive agent; instruction following; learning

Introduction

Autonomous systems are a new frontier for socio-technical advancement. Such systems will be required to team with humans, potentially operating at the level of peers and not just subordinates. One such *autonomous synthetic teammate* (AST) demonstrated that they can be included in teams without detriment to teams' or team members' performance (McNeese, Demir, Cooke, & Myers, 2017; Myers et al., 2019). Nonetheless, there remain two significant obstacles to the wider adoption of synthetic agents operating as peers or subordinates in complex environments: 1) their development and evaluation time and 2) their limited scope of transfer once developed. The AST took approximately nine years to develop plus an additional year to evaluate (Ball et al., 2010; Rodgers, Myers, Ball, & Freiman, 2013), and yet it would require further research and development to adapt it to perform a different task within the same domain and even more to adapt it to an entirely new domain.

Instructions are a ubiquitous part of the human experience. They provide guidance through a space of potential states to a solution (i.e., the problem space; Newell and Simon 1972). Without instruction, one is free to roam about the problem space freely in an attempt to find, or discover, the solution. Instruction also plays a critical role in our ability to advance as a civilization: calculus, laws of physics, and other advanced domains do not need to be re-discovered by each generation, but are taught through instruction. Although learning is sometimes characterized as the acquisition of all skills

needed for a given task, in fact, the learning of complex tasks more typically reflects the integration of already-known processes (e.g., interactive routines; Gray 2008) in novel ways (Gray, Sims, Fu, & Schoelles, 2006) – of which one way is through instruction (Salvucci, 2013).

Recent advances have demonstrated the ability to turn a set of instructions into declarative knowledge that is then used to enable performance across paradigms of different complexity: ranging from those typically used by experimental psychology to dialing while driving an automobile (Salvucci, 2013). In a similar vein, Kirk, Mininger, & Laird (2016) have successfully demonstrated the ability to train robots on novel tasks through direct interaction. The objective associated with the presented research is to leverage past work on instruction learning to address the development and transfer issues, simultaneously. Specifically, we propose a generalizable, *undifferentiated agent* (uAgent) that can learn a new task relatively independently through written instruction and be trained to a desired level of proficiency with reduced developer intervention.

To achieve these goals, the uAgent was developed with a modular architecture, to allow for expansion into other tasks and fields with minimal burden to other researchers. The components of the architecture are instruction parsing, an ontology of instruction, declarative memory representation, and procedures for accomplishing the instructed tasks. Each of the uAgent components are discussed in the following sections, followed by their integration as a single system. Finally, a case study on development times relative to current approaches to model development times is presented.

In the following sections, we will discuss this development in each of the discretized modules that together represent the uAgent, along with the general specifications for said modules to allow for future revision and expansion. Second, the uAgent will be specialized to desired levels of proficiency using AFRL's Autonomous Research System (ARES; Nikolaev et al., 2016).

Toward Undifferentiated Cognitive Models

The undifferentiated agent, or uAgent, is a system capable of learning new tasks through instruction. The process involves: (1) parsing the instructions in to a structure that can be (2) integrated with prior information within a declarative memory system through an ontology of instruction. Given

an integrated declarative system, the uAgent (3) associates it newly acquired knowledge with existing interactive routines through a controlled vocabulary. Finally, the uAgent is ready to (4) perform the task based on its knowledge of instructions. Components associated with each step will be discussed in detail, below.

Instruction Parsing

Though many advances have been made in the field of natural language processing (NLP), it still remains a challenging problem to extract complex rules and meanings out of text. To make the problem of parsing text more tractable, we use a controlled natural language - Attempto Controlled English (ACE; Fuchs, Kaljurand, & Kuhn, 2008).

A controlled natural language is a language that permits only a subset of grammatical constructions available in natural language (in this case, only present perfect tense, no use of second person, and a very specific syntax for commands). These restrictions make it possible for software (e.g., the Attempto Parsing Engine, APE; Fuchs et al., 2008) to automatically parse sentences written in the controlled language into logical statements called discourse representation structures (DRS). These structures approximate first-order logic.

The requirements of the Instruction Parsing module are as follows. First, the language requirements of incoming instructions must be specified (here, we use the ACE controlled language). Second, these instructions must be processed into a form compatible with the target declarative memory system structure to be used by the acting uA-

In the current system, we begin with plain English instructions of a task of interest. Two types of tasks we are currently working with include basic experimental psychology tasks - psychomotor vigilance (Dinges & Powell, 1985) and visual search (Treisman & Gelade, 1980) - and a material engineering task in which an individual guides a set of experiments with a 3D printer (Nikolaev et al., 2016). In both cases, instructions are re-written by hand into sentences that follow the rules of the ACE language. Then, the instructions are provided to APE¹ to translate the ACE sentences into DRS structures, which are then integrated into a declarative memory structure based on an ontology of instruction.

Ontology of Instruction

The primary function of the Ontology module is to directly formalize the structure of information necessary to complete the desired tasks and goals of the uAgent. This furthers the goal of the uAgent as a whole, as it provides the foundation for the structure and relationships within the declarative memory system used by the uAgent (see Figure 1).

In order to create a system that is both generalizable and able to correctly handle diverse types of instructions, an ontology was created that is capable of representing instructions for a cognitive agent task. This instruction ontology can be

¹<http://attempto.ifi.uzh.ch/site/resources>

English:

You will be seated in front of a computer screen.
A letter will appear in the middle of the screen.
When you see the letter, press the spacebar.

ACE:

p:psychomotorVigilance is a task.
There is a screen.
There is a letter.
There is a subject.
The n:spacebar is a button.
If the task is active then the subject v:watchesFor the letter and the letter v:appearsOn the screen.
If the letter v:appearsOn the screen then the subject presses the n:spacebar. The task is active.

Table 1: PVT instructions in English and ACE.

used to directly inform relationships among tokens of knowledge within a cognitive agent performing tasks. Further, it can be leveraged to derive a semantically-anchored declarative memory system for long-term storage for knowledge, such as a knowledge graph (Noy et al., 2019). It can also support experiment design, irrespective of any agent, by providing a structured basis for evaluating the content and design of similar tasks. Additionally, because an ontology contains a precise axiomatization of the knowledge it is supposed to represent, deductive reasoning techniques can be applied to detect possible gaps or errors in instructions. Further information regarding the ontology can be found in (Eberhart et al., 2020).

The ontology was developed to represent the relationships between steps, items, actions, instructions associated with tasks relying on a graphical user interface. To ensure a potentially high degree of complexity in instructions, the multi-stage Intelligence, Surveillance, & Reconnaissance Multi-Attribute Task Battery (ISR-MATB) task (Frame et al., 2019) was used as an example task when developing the ontology. Because it has multiple interconnected cognitive tasks, using the ISR-MATB aids in the development of a undifferentiated representation of instruction knowledge. The ontology was produced by following the Modular Ontology Modeling (MOMo) methodology, outlined in (Krisnathi & Hitzler, 2016; Hitzler & Krisnathi, 2018; Shimizu, Hammar, & Hitzler, 2021), and is designed to ensure high quality and reusability of the ontology. The adaptability required to model the ISR-MATB task, together with the modular techniques used to create it, mean that the ontology can very easily be adapted for use in new tasks.

Currently, DRS items from instruction are obtained as input to the ontology whenever an agent begins learning through instruction (see Figure 1). The DRS structured information is then available to an agent during a task, and additional knowledge that the agent acquires can be added to supplement this. As new tasks are implemented and tested

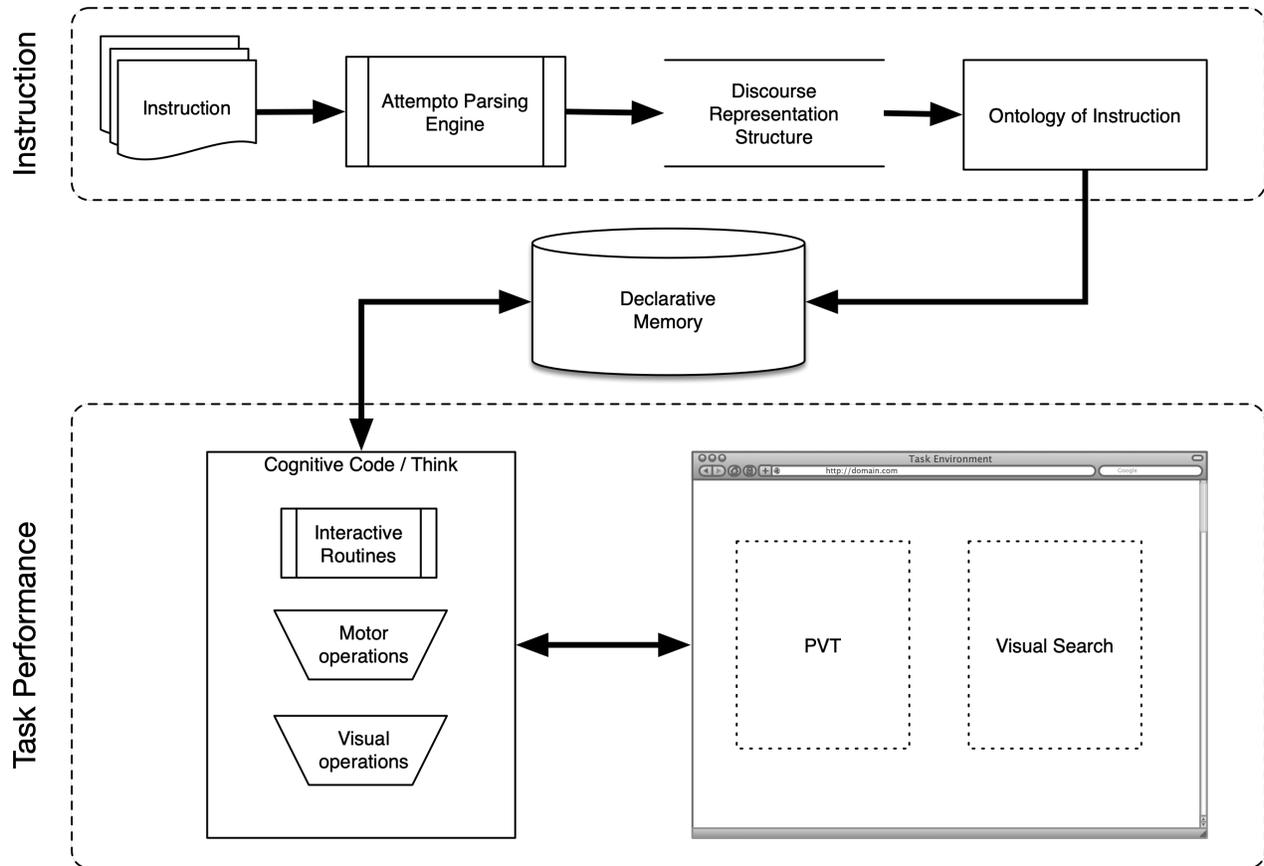


Figure 1: Architectural components of the undifferentiated cognitive agent (uAgent).

this process is simple to extend to encompass new types of knowledge, since the structure of the ontology and the format of input data is agnostic to the actual content of the knowledge represented.

To summarize, the Ontology leverages information theory and formal logical structures to ensure that pertinent information is assimilated in the most reasonable, orderly fashion possible from a theoretical standpoint. Importantly, this ensures that any future expansions of the uAgent into additional research fields and tasks will be expedited, as any additional pertinent information can be distilled directly into the most useful form through the ontology and into the uAgent’s declarative memory system.

Declarative Memory System

The declarative memory system of the uAgent contains information associated with parsed instructions, prior knowledge, and a controlled vocabulary connecting verbs to known procedures in the procedural system. The approach taken to represent the declarative memory system was a knowledge graph (Noy et al., 2019), which describes facts, actions, objects of interest, and the relationships between them.

The uAgent here stores a knowledge graph built from declarative chunks. Specifically, it incorporates chunks that,

using ACT-R-like slot-value pairs (Anderson, 2007), links knowledge together in a graph by having one chunk’s slots include other chunks as values for those slots. As such, the representation is flexible enough to incorporate all the declarative knowledge needed in the instructions for our purposes, including not only basic actions but also conditionals and sequences of actions.

A declarative memory system structured as a knowledge graph requires connections to real action/observable behavior to ground the information in the actions available within the instructed task. Without grounding, the agent can have all of the available information about the task but no way to observe or interact with its task environment. To this end, a controlled vocabulary (CV) was introduced to map verbs onto concepts or actions. For example, a CV entry for “search” could map onto an interactive routine (Gray, 2008) instructing the agent to attend a location, locate an item there, and encode it. Within the ACT-R paradigm, this led to creating a new class of chunks for CV entries. These chunks contain the CV term and map to a production or set of procedures built into the agent prior to instruction, thereby grounding that term onto a known set of actions (Ji, van Rij, & Taatgen, 2019).

Novel task strategies can be constructed using these grounded interactive routines, thereby allowing the agent to

interact with an environment for which it was not specifically designed and perform tasks without needing to have a whole set of bespoke procedures and knowledge built into it.

Altogether, the CV defines the set of verbs which are already grounded to behavior(s); in essence, it represents the agent’s knowledge of general behaviors *a priori*. Accordingly, by relating the CV to task appropriate interactive routines, we ensure the knowledge is inherently grounded to the environment.

As a module, the knowledge graph serves as the basis of the uAgent memory: it contains information pertinent to the instructions given, but processed through the lens of overall task knowledge it should have before hand (i.e. the Ontology). It must follow the format of the structures provided within the formal ontology, and further, should use a defined controlled vocabulary to map those terms onto active agent behaviors where appropriate.

Procedural Memory System

Given the above declarative memory structures for representing instructions, the system needs to ground concepts to simulated actions via interactive routines (i.e., embedded or learned procedural knowledge). Models developed in cognitive architectures such as ACT-R (Anderson, 2007) or Soar (Laird, 2012) typically use production systems to represent this procedural knowledge. Here, we take a different approach, using *cognitive code* (Salvucci, 2016) to maintain and execute procedural knowledge. Cognitive code embeds procedural knowledge into a common programming language, facilitating the development of model code while maintaining the most important properties of human-like abilities and limitations inherent to any cognitive architecture. Specifically, we are using the *Think* architecture ², which incorporates declarative and procedural concepts taken primarily from ACT-R and provides them for easy use via the Python programming language.

Several components of this project have led to important extensions of Think’s code base. One extension involves the integration of traditional declarative memory with Think execution. The default Think code base includes a declarative memory module that embodies ACT-R’s core theory of memory (Anderson, 2007). For this project, we bypass this traditional memory module, and instead use the ontology and knowledge graph described earlier as the model’s primary long-term declarative storage. The Think procedures still maintain short-term declarative items, namely those that comprise the current “context” during execution (i.e., information that would traditionally be stored in ACT-R’s *imaginal* buffer).

Besides this integration of a new type of declarative memory, the other critical extension of Think’s code base relates to the realization of procedural learning. Although cognitive code can often be made to operate in ways very similar to traditional production systems, a critical difference is that cogni-

tive code cannot (in most cases) be constructed during simulation as some architectures have done with procedural learning. For example, ACT-R’s production compilation mechanism (Taatgen & Lee, 2003)) transforms declarative instructions into procedural form which eventually leads to gradual learn of new procedures; the most critical aspect of this learning is that, at first, a model must perform a declarative retrieval to remember the learned instruction before executing it, but later, the compiled instruction (in the form of a production rule) skips the retrieval and simply executes the associated action. Although Think does not create new code on the fly in the same way, we have augmented its capabilities by adding procedural learning that captures the essence of ACT-R’s production compilation—specifically, in performing declarative retrievals early in learning (which take additional time and may fail), and then skipping these retrievals later in learning (leading to gradual speedup and eventually fast performance).

As a module, the cognitive code contained with Think serves as the “actual” uAgent, so to speak – it represents the system which is deciding and acting upon the best course of behavior during any task. In theory, this could be replaced with any number of cognitive architectures, provided they are capable of using the prespecified knowledge graph structures to serve as the basis of memory, and further, have a correctly specified controlled vocabulary to map that knowledge graph onto the behaviors known to the system *a-priori*.

System Integration

To develop the uAgent with an adaptable framework going forward, we used a modular design approach (Bryson, 2000). In particular, this capitalizes on the interdisciplinary nature of the researchers involved while simultaneously minimizing the overall burden of coordination. To that end, during development the fundamental uAgent capabilities were segregated into discrete modules. Overall integration of these modules was then assigned to a few individuals, with the entire research team meeting to discuss overall design strategies as appropriate. Of note, this approach also allowed for a degree of asynchronous development across the research teams involved, thereby reducing the project coordination burden significantly. In addition, the modular approach ensures that the uAgent will be adaptable to other fields of research and task performance, as future research can adapt the uAgent by focusing on a specific uAgent module where appropriate. We now move on to discuss the primary modules of interest in the uAgent.

Given the interdisciplinary nature of this research, we first settled on the use of the open source Python as the primary programming language, integrating each individual uAgent module into one coherent system. In particular, this allows us to utilize the Think system (Salvucci, 2021) in order to simulate both the uAgent behavior, and the environment in which it is actively behaving. Further, whenever these separate modules are expected to interact directly, we worked to determine the best overall form of interface and information exchange

²<https://github.com/salvucci/think>

to facilitate ease of integration and future expansion. To that end, we now note the primary interface decisions we made during said development.

First, we concluded that the Ontology of instruction should serve as a form of blueprint for the knowledge graph. This ensures that the Instruction Parsing module will produce structures that can be assigned to knowledge graph structures where appropriate. Effectively we are leveraging the relations inherent to the Ontology in order to improve the capabilities of the instruction interpretation; in essence, the uAgent can make informed assumptions about the informational structure while processing any incoming instructions.

Similarly, to ensure the agent is capable of acting on those instructions, we concluded that the knowledge graph module should also consider a controlled vocabulary representing the behaviors found within the Think cognitive agent. This controlled vocabulary is essentially the actions that the think uAgent is capable of performing in the current environment. In essence, we ensure that the knowledge graph structures which serve as the basis of the Think agent memory also have a direct mapping onto Think behaviors where appropriate.

Altogether, we integrate each of the uAgent modules into a coherent end-to-end system, and explicitly define the interface requirements necessary to ensure the system can take instructions as input, and produce human behavior with high fidelity.

Case Study

As a proof-of-concept for the approach, we built an end-to-end system that takes ACE instructions of cognitive tasks commonly used in basic research - psychomotor vigilance and visual search - converts the instructions into a knowledge representation capable of performing the task, and then performs the task in a simulated environment.

As an exercise to determine if the current approach could save time with respect to building a traditional ACT-R model, we compared the amount of time it took to build a model of a set of cognitive tasks with the amount of time it took to write ACE instructions of the same task. The task we used was a novel task battery that includes a set of commonly used experimental psychology tasks (Frame et al., 2019; Eberhart et al., 2020). This battery includes four subtasks - psychomotor vigilance, visual search, auditory search, and multi-cue decision-making. We built a model of the task in a Java implementation of ACT-R 6 and wrote a set of ACE instructions for it.

It took approximately 120 hours to build the ACT-R model, but only approximately 30 hours to write the ACE instructions. This exercise suggests that the present method could potentially save a substantial amount of time in developing new models and agents. Moreover, writing the ACE instructions required only a brief reading of publicly available tutorials on the ACE language, and not training and experience in writing ACT-R models, the latter of which can be substantial. In our proof-of-concept system, we showed that the ACE

instructions of the PVT and Visual Search subtasks could be successfully integrated into the ontology and the agent could use this resulting knowledge to perform the task. We are working toward end-to-end demonstrations of the other two subtasks.

Conclusions and Future Work

Progress toward a modeling framework capable of being taught new tasks through written instruction was presented. As evidenced in the uAgent case study, such an approach will likely significantly reduce model and agent development times. Further, the modular-based approach toward uAgent development will facilitate the integration of other cognitive architectures by using the uAgent declarative memory as its knowledge repository.

While the uAgent shows promise as a means for teaching models how to perform new tasks, multiple challenges remain. For example, it is unreasonable to assume that the union of instruction and prior knowledge is sufficient for completing an instructed task. As a result, we have begun developing approaches for detecting and resolving gaps in a uAgent's knowledge base. This work will require multidisciplinary approaches to model development coupled with empirical investigations into when and how humans detect and resolve knowledge gaps.

In order to better understand how humans form representations from instructions and identify and resolve gaps in understanding from those instructions, we plan to conduct a human-subjects experiment using the task battery described above. We plan to teach participants to perform the tasks in the battery using either a complete set of instructions, or a set with ambiguities with respect to certain types of knowledge. We plan to use think-aloud protocols to track how participants extract knowledge from these instructions and how they detect and resolve uncertainty. We believe this will provide insights in how to improve the undifferentiated model's knowledge acquisition.

Acknowledgments

This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-18-1-0386 to Kansas State University, 19RHCOR089 3003B & 3003C to AFRL, and FA9550-18-1-0371 to Drexel University.

References

- Anderson, J. R. (2007). *How can the human mind exist in the physical universe?* Oxford University Press.
- Ball, J., Myers, C., Heiberg, A., Cooke, N. J., Matessa, M., Freiman, M., & Rodgers, S. (2010, aug). The synthetic teammate project. *Computational and Mathematical Organization Theory*, 16(3), 271–299. doi: 10.1007/s10588-010-9065-3
- Bryson, J. (2000). Cross-paradigm analysis of autonomous agent architecture. *Journal of Experimental & Theoretical Artificial Intelligence*, 12(2), 165–189.

- Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual rt task during sustained operations. *Behavior research methods, instruments, & computers*, 17(6), 652–655.
- Eberhart, A., Shimizu, C., Stevens, C., Hitzler, P., Myers, C. W., & Maruyama, B. (2020). A Domain Ontology for Task Instructions. In B. Villazón-Terrazas, F. Ortiz-Rodríguez, S. M. Tiwari, & S. K. Shandilya (Eds.), *Knowledge graphs and semantic web. second iberoamerican conference and first indo-american conference, kgswc 2020* (pp. 1–13). Mérida, Mexico: Communications in Computer and Information Science, vol. 1232.
- Frame, M., Lopez, J., Myers, C., Stevens, C., Estep, J., & Boydston, A. (2019). Development of an autonomous management system for human-machine teaming with multiple interdependent tasks. In *Presented to the annual meeting of the psychonomic society conference, montreal, qc, canada, november 2019*.
- Fuchs, N. E., Kaljurand, K., & Kuhn, T. (2008). Attempto controlled english for knowledge representation. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 5224 LNCS, pp. 104–124). doi: 10.1007/978-3-540-85658-0_3
- Gray, W. D. (2008). The Interactive Routine as Key Construct in Theories of Interactive Behavior. In V. Sloutsky, B. Love, & K. McRae (Eds.), *30th annual meeting of the cognitive science society* (p. 127). Austin, TX.
- Gray, W. D., Sims, C. R., Fu, W.-T., & Schoelles, M. J. (2006). The soft constraints hypothesis: a rational analysis approach to resource allocation for interactive behavior. *Psychological Review*, 113(3), 461–482.
- Hitzler, P., & Krisnadhi, A. (2018). A tutorial on modular ontology modeling with ontology design patterns: The cooking recipes ontology. *CoRR*, abs/1808.08433. Retrieved from <http://arxiv.org/abs/1808.08433>
- Ji, M. Y., van Rij, J., & Taatgen, N. A. (2019). Discoveries of the algebraic mind: A PRIMS model. *Proceedings of ICCM 2019 - 17th International Conference on Cognitive Modeling*, 71–76.
- Krisnadhi, A., & Hitzler, P. (2016). Modeling with ontology design patterns: Chess games as a worked example. In P. Hitzler, A. Gangemi, K. Janowicz, A. Krisnadhi, & V. Presutti (Eds.), *Ontology engineering with ontology design patterns – foundations and applications* (Vol. 25, pp. 3–21). IOS Press.
- Laird, J. E. (2012). *The Soar Cognitive Architecture*. MIT Press.
- Laird, J. E., Gluck, K., Anderson, J., Forbus, K. D., Jenkins, O. C., Lebiere, C., ... Kirk, J. R. (2017). Interactive Task Learning. *IEEE Intelligent Systems*, 32(4), 6–21. doi: 10.1109/MIS.2017.3121552
- McNeese, N. J., Demir, M., Cooke, N. J., & Myers, C. (2017). Teaming With a Synthetic Teammate: Insights into Human-Autonomy Teaming. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 001872081774322. doi: 10.1177/0018720817743223
- Myers, C., Ball, J., Cooke, N., Freiman, M., Caisse, M., Rodgers, S., ... McNeese, N. (2019). Autonomous Intelligent Agents for Team Training. *IEEE Intelligent Systems*, 34(2), 3–14. doi: 10.1109/MIS.2018.2886670
- Newell, A., & Simon. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice Hall.
- Nikolaev, P., Hooper, D., Webber, F., Rao, R., Decker, K., Krein, M., ... Maruyama, B. (2016). Autonomy in materials research: a case study in carbon nanotube growth. *Nature Partner Journal: Computational Materials*, 2(1), 16031. Retrieved from <http://www.nature.com/articles/npjcompumats201631> doi: 10.1038/npjcompumats.2016.31
- Noy, N., Gao, Y., Jain, A., Narayanan, A., Patterson, A., & Taylor, J. (2019). Industry-scale knowledge graphs: Lessons and challenges. *Communications of the ACM*, 62(8), 36–43. doi: 10.1145/3331166
- Rodgers, S., Myers, C., Ball, J., & Freiman, M. (2013). Toward a situation model in a cognitive architecture. *Computational and Mathematical Organization Theory*, 19, 313–345.
- Salvucci, D. D. (2013). Integration and reuse in cognitive skill acquisition. *Cognitive Science*, 37(5), 829–860. doi: 10.1111/cogs.12032
- Salvucci, D. D. (2016). Cognitive Code: An Embedded Approach to Cognitive Modeling. In *Proceedings of the 14th international conference on cognitive modeling* (pp. 15–20).
- Salvucci, D. D. (2021). Interactive Grounding and Inference in Instruction Following. *To appear in Topics in Cognitive Science*.
- Shimizu, C., Hammar, K., & Hitzler, P. (2021). Modular ontology modeling. *Semantic Web*. (Under Review.)
- Taatgen, N. a., & Lee, F. J. (2003). Production compilation: a simple mechanism to model complex skill acquisition. *Human factors*, 45(1), 61–76. doi: 10.1518/hfes.45.1.61.27224
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, 12, 97–136.

Exploring the Decision Component of the Activation-Decision-Construction-Action Theory for Gain and Loss Facing Scenarios

Tei Laine (teihlaine@gmail.com)

Tomi Silander (tomi.silander@naverlabs.com)

NAVER LABS, 6 Chemin de Maupertuis
38240 Meylan, France

Abstract

Inspired by Masip et al.'s (2016) test of ADCAT model's decision component, we wanted to see if we could replicate their findings using different data from a similar scenario-based study. They found that expected value of telling the truth predicted the decisions to lie or tell the truth more accurately than the expected value of lying, and even better than the motivation to lie, which they defined as a difference between these two expected values. In contrast, in our modeling study the motivation to lie was the best predictor of choices for both actual liars and truth tellers in conditions involving gains and large losses, whereas only in the condition involving large losses the expected value of telling the truth outperformed the expected value of lying. We conclude that whether the participants could gain something or avoid losing something by deceiving determined if they focused on benefits of lying or costs of telling the truth.

Keywords: Deception; Motivation; Social cognition; Social-cognitive theory; Risk taking; Lie aversion

Introduction

Not all deception is bad or reprehensible, or socially unacceptable; people lie in various ways and for various reasons which can be pro-social (e.g., to avoid conflict or help someone), or selfish but do not hurt others (e.g., saving face or making an impression) (dePaulo, Kashy, Kirkendol, & Wyer, 1996; Erat & Gneezy, 2012; Gupta, Sakamoto, & Ortony, 2013). At the same time, the costs and benefits that drive decisions to lie in everyday life may reside in different domains, for instance they can be financial, material, reputational, or psychological (Sakamoto, Laine, & Farber, 2013). Alternatively, they may lie in the same domain and even carry the same absolute (objective) value, but differ in psychological (subjective) value. For example, getting an undeserved discount on a purchase and avoiding paying for damages one has caused may both carry an equal (monetary) value, but most people may choose to deceive in one situation but not in the other; they may be able to justify that deceiving for the discount is only fair ("The product is overpriced, anyway"), but feel good about admitting the guilt of accidentally breaking something and paying for the damages.

Standard economic theories of rational behavior approach decisions to lie as cost-benefit analysis. They posit that whenever the expected benefit from lying exceeds the outcome of being honest a selfish individual, a "homo economicus," should lie, and the decision should be determined solely by the trade-off between the gain from lying and the penalty

incurred if detected (Abeler, Becker, & Falk, 2014). Along these lines, Levine, Kim, and Hamel (2010) posit that people "lie for a reason"; they tell the truth if it does not prevent them from attaining their goals, and only when it does they may consider deception. In other words, people lie only when it is more beneficial or less harmful with respect to their goal attainment than truth. Relatively recently economists have started to acknowledge that the act of lying may have an intrinsic cost that deters people from lying even if it would be beneficial (Gneezy, Rockenbach, & Serra-Garcia, 2013).

However, even for social goals people lie a lot less than the economic models predict, and the discrepancy cannot be solely explained by unusually strong risk aversion or pure lie aversion (Dhimi & al-Nowaihi, 2007; Gneezy et al., 2013; López-Pérez & Spiegelman, 2012). Cappelen, Sørensen, and Tungodden (2013) showed that non-economic aspects of a choice situation have role in decisions to lie. For instance, in the context of taxes, emotions have been found to influence the propensity to take risk and evade, so that the act of tax reporting may elicit anticipated emotions of how one would feel if audited and punished, and consequently these emotions drive behavior and future tax compliance (Coricelli, Rusconi, & Villeval, 2014). Furthermore, Maciejovsky, Schwarzenberger, and Kirchler (2012) have argued that emotions can moderate the relative effectiveness of economic variables such as audit probabilities and fines in tax ethics.

Walczyk, Harris, Duck, and Mulay (2014) have proposed a quasi-rational model of deceptive decision making, called *Activation-Decision-Construction-Action Theory (ADCAT)*, that combines costs and benefits from different domains, such as material outcomes, and affective responses ranging from apprehension of being detected to thrill of successfully deceiving. According to the model, after considering utilities and probabilities of the most important consequences of all choice options, a decision maker chooses the option that best achieves her goal.

Using everyday scenarios to test ADCAT model's decision component, Masip, Blandón-Gitlin, de la Riva, and Herrero (2016) showed that the decision to lie vs. tell the truth was associated with expected consequences of stating the truth, but not with expected consequences of deceiving. Cassidy, Wyman, Talwar, and Akehurst (2019) studied the relationships between Walczyk's model variables (expectations from lying and truth telling) and decisions to lie vs. tell the truth

varying motives for lying (either benefit to oneself or another person) and potential costs (either to oneself or another person). Contrary to Masip et al. (2016), they found no relation between the expected value of truth telling and decision to lie in self-oriented lies,¹ whether or not there was a cost to another, whereas they found a significant negative relationships between the expected value of truth telling and decision to lie, when the lie was other-oriented. Overall, when there was an additional cost of lying — implicating the other person of wrongdoing that self had conducted, or implicating self for wrongdoing the other had conducted — the participants were less prone to lie.

Building on this previous research, using the data from the scenario-based study by Sakamoto et al. (2013), which varied riskiness, domains of costs and benefits, severity of loss when detected, and motives for deceiving, we wanted to test if ADCAT model would make the same predictions as their results suggested, namely that the potential outcomes from (successfully or unsuccessfully) deceiving do matter in decisions to deceive. Indeed, we found that in most conditions the expected value of lying both correlated with and predicted decisions to lie better than the expected value of telling the truth.

The rest of the paper is organized as follows. We start by briefly discussing Walczyk’s ADCAT model and its empirical test by Masip et al. (2016). Then we review the experimental settings and the data of the Sakamoto et al. (2013) study. We continue by describing our modeling approach and presenting the results. We conclude with a discussion of possible reasons why our findings differ from those of Masip et al. (2016).

The ADCAT model

The *Activation-Decision-Construction-Action Theory (AD-CAT)* by Walczyk et al. (2014) is a cognitive model of deception (with high stakes) specifying the roles of cognitive, emotional, motivational, and social processes in the decisions to deceive, explicitly accounting for constructs such as working memory and theory of mind. The theory is more elaborate than is required for the current purposes, so it is presented here only to the extent that it applies to “quasi-rational” (Walczyk’s term) decision making in deception context, omitting for instance discussion on cognitive load or lie construction.

Quasi-rational decision component

According to the model the decision to deceive is influenced by both the emotional reaction to the choice options and the social context, and the decision maker chooses an option that best achieves her goal given her estimates of utilities and likelihoods of the outcomes, i.e., expected costs and benefits of choosing a particular option. She estimates the expected value of action a (either lying or telling the truth) based on action’s n_a different possible outcomes using the equation:

$$EV(a) = \sum_{i=1}^{n_a} p_i(a)v_i(a),$$

¹Masip et al. (2016) used mostly self-oriented lies in their scenarios.

where $v_i(a)$ is the value of i th outcome (gain or loss) of action a , and $p_i(a)$ its probability. She then chooses the action with the highest expected value. Finally, she uses these estimates to assess an overall level of motivation to lie, M . This is done as if she were intuitively following the equation:

$$M = EV_{\text{lying}} - EV_{\text{truthtelling}}.$$

The higher the value of M , the more likely she will lie, and spend cognitive resources in constructing the lie. Basically, M determines which particular lie she will tell; for instance, whether she chooses to fabricate a story rather than just omit a crucial piece of information, the latter consuming less resources (Walczyk et al., 2014).

Walczyk et al. (2014) emphasizes that cognitive processes underlying deception and truth telling do not differ, but memory processes, decision making, and problem solving are essential in both. Particularly, in lying, the truth, goals, and the social context are activated in the working memory, which in turn bring in relevant life memories of previous decisions, which then control the motivation for dishonesty, while means-end problem solving is used by the liar to move from the current state to the desired goal state.

Masip et al. (2016)’s study

To empirically test the ADCAT model’s decision component, Masip et al. (2016) administered two separate questionnaires. In Questionnaire 1, the participants read ten scenarios and made binary choices between lying and telling the truth in those scenarios. The authors correlated these choices with the expected values M , EV_{lying} , and $EV_{\text{truthtelling}}$ calculated from the participants’ responses in Questionnaire 2, in which they read again the same ten scenarios, and for each scenario generated a possible consequence of telling the truth, probability of that consequence, and how good or bad it would be. They were also asked to come up with an alternative consequence of telling the truth and indicate its valence. Finally, they were asked to think about what kind of lie they would tell to avoid the negative consequences of disclosing the truth, and how likely they expected it to go undetected, and the consequence of not getting detected.

They classified the participants into liars and truth tellers based on the expected values, and tested how well the classification matched the actual choices in Questionnaire 1. Their results showed that for both actual liars and truth tellers (in Questionnaire 1) the expected value of telling the truth was a better predictor of their choices than the expected value of lying, and it was even slightly better predictor than the motivation score M . From this they concluded that the expected outcome of successfully or unsuccessfully lying may not play a role in decisions to lie but what matters are the consequences of revealing the truth.

Current experiment

Based on their findings, Masip et al. (2016) suggested that Walczyk’s model could be made more parsimonious by omit-

ting the expected value of lying from the equation, and equating motivation to lie with expected value of telling the truth. They supported their argument with Levine et al. (2010)'s *veracity principle*, namely that people usually tell the truth unless it interferes with their goal attainment, and studies on pure lie aversion (Gneezy et al., 2013; López-Pérez & Spiegelman, 2012). According to the former, lying requires justification whereas telling the truth does not, and according to the latter, the act of lying has a cost regardless of its consequences. However, Sakamoto et al. (2013) showed that the perceived benefit of successfully deceiving predicted the deceptive choices, but only in loss-facing scenarios.² On the other hand, tax payers have been found to refrain from cheating in their taxes for fear of being audited and penalized (Alm, 2012; Slemrod, 2007). Using another dataset we wanted to study if Walczyk's model will support Masip et al. (2016)'s conclusion or if it would replicate findings of Sakamoto et al. (2013), deeming Masip *et al's* suggestion of dismissing the expected value of lying premature.

Method

Participants In the online study conducted by Sakamoto et al. (2013) on Amazon Mechanical Turk 492 participants (276 men, 214 women, 2 unknown, median age 29, age range 18-77 years) read a single scenario of a common everyday life situation, and answered several questions pertaining to the scenario. They were also asked demographic information including age, income, and education.

Data Data comprises of participants' responses to eight questions about potential communicative messages — either deceptive or truthful — that could be exchanged in the scenario they read. The participants were asked to imagine themselves as the protagonist in the scenario situation, and indicate the likelihood with which they would choose the deceptive message over the honest one, and to evaluate several aspects of the scenario. They gave responses on a continuous Likert type scale with only the end points labeled (e.g., very unlikely - very likely, very bad - very good): for questions Q1, and Q3 - Q8 the scale ranged from -5 to 5, and for question Q2 from 0 to 1. The questions were (with simplified wording): Q1. How likely would you lie rather than tell the truth in this situation? Q2. How likely would your lie be detected? Q3. If it was detected, how good or bad would the result be for you? Q4. If it was not detected, how good or bad would the result be for you? Q5. If you told the truth, how good or bad would the result be for you? Q6. How would telling the truth make you feel? Q7. How would lying make you feel? Q8. How truthful is the deceptive message in this situation? The question Q1 was asked first and then the remaining questions Q2 - Q8 were presented in a random order.

Material The study used both gain and loss facing scenarios depicting a situation in which the speaker had an incentive to deceive the hearer or hearers. The two loss conditions

varied the magnitude of loss the speaker could expect to incur if getting detected (large vs. small). This was done by adding some extra context to scenarios. Example scenarios are shown in Figure 1.

In both gain and loss facing scenarios a deceptive communication option was paired with an honest option. In gain-facing scenarios — situations in which making the hearer believe something false benefited either the speaker or hearer — the deceptive choice was associated with probabilistic gain, whereas the honest choice was associated with certain no-gain. In contrast, in loss-facing scenarios — situations in which telling the truth harmed either the speaker, hearer, or someone else — the honest choice was associated with certain loss, whereas the deceptive choice was associated with probabilistic no-loss.

In addition to asymmetric outcome structures, there were other differences; in gain scenarios the non-deceptive message was formulated as an act of omission (i.e., “say nothing” as no response was required from the speaker), but in loss scenarios as an act of commission (i.e., as a truthful message that was sufficiently informative, clear, and appropriate (Grice, 1975)). This was necessary in order to keep the gain scenarios in the domain of pure gain, since for instance overtly stating the harmful truth would have introduced an outcome in the domain of losses.

Although both Sakamoto et al. (2013) and Masip et al. (2016) used quite similar scenarios their protocols had some important differences:

- While Masip et al. (2016) administered two separate questionnaires for the participants to decide between deceiving and telling the truth, and to evaluate the potential outcomes, Sakamoto et al. (2013) asked the participants to complete these two tasks in a single questionnaire.
- Sakamoto et al. (2013) used predefined consequences, and asked participants to evaluate their valences, in addition to estimating the detection probability. This design assumed that truth telling resulted in one certain (non-positive) outcome and telling a lie had two potential outcomes, one non-negative when going undetected (with probability $1 - p$) and one negative when detected (with probability p) (the payoff structure is presented in Table 1). Masip et al. (2016) in turn asked the participants to generate the possible consequences of both telling the truth and successfully lying in addition to indicating their probabilities and valences. These differences are summarized in Table 2.
- Finally, while Masip et al. (2016) asked each participant to respond to several scenarios, Sakamoto et al. (2013) presented each participant a single scenario randomly picked from a pool of 48 scenarios. The former also analyzed their data scenario-wise, whereas the latter pooled the responses within each condition into one set for the analyses.

²Most scenarios used by Masip et al. (2016) were loss-facing.

Gain scenario. Harry is in a job interview. He is not familiar with the field, and only moderately interested in it. Unsurprisingly, the job is beyond his skills. He is currently unemployed and desperately needs to find a job so he can support his family. When the interviewers offer Harry an opportunity to provide any additional information, he can't admit he would take any job he gets. If the interviewers find out that Harry deceived them, they won't hire him. Imagine yourself as Harry.

[Honest option] Say nothing.

[Deceptive option] "I think I am the right person to meet the challenges of this position, for which I am highly qualified."

Loss scenario. Stephanie belongs to a "popular" group in her high school. One day, she happens to chat with Nancy, who is associated with the "nerd" group. They discover that they share an interest in ancient South American cultures. Nancy invites Stephanie to come over on Saturday to see her collection of books and collectibles, and Stephanie agrees. However, the popular kids are also planning to meet on Saturday. When asked why she is not planning to join them, Stephanie is reluctant to admit that she is meeting one of the nerds. **(Low loss)** The popular group's unwritten rules allow some interaction with the nerds for things like getting help with homework, but it's still seen as not a very cool thing to do. If the other group members find out that Stephanie deceived them, Stephanie will be mildly embarrassed. **(High loss)** The popular group's unwritten rules forbid any interaction with the nerds. If the group members find out that Stephanie deceived them, she will be expelled from the group. Imagine yourself as Stephanie when one popular group member asks, "Why aren't you joining us this Saturday?"

[Honest option] "I'm meeting Nancy on Saturday."

[Deceptive option] "Guess what? We're getting a new dog on Saturday. I'm so excited!"

Figure 1: Examples of scenarios and response options.

Table 1: Payoff valences of telling the truth and lying for gains and losses in the Sakamoto et al. (2013) study.

Condition	Telling truth	Successful lie	Detected lie
Gain	No gain (0)	Gain (+)	Loss (-)
Loss	Loss (-)	No loss (0)	Loss (-)

Table 2: Outcome structure of telling the truth and lying in the studies by Masip et al. (2016) and Sakamoto et al. (2013).

Study	Telling truth	Lying
Masip et al. (2016)	Outcome 1 (-) p Outcome 2 (+) $1 - p$	Outcome (+) $1 - p_d$
Sakamoto et al.(2013)	Outcome (0/-) $p = 1$	Outcome 1 (+/0) $1 - p_d$ Outcome 2 (-) p_d

p is the probability of the outcome, and p_d is the detection probability. (+/0/-) marks the valence of the outcome.

Procedure In the study by Sakamoto et al. (2013) the participants indicated their likelihood of lying vs. telling the truth using a continuous scale with the ends marked with the honest message at the left end and the dishonest message at the right end of the scale. This scale was interpreted such that any choice to the right of the mid-point of the scale (indifference) meant that they would more likely lie than be honest, and the choice at the extreme right end of the scale meant that they would definitely lie (i.e., with 100% probability). For the current analysis purposes these continuous choices were discretized so that values above zero were coded as 1 (lie), and values at zero and below, were marked as 0 (tell truth).³

We calculated the expected values of lying and telling the truth using the utilities and probabilities of the outcomes (assessed in questions Q2-Q5) with the following equations:

$$EV_{lying} = (1 - p_d)v_{no-d}(lie) + p_d v_d(lie),$$

$$EV_{truthtelling} = v(truth),$$

where p_d is the detection probability, and $v_d()$ and $v_{no-d}()$ values of outcomes when detected and not detected. For the sake of simplicity, like Walczyk et al. (2014) and Masip et al. (2016), we assumed that any affective reactions (e.g., guilty feeling) were included in the outcome utility, instead of incorporating them explicitly in the equations (with values obtained from questions Q6 and Q7).

In order to test how well each of the three expected values could tell apart liars from truth tellers (i.e., predict the actual lying decisions from the expected values and M) we first dichotomized these values following the methodology used by Masip et al. (2016). Briefly, we first calculated the proportion (say $X\%$) of participants who chose to lie based on earlier discretized choice values. In our predictions, we

³Basically, we coded as liars those participants who indicated that they will more likely lie than tell the truth.

then matched this true proportion by predicting $X\%$ of participants with lowest $EV_{truth\ telling}$ (and with highest EV_{lying} and M) to be liars. Finally, in each condition, we compared these predictions to the actual choices, and calculated how many predictions were correct. Like Masip, we also computed this prediction accuracy separately for actual liars and truth tellers.

Results

For the data analysis we pooled all participants' responses together by the three scenario conditions: gain (N=161), large loss (N=162), and small loss (N=169).⁴

We started by correlating the participants' decisions to deceive with the motivation to deceive, and expected values of successfully deceiving and telling the truth. These correlations are shown in Table 3. While expected value of truth telling had the weakest correlation with the decisions to lie, M had the strongest in all three conditions, expected value of lying being in between.

Table 3: Correlations between decisions to lie and the expected values calculated from ratings in questions Q2-Q5.

Condition	$EV_{truth\ telling}$	EV_{lying}	M
Gain	-0.3760	0.4684	0.5360
Large loss	-0.3127	0.3268	0.4128
Small loss	-0.3271	0.5236	0.5733

The percentages of correctly identified liars and truth tellers, using the dichotomized expected values as described above, are shown in Table 4. Just like Masip et al. (2016), we did not achieve perfect identification: all individuals who indicated that they would lie (tell the truth) were not coded as liars (truth tellers) by their dichotomized expected values. However, even if the scenarios and conditions were not exactly similar in these two studies, the overall identification accuracies were quite close: the identification rates averaged over their four scenarios retained for analyses were 67.11%, 62.40%, and 68.07%, for M , EV_{lying} , and $EV_{truth\ telling}$, respectively, whereas our rates, averaged over the three conditions in Sakamoto et al. (2013)'s data, were 75.59%, 69.83%, and 66.06%, respectively.

For both the actual liars and truth tellers, and overall, the motivation to lie was the best predictor of their choices in the conditions involving gains and large losses. For small losses, the expected value of lying was the best predictor. While the motivation to lie predicted the choices best in large losses, that was the only condition in which the expected value of truth telling was more accurate than the expected value of lying, although the differences were not large.

Interestingly, while regression analysis run by Sakamoto et al. (2013) indicated that the outcome of successful deception

⁴We could have analyzed the data per given scenario, but this would have resulted quite small sample sizes, with about ten data-points per scenario.

Table 4: Correct identification rates of liars and truth tellers.

Predictor	Condition	Percentile for cutoff	Identification %		
			Liars	Truth tellers	Overall
M	Gain	$P_{50} = -0.37$	77.78	77.50	77.64
	Large loss	$P_{51} = 0.02$	68.75	69.51	69.14
	Small loss	$P_{60} = 0.49$	67.16	78.43	74.00
EV_{lying}	Gain	$P_{50} = -0.38$	71.60	71.25	71.43
	Large loss	$P_{51} = -0.96$	61.25	62.20	61.73
	Small loss	$P_{60} = -0.47$	70.15	80.39	76.33
EV_{truth}	Gain	$P_{50} = 0$	63.00	73.75	68.32
	Large loss	$P_{49} = -1.25$	62.50	63.41	63.00
	Small loss	$P_{40} = -1.54$	56.72	73.53	66.86

Table 5: Comparison of logistic regression models

Condition	$\beta_L \ \beta_T$		ΔBIC			LLR p-values		
	EV_L	EV_T	EV_L	EV_T	M	EV_L	EV_T	M
Gain	1.9	-1.8	13	26	-5	2×10^{-5}	3×10^{-8}	0.8
Large loss	1.3	-1.1	8	12	-5	2×10^{-4}	4×10^{-5}	0.6
Small loss	2.9	-1.6	15	62	4	7×10^{-6}	3×10^{-16}	0.003

predicted the decisions to lie in losses, but not in gains, our results indicated the opposite; the expected value of deception — incorporating outcomes of both successful and failed deception — predicted the decisions to lie and tell the truth more accurately in gains than in losses. Also the expected value of truth telling was slightly better predictor of both decisions in gains than in losses.

Since M is the difference between EV_{lying} and $EV_{truth\ telling}$, one might argue that it constitutes a more complex model. We therefore also conducted logistic regression analyses and used the Bayesian Information Criterion (BIC) to compare the evidence for a model with two predictors,

$$P_{\beta_L, \beta_T}(lying) \propto e^{\beta_L EV_L + \beta_T EV_T},$$

against the models using only one of these predictors, i.e., models in which one of the β -coefficients was forced to be zero, or in case of M , forcing $\beta_T = -\beta_L$. Models being nested, we also computed the statistical significance of the difference of log-likelihoods of the models.

Our model selection analysis (Table 5) clearly favors using both the value of lying and the value of truth telling in explaining decisions to lie. The BIC difference (ΔBIC) greater than 6 is generally seen as a strong evidence against simpler hypothesis and the difference more than 10 as very strong, while negative values of BIC favor simpler hypotheses. The likelihood ratio test (LLR) for comparing logistic regression models show that the differences in model fits are statistically very significant (see Table 5).

The β coefficients in fitted two-predictor models also automatically recovered the structure highlighting the role of the difference $EV_{lying} - EV_{truth\ telling}$. In the gain and large loss

conditions these two predictors are almost equally weighted, while in the small loss condition the expected value of lying is twice as important as the value of telling the truth. This is also highlighted in Table 5 where the negative Δ BIC suggests using M instead of P_{β_L, β_T} in the gain and large loss conditions. The small loss condition is better modeled using separate weights for EV_{lying} and $EV_{\text{truth telling}}$.

Our findings are somewhat contrary to what Masip et al. (2016) found, and it may be because the negative outcomes from truth telling that we gave our participants were not judged severe enough to lie, but instead the lying behavior was driven by the expected positive outcome from successfully deceiving. In turn, Masip et al. (2016)'s participants either may not have been optimistic about their lie succeeding and being helpful, or did not find the outcome from lying (that they themselves gave in Questionnaire 2) attractive enough, and therefore were driven by the very bad outcome from telling the truth, which they wanted to avoid.

Discussion

Decisions to deceive may be driven by two “opposite” motives: an attempt to avoid a loss from harmful truth or an attempt to gain something by lying. These two perspectives may explain the asymmetries between the two studies by Masip et al. (2016) and us. They started with the consequences of being honest by asking what are the possible outcomes if the harmful truth is revealed, contrasting them to potential consequences of successfully deceiving. In contrast, Sakamoto et al. (2013) focused on the risky aspect of decisions to deceive, and asked participants to evaluate benefits of successfully deceiving and costs of getting caught, while assuming that telling the truth only had bad consequences. In fact, in their scenarios truth telling and detection shared the same outcome (e.g., not being hired), which the decision maker tried to avoid by lying. In both cases the truth got revealed, which resulted in either a loss (loss scenarios) or no gain (gain scenarios), but in case of detection there was an additional (implicit) cost of being stigmatized as a liar. In turn in gain scenarios status quo always persisted (e.g., the hearer still felt bad about her looks), but the liar incurred a cost of getting caught.

Furthermore, since Masip et al. (2016) asked their participants to focus on truth and its consequences, they may have judged its harmfulness more severely than benefits of undetected deception. In turn, Sakamoto et al. (2013) gave the outcomes and their valences to their participants, focusing on the benefits of deception as opposed to cost of telling the truth, which may have lead the participants to judge benefit of successful deception more extremely than harmfulness of truth (or failed deception). This could be the reason why in our study the motivation to lie (difference between value of lying and value of telling the truth) predicted participants' choices the most accurately. While both studies contrasted the outcome from telling the truth (bad or very bad) to successful deception, in Masip et al. (2016)'s study the outcome from

undetected deception could still have been somewhat bad, whereas in Sakamoto et al. (2013)'s study it was assumed to be very good. Finally, it seems that it also mattered if by deceiving the participants could gain something or avoid losing something, and that determined whether they focused on benefits of lying or costs of telling the truth, in other words, the reference point they adopted, as suggested by the prospect theory (Kahneman & Tversky, 1979).

References

- Abeler, J., Becker, A., & Falk, A. (2014). Representative evidence on lying costs. *Journal of Public Economics*, *113*, 96–104.
- Alm, J. (2012). Designing alternative strategies to reduce tax evasion. In A. P. Michael Pickhardt (Ed.), *Tax evasion and the shadow economy* (pp. 13–32). Edwar Elgar Publishing.
- Cappelen, A. W., Sørensen, E. Ø., & Tungodden, B. (2013). When do we lie? *Journal of Economic Behavior & Organization*, *93*, 258–265.
- Cassidy, H., Wyman, J., Talwar, V., & Akehurst, L. (2019). Exploring the decision component of the activation-decision-construction-action theory for different reasons to deceive. *Legal and Criminological Psychology*, *24*(1), 87–102.
- Coricelli, G., Rusconi, E., & Villeval, M. C. (2014). Tax evasion and emotions: An empirical test of re-integrative shaming theory. *Journal of Economic Psychology*, *40*, 49–61.
- dePaulo, B. M., Kashy, D. A., Kirkendol, S. E., & Wyer, M. M. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, *70*(5), 979–995.
- Dhami, S., & al-Nowaihi, A. (2007). Why do people pay taxes? Prospect theory versus expected utility theory. *Journal of Economic Behavior & Organization*, *64*, 171–192.
- Erat, S., & Gneezy, U. (2012). White lies. *Management Science*, *58*, 723–733.
- Gneezy, U., Rockenbach, B., & Serra-Garcia, M. (2013). Measuring lying aversion. *Journal of Economic Behavior & Organization*, *93*, 293–300.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics volume 3: Speech acts* (pp. 41–58). New York: Academic Press.
- Gupta, S., Sakamoto, K., & Ortony, A. (2013). Telling it like it isn't: A comprehensive approach to analyzing verbal deception. In F. Paglieri, L. Tummolini, R. Falcone, & M. Miceli (Eds.), *The goals of cognition. Essays in honor of Cristiano Castelfranchi*. London: College Publications.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.
- Levine, T. R., Kim, R. K., & Hamel, L. M. (2010). People lie for a reason: Three experiments documenting the principle of veracity. *Communication Research Reports*, *27*(4), 271–285.

- López-Pérez, R., & Spiegelman, E. (2012). Why do people tell the truth? Experimental evidence for pure lie aversion. *Experimental Economics*, *16*(3), 233–247.
- Maciejovsky, B., Schwarzenberger, H., & Kirchler, E. (2012). Rationality versus emotion: The case of tax ethics and compliance. *Journal of Business Ethics*, *109*, 339–350.
- Masip, J., Blandón-Gitlin, I., de la Riva, C., & Herrero, C. (2016). An empirical test of the decision to lie component of the Activation-Decision-Construction-Action Theory (ADCAT). *Acta Psychologica*, *169*, 45–55.
- Sakamoto, K., Laine, T., & Farber, I. (2013). Deciding whether to deceive: Determinants of the choice between deceptive and honest communication. *Journal of Economic Behavior & Organization*, *93*, 392–399.
- Slemrod, J. (2007). Cheating ourselves: The economics of tax evasion. *Journal of Economic Perspectives*, *21*(1), 25–48.
- Walczyk, J. J., Harris, L. L., Duck, T. K., & Mulay, D. (2014). A social-cognitive framework for understanding serious lies: Activation-decision-construction-action theory. *New Ideas in Psychology*, *34*, 22–36.

Understanding Human Social Communication: A Computational Model of Gossip

Jeungmin Lee (jeungminlee@kaist.ac.kr)

Department of Bio and Brain Engineering, Program of Brain and Cognitive Engineering
College of Engineering, KAIST, Daejeon, 34141, South Korea

Jerald D. Kralik (jerald.kralik@gmail.com)

Department of Bio and Brain Engineering
College of Engineering, KAIST, Daejeon, 34141, South Korea

Jaeseung Jeong (jsjeong@kaist.ac.kr)

Department of Bio and Brain Engineering, Program of Brain and Cognitive Engineering
College of Engineering, KAIST, Daejeon, 34141, South Korea

Abstract

Updating people about the actions of others—social communication—is a powerful means by which humans learn about the world and maintain stable societies. However, how the mind/brain achieves this ability computationally remains unclear. Our goal is to model when, how, and why people choose to communicate information about others to others. Here we present current progress. We first describe our social communication framework, the test paradigm for model development and assessment, and an empirical experiment we conducted to obtain novel data to test model predictions. We then present our model, and compare it with two others. Our model outperformed the others, capturing the main patterns of the empirical data and matching the specific results most closely (i.e., percent of cases deciding to communicate about a target individual). Thus, our model successfully simulates human social decision-making, helping to understand how it is achieved by the human mind/brain.

Keywords: evolution of social cognition; theory of mind; communication; decision-making; computational model

Introduction

Observing the actions of others is a principal means by which humans learn and update knowledge—both about the world as well as the person performing the act—greatly extending our reach beyond our own individual experiences. Moreover, learning from and about others ratchets up even further with communication, not only from the performer to the observer, but in turn from the observer to someone else, and so on. In this way, information quickly disseminates across the social network (in turn enabling social networks to scale). Additionally, the ability to influence a person’s future actions increases dramatically (e.g., via social influence or appeals to authority when someone’s actions are in question). Communication about the action of others, then, provides extraordinary value to social groups, and likely played a leading role in the evolution of the human brain (Dunbar, Marriott & Duncan, 1997; Dunbar, 2004). To date, however, how we achieve this capacity at a computational level remains unclear.

This is because the ability is deceptively elaborate and complex. For example, focusing on the observer as the *central agent*, there are multiple critical factors that determine whether to transmit information about the actions of a *target* individual to someone else, i.e., a *receiver*. In general, it requires assessing the significance of the target’s action, and whether a receiver would be interested in learning of it (and/or whether the information could likely feedback to the target person and influence future behavior). On first order, the significance of the action can be measured in terms of potential *benefits* and *costs* to self and others. Making this evaluation requires the central agent to have an internal *value scale* that assigns the degree of significance to particular target actions. In other words, the central agent must possess a minimal *affective* apparatus (Gazzaniga, Ivry & Mangun, 2013). In assigning value, a target’s actions can again be categorized in two general categories: whether significant as world knowledge independent of the target, or valuable knowledge about the target him/herself. For the latter, given that social interactions comprise such a significant portion of our daily lives (whether at home, workplace, or almost anywhere else), information about others (i.e., their locally stable traits and behaviors) is critical. Indeed, Dunbar and colleagues (1997) found that over 60% of conversations involve discussing others. Because of this significance, and the general lack of development to date, our model currently focuses on this *social knowledge*.

Intriguingly, this type of social communication—telling others about someone else—has normally been defined as *gossip*. Although gossip may seem superficial, it is in fact an important mechanism underpinning society (Dunbar, Marriott & Duncan, 1997; Dunbar, 2004; Foster, 2004). For shorthand, we thus also use the term “gossip” to refer to this form of social communication: in which one informs another individual about events involving someone else not present.

In fact, disseminating knowledge about an individual’s actions across a social network (boot-strapping culture and societies) belies a formidable affective-sociocognitive engine

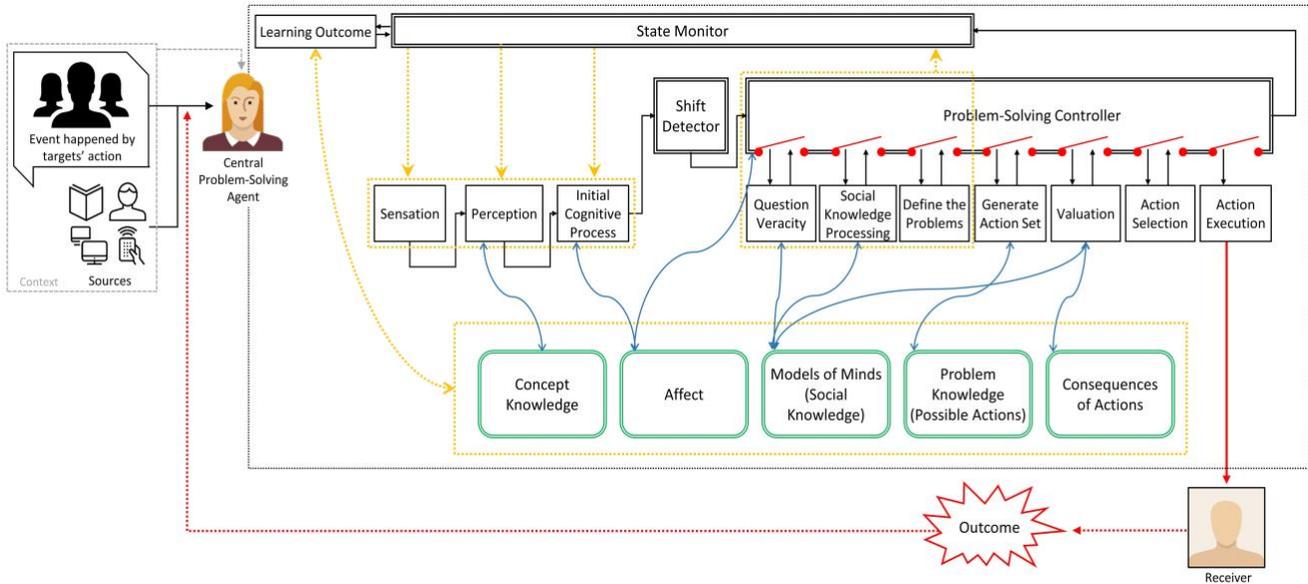


Figure 1. The complete framework of internal process regarding social information and gossip decision. The central agent goes through a series of internal processes (black boxes) managed by higher metacognitive processes (double-lined boxes) by accessing the knowledge (green rounded boxes). See text for details (Lee, Kralik, & Jeong, 2018; 2019).

under the hood: one that includes not only the affective valuation assessment, but also mind-reading (i.e., interpreting intentions underlying someone’s actions and whether others know or would care about it), and social accounting (i.e., a social currency, based on the value the information provides to others) (Cosmides & Tooby, 1992; Gazzaniga, Ivry & Mangun, 2013; Lee, Kralik, & Jeong, 2018; 2019). Our goal, then, is to model when, how, and why people choose to communicate social information to others.

The current paper presents our progress, by first describing our overall social communication framework, as well as the test paradigm for model development, and the empirical experiment we conducted in our cognitive neuroscience laboratory to obtain novel data to test model predictions (Anonymous, *submitted*). We then present the computational development of our main model, together with two alternative models. We then compare the models on how well they fit the empirical findings.

Framework, Methods, and Models

In this section we describe the framework and test paradigm used for the empirical experiment and model development. We then describe our main computational model, and then two others based on simpler versions of the main model.

Framework for Social Communication

We developed the general framework to capture the fundamental components of social intelligence and communication (Figure 1) (Aronson, Wilson, Akert & Sommers, 2016; Cosmides & Tooby, 1992; Dunbar, Marriott & Duncan, 1997; Dunbar, 2004; Foster, 2004; Gazzaniga,

Ivry & Mangun, 2013; Haidt, 2007; Kralik, 2017; Kralik et al., 2018; Lee, Kralik, & Jeong, 2018; 2019). It is based on a *target* individual being involved in some event, such as hitting a coworker, caught cheating on an exam, helping people escape a burning building, or going to the movies. The *Central Agent* then learns of the event, and subsequently processes it via a series of subprocesses (black boxes in Figure 1) that also utilize specific knowledge stores (green boxes).

The overall process begins with the initial sensory input passed from *sensation* to *perception* to *initial cognitive processing*, in which for example, auditory input is transformed to sentences, and initial meaning is ascertained, including the event content evoking an initial affective (i.e., emotional) response, such as cheating on an exam being bad. Although we have been developing these early modules, the current model simply begins with a three-element vector representing the event involving a target individual, which then is mapped to an *affect score* (described further below). The affect score reflects the level of initial interest or concern the target’s action evokes, with the overall affect assignment process being the affective/emotion center of our framework and model (Gazzaniga, Ivry & Mangun, 2013). Next, if the affect value is sufficiently high, the *shift detector* activates the *problem-solving controller*, which in turn activates individual subprocesses by closing the corresponding affect gates.

The subprocesses include determining the reliability of the information source, updating the model of the target’s mind, determining the specific problem(s) at hand, generating the relevant action set (such as whether or not to communicate to another person as *receiver*), valuation of the possible actions

(i.e., assessing the benefits versus costs), action selection based on these action valuations, action execution, and then monitoring of the action outcome for potential learning (though learning is not yet explicitly modeled) (see Figure 1). Computational development has thus far focused on *valuation*, the central process that determines whether to take action based on the event, described next.

Test Paradigm and Behavioral Experiment

Input to the system is an event or *scenario* involving someone, such as someone caught cheating on an exam or verbally abusing a coworker. Each event is then represented by three fundamental factors: the *target* individual, the event *content*, and the *valence* of the content, that is, whether positive or negative. Each independent variable (i.e., our three event factors) is then further divided into a number of fundamental categories. As stated, content *valence* is subdivided into positive and negative events. For *target*, we examined *ingroup* versus *outgroup* versus *celebrity*, in order to test the important social factors of contact, caring, and status (described further below). Finally, for *content*, our intention was to produce a comprehensive set of social events that occur in daily life (either rarely or frequently). Based on theoretical considerations and literature review, this resulted in eight content domains that can be roughly aligned according to how much we care about them, that is, how much affect or emotion they evoke, represented by an affect score, listed in Table 1 (Aronson et al., 2016; Cosmides & Tooby, 1992; Dunbar, Marriott & Duncan, 1997; Dunbar 2004; Foster, 2004; Gazzaniga, Ivry & Mangun, 2013; Haidt, 2007; Kralik, 2017; Lee, Kralik & Jeong, 2018; 2019). Based on all combinations of these factors, we developed 48 different scenarios (3 target · 2 valence · 8 content). Using this comprehensive set of scenarios, we developed our main computational model and generated a set of predictions of how the three independent variables (target, content, and valence) influence gossip spreading.

Table 1. Eight content domains and their affect score (where *m* signifies morality domain: Haidt, 2007).

<i>Content</i>	<i>Affect Score</i>
Prosociality (care/harm) ^m	7
Fairness (fair/cheating) ^m	6
Competition (positive/negative)	5
Social-oriented (altruism/selfishness)	5
Community (loyalty/betrayal) ^m	4
Respect (authority/subversion) ^m	4
Purity (sanctity/degradation) ^m	3
General social affairs (positive/negative)	1

To collect the empirical data, we recruited 102 participants (59 females and 43 males, mean age 23.8 years, range 20-32), and each participant was shown a gossip scenario and asked if they wanted to spread it to other people (i.e., receivers) (Lee, Kralik & Jeong, *submitted*). The 48 scenario types were

replicated three times. As a result, 144 different gossip scenarios were used per participant in the experiment. Gossip rates of the scenarios were calculated by taking the mean across replications and participants.

Test Paradigm and Behavioral Experiment

We now describe the computational models, developed and tested using Matlab (The Mathworks, Natick, MA, USA). For all three models, on every trial, that is, when one of the 48 scenario events occurs, the given event’s *content* is converted into (a) an *affect score* (Table 1) and (b) a valence flag (i.e., 0 for positive, 1 for negative), which occurs in the *Initial Cognitive Process* module in Figure 1. The *target* for the given event is then identified as *ingroup*, *outgroup*, or *celebrity* in the *Social Information Processing* module and then converted to *contact*, *caring*, and *status* values shown in Table 2.

Table 2. Parameter values of the best fits of the three components of target (contact, caring, and status) for ingroup, outgroup, and celebrity for the three models.

Model 1			
	t _{contact}	t _{care}	t _{stat}
Ingroup	1	0.9	0.2
Outgroup	0.1	0.1	0.1
Celebrity	0.1	0.3	1
Model 2			
	t _{contact}	t _{care}	t _{stat}
Ingroup	X	0.8	X
Outgroup	X	0.05	X
Celebrity	X	0.5	X
Model 3			
	t _{contact}	t _{care}	t _{stat}
Ingroup	X	0.8	X
Outgroup	X	0.05	X
Celebrity	X	0.5	X

This content, valence, and target information are all then sent to the valuation module, the critical subprocess in which the central agent decides whether or not to take a particular action (in our case, gossip) among a set of possible choices. To maximize expected outcome, the central agent needs to carefully consider all the pros and cons of the possible actions. We next describe this valuation process (and then action selection) further.

Our Main Model: Model 1 To decide which action to take, the central agent needs to estimate the outcome of every possible action based on *benefits* and *costs*. Because our current focus is the conditions under which one would or would not communicate to someone (a receiver) about someone else (the target), our model considers two actions: (1) gossiping or (2) not gossiping. In this case, *the costs of*

gossiping become the benefits of not gossiping. The equations, then, are the following:

$$Value_{Gossip} = A \cdot B_{Total} \quad (1)$$

$$Value_{Not\ Gossip} = A \cdot C_{Total} \quad (2)$$

where A is the affect score of the given information (listed in Table 1), such that the given action value increases with the affective response.

Based on literature review and our own theoretical development, Table 2 shows the list of potential benefits and costs involved in gossip (Aronson et al., 2016; Cosmides & Tooby, 1992; Dunbar, Marriott & Duncan, 1996; Dunbar 2004; Foster, 2004; Gazzaniga, Ivry & Mangun, 2013; Haidt, 2007; Kralik, 2017; Kralik et al., 2018; Lee, Kralik, & Jeong, 2018; 2019). The most obvious benefit of gossip is that one can avoid facing the target directly (B1), especially if an unfavorable outcome is expected (e.g., the target becoming upset). Additionally, the central agent may obtain further information from the receiver about the incident or target (B2). Third, and especially critical, the information provided to receivers can update their ‘broken’ models about the target and the world (B3). Fourth, gossip can also promote fairness balance and societal stability by rewarding positive actions and punishing negative ones. This can potentially be accomplished by influencing the social status of the target, via affecting their reputation (B4). Additionally, fifth, receivers may be in better position to directly contact the target to reward or punish the behavior (B5). Finally, the target and receivers may enjoy entertaining target activities and learn from them (B6).

Although this indirect form of communication has many benefits, there are obvious costs. Although gossip allows avoiding direct contact with the target (B1), the target might yet ascertain the source of the gossip (i.e., the central agent) and retaliate (C1). Moreover, there is a risk that event information (and thus, about the target) is incorrect (C2), resulting in deleterious effects such as altering the target’s social status. Because maintaining accurate models of others’ minds is critical within a multi-agent society, sharing false information sows confusion. Third, the central agent may also earn a bad reputation as a gossip (C3). Fourth, related to B2, additional information from receivers may be wrong or misleading (C4). Fifth, because of its indirectness, influencing the target via gossip may not match what the central agent intended (C5). Finally, choosing to gossip, like any action, requires both cognitive and behavioral effort, both to take the action and monitor its effects (C6) (Lee, Kralik & Jeong, 2019).

Total benefits and costs (B_{Total} and C_{Total} in the equations) are the summation of all potential benefits and costs. That is:

$$B_{Total} = B_1 + B_2 + \dots \quad (3)$$

$$C_{Total} = C_1 + C_2 + \dots \quad (4)$$

where each benefit B_i and cost C_i are then calculated as a function of the target factors T (i.e., contact, care, and status), valence V , and benefit-cost weighting factor w thus:

$$B_i = T_{contact, Bi} \cdot (T_{care, Bi} + T_{status, Bi}) \cdot V_{Bi} \cdot w_{Bi} \quad (5)$$

$$C_i = T_{contact, Ci} \cdot (T_{care, Ci} + T_{status, Ci}) \cdot V_{Ci} \cdot w_{Ci} \quad (6)$$

The benefit-cost weighting factors w are listed in Table 3, and are again based on theoretical considerations, literature review, and adjusting for model fitting (with yet maintaining the general relative positions among them).

To calculate target and valence factors (i.e., $T_{contact}$, T_{care} , T_{status} , and V), we combine the target and valence values input to valuation — i.e., those in Table 2 and the valence 1 (positive) or 0 (negative) flag — with gating values, g , as 0 (irrelevant) or 1 (relevant), based on whether each factor is relevant to the given benefit or cost. The gate values are listed in Table 3.

Table 3. Table 3. Benefits and costs of gossiping. Cells contain weighting (w) and gating values. The gates are social or valence filters and used in the valuation equation.

Valuation Categories	w	Target (T)			Valence (V)
		$g_{contact}$	g_{care}	g_{status}	v
B_1 : Avoid direct contact with the target	1	1	0	0	1
B_2 : Feedback to the gossip from receiver	0.4	0	1	0	0
B_3 : Update receiver’s knowledge	1	0	1	0	0
B_4 : Influence target’s social status	0.9	0	0	1	1
B_5 : Receiver influences target’s behavior	0.8	0	1	0	1
B_6 : Entertainment and social learning	0.4	0	1	0	1
C_1 : Potential direct contact from the target	0.6	1	1	0	1
C_2 : Risk of spreading wrong information	0.4	0	0	1	1
C_3 : Earn bad reputation as a gossip	0.7	0	0	0	1
C_4 : Get wrong/misleading information from receivers	0.9	0	0	0	0
C_5 : Influence target improperly	0.8	0	1	0	1
C_6 : Cost effort both cognitively and behaviorally	0.6	0	1	0	1

Contact is based on whether the target can actually reach the central agent; *care* indicates how much the central agent cares about the target, such that the parameter is high if the

central agent is invested in the outcome; and *status* represents the target's position within the social hierarchy, such that celebrities are high, ingroup low, and outgroup the lowest. The input values and gates are then combined thus:

$$T_{contact} = I + g_{contact} \cdot (t_{contact} - I) \quad (7)$$

$$T_{care} = g_{care} \cdot t_{care} + (I - g_{care}) \cdot ave - t_{care} \quad (8)$$

$$T_{status} = g_{status} \cdot t_{status} \quad (9)$$

$$V = [(I - v_{default}) \cdot g_{valence} \cdot v] + v_{default} \quad (10)$$

In general, for contact, if the central agent and target cannot directly contact each other (i.e., contact gate is 0), then the particular benefit or cost has no effect for the given scenario; for care, if caring is relevant for the given benefit or cost (gate=1), then it becomes the value for the target group in Table 2, otherwise (gate=0) it is the average of the three target group values. Finally, for valence, some benefits and costs would be expected to have a greater impact for negative events (such as directly contacting the target when they've done something egregious), and we represented this by having a default value for positive valence which increases for negative valence if the gate is 1.

Model 2 Although we believe the factors and values for our main model are well justified, it nonetheless is important to test their significance in the model. We thus developed two competing models that simplified prominent factors. For Model 2, gating values for the *contact* and *status* components of the *target* were all set to 0 while the gates for *care* were all set to 1. That is, here we collapsed the target components of contact, care, and status into one general factor that represented the difference among the three target categories of ingroup, outgroup, and celebrity. Table 2 shows the best fit values for this vector.

For *valence*, Model 2 does not consider cases where negative events may be more impactful than positive ones, and thus sets all valence gates to 0, using only the default value for all events. The best fit $v_{default}$ was 0.7.

As $g_{contact}$, g_{status} , and $g_{valence}$ are all set to 0, the valuation equations for Model 2 become:

$$B_i = T_{care, Bi} \cdot v_{default} \cdot w_{Bi} \quad (11)$$

$$C_i = T_{care, Ci} \cdot v_{default} \cdot w_{Ci} \quad (12)$$

This equation in turn means that the individual benefits and costs are considered as one total value for each. To best fit these benefit and cost weights, then, one parameter was used for benefits (α_{gossip}) and one for costs ($\alpha_{not-gossip}$). For Model 2, the best fit values were $\alpha_{gossip}=1.05$; and $\alpha_{not-gossip}=1$.

Model 3 The second competing model, Model 3, was the same as Model 2 except for *valence*. For this model, for all benefits and costs, all negative events were considered more impactful than positive ones; and thus, all valence gates were

set to 1. The Model 3 valuation equations then are expressed as:

$$B_i = T_{care, Bi} \cdot [(I - v_{default}) \cdot v + v_{default}] \cdot w_{Bi} \quad (13)$$

$$C_i = T_{care, Ci} \cdot [(I - v_{default}) \cdot v + v_{default}] \cdot w_{Ci} \quad (14)$$

Table 3 shows the best fit weights for the singular target effect. The best fit values for the relative benefit to cost effects were again $\alpha_{gossip}=1.05$; and $\alpha_{not-gossip}=1$.

Results

We first examine the main effects for content *valence* and *target*. We then examine the *target* by *valence* interactions.

Figure 2 shows how the three competing models compare to the empirical findings for percent gossip for the *valence* and *target* main effects. For *valence*, percent gossip in the empirical data was significantly higher for negative than for positive content. All three models obtained the same pattern. Additionally, Model 1 had a gossip percentage closer to the empirical results than Models 2 and 3.

For *target*, percent gossip in the empirical data was significantly higher for celebrities, followed by ingroup, then outgroup. The same pattern was obtained by Model 1, whereas Models 2 and 3 showed a different pattern, with ingroup the highest. Thus, our main model provided better fits, showing that the target effects indeed appear to be a

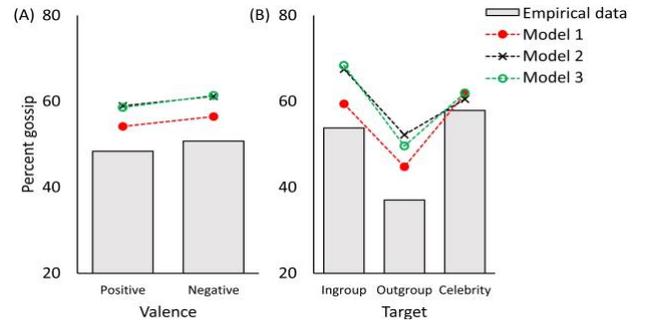


Figure 2. Results of the three models compared to the empirical data for gossiping with respect to (A) valence and (B) target main effects.

function of contact, care, and status, and both target and valence have distinct relevance to different benefits and costs.

We next examined the *target* by *valence* interactions. Figure 3 shows the results of the three competing models compared to the empirical data for percent gossip for ingroup, outgroup, and celebrity targets, broken down by positive and negative valence. For ingroup targets, as opposed to the general main effect finding of more gossiping about negative events, we have the opposite: percent gossip for positively valenced events was significantly higher than for negatively valenced ones in the empirical data. Indeed, Model 1 obtained the same pattern, whereas Model 2 and 3 showed the opposite pattern, similar to the main effect result.

For outgroup and celebrity targets, the empirical data showed higher percent gossiping for the negatively valenced events than for the positive ones (as for the main effect). In these two cases, all three models obtained the same pattern. Thus, Models 2 and 3 were unable to obtain different patterns across the three target groups, while Model 1 was able to explain the flipped relationship for ingroup targets.

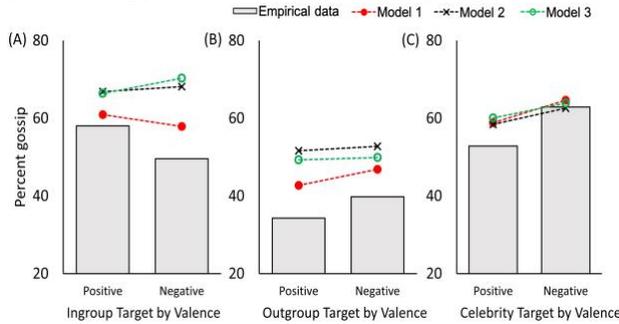


Figure 3. Results of the three models compared to the empirical data for gossiping about (A) ingroup, (B) outgroup, and (C) celebrity targets for different valence.

The main effect in the results, captured by Model 1, is that both the relative benefits of gossiping positively and costs of gossiping negatively are heightened for ingroup members. At the same time the relative lower cost and higher interest in justifying status led to increased negative gossiping about celebrities (Foster, 2004). For outgroup, a relative lack of interest dominated the findings, although the relative lower cost in negative gossiping was also observed.

We next examine the results for the *target by valence* interaction by examining positive and negative events for the three different target groups. For positively valenced events, Figure 4A shows that the percent gossiping in the empirical data was highest for ingroup targets, followed by celebrity, and finally outgroup. All three models obtained the same pattern, however, Model 1 obtained values closer to the empirical data. In contrast, for negative events, Figure 4B shows that percent gossiping in the empirical data was highest for celebrity targets, followed by ingroup, then outgroup. Model 1 again obtained the same pattern, whereas percent gossip about ingroup targets was higher than for celebrities for both Models 2 and 3. Thus, once again Models 2 and 3 showed the same patterns for positive and negative events, while Model 1 found different patterns for the two valences, matching the empirical data.

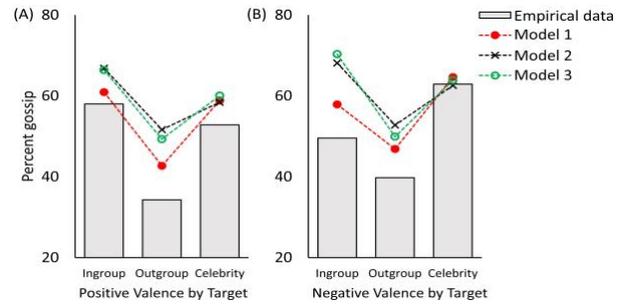


Figure 4. Results of the three models for (A) positive and (B) negative valence with respect to the different targets.

In sum, our main model, Model 1, was superior in not only the level of gossiping predicted, but in importantly capturing the general patterns in the findings, especially flipping the prediction for ingroup, with an increase in positive gossiping and decrease in negative for ingroup members.

Discussion

To understand how the human mind/brain has harnessed the ratcheting power of social information exchange, it is important to model how people process social events, and when, how, and why they choose to communicate the information to others. To this end, we have developed a general framework for human social intelligence and communication based on literature across the social sciences, and have begun developing a computational model detailing the processes. Here we presented a significantly elaborated computational version of our model, based on this literature and our own evolutionary and affective-sociopsychological theoretical considerations of why people should choose to communicate this information (Aronson et al., 2016; Cosmides & Tooby, 1992; Dunbar, Marriott & Duncan, 1997; Dunbar, 2004; Foster, 2004; Gazzaniga, Ivry & Mangun, 2013; Haidt, 2007; Kralik, 2017; Kralik et al., 2018; Lee, Kralik, & Jeong, 2018; 2019). We have also conducted an experiment in the laboratory in which we collected data on whether people would choose to communicate to others in order to test the predictions of our model (Lee, Kralik, & Jeong, *submitted*). Indeed, our model predictions were supported, successfully capturing the main patterns of results.

There are, of course, multiple avenues for future development. These include a more detailed consideration by the central agent of why the target acted as he/she did. The answers to “why” will require a richer set of social factors (beyond contact, care, and status), which will in turn be compared against the central agent’s own model of the target’s mind. For possible responses, we also plan to include communicating directly to the target (rather than to others about them). Eventually, more action detail also needs to be included, to carry out, for example, an extended conversation with the target and/or receiver. For event content, more activities are needed. We also intend to add learning to our

model, including the need for active monitoring of outcomes to assess actual action effectiveness, especially with indirect communication. Learning can also potentially capture cultural influences on moral dimension weightings.

Evidence shows that social intelligence and communication are comprised of relatively hard-wired components (and thus to some extent expert-like), together with more malleable ones, with their combination enabling *general* social intelligence across multiple content domains. Moreover, a comprehensive integration of the relevant literature across multiple fields of study shows that human social ability is indeed elaborate and complex. We thus believe that approaches such as ours that face this challenge head-on are necessary to ultimately understand human social processing, and endow artificial systems with the affective-sociocognitive processing machinery that truly leverages the power of sociality.

References

- Aronson, E., Wilson, T. D., Akert, R. M., and Sommers, S. R. (2016). *Social Psychology*. Pearson: NYC.
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In: J. H. Barkow, L. Cosmides, J. Tooby (Eds.). *The Adapted Mind*.
- Dunbar, R. I. M. (2004). Gossip in evolutionary perspective. *Review of General Psychology*, 8(2), 100–110.
- Dunbar, R. I. M., Marriott, A., and Duncan, N. D. C. (1997). Human conversational behavior. *Human Nature*, 8(3), 231–246.
- Foster, E. K. (2004). Research on gossip: taxonomy, methods, and future directions. *Review of General Psych.* 8, 78–99.
- Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2013). *Cognitive neuroscience: the biology of the mind*. WW Norton & Company.
- Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316(5827), 998–1002.
- Kralik, J. D. (2017). Architectural design of mind & brain from an evolutionary perspective. *Proc. AAAI 2017 Fall Symposium: A Standard Model of the Mind*.
- Kralik, J. D., Lee, J. H., Rosenbloom, P. S., Jackson, Jr., P. C., Epstein, S. L., Romero, O. J., Sanz, R., Larue, O., Schmidtke, H. R., Lee, S. W., McGreggor, K. (2018). Metacognition for a Common Model of Cognition. *Procedia Computer Science*: 145, 730–739.
- Lee, J., Kralik, J. D., & Jeong, J. (2018). A Sociocognitive-Neuroeconomic Model of Social Information Communication: To Speak Directly or To Gossip. *Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci)*.
- Lee, J., Kralik, J. D., & Jeong, J. (2018). A General Architecture for Social Intelligence in the Human Mind and Brain. *Procedia Computer Science*: 145, 747–756.
- Lee, J., Kralik, J. D., & Jeong, J., Submitted.

Modelling Visual Decision Making Using a Variational Autoencoder

Tyler Malloy (mallot@rpi.edu)

Department of Cognitive Science
110 8th Street, Troy, NY 12180

Chris R. Sims (simsc3@rpi.edu)

Department of Cognitive Science
110 8th Street, Troy, NY 12180

Abstract

Due to information processing constraints and cognitive limitations, humans necessarily form limited representations of complex visual stimuli when making utility-based decisions. However, it remains unclear what mechanisms humans use to generate representations of visual stimuli that allow them to make predictions of utility. In this paper, we develop a model that seeks to account for the formation of representations in utility-based economic decision making. This model takes the form of a β -variational autoencoder (β -VAE) trained with a novel utility-based learning objective. The proposed model forms representations of visual stimuli that can be used to make utility predictions, and are also constrained in their informational complexity. This representation modelling approach shares common features with related methods, but is unique in its connection to utility in economic decision making. We show through simulation that this approach can account for several phenomena in human economic decision making and learning tasks, including risk-averse behaviour and distortion in the calculation of expected utility.

Keywords: Cognitive Modelling, Decision Making, Information Theory

Introduction

In the context of decision making, a representation refers to the internal mental state of an agent, encompassing features from the external environment that are relevant to the decision task and the agent's objectives. The mechanisms of this representation formation must depend on the task being performed, as decision makers should seek to efficiently represent task-relevant information while ignoring or abstracting across irrelevant information. One important class of tasks which we study in this paper is that of economic decision-making based on visual stimulus.

Neuroeconomics has sought to further the understanding of the neurological underpinnings of economic decision making, though relatively little work in this area has focused on the mechanisms behind representation formation from visual stimuli. One potential mechanism for modelling this cognitive process is the variational autoencoder (VAE), a method that learns informationally limited representations of input that can be used to form lossy reconstructions (Pu et al., 2016). In this paper, we present an extension of the VAE framework that produces task-relevant representations of economic decision tasks, and predicts human-like decision making.

Traditional variational autoencoder models incorporate an information constraint based on the structure of the neural

network that implements them (e.g., limiting the number of nodes in a hidden layer), but the capacity of the autoencoder is not easily controlled. β -VAEs are a variant incorporating an additional parameter that controls this information bottleneck, encouraging the model to learn more informationally compact representations. The novelty of our work lies in the application of β -VAEs onto economic decision making. Our model also differs from related methods in machine learning through the use of a novel loss function that balances stimulus reconstruction error and loss in expected utility. The model predicts at a qualitative level the decision making of an individual who has limited information processing capacity, but is otherwise rational in seeking to maximize expected utility.

The main feature of β -VAEs in limiting the amount of information used to form internal representations shares a connection to other information theoretic methods such as the information bottleneck approach and rate-distortion theory (Burgess et al., 2017). The latter has been used to conceptualize human perception as optimizing task performance subject to a constraint on channel capacity (C. R. Sims, 2016).

Rate-distortion theory has also been used to model generalization of perception as resulting from the encoding of perceptual information in an efficient way that can then be used to generalize over novel experience (C. R. Sims, 2018). This effect has been shown to produce biases in statistical and categorical learning from visual features that mimic effects present in human perception (C. J. Bates, Lerch, Sims, & Jacobs, 2019). The model described in this paper seeks to achieve the same goals as these previous methods while also producing an internal representation of a perceived stimulus and a method of explicitly estimating task-relevant utility based on these representations.

Related Methods

β -VAEs have previously been used to model the formation of task-relevant representations in visual categorization and change detection tasks (C. Bates & Jacobs, 2019). In the work of Bates et al., the model consisted of a β -VAE which formed internal representations of stimulus, and a decision module that completed the task being performed based on these representations. Results from this experimentation demonstrated a categorical bias in reconstruction depending on which task is being modelled. This supports the use of the β -VAE framework in modelling the formation of task-relevant visual stim-

ulus representations.

The model and experimentation presented in this paper is an extension of this method onto the domain of economic decision modelling from visual stimulus, which requires adjustments in model structure and training. This is an important extension, as utility is the basis of economic decision making in cases where agents have access to utilities and probabilities required to determine optimal actions. Utility can also serve as the basis for models of human reinforcement learning used to make predictions of human decisions in learning tasks (Niv, 2009; Niv et al., 2015; Collins & Frank, 2012). The ability of extending the proposed model into the domain of reinforcement learning modelling will be further investigated in the discussion section.

The function and motivation behind β -VAE models shares a close connection with the information bottleneck approach (Burgess et al., 2017; Alemi, Fischer, Dillon, & Murphy, 2017). This method has been applied to modelling cognitive mechanisms that share similarities with economic decision making, such as predictive inference (Still, 2014) and information-constrained behaviour (Lai & Gershman, 2021; Malloy & Sims, 2020). One key feature of our proposed model is that it makes predictions on the formation of task-relevant representations under information constraints. Previous methods applying the information bottleneck approach to decision modelling have either not included representation formation, or done so in a task that did not involve utility predictions.

Within the field of economic decision modelling, sub-optimality in human decision making is understood within the frameworks of bounded rationality (Simon, 1990; Camerer, 1998) and rational inattention (C. A. Sims, 2003; Mackowiak, Matejka, Wiederholt, et al., 2020). Models developed under these frameworks can be used to predict how humans make decisions relative to their information processing limitations. As with previously discussed methods, these too do not explicitly model the formation of task-relevant representations of stimuli.

In this paper we present experimentation utilizing our proposed model resulting in similar predictions of sub-optimality in decision making as these related methods, while additionally modelling visual representation formation. As with all cognitive models based in neural networks, these representations are a metaphor for the contents of human cognition. However, through its novel structure and training method our model makes implications for how constrained representation formation can lead to sub-optimal performance.

Modeling Representations using β -VAEs

β -Variational Autoencoders

Variational autoencoders consist of a neural network which compresses an input into a lower dimensional representation that is then expanded back into a reconstruction of the input. The first half of this network structure is referred to as the encoder, while the second half is referred to as the decoder.

Both portions of the network are trained simultaneously as the network takes in some input and produces an output, and through training learns to reconstruct the input as faithfully as possible.

The closeness of this reconstruction is defined by a *loss function* which determines how similar the reconstruction is to the original input. Typically in VAEs the loss function is an error between the model input and output, such as the mean-squared-error.

In β -VAEs, an additional parameter (β) is introduced to control the information capacity of the lower dimensional representation, which results in an adjustable information bottleneck (Higgins et al., 2017; Mathieu, Rainforth, Siddharth, & Teh, 2019). The loss function used to train a β -VAE is as follows:

$$\mathcal{L}(\theta, \phi; x, z, \beta) = \mathbb{E}_{q_{\phi}(z|x)}[\log p_{\theta}(x|z)] - \beta D_{KL}(q_{\phi}(z|x)||p(z)) \quad (1)$$

In the above, ϕ represents the parameters of the encoder $q_{\phi}(z|x)$, which defines the probability distribution over latent representations z given the stimulus x . Additionally, $p_{\theta}(x|z)$ can be understood analogously with the decoder, as it defines the probability that a stimuli x can be produced from the latent representation z . The desired decoder is one where $p(x|z) \approx p(x|v, w)$ where v are the conditionally independent generative factors responsible for producing the stimulus, and w are the conditionally dependent factors. When $p(x|z) \approx p(x|v, w)$ the latent representation z is an adequate representation of the generative factors responsible for producing the stimuli x , as the probability of observing the original data given the latent representation is maximized. For a more complete description, see (Higgins et al., 2017).

The first term in this loss function represents the reconstruction error between the model input and output. The second term $D_{KL}(q_{\phi}(z|x)||p(z))$ represents the informational complexity of the internal representations that the model generates. When $\beta = 0$ the model seeks only to minimize reconstruction error, and as β increases the amount of information used in internal representations decreases. The loss-function used in VAEs corresponds to $\beta = 1$ and as β increases a more constrained information bottleneck is applied.

The usefulness of the β -VAE method in modeling human decision making over the traditional VAE approach is in its ability to adjust the information constraint on the latent representation. For modeling human decision making, it should be possible to determine an individual participant's information processing constraint and fit the β parameter to match that. This suggests that the proposed β -VAE model might better capture the way that individuals form representations of decision making tasks given their individual capacity for storing and processing information.

In our proposed model, the β -VAE represents a Working Memory Module (WMM) of an agent when they are making a decision based on visual stimulus. However, because stimulus representations should be domain specific due to information processing constraints, we must train this model

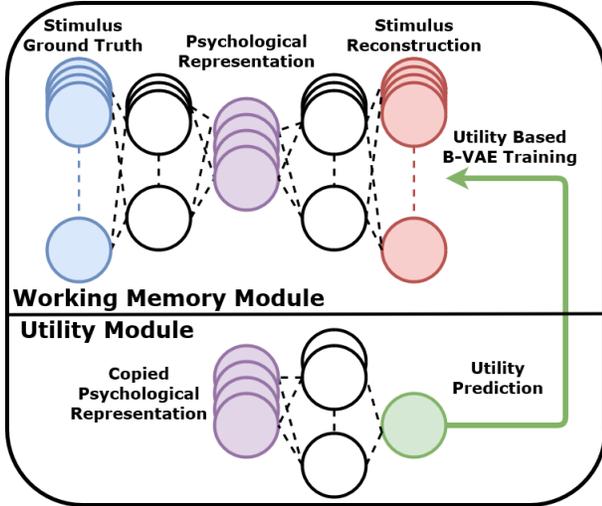


Figure 1: The Working Memory Module is a β -VAE structure that learns to reconstruct the stimulus ground truth and Utility Module prediction accuracy. Colors highlight the stimulus ground truth presented to a decision maker (blue), the internal representation that they use as working memory of the decision making problem (purple), the reconstructed stimulus (red), and the predicted utility of the stimulus (green).

to not only make accurate reconstructions, but also to allow for accurate utility predictions. This is done through the addition of a utility prediction module that allows for utility based training of the WMM.

Utility Based Training

As we are interested in modeling task-relevant representation formation and decision making, the training method of the proposed β -VAE model is adjusted to incorporate the utility of the learned representation. An additional Utility Module learns to predict the utility associated with a stimulus based on the internal representation of a stimulus that is learned by the WMM (Fig 1).

This utility module consists of a neural network that takes as input a copy of the psychological representation of a stimulus, and outputs a prediction of the utility associated with that stimulus. The network is fully connected with 2 layers of 64 units, and the output is trained based on a mean squared error loss between the prediction and the ground truth utility. This utility module is trained alongside the WMM on the same data, with the additional ground truth utilities. The utility predicted estimate is then fed into the loss function of the WMM which is trained to balance the accuracy of the stimulus reconstruction and the utility prediction as follows:

$$\mathcal{L}(r, S) = \mathcal{L}(\theta, \phi; x, z, \beta) + \nu (r(Z) - \mathbb{E}[S])^2 \quad (2)$$

Where $r(Z)$ is the utility prediction output of the Utility Module, and $\mathbb{E}[S]$ is the ground truth expected utility associated with the stimulus input to the WMM. Using this altered train-

ing method, the WMM learns to reconstruct the stimulus accurately, while reducing the squared error between the utility prediction and the stimulus ground truth utility. This structure is similar to the β -VAE based method described in (C. Bates & Jacobs, 2019), though our model explicitly predicts expected utility associated with a stimulus and uses that prediction within the WMM learning objective. These alterations allow for both predictions of utility, as well as working memory representation formation to take utility into account.

Adjusting the utility-loss weight parameter ν controls the relevance of the expected utility in calculating the loss of the model's reconstruction. For example, when the ν parameter is 0, the model learns to reconstruct stimulus as faithfully as possible without accounting for the accuracy of the utility module prediction. As ν increases, the model learns to prefer representations that allow for more accurate utility predictions. Comparing the predicted decisions of the β -VAE model with human selections in economic decision making tasks can allow for a better understanding of how humans balance task-relevance and memory reconstruction accuracy to the original stimulus when forming representations.

Economic Decision Making

Maximum Expected Utility

Expected utility is defined for a decision alternative \mathbf{x} based on the different outcomes that can occur as a result of selecting that alternative $[x_1, x_2, \dots, x_i]$, the utility of those outcomes $[u(x_1), \dots, u(x_i)]$ and the probability of those outcomes occurring $[p(x_1), \dots, p(x_i)]$ given the option that was selected by the agent. This results in the following equation for expected utility:

$$\mathbb{E}[u(\mathbf{x})] = \sum_{i=1}^n p(x_i) u(x_i) \quad (3)$$

The proposed model takes as input a single option within in a decision problem X , corresponding to outcomes and outcome probabilities, and reconstructs that input as faithfully as possible given the information constraint. However some features of the stimulus are more relevant for maximizing utility than others, so the utility-loss method is introduced to incorporate the difference between the predicted utility and the ground truth.

The utility value of the original stimulus $\mathbb{E}[S]$ in the utility-based loss function in Eq.2 is equal to the true expected utility of the original stimulus. This differs from the expected utility of the reconstruction stimulus $r(Z)$ due to the information constraint applied to the internal representation of stimuli. The calculation of these ground-truth utilities are specific to the task being performed which will be fully detailed in the experimentation section.

Sub-optimal Decision Making

A well-studied form of sub-optimality in human economic decision making is risk-aversion, which is characterized by the undervaluing of risky prospects and overvaluing of safe

prospects, relative to their true expected utility (Pratt, 1978; Holt & Laury, 2002). Traditionally, this phenomenon is accounted for by introducing an adjusted utility function that treats outcomes differently based on their value or probability (Rabin & Thaler, 2001). An example of this approach is Cumulative Prospect Theory (CPT) (Kahneman & Tversky, 1979) which has been used to model risk-aversion like effects in economic decision making (Schmidt & Zank, 2008). CPT can account for the effect of risk-aversion by weighting the utility of an outcome based on its value or probability, such as reducing the weight for outcomes that are unlikely and increasing it for likely outcomes (Schmidt & Zank, 2008).

In the following section on experimentation, we will demonstrate that our proposed model makes similar predictions of risk-averse behaviour in an economic decision making task. Importantly, the input to this model will be a visual stimulus representation of a decision making task. This makes it unique from related methods like CPT which take in as input the probabilities and outcomes associated with different options in a decision making task. Additionally, the β -VAE module within our model allows for the formation of psychological representations and stimuli reconstructions of these input that are not present in previous models of risk-aversion in economic decision making.

Experimentation

Previous methods have shown that β -VAEs can be used to produce task-relevant biases in representation formation similar to what can be expected from humans based on their behaviour (C. Bates & Jacobs, 2019). Through experimentation, we show similarly human-like behaviour when modelling visual stimulus representation formation in a utility-based economic decision making task. This is done by showing a risk-aversion effect present in the utility predictions of our model that correspond to expectations of human decision making in similar tasks. For all β -VAE models described in this paper, the model structure, hyper-parameters, and training procedure follow the original implementation described in (Higgins et al., 2017), apart from the β information constraint which is adjusted to compare different information processing constraints as described in the following sections.

Decision Making Task

We examine the behavior of our model in a “marble jar” selection task. In this task, the agent is presented with a choice between two jars of 16 marbles, where the contents of each jar are fully visible. After selecting a jar, one marble is randomly sampled from the chosen jar, and the agent receives a reward based on the color of the (randomly) selected marble.

To compare the impact of information constraints and utility-weight parameters on choice behavior, we vary these parameters and report the utility prediction and reconstruction accuracy of models at the end of training. All models are trained on 1K epochs of 1K stimuli and utility values. Marble jar stimuli are generated using a Dirichlet distribution

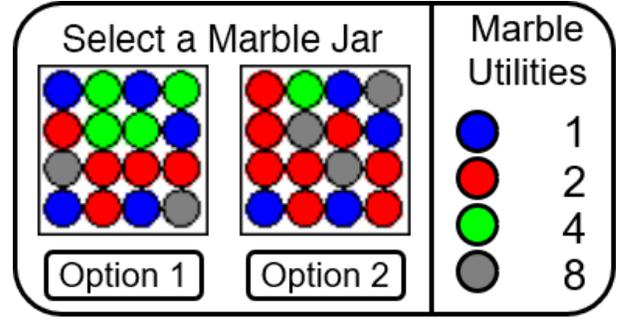


Figure 2: An example of the marble jar selection task. The decision maker chooses one of the marble jars and a single marble chosen at random will be given to the agent with the goal of maximizing their observed utility. Each color marble has a different utility and each marble jar has an expected utility and utility variance which are the mean and variance of the marble utilities. Marble ratios are defined by a Dirichlet distribution $Dir([2, 4, 6, 8])$ for grey, green, red, blue.

$Dir([2, 4, 6, 8])$ for grey, green, red and blue colored marbles, which have utilities 8, 4, 2, 1 respectively.

The task described in Figure 2 will be used to demonstrate how our model predicts human-like decision making, specifically risk-aversion, through the use of our β -VAE model trained with the utility-based learning objective.

Modelling Results

Using the example decision task shown in Figure 2 we can see how risk-aversion could be demonstrated by a decision maker. The first and second options have total utilities of 46 and 48 and variances of 5.26 and 6.66 respectively. While the second option has a slightly higher expected utility, it also has a higher variance which may impact the choice of the decision maker. In this example, a bias in choosing option 1 over option 2 would reflect an instance of risk-aversion, as the decision maker is preferring certainty in outcome over a increase in expected utility.

risk-aversion: We can use this decision making task to investigate how the utility weight parameter υ and information constraint parameter β impact the decisions of our model. We additionally include for comparison MEU and CPT calculations for these probabilities. Because of the flexibility of CPT, a wide range of possible values for predicted utility are possible, and these values are selected to reflect the risk-averse effect observed in human decision making.

Model	Utility Predictions	Recon. Error
MEU (normative)	(46, 48)	N/A
β -VAE + Utility	(48.2, 42.1)	1611.4736
β -VAE	(45.6, 47.8)	1089.0422
VAE + Utility	(46.6, 47.8)	2810.2334
VAE	(41.3, 44.8)	1803.3821
CPT	(47, 45)	N/A

The results shown in the table above indicate that the β -

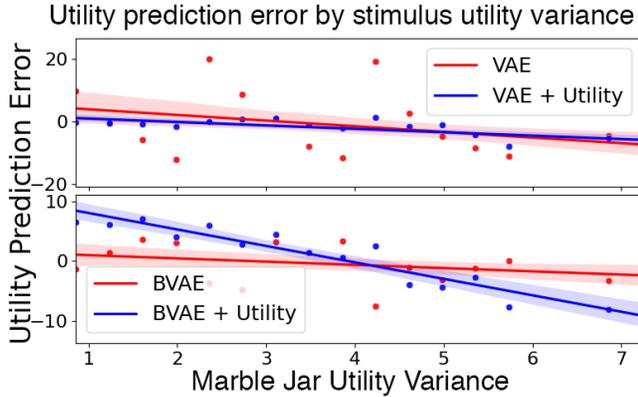


Figure 3: Utility prediction error based on marble jar utility variance of an ablation of β -VAE and VAE models with and without utility based training. Points represent mean utility predictions of all marble jars with the same utility variance. Lines represent a linear regression of all predicted utilities, calculated with the Seaborn Python library (Waskom, 2021). VAE models have $\beta = 1$, β -VAE models have $\beta = 100$, utility models have $\nu = 1000$ and non-utility models have $\nu = 0$.

VAE + Utility model demonstrates a risk-aversion effect for the example stimulus in Figure 2. Although the second jar has higher expected utility, the model values the first jar more highly. This effect can also be observed using a cumulative prospect theory model, as it has decision weight parameters that can be adjusted to produce a similar effect. However, it is important to note that the CPT model would need to fit an individual parameter for each possible outcome, whereas our proposed model is parameterized only with the utility-weight and information constraint. Additionally, CPT functions by altering the utility maximization method, whereas our approach assumes decision makers maximize utility, but doing so with limited information processing ability.

Comparing the reconstruction errors for each model demonstrates the improved generalizability to out of training stimuli afforded by the β -VAE models, which is one main justification for their use (Burgess et al., 2017). Additionally, each model trained with the utility prediction method has a lowered reconstruction accuracy. This corresponds with the expectation of a model with limited information processing, as information used to represent utility can lower the amount of information available to accurately reconstruct the original stimulus. An interesting result is that the loss-aversion like effect is observed as a result of the differences between differently colored marble proportions and utilities, as opposed to an imposed preference as is the case with the CPT model.

Note that the difference in predicted utility is exaggerated from what a human would likely determine for this task, as the information constraint ($\beta = 100$) and utility weight ($\nu = 1000$) are larger than values that would better reflect human behaviour. In practice, it is possible to fit these parameters based on observations from individuals, as has been

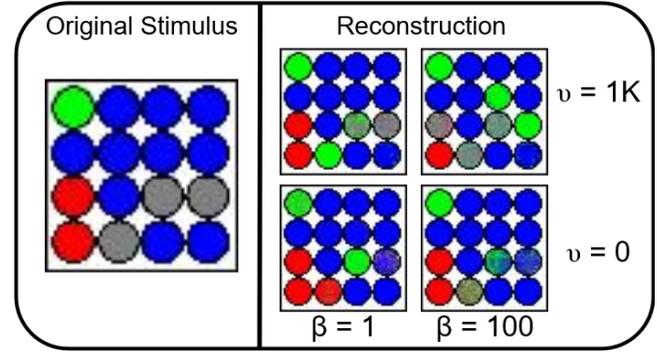


Figure 4: Comparisons of original stimulus and reconstruction for an example marble jar with the same ablation of different model types as described in Figure 3. Marble jar original stimulus is from the constructed set of marble jars with utility 42 that were not part of the original training data set.

done in related methods (Malloy & Sims, 2020; Niv et al., 2015; Collins & Frank, 2012). This could result in utility predictions matching the behaviour of individual participants, though this is outside of the scope of the present work.

Utility Estimate Bias: In order to understand the basis for the risk-aversion effect, we examined how the utility training method and information constraint impact utility predictions as the risk associated with a stimulus changes. Additionally, we sought to examine the generalizability of our utility prediction module to stimuli that have not been seen previously. To allow for this, we constructed a new data set consisting of every possible marble jar with the same total utility (42) but with different variances. The following figure compares the bias in predicted utility for models with and without an information constraint and utility-based training.

These results show that the β -VAE model alone demonstrates a risk-aversion like effect, shown by a positive utility error for low variance marble jars, and a negative utility error for higher variance jars. This corresponds with our understanding of the risk-aversion effect which increases the distortion of preference as outcomes become more or less probable.

Generalization: These utility prediction error results additionally demonstrate the high generalizability of utility prediction in models trained using the utility-based learning objective (blue). Models trained without utility included in their learning objective (red) have a wider range of prediction errors above and below the regression trend. This reflects the similarly high generalizability of human economic decision making in these types of utility-based tasks.

The final comparison of different model types is in the reconstruction accuracy of a new stimulus not used on training. Figure 4 shows a stimulus not used during training. The right hand side compares stimulus reconstructions of the same 4 model ablation described previously, with and without an information constraint and utility-based training.

These reconstruction examples demonstrate the impact that

information constraints and utility-based training have on generalized stimulus reconstruction. Models trained with the utility-based learning objective are better able to reconstruct the higher utility grey marbles. Interestingly, these grey marbles are not necessarily in the same location, as this is not relevant to the expected utility of a marble jar. The model thus demonstrates that its latent representations have acquired a useful invariance (the exact position of a marble in the jar). This corresponds to intuition from human perceptual memory in this type of task, as the location of marbles is irrelevant to predicting utility.

Discussion

Risk-Averse Representations

Results from utility prediction errors of our proposed model demonstrated a risk-aversion like effect. These results suggest that one aspect of risk-aversion is the formation of informationally compressed representations of visual stimulus in economic decision making tasks. However, alone these results do not fully explain the source of risk-averse representation formation. This can be better understood by considering the relative abundance of marbles and their utilities.

Marble piles with more grey marbles have higher variance, and one possible interpretation of underestimating these utilities would be a decision maker determining the utility of a stimuli with grey marbles by counting only the utility of those grey marbles and nothing else, leading to an under estimate. However, this is one possible explanation and additional comparisons would need to be made to more fully understand the precise ways in which these utility estimates are risk-averse. Generally the risk-averse behaviour should be understood as resulting from different probabilities and utilities of marbles, and how the information constraints and utility-weights impact representations of these stimuli.

An important implication of this explanation is that it would be unlikely to observe the same risk-averse behaviour in an alternate version of the marble task that is very uniform in marble probability or utility. While this is a slight weakness to our proposed model, these types of stimuli would also likely result in only a slight risk-averse behaviour in humans, since marble piles would on average have utilities much closer together. Additionally, the proposed model seeks to account for one source of risk-aversion, though others are likely to exist. It is possible that the type of risk-averse behaviour that results from forming informationally-compressed and utility-based representations only occurs when there is a considerable difference between stimuli utility.

Human Representation Formation

The modeling experiments presented in this paper sought to examine the properties of stimulus representations learned when facing constraints on the ability to encode and represent task features. To do this, we represented visual decision making in a similar manner as previous approaches, with a β variational autoencoder trained to learn internal representa-

tions, and a separate module trained to perform a task based on those representations. The novelty of our method is in its learning to explicitly predict utility based on task-relevant visual representations in an economic decision task.

Our results demonstrate that when agents face constraints on the ability to encode information veridically, systematic distortions are introduced in the representation of the probability and utility of decision alternatives. In particular, stimuli with a higher utility variance have a lower predicted utility, with the opposite being true for stimuli with a low utility variance. This corresponds to observations of human behaviour in economic decision making tasks. Importantly, our model makes similar predictions as existing methods while taking the input to be the visual task stimuli, and producing a psychological representation that can be used to reconstruct the original stimulus and make utility predictions.

Modelling Human Learning

As mentioned previously, the inclusion of utility predictions and a utility-based training method within our proposed model can allow for the modelling of reinforcement learning in humans. This can be done by adjusting the training method of the utility prediction module. In our experiments, decision makers were assumed to have knowledge of outcome probabilities and utilities. In the learning setting, these values would not be known and instead learned by making decisions and observing outcomes. Thus our utility prediction model would make a prediction and observe an outcome after the decision has been made, and update their prediction based on this observed outcome. This can be done using the standard temporal difference equation used in reinforcement learning, which is motivated by human biological processes implicit in learning (Niv, 2009; Niv et al., 2015).

References

- Alemi, A. A., Fischer, I., Dillon, J. V., & Murphy, K. (2017). Deep variational information bottleneck. *International Conference on Learning Representations*.
- Bates, C., & Jacobs, R. (2019). Efficient data compression leads to categorical bias in perception and perceptual memory. In *Cogsci* (pp. 1369–1375).
- Bates, C. J., Lerch, R. A., Sims, C. R., & Jacobs, R. A. (2019). Adaptive allocation of human visual working memory capacity during statistical and categorical learning. *Journal of vision*, *19*(2), 11–11.
- Burgess, C. P., Higgins, I., Pal, A., Matthey, L., Watters, N., Desjardins, G., & Lerchner, A. (2017). Understanding disentangling in β -vae. *NIPS Workshop on Learning Disentangled Representations*.
- Camerer, C. (1998). Bounded rationality in individual decision making. *Experimental economics*, *1*(2), 163–183.
- Collins, A. G., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? a behavioral, computational, and neurogenetic analysis. *European Journal of Neuroscience*, *35*(7), 1024–1035.

- Higgins, I., Matthey, L., Pal, A., Burgess, C., Glorot, X., Botvinick, M., . . . Lerchner, A. (2017). beta-vae: Learning basic visual concepts with a constrained variational framework. In *International conference on learning representations*.
- Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, *92*(5), 1644–1655.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292.
- Lai, L., & Gershman, S. J. (2021). Policy compression: An information bottleneck in action selection. Academic Press. doi: <https://doi.org/10.1016/bs.plm.2021.02.004>
- Mackowiak, B., Matejka, F., Wiederholt, M., et al. (2020). Rational inattention: A review. *CEPR Discussion Papers*(15408).
- Malloy, T., & Sims, C. R. (2020). Modelling human information processing limitations in learning tasks with reinforcement learning. *Proceedings of the international conference on cognitive modelling*, *18*, 159–160.
- Mathieu, E., Rainforth, T., Siddharth, N., & Teh, Y. W. (2019). Disentangling disentanglement in variational autoencoders. In *International conference on machine learning* (pp. 4402–4412).
- Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154.
- Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. *Journal of Neuroscience*, *35*(21), 8145–8157.
- Pratt, J. W. (1978). Risk aversion in the small and in the large. In *Uncertainty in economics* (pp. 59–79). Elsevier.
- Pu, Y., Gan, Z., Heno, R., Yuan, X., Li, C., Stevens, A., & Carin, L. (2016). Variational autoencoder for deep learning of images, labels and captions. In *Advances in neural information processing systems*.
- Rabin, M., & Thaler, R. H. (2001). Anomalies: risk aversion. *Journal of Economic perspectives*, *15*(1), 219–232.
- Schmidt, U., & Zank, H. (2008). Risk aversion in cumulative prospect theory. *Management Science*, *54*(1), 208–216.
- Simon, H. A. (1990). Bounded rationality. In *Utility and probability* (pp. 15–18). Springer.
- Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, *50*(3), 665–690.
- Sims, C. R. (2016). Rate–distortion theory and human perception. *Cognition*, *152*, 181–198.
- Sims, C. R. (2018). Efficient coding explains the universal law of generalization in human perception. *Science*, *360*(6389), 652–656.
- Still, S. (2014). Information bottleneck approach to predictive inference. *Entropy*, *16*(2), 968–989.
- Waskom, M. L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021.

Predicting Spatial Belief Reasoning: Comparing Cognitive and AI Models

Johannes Mannhardt¹ (johannes.mannhardt@students.uni-freiburg.de)

Leandra Bucher² (Leandra.Bucher@uni-siegen.de)

Daniel Brand¹ (daniel.brand@cognition.uni-freiburg.de)

Marco Ragni^{1,3} (ragni@sdu.dk)

¹ Cognitive Computation Lab, University of Freiburg, Germany, ² University Siegen, Germany,

³ Danish Institute of Advanced Studies, Denmark

Abstract

Spatial relational descriptions in everyday life sometimes need to be revised in the light of new information. While there are cognitive models for reasoning about spatial descriptions, there are currently no models for belief revision for the spatial domain. This paper approaches this need by (i) revisiting existing models such as verbal model (Krumnack et al., 2010) and PRISM (Ragni and Knauff, 2013) and adapt them to deal with belief revision tasks, (ii) evaluate these models by testing the predictive accuracy for the individual reasoner on a previously conducted experiment by Bucher et al. (2013), (iii) provide baseline models and machine learning models, provide user-based collaborative filtering and content-based filtering methods, and provide an analysis on the individual level. Implications for predicting the individual and identifying strategies and shared similar reasoning patterns are discussed.

Introduction

Belief revision refers to the cognitive ability of reasoners to change existing beliefs to eliminate contradicting beliefs. Imagine you believe “London lies further north than Berlin. Berlin is north of Krakow”. As you are in Krakow, you assume that it is no detour to first head towards Berlin. An expert in geography, however, tells you “Berlin actually lies north of London”. How do you revise your beliefs, given that you trust the expert? In what way do you change your previous assumption? Do you try to preserve as much of your preexisting beliefs as possible? You can assume that there are differences between individuals in doing so. What are the underlying cognitive mechanisms? Is it possible to predict what different reasoners will do?

Human spatial relational reasoning has been well-studied from a neuro-cognitive, computational, and experimental perspective (Knauff, 2013; Ragni and Knauff, 2013). Since it is a field with a lot of practical relevance, it is well suited for the exploration of belief revision, and by modeling, it should be possible to identify the factors influencing it. While for spatial relational reasoning there exists a vast literature, there are only a handful of studies on belief revision (Knauff et al., 2013; Krumnack et al., 2010). Still, these studies have identified systematic and surprising effects on revision preferences, including aspects such as plausibility and the properties of relations such as visualizability (Bucher et al., 2013). This provides enough variation to possibly explain individual variation.

The paper is structured as follows: In the next section, we will introduce the experiment, its results, and the resulting data set. In section 3 we will introduce existing models for spatial relational reasoning and how they have been adapted for belief revision. This

will be accompanied by models from machine learning such as recommender systems. We will present the evaluation in section 4 and finally discuss the results and their implications in Section 5.

The Data

The experimental data stems from Bucher et al. (2013). To determine the visualizability of the problems, a pilot study was carried out. 30 volunteers (14 male; aged from 19 to 55) evaluated 72 binary spatial and non-spatial statements related to their visualizability. Based on the results, a total of 192 problems were selected, 64 each for the categories visual, neutral, and spatial. Examples can be seen in Table 1.

Table 1: Examples for the three task categories.

Category	Examples
visual	The cucumber is thinner than the pumpkin. The asparagus is thinner than the cucumber.
neutral	The bird is weaker than the dog. The dog is weaker than the polar bear.
spatial	Russia is further east than Poland. Poland is further east than Germany.

The procedure of the experiment by Bucher et al. (2013) is now briefly described. Twenty volunteers (8 male; age 20-35; German as native language) were tested individually on 192 problems. Each participant was first presented with two statements (called premises), e.g.

Asparagus is thinner than cucumber
Pumpkin is thinner than asparagus

Each premise consisted of a reference object (RO) and a located object (LO). With the premise “asparagus is thinner than cucumber”, asparagus serves as the LO and cucumber as the RO. The distinction between LO and RO is common in this field of research and was proposed by Landau and Jackendoff (1993), among others.

The subject was then presented with two arrangements of the three presented objects, so-called “models”, e.g.,

Pumpkin-Asparagus-Cucumber
Cucumber-Asparagus-Pumpkin

one on each side of the monitor. Out of those two choices, only one was “correct”, meaning that it was in accordance with the

relation that the two premises established between the three objects. The task only continued if the correct model, Pumpkin-Asparagus-Cucumber, in this case, was recognized. Now each subject was confronted with a new premise, a counterfact, e.g.

Cucumber is thinner than pumpkin

Subjects were advised to treat this counterfact as indisputable. In half the cases, it was in accordance with the initial correct model, in the other half it wasn't. Only if the subject recognized an inconsistent fact as such, the experiment continued. Otherwise, the next problem started. In case the subject detected the inconsistency, the stage of belief revision followed. This stage was only reached if the subject identified the correct model in phase three and the counterfact from the last phase to be inconsistent with it. The subject was now presented with two new models, e.g.

Cucumber-Pumpkin-Asparagus
Asparagus-Cucumber-Pumpkin

one on each side of the monitor. Both of these models were created out of the correct model from phase three by the inclusion of the counterfact. One of these models was always plausible and the other one always implausible. In half of the cases, the plausible model was created by relocating the LO of the counterfact and the implausible model was created by relocating the RO. In the other half, it was reversed.

For plausibility, Bucher et al. (2013) relied on common knowledge. A premise such as “the tree is bigger than the flower” would generally be considered to be plausible, the premise “feather heavier than nail” to be implausible. The mental model “Feather-Nail-Hammer” is plausible with regard to the attribute weight. The mental model “Father-Son-Grandpa” is implausible with regard to the attribute age.

Now the participants had to choose the model that matched their expectation about the inclusion of the counterfact into the initial model. The experimental procedure included the stage of inconsistency detection because, as stated by Bucher et al. (2011), inconsistency detection is a prerequisite for belief revision. For revising one's assumptions, one, first of all, needs to recognize an inconsistency between initial assumptions and newly learned information. That's why the first phases of the experiment were conducted - to ensure that a participant was able to conclude from the two premises and then recognizes an inconsistency with that conclusion. This is when the process of belief revision happens. The approaches presented here aim to understand how preferences for a revised model are composed by comparing different modeling approaches in their accuracy of prediction. Different cognitive models were implemented/adapted and compared, including many simple models e.g. LO-preference, relocation of the object added last to the mental model, preference for the plausible model, etc., just to name a few. Also, the more advanced cognitive models PRISM (Ragni and Knauff, 2013) and the verbal model (Krumnack et al., 2010) were adapted. Following the cognitive models, four models from the machine learning area were implemented - content-based filtering (CBF), user-based collaborative filtering (UBCF), a multilayer perceptron (MLP), and an ensemble model.

In the data set used for modeling, the objects, e.g. cucumber, asparagus, etc., were replaced with A, B, and C. Also the

premises have been reformulated, e.g. “A is to the left of C” was changed to “Left;A;C”. A problem as it was presented in the data set is shown in Table 2.

Table 2: Structure of an experimental test problem.

Sequence	Task	Choices
1 Premises	Left;A;B/Right;C;B	CBA ABC
2 Model	ABC	Left;C;A
3 Decision	ABC/Left;C;A	CAB BCA

It consisted of three sequences, one for each time the participants had to make a decision. The first two sequences were largely ignored since we focused on modeling the sequence of belief revision. Predicting the first two sequences of constructing a model and detecting the inconsistency did not add any value since they only served as preliminary work for the last - the belief revision - sequence. Predicting the logically correct answer was in any case the most common one. All models did achieve the same accuracy for them, namely 0.927 for sequence one and 0.892 for sequence two.

Methods

Before presenting the cognitive and machine learning models, let's quickly compare both approaches. Cognitive models are trying to predict human behaviour by recreating the underlying cognitive process as best as possible. They are useful because they provide an accurate indication of the quality of a cognitive theory. Compared to machine learning models, however, which solely rely on statistical data, they are much less accurate in their prediction. While the big advantage of machine learning models is their great accuracy, for our purpose they have two big disadvantages. Firstly, machine learning models are often black boxes, meaning that it is not visible from the outside which patterns have been learned. This holds true especially for neural networks, but even with recommender approaches such as CBF, great effort is needed to find out exactly what has been learned. This is irrelevant for many applications, but since our use case is not only about high predictive power, but also about an understanding of the underlying cognitive processes, this is a core problem of the machine learning models. In addition, large amounts of data are required for training. Very large data sets with a large number of participants are rarely available in cognitive science since such extensive experiments are difficult to conduct.

To compare the performance of all models, we used CCOBRA¹, a python framework specifically created for behavioural reasoning analysis. This framework proposes few constraints to how the models need to be implemented, all that matters is the prediction. Accuracy is simply determined by dividing the number of correct predictions by the number of all predictions. This is done for each subject individually. For this purpose, CCOBRA provides different benchmark types. With “adaption”, the model gets the know the correct answer after each prediction

¹github.com/CognitiveComputationLab/ccobra

in order to gradually adapt to the current subject. With “coverage”, a model gets to know all answers of the subject before prediction starts. This can be used to find out how well a model is generally able to represent an individual.

Simple Theories

To give an introduction to how a cognitive theory might be established and implemented, the following simple theories were instantiated and compared. These were also the models used for the ensemble model explained in the machine learning chapter.

LO-Preference: There exists a strong cognitive effect called LO-preference described by Bucher et al. (2011), where it was discovered that participants relocate the LO of the counterfactual 87.78% of the time. It followed that the way the counterfactual is formulated strongly influences the model revision process.

Relocation of the object added last to the model: The last inserted object was, in our case, the object introduced newly by the second premise (therefore always the right object C). According to Payne et al. (1993), the way the model is constructed plays a crucial role in the way the model is saved in the reasoner’s mind and therefore has an influence on the way the model is revised when contradictory information is obtained. The hypothesis that the object added last to the model is the most likely one to be relocated is based on the assumption that it is the starting point for inspection of the mental model (Bucher et al., 2011; Knauff et al., 1998).

Preference for the plausible model: As explained in the experiment chapter, one of the revised models from sequence three was plausible while the other one was implausible. This theory states that there exists a preference for the plausible model. While Bucher et al. (2013) did find out that the preference for the plausible model was almost completely overwritten by LO-preference, it might still exist, especially without a strong contradicting effect.

Preference for the mental model presented on the left/right side of the monitor: In sequence three, the subjects were, as explained in the experiment chapter, presented with a choice for a model, one on the right side and one on the left side of the monitor. Perhaps some subjects, out of various reasons, did not construct the revised mental model in their mind before being presented with the choices but instead made a decision only after being shown both possibilities. This might possibly lead to either a left model or a right model preference.

First/Second premise rejection: Subjects had to reject one of the premises to include the counterfactual. The question is if participants had a preference regarding the premise they wish to reject.

Verbal Model

Following Polk and Newell (1995) reasoning does not necessarily depend on domain-specific cognitive processes but on more general cognitive mechanisms. The authors introduced an approach which they called verbal reasoning that makes use of the cognitive mechanisms underlying verbal language (language comprehension) to draw conclusions from premises. The way in which an inference is made therefore depends to a large extent on the decoding of the verbal information. According to Polk and Newell (1995), this decoding of verbal information plays a crucial role in reasoning rather than domain-specific events in the

brain. To instantiate the assumptions from the verbal model into a cognitive model that can be tested and evaluated, Krumnack et al. (2010) introduced a queue in which the objects of the premises are inserted. This queue displays an implicit direction that is, according to Maass and Russo (2003), determined by the cultural left/right difference, e.g. the direction in which scripture is read. In addition, there seems to be a natural tendency for a left-to-right direction when imagining spatial events since a right hemisphere dominance for attention often leads to slightly pronounced processing of objects in the – contralateral – left visual hemi-field (de Schotten et al., 2011). The verbal model implementation from Ragni et al. (2019) dealt with this personal preference by providing a compare-function, that tests the outcome of both possible implicit directions and matches the result with the actual subject’s answer. For the purpose presented here, this implementation was adapted to fit belief revision in the following way: Inclusion of the counterfactual happens by moving forward through the queue until the first object contained in the counterfactual is found. It is then moved to the end of the queue. Following this process, the counterfactual is included in any case. With the following example:

$$A^* \rightarrow B \rightarrow C$$

where the star marks the beginning of the queue and the arrows show the implicit direction, the counterfactual “C is to the left of A” results in the model B-C-A, while with the queue

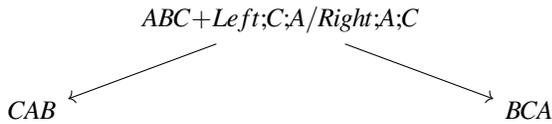
$$A \leftarrow B \leftarrow C^*$$

the same counterfactual results in the model C-A-B. The preferred implicit direction of the current subject was determined before prediction started. This was possible with the CCOBRA benchmark type coverage.

PRISM

PRISM stands for “preferred inferences in reasoning with spatial mental models”. It is a computational model that is based on the assumptions of the theory of preferred mental models, put forward by Ragni and Knauff (2013). The premises “Asparagus is thinner than cucumber” and “Pumpkin is thicker than cucumber” are ubiquitous, meaning that they only induce a single model, namely Asparagus-Cucumber-Pumpkin. When premises lead to multiple possible models like in the case of “Asparagus is thinner than Pumpkin” and “Pumpkin is thicker than cucumber”, for which the two possible models are Asparagus-Cucumber-Pumpkin and Cucumber-Asparagus-Pumpkin, these problems are described as indeterminate. The theory of preferred mental models suggests that reasoners have a preference when deciding for one out of multiple possible models. The cognitive model that emerged from this theory was implemented in 2013 and then re-implemented in 2019 for Ragni et al. (2019). It is the later implementation the model presented here is based upon. The similarity between different models is determined by the number of swap operations needed to create one from the other. This approach was not applicable to our task, since both revised models required two swap-operations to get back to the initial model. The following, adaptive approach was used: in one of the variations, the relation of the two left objects to each other has remained the same, in the other the relation of the two right objects. Perhaps different

participants had preferences with regard to which revised model "feels" more similar to the initial model.



For some people, keeping the relation of the two left objects as it is (B is to the right of A) might feel more similar to the initial model than keeping the relation of the two right objects (C is to the right of B). For other people, the opposite might hold. This theory was implemented through an adaptive approach, in which PRISM gradually learned which of the two similarity-functions fit best to the current subject.

Machine Learning Models

To achieve the best possible prediction, various machine learning models were tested on the data set, namely CBF, UBCF, an MLP, and an ensemble model. Additionally, another simple data-driven model that is used as a baseline is the most-frequent-answer model (MFA) which relies on predicting the most commonly selected response for each task. Generally, MFA can be considered to be an upper bound for models that are not able to adapt to individual participants.

Both CBF and UBCF are so-called recommender systems. While these approaches are usually used to suggest videos, products, images or other content to a user, they are used here to suggest/predict the most appropriate response based on the previous user's behaviour.

CBF (content-based filtering) is about making a decision in a situation in the same way a subject responded to a previously experienced, similar situation. For an online shop, CBF would suggest products similar to previously bought products. For a video portal, CBF would suggest videos similar to previously watched videos. To use CBF for our purpose, similarity between different tasks had to be determined. Since all objects of the premises/models were replaced with A, B, and C, the same tasks were repeated over and over again e.g. all tasks that looked like ABC/Left;C;A with the choices CAB/BCA (in that order) were treated as similar to one another, no matter the original objects behind A, B, and C (boxer, car, tree, etc.). Similarity in content between the different problems was therefore determined solely by the relations, by the side on the monitor on which both choices were located, and by the side on which the plausible model was located. CBF was tested with the benchmark types adaption and coverage.

UBCF (used-based collaborative filtering) is about finding similar users to the current user, and then take the behaviour of those users for prediction. For an online shop, UBCF would suggest those products that are bought by users who generally buy the same products as the current user. For a video portal, UBCF would suggest those videos that are watched by users who generally watch the same videos as the current user. For our purpose, the similarity of all subjects to the current subject was determined to form a subject neighborhood. Then, for prediction, the answers of the similar subjects were weighted more heavily than the answers of the not-so-similar subjects. This resulted in a so-called similar-

ity matrix, in which pairwise similarity was determined between all participants. Simply put, UBCF worked similarly to MFA with the difference that the responses of participants who had previously behaved similarly as the current participant were weighted more heavily. This approach was interesting to find out to what extent different users shared similar reasoning patterns when confronted with the same tasks. UBCF was tested with the benchmark types adaption and coverage. In the first case, the subject-neighborhood gradually formed. In the second case, the subject-neighborhood was fully formed before the prediction started.

An MLP (multilayer perceptron) is a basic feed-forward neural network. Riesterer et al. (2020) compared multiple methods for predictive modeling. This comparison included, amongst various cognitive and statistical modeling approaches, a multilayer perceptron, which achieved the highest accuracy, outperforming MFA and an auto-encoder model. Although this comparison was done in the syllogistic domain, testing this MLP in our domain seemed promising. The MLP featured a topology of 10-256-256-4. The 10-dimensional input-layer one-hot-encoded the task presented in sequence three. The 4- dimensional output-layer encoded the response as a one-hot-encoded vector. Fig. 1 shows the topology of the neural network together with an example task.

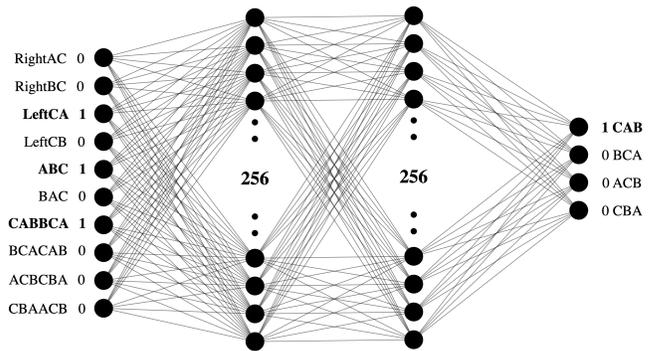


Figure 1: Neural network with the example task ABC/Left;C;A, encoded as a one-hot-encoded vector.

Before prediction, the MLP was trained on all 19 subjects for 30 epochs with a batch size of eight. Once prediction started, the MLP, in order to adapt to the current subject, trained for two epochs after every given answer.

Ensemble modeling is a machine learning technique about combining multiple models by aggregating all predictions into a single prediction. This can be done via multiple methods e.g. averaging/weighting the different predictions. For this purpose, all cognitive theories described in the simple theories section were adaptively combined by taking the model for prediction that so far performed best for the current subject. This approach was chosen to find out exactly which strategies underlie the reasoning process for every individual reasoner. The ensemble model was also tested with the benchmark type coverage, in which case the most fitting strategy was determined before prediction started.

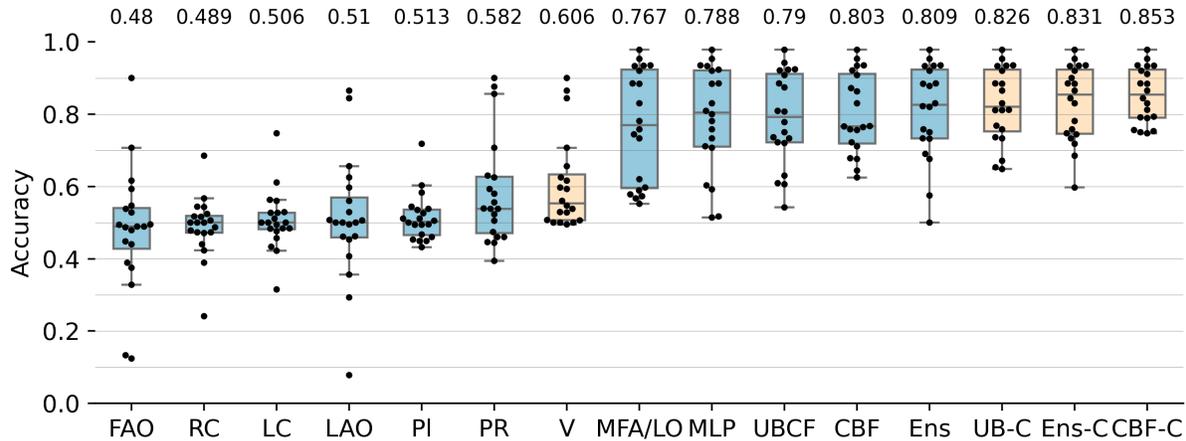


Figure 2: Accuracies of models for belief revision. Only sequence three is considered. Included are MFA (corresponds to LO-relocation), relocation of the object added first/last to the model (FAO, LAO), preference for the left/right choice (LC, RC), preference for the plausible model (PI), PRISM (PR), the verbal model (V, coverage), MFA, CBF, UBCF, the MLP and the ensemble model (Ens). UB-C, Ens-C, and CBF-C are the same models as UBCF, Ens and CBF, but with the CCOBRA benchmark type coverage.

Evaluation

Fig. 2 shows the accuracies of all presented models. The orange boxes show the models with the benchmark type coverage, meaning that those models didn't gradually adapt to the current subject after every answer, but before the prediction started, which of course results in a higher accuracy since the model doesn't need to adapt over time. The machine learning models did surpass the cognitive models. LO-relocation performed just as well as MFA. Both models were, in any case, identical in their prediction, which is why only MFA is shown in Fig. 2. It stood out that while the simple cognitive theories did achieve accuracies not much higher than the random model, they had a high variance. Single few subjects were predicted very well, or very bad, by them, as can be seen by the few outliers. PRISM and the Verbal model lie between the simple theories and MFA. However, out of all cognitive models, they were the only adaptive ones.

All machine learning models were able to outperform MFA, but did also differ greatly from one another in their accuracies. With an accuracy of 0.8 (0.853 with coverage), CBF did perform best out of all featured models. The ensemble model was a close second. With much distance to the ensemble model/CBF, but still significantly better than MFA, UBCF and the MLP did achieve similar accuracies of 0.79 (0.826 with coverage) and 0.788.

The small difference between the performance of the ensemble model and CBF leads to the assumption that it was possible to extract the main strategies CBF was able to leverage. Therefore, the dominant individual strategies for most subjects could be extracted, shown in Tab. 3.

It can be questioned whether the left model preference and the right model preference were actual strategies since they did solely depend on the way the task was presented on the screen. They did not add value to understanding the cognitive processes related to belief revision. Rather, they showed that two participants were probably not able to integrate the counterfactual into the initial

Table 3: Dominant strategy for each subject.

Strategy	Test Subjects
LO-relocation	1,3,4,5,6,7,8,9,11,12,15,16,18,20
Last added object	14,17
First added object	2
Plausible model	13
Left model preference	10
Right model preference	19

model according to their own preferences, perhaps due to an unwillingness towards expending cognitive effort.

Effect of Task Category on Dominant Strategy

While LO-relocation was the dominant strategy for subjects 16 and 20, CBF was able to predict them even better than MFA, which indicates that both subjects used different revision tactics for different tasks. To find out whether task characteristics had an effect on the dominating strategy, the effect of the task categories regarding LO-preference was investigated. As described in the experiment section, each task was assigned to one of three categories - visual, spatial, and neutral. We examined the effect of those categories on the dominating strategy, namely LO-preference. The results are shown in Fig. 3. Although there was wide variation among the subjects in terms of the dominant tactic with respect to the task category, no clear pattern could be discerned that would have applied to all subjects. Therefore, no clear statement could be made about the influence of the categories on the dominant strategy. However, it seemed that subjects 16 and 20 pursued different strategies for different tasks. Whether the category or other task characteristics played a greater role was not investigated.

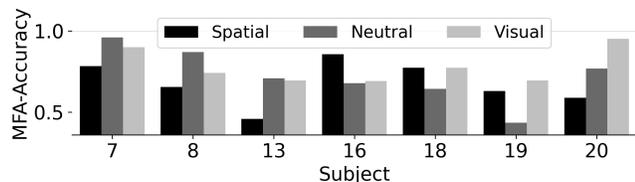


Figure 3: Effect of task category on accuracy of MFA.

Discussion

Various approaches from the fields of cognitive modelling and machine learning were adapted/implemented to test their applicability for belief revision. In addition, the behaviour of all test participants was analyzed individually. Comparing the accuracies of all models on the data set to an experiment conducted by Bucher et al. (2013) provided interesting results. PRISM and the verbal model were overshadowed not only by the statistical recommender approaches, but also by LO-relocation, which delivered the same accuracy as the upper bound MFA, making it the dominant strategy for counterfactual model variation. Adaptive approaches outperforming MFA shows that individual belief revision preferences follow individual reasoning strategies and that those strategies can be gradually learned and leveraged to enhance the quality of prediction.

Subject 12 did relocate the LO in 97.8%, subject 1 in 95.3% of the cases, which made them the most predictable participants. CBF-C reached an accuracy of 0.747 for subject 10, 0.75 for subject 11, and 0.753 for subject 19, which made them the most unpredictable participants. Fig. 4 shows the strengths and limitations of the machine learning models and provides interesting insights into their inner workings.

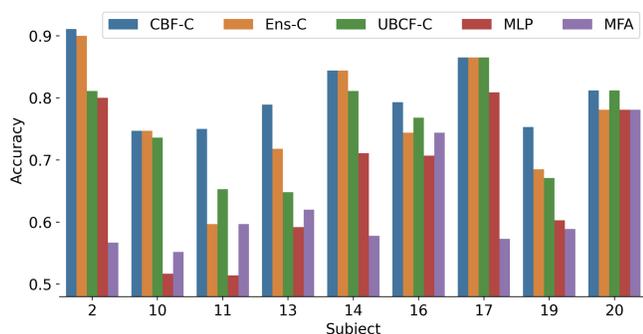


Figure 4: Accuracies of models for belief revision per subject. Only those subjects are presented for which there was a big difference in the accuracy of prediction.

CBF and the MLP performed better than UBCF, which shows that individual subjects were in fact very individual in their behaviour on the task - twenty test-subjects were too few to create a fitting user-neighborhood for each subject. The fact that some subjects were predicted very well by UBCF (2, 14 and 17) while other were not (8, 11, 13), reinforces this assumption.

As can be seen in Fig. 4, subject 11 was better predicted by MFA than by the ensemble model and UBCF. From this one can

conclude that either the dominant strategy for subject 11 changed over the course of the experiment, meaning that the adaptive models learned something that was no longer true after a certain time, or that the main strategy used by subject 11 wasn't included in the ensemble model (and wasn't used by another subject). No clear pattern could be extracted according to which subject 11 acted. Subject 13 was better predicted by the ensemble model and CBF than by the other models. This was because those two models were the only ones that could, due to the way they were implemented, learn the preference for the plausible model.

The overall findings lead to the conclusion that it might be impossible to break down behavioural patterns to a single cognitive theory, at least on this domain, no matter how advanced and all-inclusive it might be. This was nicely shown with the ensemble model. Six different strategies were needed to achieve an accuracy close to CBF. Some test subjects deviated from the usual, dominant strategies and pursued their own tactics, which has led to four out of the six strategies representing only a single subject best.

A limitation lies in the underlying experiment and different approaches leading to the same result. The experiment was originally designed with the goal to investigate the effect of visualizability, of LO-preference and to see if it can be overwritten by plausibility - and not for differentiation between different cognitive models. Hence, left object relocation directly corresponded to first premise rejection and in some cases also to the verbal model and PRISM, while right object relocation directly corresponded to second premise rejection, relocation of the last added object, and, again, in some cases also to the verbal model and PRISM. This made it impossible to differentiate between the tactics the subjects used whose responses were better predicted by models other than LO-preference/MFA. Therefore little could be said about the applicability of PRISM and the verbal model to belief revision. While both models did perform well the way they were adapted to fit counterfactual model variation, they essentially did the same what much simpler and less polished models did and could not leverage their strengths. It would therefore be necessary to put the results of this work to the test in an experiment carried out differently e.g. with more than three different objects, three or more premises to form the initial model, more than two choices for the integration of the counterfact, more than 20 subjects, etc., to see whether the results still hold.

In conclusion, there exists no model complex enough that it's parameters are able to fully represent an individual, not least because of the inconsistencies inherent in human nature. Human spatial relational reasoning is sometimes illogical, contradictory, and subject to strong fluctuations. Nevertheless, the results of this work show that it's possible, at least within the presented domain, to extract many of the underlying behavioural patterns. Cognitive theories and the models derived from them provide interesting insights into what is normally taken for granted and provide a contribution to understanding human cognition, even if the findings, like it is the case with this work, only represent a constricted and clearly defined domain like spatial-relational reasoning.

References

- Bucher, L., Krumnack, A., Nejasmic, J., and Knauff, M. (2011). Cognitive processes underlying spatial belief revision. *Proceedings of the 33rd CogSci*, pages 3477–3482.
- Bucher, L., Nejasmic, J., Bertleff, S., and Knauff, M. (2013). Plausibility and visualizability in relational belief revision. *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, pages 1946–1951.
- de Schotten, M. T., Dell’Acqua, F., Forkel, S. J., Simmons, A., Vergani, F., Murphy, D. G. M., and Catani, M. (2011). A lateralized brain network for visuospatial attention. *Nature Neuroscience*, 14(10):1245–1246.
- Knauff, M. (2013). *Space to reason: A spatial theory of human thought*. MIT Press.
- Knauff, M., Bucher, L., Krumnack, A., and Nejasmic, J. (2013). Spatial belief revision. *Journal of Cognitive Psychology*, 25(2):147–156.
- Knauff, M., Rauh, R., Schlieder, C., and Strube, G. (1998). Continuity effect and figural bias in spatial relational inference. In *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*, pages 573–578, Mahwah, NJ. Mahwah, NJ: Erlbaum.
- Krumnack, A., Bucher, L., Nejasmic, J., and Knauff, M. (2010). Spatial reasoning as verbal reasoning. *Proceedings of the 32th Cognitive Science Society*, pages 1002–1007.
- Landau, B. and Jackendoff, R. (1993). “what” and “where” in spatial language and spatial cognition. *Behavioral and Brain Sciences*, 16(2):217–238.
- Maass, A. and Russo, A. (2003). Directional bias in the mental representation of spatial events nature or culture? *Psychological Science*, 14:296–301.
- Payne, J. W., Bettman, J. R., and Johnson, E. J. (1993). *The Adaptive Decision Maker*. Cambridge University Press, Cambridge.
- Polk, T. A. and Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review*, 102:533–566.
- Ragni, M., Friemann, P., Bakija, E., Habibie, N., Leinhos, Y., Pohnke, D., Satyawan, Y., Schöchlin, M., and Turon, R. (2019). Predicting individual spatial reasoners: A comparison of five cognitive computational theories. In Stewart, T., editor, *Proc. 17th ICCM*, pages 157–162, Canada. University of Waterloo.
- Ragni, M. and Knauff, M. (2013). A theory and a computational model of spatial reasoning with preferred mental models. *Psychological Review*, 120(3):561–588.
- Riesterer, N., Brand, D., and Ragni, M. (2020). Predictive modeling of individual human cognition: Upper bounds and a new perspective on performance. *Topics in Cognitive Science*, 12(3):960–974.

Diverse Experience Leads to Improved Adaptation: An Experiment with a Cognitive Model of Learning

Chase McDonald (chasemcd@cmu.edu)

Erin H. Bugbee (ebugbee@cmu.edu)

Erin McCormick (enmccorm@cmu.edu)

Department of Social and Decision Sciences, Carnegie Mellon University
Pittsburgh, PA 15213 USA

Josh Fiechter (jfiechte@ball.com)

Leslie M. Blaha (leslie.blaha@us.af.mil)

711th Human Performance Wing, Air Force Research Laboratory
Pittsburgh, PA 15213 USA

Christian Lebiere (cl@cmu.edu)

Department of Psychology, Carnegie Mellon University
Pittsburgh, PA 15213 USA

Cleotilde Gonzalez (coty@cmu.edu)

Department of Social and Decision Sciences, Carnegie Mellon University
Pittsburgh, PA 15213 USA

Abstract

In dynamic decision tasks, the situations we confront are never the same: the world is constantly changing. Generally, our ability to generalize learned skills depends on the similarity between the learned skills and the situations in which we will apply those skills. However, in dynamic tasks, the situations we are trained in will most likely be different from the situations in which we need to apply skills. For example, in the face of emergencies, one could be trained to handle hypothetical disaster scenarios, but remain unprepared for the emergency that is actually experienced. This raises an important question: how can we best prepare for the unexpected? Cognitive science research suggests that heterogeneity during training helps people adapt to unexpected situations. However, evidence for a general diversity hypothesis is limited. In this research, we investigate this *Diversity Hypothesis* using a cognitive model of learning and decisions from experience based on Instance-Based Learning (IBL) Theory. We focus on the concept of *decision complexity* to investigate whether confronting decisions of diverse complexities results in improved adaptation to unexpected decision complexities, compared to situations of constant decision complexity. We conduct a simulation experiment using an IBL model in a Gridworld task, and expose agents to various degrees of diversity as they learn; we then observe how these agents transfer their acquired knowledge to a situation of novel decision complexity. Our results support the Diversity Hypothesis and the benefits of diversity on adaptation.

Keywords: transfer of learning; diversity hypothesis; instance-based learning; adaptation; gridworld tasks

Introduction

Most decisions we make in life are dynamic: we evaluate potential alternatives sequentially, determine the values of the options as they develop over time, and select our options in the presence of environmental uncertainties and time constraints (Gonzalez, Lerch, & Lebiere, 2003; Gonzalez, Fakhari, & Busemeyer, 2017). Unfortunately, most research on decision making today involves static situations: decisions are often studied in one-shot choice environments, with no

time constraints or high workload and where most information is provided to the decision maker (Gonzalez et al., 2003; Gonzalez, 2013). Notably, research on *heuristics and biases* has dominated behavioral decision research. For example, while demonstrating the explanatory power of Prospect Theory, one of the best known theories of risk, researchers often use monetary gambles (i.e., “prospects”) that explicitly state outcomes and associated probabilities. People are presented with a description of the alternatives and are asked to make a choice based on the conditions described (Tversky & Kahneman, 1974).

In dynamic situations, decision making is considered as a learning process, in which individuals must rely on their experience to make decisions (Gonzalez et al., 2003). Importantly, by definition, dynamic situations are unique and constantly evolving. Thus, in dynamic situations, a decision maker never confronts the same exact decision situation more than once—“you cannot step twice into the same stream” (Burnet, 1930). An important research question is therefore: how can decision makers prepare for unexpected and novel situations? This question has been addressed in the learning, skill acquisition, and transfer of skills literatures. For example, it is clear that decision makers can successfully transfer learning when the skills learned during training can be reinstated at transfer, or more generally, when transfer situations share some similarity of the procedures and skills learned during training (Healy, Wohldmann, Parker, & Bourne, 2005; Healy, Wohldmann, Sutton, & Bourne Jr, 2006). While these conditions might be possible in less dynamic situations, they might be more difficult to meet in dynamic conditions of choice.

Schmidt and Bjork (1992) argued that what works best for improving performance during training will not necessarily work well in new conditions of transfer; they suggested

that diverse training might be beneficial. This idea has been tested in some studies in which diverse training appears to be particularly important for adaptation to unexpected situations (Brunstein & Gonzalez, 2011; Gonzalez & Madhavan, 2011). For example, Brunstein and Gonzalez (2011) studied effects of diverse training in a luggage screening task. They prepared targets of various categories (e.g., knives, guns, etc.) and tested conditions in which people were trained in only one category of objects (e.g., guns) or in diverse categories (e.g., guns, knives, etc.). They observed that those individuals who were trained with diverse categories were able to classify novel items as potentially dangerous in a transfer condition, while those trained with consistent categories of weapons exhibited poor adaptation. Their conclusions suggest a general *Diversity Hypothesis*: Acquiring diverse experiences during learning will result in better adaptation to unexpected situations.

Here, we test the Diversity Hypothesis and investigate the adaptation to novel levels of decision complexity. Decision complexity is defined as in Nguyen and Gonzalez (2020): the trade-off between low-cost, low-value and high-cost, high-value alternatives. When we make decisions, we often have to handle such cost-benefit trade-offs to determine what actions to take. To test this idea, we rely on a Gridworld task developed by Nguyen and Gonzalez (2020), where agents perform a goal-seeking task under uncertainty by navigating a grid. In this situation, we test how the diversity of experienced levels of decision complexity during learning affects adaptation to unexpected levels of decision complexity. This is carried out using a cognitive model based on Instance-Based Learning Theory (IBLT; Gonzalez et al. (2003)), and we discuss the resulting predictions for human adaptation to novel decision situations.

Instance-Based Learning Theory

IBLT is a theory of decisions from experience, derived from the mechanisms proposed in the ACT-R cognitive architecture (Anderson & Lebiere, 1998), developed to explain human learning in dynamic decision environments (Gonzalez et al., 2003). IBLT provides a decision making algorithm and a set of cognitive mechanisms that can be used to implement computational models of human decision making and learning processes. The algorithm involves the recognition and retrieval of past experiences (i.e., instances) according to their relevancy to a current decision situation, the generation of expected utility of the various decision alternatives, and a choice rule that generalizes from experience. An “instance” in IBLT is a memory unit that results from the potential alternatives evaluated. These are memory representations consisting of three elements: a situation (a set of attributes that give a context to the decision, or state S); a decision (the action taken corresponding to an alternative in state S , or action A); and a utility (expected utility or experienced outcome x of the action taken in a state).

An option $k = (S, A)$ is defined by taking action A in state

S . At time t , assume that there are $n_{k,t}$ different generated instances $(k, x_{i,k,t})$ for $i = 1, \dots, n_{k,t}$, corresponding to selecting k and achieving outcome $x_{i,k,t}$. Each instance i in memory has an *Activation* value, which represents how readily available that information is in memory, and it is determined by similarity to past situations, recency, frequency, and noise (Anderson & Lebiere, 2014).

Here we consider a simplified version of the Activation equation which only captures how recently and frequently instances are activated:

$$Act_{i,k,t} = \ln \left(\sum_{t' \in T_{i,k,t}} (t - t')^{-d} \right) + \sigma \ln \frac{1 - \xi_{i,k,t}}{\xi_{i,k,t}} \quad (1)$$

where d and σ are the decay and noise parameters, respectively, and $T_{i,k,t} \subset \{0, \dots, t - 1\}$ is the set of the previous timestamps in which the instance i was observed. The rightmost term represents the Gaussian noise for capturing individual variation in activation, and $\xi_{i,k,t}$ is a random number drawn from a uniform distribution $U(0, 1)$ at each time step and for each instance and option.

The probability of retrieving an instance i from memory is a function of its activation $Act_{i,k,t}$ relative to the activation of all instances:

$$P_{i,k,t} = \frac{\exp\left(\frac{Act_{i,k,t}}{\tau}\right)}{\sum_{j=1}^{n_{k,t}} \exp\left(\frac{Act_{j,k,t}}{\tau}\right)} \quad (2)$$

where τ is the Boltzmann constant (i.e., the “temperature”) in the Boltzmann distribution. For simplicity, τ is often defined as a function of the same σ used in the activation equation $\tau = \sigma\sqrt{2}$. Importantly, the noise and temperature values add stochasticity to the model, ensuring that action selection is non-deterministic. The nature of the model allows for exploration of the option space to reduce over time, and to treat the “explore-exploit tradeoff” without hard coded exploration (e.g., as in ϵ -greedy reinforcement learning methods (Sutton & Barto, 2018)).

The expected utility of option k is calculated based on a mechanism called *blending* (Lebiere, 1999), using the past experienced outcomes stored in each instance. Here we employ the blending calculation as defined for choice tasks (Gonzalez & Dutt, 2011; Lejarraga, Dutt, & Gonzalez, 2012):

$$V_{k,t} = \sum_{i=1}^{n_{k,t}} P_{i,k,t} x_{i,k,t}. \quad (3)$$

The blending operation (Eq. 3) is the sum of all past experienced outcomes weighted by their probability of retrieval. The choice rule is to select the option that maximizes the blended value.

Experiment: Knowledge Transfer Across Decision Complexities

Gridworld Goal-Seeking Task

We use the goal-seeking Gridworld environments developed by Nguyen and Gonzalez (2020), implemented in the OpenAI

In all three conditions, the agents are required to perform under a new spawn position with decision complexity $\Delta_d = 3$, unexpectedly, after their 60th episode.

We hypothesize that the agents with the most diverse experiences—the agents in the Mixed condition—will perform better during adaptation than the agents in the Low and High complexity conditions. We expect that the agents who have been exposed to more diversity in decision complexity during learning will be able to adapt to a new decision complexity more effectively than those that learned with a consistent level of decision complexity. We also expect that agents in the Low Condition will perform better during learning than agents in the Mixed and High conditions. This is due to the variation in spawn location for the Mixed condition and increased decision complexity for the High condition.

We simulate 100 distinct grid configurations with different goal locations and obstacles. Spawn locations corresponding to the desired levels of decision complexity are generated for each distinct grid.

Our primary dependent measure is *accuracy*, defined as the proportion of episodes where the agent obtains the preferred (i.e., maximum value) goal. Using this metric, we examine agents’ performance in the learning and adaptation phases in aggregate, over time, and at the transition between phases.

Results

Overall Accuracy

The average accuracy across 60 learning episodes and 60 adaptation episodes in each condition is shown in Figure 2. The results are aggregated across all 100 grid configurations, with three independent trials in each. We observe that during learning, agents in the Low decision complexity condition perform significantly better than the agents in the High complexity and Mixed complexity conditions. During the adaptation phase, however, agents in the Low complexity condition experience only a slight improvement compared to the learning phase, while agents in the Mixed complexity condition show the largest improvement from the learning phase. Agents in the Mixed condition agent are able to use the diverse experiences acquired in the learning phase to, on average, adapt more successfully than the agents in the other conditions.

Accuracy Over Time

In addition to overall average accuracy, we plotted the learning curves of the agents, which show the average accuracy per episode, grouped by the experimental condition. The results are presented in Figure 3. We observe that although the agents in the Low complexity condition learn to perform accurately very rapidly compared to the High complexity and Mixed conditions, this is the condition where agents appear to have the most difficulty adapting immediately to the new level of complexity (more discussion on this “surprise” effect in the next section). Perhaps the most interesting observation is that agents in the Mixed condition are the only ones that

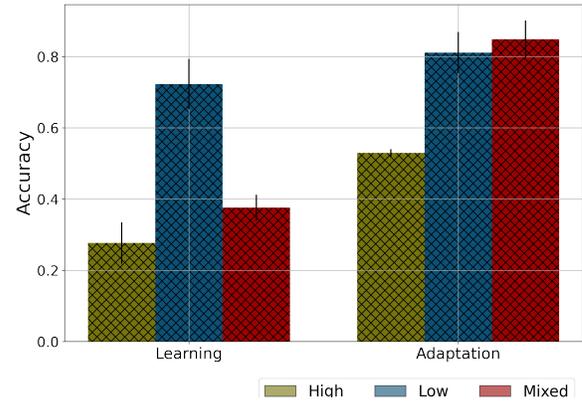


Figure 2: Average accuracy during the learning and adaptation phase for each condition.

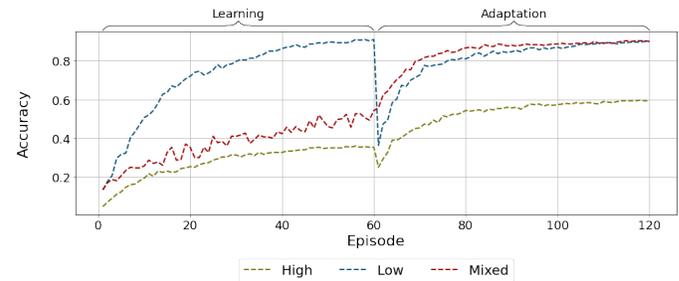


Figure 3: The learning curve for each condition. Agents transfer to the unseen decision complexity $\Delta_d = 3$ at the 60th episode, remaining there until the 120th episode.

continue to improve their performance, without an initial decrease during adaptation. During the 60 adaptation episodes, the Low complexity agents are able to match the performance of the Mixed condition agents. This contrasts with the High complexity agents, which are unable to achieve comparable levels of accuracy.

Surprise Effect

Following on the previous analysis, here we focus on the “surprise” effect per condition, characterized by both the accuracy in the first adaptation episode alone (i.e., Episode 61), as well as by the change in accuracy from the last episode of learning and the first episode of adaptation (i.e., Episodes 60 and 61). Figure 4 presents both of these measures. We observe that the agents in the Mixed complexity condition have the highest average accuracy in Episode 61. Furthermore, the difference in accuracy between Episodes 60 and 61 for the Mixed complexity agents is near zero. That is, their surprise effect is low.

In contrast, the Low complexity condition has a lower performance than the Mixed condition, but has the highest surprise effect, where the accuracy decreased more than in any of the other conditions in Episode 61. The High complexity

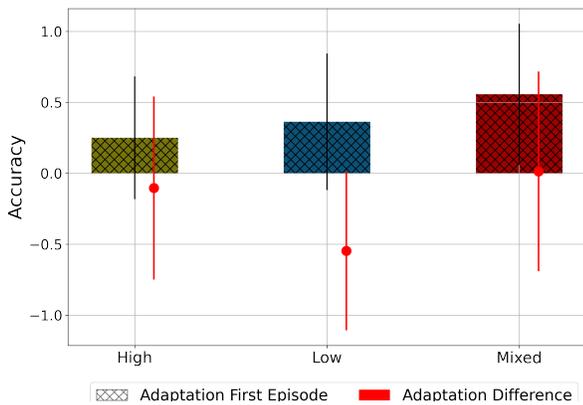


Figure 4: The accuracy in the transfer episode and the difference in accuracy between the first adaptation (Episode 61) episode and final episode of training (Episode 60).

condition has the lowest accuracy in the first episode of adaptation and a small surprise effect as it transitions to a lower decision complexity.

Explanations for the Benefits of Diversity

In this section, we dive into the mechanisms that may lead to the benefits of diversity for adaptation. A primary explanation is that the likelihood that agents will experience states during learning that are similar—or equivalent—to the states that they will experience in the adaptation phase changes across conditions.

Due to both the nature of the task and the definition of decision complexity, an agent in the High complexity condition is more likely to end up on a shorter path and fail to gain sufficient exposure to the environment to facilitate transfer. To demonstrate this, we simulate a random agent in the same training phase for the Low and High conditions and measure the average number of steps per episode. The High condition with a random agent has, on average, significantly (two-sided T -test, $p < 0.01$) shorter episodes (20.57 ± 0.74 steps) than the Low complexity condition (24.37 ± 0.59 steps). This shows that an agent is more likely to reach a goal earlier (e.g., the nearest distractor) in the High complexity condition, and thus be less exposed to the environment. This lack of exposure during the learning phase makes it more difficult for an agent to apply the instances stored in memory to new situations successfully. The memory instances will be biased towards the previously learned behavior.

As discussed, the Low complexity condition dictates a spawn location that has an increased relative distance to the nearest distractor target. The longer expected episode length in the Low condition—the same value presented above—allows agents an increased opportunity to gather diverse experiences in the environment.

Finally, the Mixed complexity condition results in a higher likelihood of experiencing states that are similar to the com-

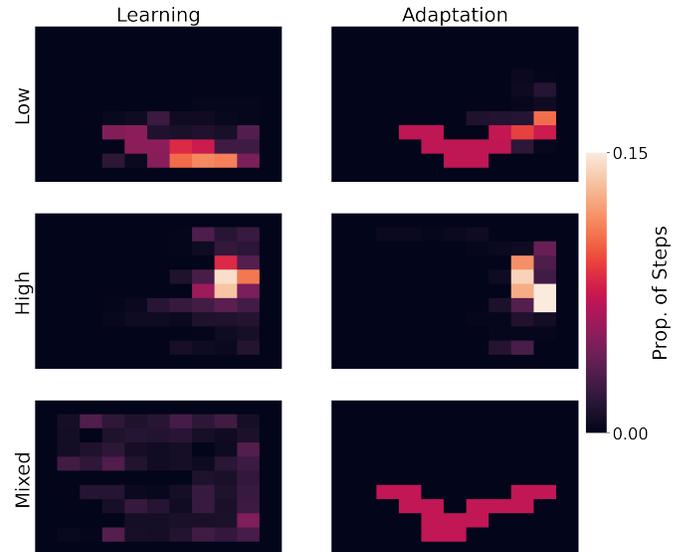


Figure 5: The proportion of times an agent in each condition visited a particular grid cell throughout the learning and adaptation phases. Each condition pictured here represents the same grid configuration.

plexity experienced during the adaptation phase. In contrast to the above cases, this is because diversity is built into the learning phase. The agent in the Mixed condition has, by definition, a diverse set of experiences. The relative levels of diversity correspond to the relative performance in adaptation.

An illustrative example of the differences in the diversity of experiences during the learning and adaptation phases for each condition is shown in Figure 5. Considering the learning phase, the Mixed complexity condition depicts the highest level of diversity in state visitations, followed by the Low complexity condition; whereas the High complexity condition has a small and focused set of highly visited states.

The behavior in the adaptation phase demonstrates how the behavior during learning translates to the unexpected situation during adaptation phase: in both the Low and Mixed conditions, the agent is able to discover a roughly equivalent policy, whereas the High condition agent fails to learn the location of the preferred goal and a policy that will allow it to reach that position.

Discussion

In past research, the notion of diversity has been applied to motor tasks (Wulf, 1991), visual discrimination tasks (Wolfe, Friedman-Hill, Stewart, & O’Connell, 1992) and classification decisions (Brunstein & Gonzalez, 2011; Gonzalez & Madhavan, 2011). Here, we expand this line of research to demonstrate the diversity of training in the context of decision complexity. We find that agents who learn in consistent decision complexity environments have poorer adaptation to novel and unexpected situations than those that learn with diverse decision complexity.

An interesting observation is that agents that learned in the Low complexity condition performed closely to agents in the Mixed condition during adaptation, while agents that learned in the High complexity condition are very far from reaching the level of performance of the Mixed complexity agents. An explanation we offer in our analyses is that the experiences of the agents in the Low complexity condition are quite diverse during the learning phase. By definition, a Low complexity decision would encourage the agents to navigate the Gridworld to find the target of higher value, because the decision trade-off is easy to resolve (Nguyen & Gonzalez, 2020). In other words, it is a “no brainer” to ignore the temptation of a distractor, because a larger value target is also close to the spawn location. These diverse experiences are thus applicable to a novel level of complexity at transfer, as shown in Figure 5.

In our immediate future work, we plan test both the robustness of the results to changes in environmental parameters, as well as the predictions of these simulations in human experiments. Are humans with diverse experiences in the Gridworld able to adapt more successfully to novel situations? Given that IBL models have been shown to emulate human behavior very closely in many tasks including the Gridworld (Nguyen & Gonzalez, 2021), we expect that the predictions of this paper will hold in human experiments. How far can we stretch the Diversity Hypothesis? That is, how different can the transfer conditions be to take advantage of the diversity of training? Answers to these questions can help us craft diverse training conditions and predict the way these conditions can result in robust decisions under changing and dynamic situations.

Acknowledgments

This research was supported by the US Department of Defense Award Number: POJN1006188. Distribution A: Cleared for Public Release AFRL-2021-1520

References

- Anderson, J. R., & Lebiere, C. (1998). The atomic components of thought.
- Anderson, J. R., & Lebiere, C. J. (2014). *The atomic components of thought*. Psychology Press.
- Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., & Zaremba, W. (2016). Openai gym. *arXiv preprint arXiv:1606.01540*.
- Brunstein, A., & Gonzalez, C. (2011). Preparing for novelty with diverse training. *Applied Cognitive Psychology*, 25(5), 682–691.
- Burnet, J. (1930). Early greek philosophy, 1892. *Bywater's no. LVIII*, 137.
- Gonzalez, C. (2013). The boundaries of instance-based learning theory for explaining decisions from experience. In *Progress in brain research* (Vol. 202, pp. 73–98). Elsevier.
- Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating decisions from experience in sampling and repeated choice paradigms. *Psychological Review*, 118(4), 523–51.
- Gonzalez, C., Fakhari, P., & Busemeyer, J. (2017). Dynamic decision making: Learning processes and new research directions. *Human factors*, 59(5), 713–721.
- Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, 27(4), 591–635.
- Gonzalez, C., & Madhavan, P. (2011). Diversity during training enhances detection of novel stimuli. *Journal of Cognitive Psychology*, 23(3), 342–350.
- Healy, A. F., Wohldmann, E. L., Parker, J. T., & Bourne, L. E. (2005). Skill training, retention, and transfer: The effects of a concurrent secondary task. *Memory & Cognition*, 33(8), 1457–1471.
- Healy, A. F., Wohldmann, E. L., Sutton, E. M., & Bourne Jr, L. E. (2006). Specificity effects in training and transfer of speeded responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(3), 534.
- Lebiere, C. (1999). Blending: An act-r mechanism for aggregate retrievals. In *Proceedings of the sixth annual act-r workshop*.
- Lejarraga, T., Dutt, V., & Gonzalez, C. (2012). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making*, 25(2), 143–153.
- Minsky, M. (1961). Steps toward artificial intelligence. *Proceedings of the IRE*, 49(1), 8–30.
- Nguyen, T. N., & Gonzalez, C. (2020). Effects of decision complexity in goal-seeking gridworlds: A comparison of instance-based learning and reinforcement learning agents. In *Proceedings of the 18th intl. conf. on cognitive modelling*.
- Nguyen, T. N., & Gonzalez, C. (2021). Theory of mind from observation in cognitive models and humans. *Topics in Cognitive Sciences TopiCS*.
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological science*, 3(4), 207–218.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, 185(4157), 1124–1131.
- Wolfe, J. M., Friedman-Hill, S. R., Stewart, M. I., & O’Connell, K. M. (1992). The role of categorization in visual search for orientation. *Journal of Experimental Psychology: Human Perception and Performance*, 18(1), 34.
- Wulf, G. (1991). The effect of type of practice on motor learning in children. *Applied Cognitive Psychology*, 5(2), 123–134.

Curiosity as Pattern Matching: Simulating the Effects of Intrinsic Rewards on the Levels of Processing

Kazuma Nagashima (nagashima.kazuma.16@shizuoka.ac.jp),
Junya Morita (j-morita@inf.shizuoka.ac.jp),
Yugo Takeuchi (takeuchi@inf.shizuoka.ac.jp)

Department of Informatics, Graduate School of Integrated Science and Technology, Shizuoka University,
3-5-1 Johoku, Naka-ku, Hamamatsu-shi, Shizuoka-ken, 432-8011 Japan

Abstract

Many studies have been conducted concerning curiosity, a type of intrinsic motivation in humans and artificial agents. However, the specifics of the correspondence between curiosity in humans and artificial agents have not yet been fully explained. This study examines this correspondence on the Adaptive Control of Thought–Rational (ACT-R) cognitive architecture by exploring situations in which curiosity effectively promotes learning. We prepared three models of path planning, representing different levels of thinking, and had them learn in multiple-breadth maze environments while manipulating the curiosity strength. The results showed that curiosity in learning an environment negatively affected the model with a shallow level of thinking. However, it positively affected the model with a deliberative level of thinking. We consider that the results show some commonalities with human learning.

Keywords: cognitive modeling; intrinsic motivation; curiosity; ACT-R

Introduction

Although curiosity is assumed to be an effective source of motivation that encourages humans to engage in long-term learning, it does not always work effectively. To explore the conditions in which intrinsic motivation works well, we examine the influence of the levels of the thinking process that many cognitive scientists have discussed (e.g., Brooks, 1986; Kahneman, 2011). This study does not go into the details and differences of such theories but assumes broad distinctions between the shallow automatic level in which a person does not think carefully (fast process) and the deep deliberative level in which a person takes time to think carefully (slow process).

This study aims to clarify the role of intrinsic motivation in such levels of thinking. To accomplish this, we prepared models that instantiate the information processing of each thinking level. The prepared models were constructed based on a cognitive architecture, which is a structure enabling cognitive functions in various tasks by various individuals (Anderson, 2007). By assuming a common structure, differences in the thinking levels are represented as combinations of primitive processes provided by the architecture.

Of the several cognitive architectures developed to date, this study uses Adaptive Control of Thought–Rational (ACT-R; Anderson, 2007) because this architecture has the most publications showing the details of the models for various tasks (see Kotseruba & Tsotsos (2018) for a quantitative review). By referring to these models, we can implement several thinking levels with the validation made by the previous

studies. Furthermore, ACT-R has two types of knowledge (declarative and procedural), which seem useful to represent different levels of thinking.

As a representation of curiosity in ACT-R, this study uses a mechanism proposed in our previous study (Nagashima, Morita, & Takeuchi, 2020). Although there are other options for motivation theory in ACT-R (e.g., Juvina, Larue, & Hough, 2018), our previous proposal has the advantage of implementing curiosity as rewards accompanied by pattern matching. We consider this characteristic effective to examine the complex relations between curiosity and levels of thinking. However, our previous study failed to demonstrate that the mechanism relates to human learning. Therefore, the current study newly implements models of different processing levels and tries to find common features with human learning by examining the relation between those levels and the mechanism of intrinsic motivation.

In order to clearly present the goal of this study, the following section shows previous studies concerning intrinsic motivation and ACT-R. Following this, we briefly introduce a curiosity mechanism proposed by Nagashima et al. (2020). We then discuss this mechanism’s implementation and run simulations of a specific task. Finally, we summarize the implications of the study and indicate future directions.

Related Works

As noted above, curiosity is regarded as a type of intrinsic motivation. Therefore, this section introduces studies concerning intrinsic motivation to explore situations in which curiosity works effectively. Following this, a brief introduction of ACT-R is presented, focusing on the relationship with levels of thinking.

Intrinsic Motivation in Humans and Artificial Intelligence

To date, a large body of studies has been created concerning learning as facilitated by intrinsic human motivation. For example, Malone (1981) categorized intrinsic motivation into three types: “challenge,” which comes from goals of appropriate difficulty; “fantasy,” which leads to the imagination of unrealistic experiences; and “curiosity,” which is stimulated by something surprising, interesting, or fun. These types are not independent but interrelated. Therefore, reviewing the categories other than curiosity can also help to place the study in a broader context.

Malone’s classification of motivation as challenge has been related to the discussion of the optimal level of intrinsic motivation (Csikszentmihalyi, 1990; Yerkes & Dodson, 1908). In humans, there are appropriate levels of task difficulty at which intrinsic motivation is stimulated. Based on this idea, Baranes, Oudeyer, and Gottlieb (2014) found through experiments that intrinsic motivation that is neither too high nor too low for a task is effective. Furthermore, the appropriate level of difficulty for an individual depends on the individual’s preferred level of thinking. Based on this discussion, we assumed the dependency of the appropriateness of the challenge on an individual’s level of thinking. In other words, intrinsic motivation can be enhanced by providing tasks that are suitable for the level of thinking that the individual prefers.

We consider that the above discussion of challenge cannot be separated from a discussion of curiosity. Rather, we treat curiosity as a mechanism of intrinsic motivation evoked by the appropriate difficulty of a task. Various studies of artificial agents have addressed the mechanisms of curiosity. The key principle of modeling curiosity can be obtained from the theory of prediction error (Friston, 2010). The emotions of surprise, interest, and enjoyment that trigger curiosity are caused by discrepancies between perceptions of the external world and predictions derived from experience. Based on this theory, autonomous agents have been constructed to learn an environment based on curiosity (Aubret, Matignon, & Has-sas, 2019; Schmidhuber, 2010; Singh, Barto, & Chentanez, 2005). In contrast to conventional reinforcement learning, in which one receives a reward directly from the external environment (Sutton & Barto, 1998), the rewards generated from intrinsic motivation fluctuate depending on the state of the internal environment.

In recent years, this topic has progressed remarkably with a framework for deep reinforcement learning through an end-to-end approach (Burda et al., 2018; Mnih et al., 2015; Pathak, Agrawal, Efros, & Darrell, 2017). In particular, Burda et al. (2018) have shown that agents with curiosity can learn a wide range of environments and improve their game scores without explicit extrinsic rewards.

Levels of Thinking in ACT-R

The studies presented in the previous section implemented curiosity-based agents using a reinforcement-learning framework. However, with the framework alone, it is difficult to explore situations in which intrinsic motivation functions effectively. Thus, a framework that seamlessly connects the learning algorithms and the process of inference in a task is needed.

As noted previously, we use ACT-R as such a framework to connect multiple levels of thinking and curiosity-based learning. ACT-R has modules corresponding to brain regions. For example, the declarative module (prefrontal cortex) retains experience and knowledge, and the goal module (anterior cingulate cortex) manages states in tasks. The production rules stored in the production module (basal ganglia) are selected based on the status of such modules, and they send commands

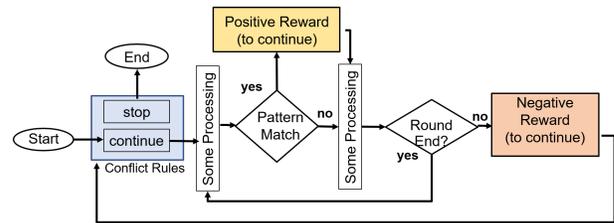


Figure 1: Flowchart of the task continuation framework presented in (Nagashima et al., 2020). A positive reward is generated by “pattern matching” accompanied by memory retrieval.

to the modules as actions (e.g., they search for knowledge that meets the conditions and update the current state of the task). These rules include variables that realize flexible correspondence (pattern matching) with module states. According to the ACT-R theory, pattern matching is realized in the cerebral cortex, specifically the prefrontal cortex. If knowledge retrieval from the declarative module becomes unnecessary in the task, the role of the basal ganglia becomes dominant in the process of proceduralizing the task. Therefore, we consider pattern matching accompanied by memory retrieval the criterion that distinguishes a deliberative level of thinking from a shallow level of thinking in ACT-R. In other words, the shallow level of thinking involves little pattern matching, while the deliberative level of thinking involves extensive pattern matching.

Mechanism of Curiosity in ACT-R

This section presents the mechanism of intrinsic motivation proposed by Nagashima et al. (2020). According to their basic idea, curiosity, especially that involving a higher cognitive function, is connected with pattern matching. Several authors have stated that enjoyment, a source of curiosity (Friston, 2010), is related to discovering novel patterns in the environment (Caillois, 1958; Csikszentmihalyi, 1990; Huizinga, 1939; Koster, 2004; Schmidhuber, 2010). Following such discussion, Nagashima et al. (2020) focused on a mechanism of pattern matching by computers as a concept that corresponds to pattern discovery by humans.

Moreover, based on the correspondence between human curiosity and pattern matching, Nagashima et al. (2020) proposed a framework for task continuation in a general environment (Figure 1). This framework assumes a task that consists of several rounds. At the start of each round, the model decides whether to continue or stop the task by selecting production rules corresponding to each option. After it decides to continue the task, the model proceeds with the round. When the model encounters a condition that ends the round, a new round begins.

The selection of production rules is controlled by utility learning in ACT-R (Wai-Tat & Anderson, 2006). In the above process, the initial utility value of the *continue rule* is considered higher than that of the *stop rule*. The process of becom-

ing bored from this initial state can be modeled by assigning a trigger of a negative reward to the rule that recognizes the end of each round.

To prevent boredom and consider the conditions that result in positive rewards and continued learning, curiosity is required. In this mechanism, rules that trigger positive rewards are defined as rules that fire as a result of the successful retrieval of declarative knowledge in the task. The search for declarative knowledge requires pattern matching between the conditional clauses of the rule (the current situation) and the memory in declarative knowledge. However, this rule gradually becomes used for repeated executions; that is, “production compilation” in ACT-R integrates the two rules and generates a compressed hierarchical rule. After integration occurs, it becomes routine and cannot be related to a reward. Then, the utility value of the *continue rule* decreases, and the *stop rule* fires.

Therefore, the framework represents the decrease in curiosity that comes from the discrepancy between the model’s predictions (routine compiled knowledge) and the results of the action. In short, long-term task continuation is achieved by keeping the model engaged in pattern matching between the conditional clauses of production rules and declarative knowledge. Thus, the mechanism is consistent with the key principle of curiosity (Friston, 2010), while it utilizes the distinction between declarative and procedural knowledge in ACT-R.

Simulation

To examine the conditions in which curiosity functions effectively, we conducted a simulation study using the mechanism presented above. In this section, we first clarify the purpose of the simulation. Following this, the actual manipulations of the simulation are defined, and the results are presented.

Aims and Indicators

The purpose of the simulation was to address the following two successive questions:

1. What kinds of factors stimulate curiosity?
2. How does stimulated curiosity affect task learning?

To address the first question, this simulation manipulated the learning factors from both the internal and external viewpoints. The external factor can be considered the breadth of the learning environment (difficulty of the task), while the internal factor corresponds to the cognitive strategies (levels of thinking) implemented in which the model can be used. The influence of these internal/external factors on curiosity is measured as (a) the number of continuations of a task (number of firings of the *continue rule* in Figure 1).

The second question is explored because task continuation does not always contribute to task learning. To assess the effects of intrinsic motivation on task learning, we examined (b) the goal achievement rate, (c) the behavior pattern of the environment search, and (d) the number of newly generated

rules. The index (b) is the outcome of task learning, and the index (d) indicates the internal changes in the model caused by task continuation. Regarding the connection between outcomes and internal changes, the present study computed the behavioral index (c) as the information entropy of the environment search:

$$Hr = \frac{-\sum_{i \in n} p(x_i) \log p(x_i)}{\log n} \quad (1)$$

where x_i and n indicate each location and the number of locations in the map, respectively. This index increases if the model explores the environment extensively but decreases if the model insists on the same behavioral pattern during the task.

Simulation Conditions

We used maze searching as a task and manipulated the external factor by changing the size of the map; the internal factor was searching strategies as the model’s level of thinking. Figure 2 shows the overview of the manipulations.

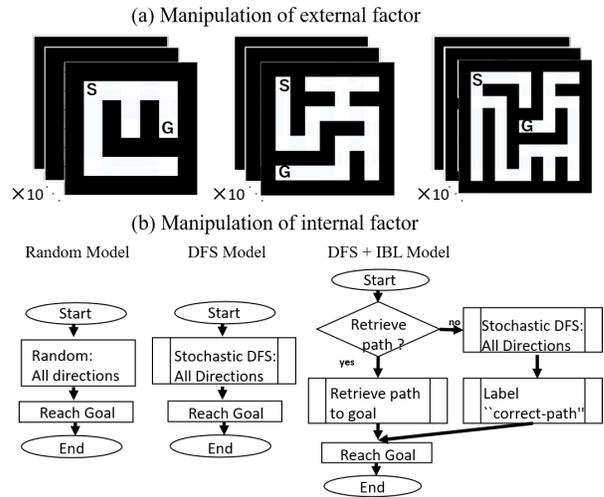


Figure 2: Manipulations of external and internal factors.

Manipulation of the External Factor Figure 2a depicts several maze maps, which were automatically created in a grid world using a maze generation algorithm that randomly changed wall positions with the constraint of avoiding a loop structure. As a parameter for the algorithm, we varied the map sizes between 5×5 , 7×7 and 9×9 . Ten maps were prepared for each size. These maps were given a start and a goal by choosing two positions having the largest number of intersections to traverse. The map was included in the model as location information that uses declarative knowledge to construct a topological map (Reitter & Lebiere, 2010). The model’s current location is maintained as a slot content of the goal buffer. In each round, the model was first located at the start position and moved to the next location by retrieving the path from the declarative module. When the current location matched the goal, a new round began. The model repeated

this for each map until it became bored (fired the *stop rule* in Figure 1).

Manipulation of the Internal Factor To manipulate the internal factor, we prepared three models corresponding to the different levels of thinking in the above environment. Figure 2b illustrates the abstract flow of the models, presenting the shallowest model (the left model), the most deliberative model (the right model), and the intermediate model (the middle model). Because these models were developed based on Nagashima et al. (2020), please refer to the original literature for the details of the base model. Brief descriptions of each model are as follows.

1. *Random model (random)*

At each point in the movement, the model stochastically fires a rule representing the next direction (east, west, south, or north). Based on the direction and current position, the model searches for a path from declarative knowledge. If the model succeeds in finding a path, it moves to the location according to the direction searched (changing the state of the goal buffer). If the model fails to find a path, it repeats the same procedure. As the rounds proceed, the model compiles such retrieved declarative knowledge into procedural rules.

2. *Stochastic DFS model (DFS)*

This model uses a stochastic depth-first search (DFS), as presented in Reitter and Lebiere (2010). This strategy determines a path by backtracking with the stack structure implemented by ACT-R’s declarative and imaginal modules. As with the random model, this model first stochastically determines the direction of movement. After successfully retrieving a path linking the current location to the directed location, the model creates a new chunk linking the two locations as “already searched” and stores it in the declarative module. The model repeats this process until it reaches a goal or fails to retrieve a path. When the model fails to retrieve a path (reaches a dead end), it reverts to the previous location using a memorized chunk (already searched). Like the random model, the stochastic DFS model learns new rules by compiling declarative knowledge on paths, but it can repeat more rounds because it has internal resources that allow it to reach a goal effectively.

3. *Stochastic DFS and IBL model (DFS+IBL)*

This model is the same as that presented in Nagashima et al. (2020). The model performs a combination of the probabilistic DFS (Reitter & Lebiere, 2010) and instance-based learning (IBL: Gonzalez, Lerch, & Lebiere, 2003).¹ At the beginning of the task, the model uses the stochastic DFS to explore the maze. Each time the model reaches a goal, it labels all the chunks used in the current round as the “correct path.” In the next rounds, if the model can retrieve the

knowledge, it uses it. If it cannot retrieve it, the model uses a probabilistic DFS to reach the goal from the current position. Among the three models, this model has the most deliberative and costly strategy. It always tries to memorize chunks and retrieve correct paths from its memory. As the round proceeds, however, the model accumulates the “correct path” and eventually compiles it into procedural knowledge, which leads to the most effective goal achievement behavior.

Parameters The simulation used the default ACT-R 7.14. The initial utility value of the *continue rule* was set to 10, and the initial utility value of the *stop rule* was set to 5.² We also assigned negative reward triggers ($r = 0$), which were lower than the initial utility value of the *stop rule*, to rules that recognized the end of the round (reaching a goal or recognizing that the time limit of each round had passed) and assigned positive reward triggers to rules that included pattern matching as curiosity. In this study, we selected the *path finding rule* accompanied by the pattern matching to the declarative memory as a positive reward trigger and varied the value from 1 to 20 as the simulation conditions (strength of curiosity). For each condition of the positive reward value, the model ran the task 1000 times at a maximum of 80 rounds each time. In addition, we set the time limit for each round to 100 s in ACT-R simulation time. When the time limit was reached, the model was forced to move to the next round.

Results

Figure 3 displays the results of the simulation as a function of the reward values for path finding. Each point in the graphs indicates the average scores of the four indices ($n = 10000$). The effects of the external factor (the map sizes) are shown in the difference of the three lines in each graph, and the influence of the internal factor can be seen by comparing the three graphs vertically aligned in the figure. The horizontal alignment of the graphs corresponds to the four indices presented at the beginning of this section, and the rightmost figures are correlation matrices of the indices and the two dependent variables (rewards and map size). In the following section, we examine the details of the results according to the two questions presented as the aims of this simulation.

Factors Stimulating Curiosity The left three graphs in Figure 3-a indicate the number of continued rounds. The strong effects of the internal factor on this index are clearly seen. In the upper two models (DFS+IBL and DFS), greater intrinsic rewards increased the number of task continuations (DFS+IBL: $r = .94$; DFS: $r = .97$). In contrast, the random model indicated a weaker correlation between the rewards and the number of rounds ($r = .67$), exhibiting an inverted U shape. Greater intrinsic rewards promoted task continuation until approximately 14 and then decreased task continuation. This inverted U shape suggests the existence of an

¹Although the original IBL used a blending mechanism, the current model does not use the mechanism; it only utilizes the learning in declarative memory.

²Following Anderson et al. (2004), noise parameters were set as follows: ans (activation noise level) = 0.4 and egs (production noise level) = 0.5.

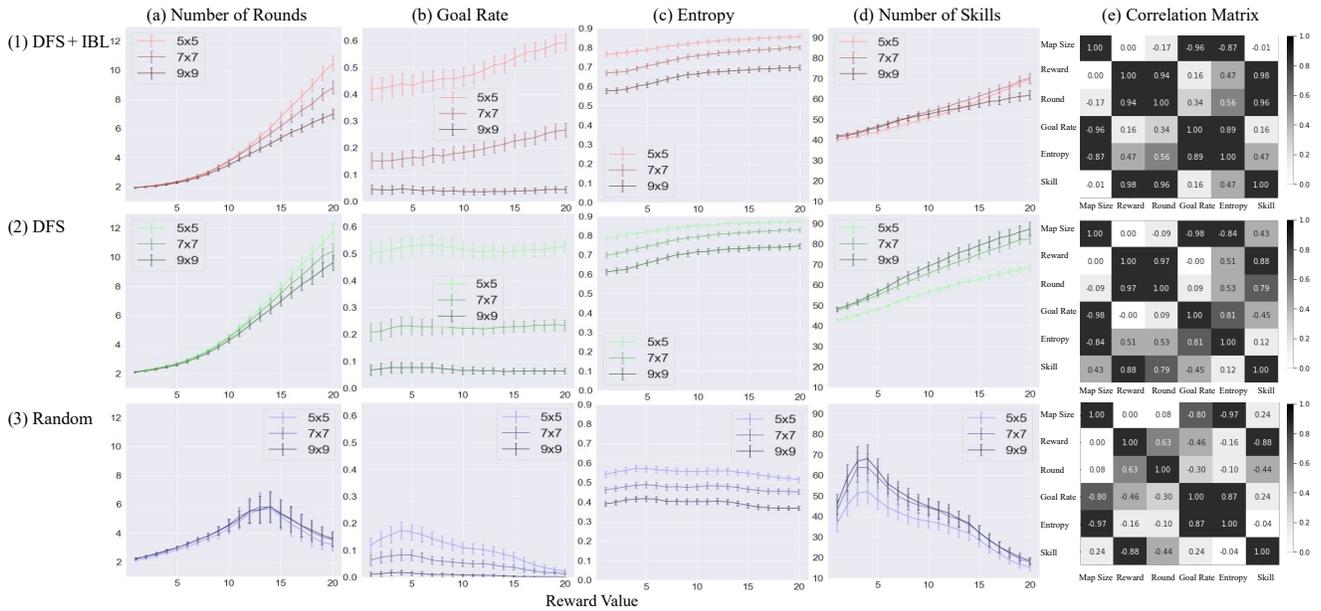


Figure 3: Results of simulation. The panels distinguish models vertically and indicators horizontally. The error bars in each graph represent the standard deviation scaled by 10^3 . The vertical axes of the graphs show the titles of the graphs. a: number of rounds continued, b: goal rates, c: entropy, and d: number of new rules generated by the production compilation. The rightmost graphs show e: correlation matrices between each variable.

optimal level, which is frequently pointed out in theories of intrinsic motivation. For example, according to Yerkes and Dodson (1908)’s classical theory, as a learner’s arousal level increases, performance increases to a certain point and decreases beyond that point.

The effects of manipulating the external factor also varied between the three models (DFS+IBL: $r = -.17$, DFS: $r = -.09$, random: $r = .08$). In the upper two models, we observed the negative effects of the size of the external factor, especially in high-reward conditions. Greater intrinsic rewards were more effective in smaller maps, where the model could more easily reach a goal. In other words, it suggested that applying DFS to a wide range of environments was not effective in obtaining internal rewards. The DFS strategy needed to backtrack to find a path, but the time cost of backtracking increased with the size of the map, which can be interpreted as decreasing the chances of finding a path within a limited time period.

Effects of Task Continuation on Learning The remaining of the indices in Figure 3 indicate how the stimulated intrinsic motivation affected task learning. As can be seen in the upper two models, larger rewards increased the entropy (IBL+DFS: $r = .47$, DFS: $r = .51$) and the number of learned rules (IBL+DFS: $r = .98$, DFS: $r = .88$), indicating that the deliberative strategy made use of intrinsic rewards to expand the search of the environments. Furthermore, the learning outcome results reveal the effects of IBL. The model with IBL (the upper model) showed the positive effects of intrinsic

rewards ($r = .16$), especially in the smaller maps (5x5: $r = .97$, 7x7: $r = .97$, 9x9: $r = -.17$). However, the intrinsic rewards in the model without IBL had no effect on goal achievement ($r = -.00$).

Note that the IBL itself did not always work effectively in terms of goal rates. In the smaller-reward conditions, the IBL model had lower goal rates than the DFS-only model. However, when greater intrinsic rewards were given, the performance of the model with IBL exceeded that of the model without IBL. As described previously, IBL is a costly and slow strategy that always tries to retrieve a correct path. Therefore, it takes time to make use of such experiences to improve performance. However, it can be assumed that it is difficult to learn to reach a goal without labels of the correct paths. As the flat pattern of the goal rates in the DFS model without IBL (Figure 3-2-b) indicates, the lack of explicit correct labels led to disoriented wandering behaviors in the environment.

Unlike the other two models, the random model with a shallow strategy had a different overall trend. Like the results of the number of rounds, the number of skills (correlation with the reward: $r = -.88$) and goal rates (correlation with the reward: $r = -.46$) exhibited inverted U-shaped trends. Furthermore, in Figure 3-3-a, the peaks of the inverted U shapes in these two indices are smaller than that in the number of rounds, reflecting negative correlations of the two indices with the intrinsic rewards. More critically,

³The standard deviation rather than the standard error, which varies with n , is used to indicate the variability in the data.

in Figure 3-3-c, intrinsic reward has a negative effect on entropy ($r = -.16$). These results suggest that higher intrinsic rewards triggered by path finding strengthen irrational low-level behaviors (repetitive visits of the same locations without expanding the search) rather than leading to the creation of additional rules to achieve the goal. The compiled rule in the random model can be found in the appendix.

Conclusion

The purpose of this study was to examine the conditions in which intrinsic motivation affects learning. To achieve this purpose, we modified the previous model for intrinsic motivation in ACT-R (Nagashima et al., 2020) to represent different levels of thinking. Unlike the conventional methods for reinforcement learning (Aubret et al., 2019; Schmidhuber, 2010; Singh et al., 2005), the ACT-R architecture makes it possible to represent a detailed strategy for different thinking levels with realistic time constraints. We consider those features of ACT-R (different knowledge representations and assumptions of simulating reaction time) useful for representing the distinctions of levels of thinking and examining complex interactions with curiosity.

In the simulation, we manipulated the external and internal factors. As a result, the deliberative models showed the positive effects of intrinsic motivation on task continuation, learning skills, and searching behaviors. Regarding the outcomes of learning, however, only the slowest and most costly model benefited from intrinsic motivation. The model that did not evaluate the correctness of retrieval exhibited disoriented wandering through the environment. Moreover, the model that did not memorize the environment was negatively influenced by intrinsic motivation.

Summarizing these findings, we were able to characterize the effects of curiosity on behaviors in different levels of thinking. There are claims that intrinsic motivation works well with deliberative thinking, which requires “autonomy,” “mastery,” and “purpose,” and that extrinsic motivation works well with shallow thinking, which is usually used in routine work (Pink, 2011). Our model’s behaviors follow this idea, thereby corresponding to the human learning process.

In the future, we will analyze the causal relationship between each variable in detail to disentangle the complexities of the results presented in Figure 3. In addition, we will arrange the task setting to include the process of obtaining initial motivation. In the present study, we assumed that humans start with high motivation for a task. However, in reality, a person’s motivation for a task is likely to vary depending on the difficulty and contents of the task (Malone, 1981), as presented in the DFS and IBL model in the 9×9 condition. We considered that those results indicating the relative ineffectiveness of curiosity in difficult tasks were caused by the time limit. Therefore, we also need to examine the effect of time limits on the relation between the level of thinking and the strength of curiosity. By conducting studies addressing such limitations, we can explore more detailed conditions of task

continuation, especially those before a model reaches optimal levels.

Appendix

Listing 1 presents rules relating the movement of locations in the random model (Check-Path and Check-Goal) and a rule that was generated through the production compilation of those rules (Check-Path-And-Check-Goal).

Listing 1: Productions rules in the random model. Strings in brackets indicates variables.

```

Check-Path
If
  The current task status is 'confirming'
  The current location is <location1>
  The retrieved path has <location2>
Then
  Change the current task status to 'check-goal'
  Change the current location to <location2>

Check-Goal
If
  The current task state is 'check-goal'
  The current location is <location>
  The goal is not <location>
Then
  Change the current task status to 'check-goal'

Check-Path-And-Check-Goal
If
  The current task status is 'confirming'
  The current location is the <location1>
  The retrieved path has the <location2>
  The goal is not <location2>
Then
  Change the current location to <location2>
  Change the current task status to 'check-goal'
    
```

Check-Path moves the current location in the goal buffer to the location described in the retrieved path. Check-Goal confirms that the moved location is not the goal in order to continue searching for the goal location. The compiled rule integrates those rules, having the condition that checks the retrieved destination is not the goal location and the action that leads to non-goal locations. This production is further integrated with rules retrieving the path with specific destinations and becomes a rule conflicting with the rule leading to the goal location. It can be considered that the inverted U shape presented in the random model of Figure 3 occurs as a result of generating such goal-avoiding rules.

References

Anderson, J. R. (2007). *How can the human mind occur in the physical universe*. Oxford Press.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036.

Aubret, A., Matignon, L., & Hassas, S. (2019). A survey on intrinsic motivation in reinforcement learning. *Corr*.

Baranes, A. F., Oudeyer, P.-Y., & Gottlieb, J. (2014). The effects of task difficulty, novelty and the size of the search

- space on intrinsically motivated exploration. *Frontiers in neuroscience*, 8, 317.
- Brooks, R. (1986). A robust layered control system for a mobile robot. *IEEE Journal on Robotics and Automation*, 2(1), 14–23.
- Burda, Y., Edwards, H., Pathak, D., Storkey, A., Darrell, T., & Efros, A. A. (2018). Large-scale study of curiosity-driven learning. *CoRR*.
- Caillois, R. (1958). *Les jeux et les hommes*.
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*.
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.
- Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, 27(4), 591–635.
- Huizinga, J. (1939). *Homo ludens versuch einer bestimmung des spielementest der kultur*.
- Juvina, I., Larue, O., & Hough, A. (2018). Modeling valuation and core affect in a cognitive architecture: The impact of valence and arousal on memory and decision-making. *Cognitive Systems Research*, 48, 4 - 24.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Koster, R. (2004). *Theory of fun for game design*. O'Reilly Media, Inc.
- Kotseruba, I., & Tsotsos, J. K. (2018, Jul 28). 40 years of cognitive architectures: core cognitive abilities and practical applications. *Artificial Intelligence Review*.
- Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science*, 5(4), 333–369.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... others (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533.
- Nagashima, K., Morita, J., & Takeuchi, Y. (2020). Modeling intrinsic motivation in act-r: Focusing on the relation between pattern matching and intellectual curiosity. In *ICCM 2020: 18th International Conference on Cognitive Modeling*.
- Pathak, D., Agrawal, P., Efros, A. A., & Darrell, T. (2017). Curiosity-driven exploration by self-supervised prediction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 16–17).
- Pink, D. H. (2011). *Drive: The surprising truth about what motivates us*. Penguin.
- Reitter, D., & Lebiere, C. (2010). A cognitive model of spatial path-planning. *Computational and Mathematical Organization Theory*, 16(3), 220–245.
- Schmidhuber, J. (2010). Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development*, 2(3), 230–247.
- Singh, S., Barto, A. G., & Chentanez, N. (2005). Intrinsically motivated reinforcement learning. In L. K. Saul, Y. Weiss, & L. Bottou (Eds.), *Advances in Neural Information Processing Systems 17* (pp. 1281–1288). MIT Press.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT Press.
- Wai-Tat, F., & Anderson, J. R. (2006, 6). From recurrent choice to skill learning: A reinforcement-learning model. *Journal of Experimental Psychology. General*, 135, 184–206.
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18(5), 459–482.

A Predictive Processing Implementation of the Common Model of Cognition

Alexander Ororbia (ago@cs.rit.edu)

Department of Computer Science, Rochester Institute of Technology, Rochester, NY 14623

M. Alex Kelly (AlexKelly@cunet.carleton.ca)

Department of Computer Science, Bucknell University, Lewisburg, PA, USA

Department of Cognitive Science, Carleton University, Ottawa, ON, Canada

Keywords: cognitive architecture; predictive processing; holographic memory; continual learning

While modern machine learning techniques based on deep artificial neural networks (ANNs) have an impressive ability to process data to uncover patterns, they do not typically model high-level cognition or more than a single task. If an ANN is trained on a series of tasks, catastrophic interference occurs, with each new task causing the ANN to forget all previously learned tasks (McCloskey & Cohen, 1989). Conversely, symbolic cognitive architectures can capture the complexities of high-level cognition but scale poorly to the naturalistic, non-symbolic data of sensory perception (e.g., images) or to big datasets necessary for modelling learning over a lifetime (e.g., corpora with hundreds of millions of words). Is it possible to provide a theory that bridges ANNs and symbolic models, a reduction of the symbolic to the neural, while retaining the strengths and capabilities of each?

We propose a cognitive architecture that is built on two biologically plausible, neural models: neural generative coding (NGC; Ororbia, Mali, Giles, & Kifer, 2020) and holographic memory (Kelly, Arora, West, & Reitter, 2020). By combining the two, we create a model of cognition that has the power of modern machine learning techniques while retaining long-term memory, single-trial learning, transfer-learning, and other cognitive capacities associated with high-level cognition. Our intent is to advance towards a cognitive architecture capable of capturing human performance at all scales of learning, from the half-hour lab experiment to skills acquired gradually over a lifetime.

Since Newell (1973) first argued that good empirical work and piecemeal theoretical work are insufficient to understand the mind, researchers in cognitive science have sought to develop functional, testable theories of cognition as a whole. Cognitive architectures serve as both unified theories of cognition and as computational frameworks for implementing models of specific experimental tasks. Forty years of research has developed hundreds of cognitive architectures with strong commonalities to each other (Kotseruba & Tsotsos, 2018) suggesting an emerging consensus on the basic principles of cognition, on the basis of which Laird et al. (2017) propose a *Common Model of Cognition*, a high-level theory of the modules of the mind and how they interact (see Fig. 1).

The Common Model of Cognition consists of perceptual

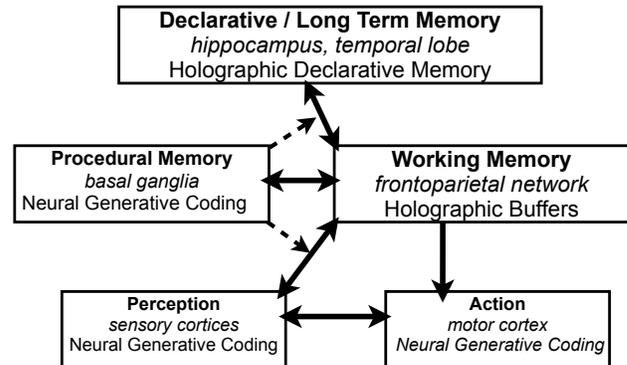


Figure 1: Common Model of Cognition (Laird et al., 2017), associated brain areas (Stocco et al., 2021), and our approach to modelling each module. Solid arrows are data passing. Dashed arrows indicate modulation of a data passing path.

and motor modules that interact with the agent’s environment, working memory to hold the active data in the agent’s mind, a declarative memory that holds the agent’s world knowledge, and a procedural memory that controls information and evaluates possible actions. An evaluation of fMRI data from 200 participants across tasks found correlations in patterns of activity across brain areas consistent with the Common Model of Cognition’s modules (Stocco et al., 2021).

Proposed Architecture

Neural Generative Coding (NGC) is a scalable instantiation of predictive processing brain theory (Clark, 2015) yielding an efficient, robust form of predict-then-correct learning.

Neural Generative Sensory Cortices use NGC for processing a specific modality of data. In Ororbia et al. (2020), we show that NGC learns a good density estimator of data (from which new samples can be sampled or “fantasized”), in conjunction with desired target functionality (e.g., classification, regression), in not only the cases of static input but also in cases of time-varying data series.

Neural Generative Motor Cortex In Ororbia and Mali (2021), we generalize NGC to the case of action-driven tasks, i.e., active NGC (ANGC), common in reinforcement learning (RL), providing evidence that NGC can be used to build a

coupled generative model and controller system that solves RL problems, particularly those when the reward signal is sparse or non-existent. ANGC will serve as the motor cortex.

Neural Generative Basal Ganglia In Ororbia, Mali, Kifer, and Giles (2019), we model the functionality of the basal ganglia in suppressing/inhibiting neural activity for the purpose of action selection and task switching (Cameron, Watanabe, Pari, & Munoz, 2010), a behavior we argue is critical in facilitating effective continual learning without catastrophic interference. This task selection model, which learns through competitive Hebbian learning, will serve as the basis for part of the basal ganglia in our architecture, acting to coordinate the exchange of information between the working memory and the sensory and long-term memory cortices.

Holographic memory (Plate, 1995) is a formalism for capturing the capacity for humans to learn and recall arbitrarily complex associations between stimuli in the environment. Holographic memory is immune to the catastrophic interference typical of more conventional ANNs (Mannering & Jones, 2021), allowing it to be used to construct models that handle multiple, unrelated tasks (Cheung, Terekhov, Chen, Agrawal, & Olshausen, 2019).

Working Memory Each buffer in working memory is a holographic vector. Holographic memory vectors have an established ability to account for memory phenomena such as serial and free recall of lists (Franklin & Mewhort, 2015).

Declarative Memory is the composition of many individual holographic vectors (each representing a distinct concept). Our model accounts for human performance in recall, probability judgement, decision-making (Kelly, Arora, et al., 2020), and learning the meaning and part-of-speech of words (Kelly, Ghafurian, West, & Reitter, 2020).

Conclusions and Future Research

Humans are capable of continual learning, deep expertise, single-trial learning and agile adaptation to dynamic environments, and transfer learning across multiple tasks. Conventional ANNs struggle to replicate these abilities. Solving the problem of lifelong learning will aid us both in understanding the human mind and in the development of intelligent agents that are better able to generalize to real world environments. Our proposed implementation of the Common Model of Cognition is composed of neuro-cognitively plausible components, i.e., holographic memory, predictive processing circuits, and competitive learning. A promising research direction is the application of our architecture to where the challenge of catastrophic interference is most prevalent: reinforcement learning across lengthy, diverse streams of tasks where knowledge retrieval, transfer, and composition are absolutely critical.

References

Cameron, I. G., Watanabe, M., Pari, G., & Munoz,

- D. P. (2010). Executive impairment in parkinson's disease: response automaticity and task switching. *Neuropsychologia*, 48(7), 1948–1957.
- Cheung, B., Terekhov, A., Chen, Y., Agrawal, P., & Olshausen, B. (2019). Superposition of many models into one. In *Adv Neural Inf Process Syst* 32 (pp. 10868–10877).
- Clark, A. (2015). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.
- Franklin, D. R. J., & Mewhort, D. J. K. (2015). Memory as a hologram: An analysis of learning and recall. *Canadian Journal of Experimental Psychology*, 69, 115–135.
- Kelly, M. A., Arora, N., West, R. L., & Reitter, D. (2020). Holographic declarative memory: Distributional semantics as the architecture of memory. *Cognitive Science*, 44(11), e12904. doi: 10.1111/cogs.12904
- Kelly, M. A., Ghafurian, M., West, R. L., & Reitter, D. (2020). Indirect associations in learning semantic and syntactic lexical relationships. *Journal of Memory and Language*, 115, 104153. doi: 10.1016/j.jml.2020.104153
- Kotseruba, I., & Tsotsos, J. K. (2018). 40 years of cognitive architectures: Core cognitive abilities and practical applications. *Artificial Intelligence Review*.
- Laird, J. E., Lebiere, C., & Rosenbloom, P. S. (2017). A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics. *AI Magazine*, 38(4), 13–26. doi: 10.1609/aimag.v38i4.2744
- Mannering, W. M., & Jones, M. N. (2021). Catastrophic interference in predictive neural network models of distributional semantics. *Comput Brain Behav*, 4, 18–33.
- McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. *The psychology of learning and motivation*, 24(109), 92.
- Newell, A. (1973). You can't play 20 questions with nature and win. In W. G. Chase (Ed.), *Visual information processing* (p. 283–308). New York: Academic Press.
- Ororbia, A. G., & Mali, A. (2021). Adapting to dynamic environments with active neural generative coding. doi: 10.13140/RG.2.2.14273.48485/1
- Ororbia, A. G., Mali, A., Giles, C. L., & Kifer, D. (2020). Continual learning of recurrent neural networks by locally aligning distributed representations. *IEEE Transactions on Neural Networks and Learning Systems*.
- Ororbia, A. G., Mali, A., Kifer, D., & Giles, C. L. (2019). Lifelong neural predictive coding: Learning cumulatively online without forgetting. *arXiv:1905.10696*.
- Plate, T. A. (1995). Holographic reduced representations. *IEEE Transactions on Neural networks*, 6(3), 623–641.
- Stocco, A., Sibert, C., Steine-Hanson, Z., Koh, N., Laird, J. E., Lebiere, C. J., & Rosenbloom, P. (2021). Analysis of the human connectome data supports the notion of a “common model of cognition” for human and human-like intelligence across domains. *NeuroImage*, 235, 118035.

Attitudinal Polarization on Social Networks: A Cognitive Architecture Perspective

Mark Orr (mo6xj@virginia.edu)

Biocomplexity Institute, University of Virginia,
P.O. Box 400298, Charlottesville, VA 22904 USA

Andrea Stocco (stocco@uw.edu)

Department of Psychology, University of Washington
Campus Box 3515525, Seattle, WA 98195 USA

Christian Lebiere (cl@cmu.edu)

Don Morrison (dfm2@cmu.edu)

Michael Martin (mkm@cmu.edu)

Department of Psychology, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213

Abstract

Polarization of attitudes is an important, and often troubling or disruptive, effect of interest in many fields. We seek to shed some light on how polarization arises by applying cognitive architectures to the problem. We created a novel embedding of individual cognitive agents, using ACT-R's declarative memory model, into social networks, simulated them communicating over time, and observed the evolution of the agents' attitudes, both collectively and individually. The primary measures we use are both Shannon entropies, of the distribution of attitudes in the final configuration of the whole social network, and of the distributions of memory traces in the individual agents at the end of the simulation. Simulations were run over ten different network topologies, using three different distributions of narrative valences, and five different values of the agents' memory decay parameter. These simulations demonstrated that polarization can be understood from a social and cognitive perspective simultaneously, each providing insights into the system's behavior.

Keywords: Cognitive Modeling; Long-Term Memory; Resting-state fMRI; Functional Network, Attitudes, Polarization, Social Networks.

Introduction

Attitudinal polarization has a long history in political science, sociology, and social psychology. It is no less relevant today than it was 50 years ago. A seemingly obvious scientific question is to ask to what extent we can understand attitudinal polarization from the perspective of cognitive architectures. This question has interest beyond the purely scientific. Understanding the structure of attitudes (as relations among beliefs) and the dynamics of attitude change can yield actionable insight for applications in, for example, public health (Orr, Thrush & Plaut, 2013). Yet, our understanding of attitudinal polarization, from a cognitive perspective, relies nearly exclusively on work in social psychology, a discipline with little intersection with cognitive architectures.

Polarization in attitudes, typically, is a valenced affair, in which an object of contention is evaluated with respect to its goodness/badness, desirability/repulsiveness, approachability/avoidability. It is typically described as a distribution across individuals (humans, bots, agents, or

even models). It is also naturally described as a networked phenomenon, where clusters of individuals can develop solidarity or polarization or other variants on a theme. Important questions include: What aspects of cognitive functioning are implicated by polarization? Do polarized minds lead to polarized social spaces (or vice versa)? Are there interesting threshold effects or other non-linear relations between mental and social scales?

If the answers to these kinds of questions seem obvious to you, then consider this. The famous Shelling segregation model (Shelling, 1971) provided somewhat shocking insight into the relation between mental states and social structure - low degrees of individual-level preferences for segregation generated strong system-level segregation. We use this example to illustrate that the relation between levels of scale is not obvious. It must be investigated rigorously. Using the perspective of cognitive architectures, as a computational, mechanistic lens, should yield a set of novel insights into polarization and other social phenomena of interest to those working on public health, security, human rights and environmental issues.

The goal of this paper is to describe an approach for studying attitudinal polarization using cognitive architectures and to show its potential value. We do this in a stylized way, with an abstract social space and the co-opting, in a highly formal way, of a specification of attitudes from social psychology. The central question we pose, but do not yet answer, is this: Can we describe the conditions, initial or otherwise, of the mental and social systems that guarantee stability in both (or either) the mental or social systems. Stability is well-understood in real-valued or binary networked systems (e.g., Bhat and Redner, 2019). But these networks are a poor abstraction for human cognitive complexity and their organization in social structures. What about socio-cognitive systems?

Toward this end, we provide a study of the relation among the distribution of attitudes and beliefs in a population and the social network structure of that population with respect to two outcomes, one in terms of external behavior and the

other in terms of internal mental representation. The former is derived from the distribution of beliefs in the population and the latter is derived from the distribution of beliefs within individuals. Thus, we capture both the mental and the social in equal, symmetrical measures. In the results section, we will tie our work to future directions in the contexts of cognitive architectures, social psychology and sociology.

Simulating Social Networks

The spread of information across social networks is difficult to study experimentally. For this reason, researchers frequently make use of either large-scale, quasi-experimental data, such as analysis of large corpora of Twitter messages, or multi-agent simulations. Such techniques also are routinely used by social media companies.

In social network simulations, agents are modeled as nodes in a network whose edges are communication channels. Agents exchange information across these channels. The spreading of information is then studied as a function of factors such as network geometry (e.g., small world networks), agent goals (e.g., reaching consensus), and communication intent.

Simulating Plausible Cognitive-Social Agents

To reduce the complexity of the simulations, most computational social science efforts use relatively simple agents, often with little or no cognitive ability. This is sufficient to capture some network-level dynamics, such as those that lead to consensus within a group (Romero & Lebiere, 2014) or the production of original ideas in science. When the goal is to understand the interplay between social interactions (at the network level) and psychological constructs (at the agent level), it is warranted to imbue the agents with cognitively plausible assumptions about their thought processes. For example, Lindstrom et al. (2019) augmented social agents with reinforcement learning capabilities to successfully capture the addictive qualities of social media behavior.

Because we are interested in the interaction between network dynamics and internal beliefs, we endowed our agents with a realistic model of declarative memory. Specifically, we used Anderson’s model of memory, reflecting the rational analysis of our environment reflected in our memory mechanisms (Anderson & Schooler, 1991). Those regularities, such as the power law of practice and forgetting, have also been observed in recently developed information environments such as social networks (e.g., Hubermann et al, 1998; Stanley & Byrne, 2016). In this model, the availability of memory m is related to its base-level activation function $B(m)$. Every time m is retrieved or re-encoded, a new trace is created. The final activation of m is the sum of the decaying activations of all its traces, with stochastic noise added to make the retrieval process stochastic:

$$B(m) = \log \sum_i t_i^{-d}$$

where t_i is the time elapsed since the creation of the i -th trace and d is the characteristic decay rate of an agent memory.

Connecting Memory and Social Behavior

In addition to receiving and internalizing information, an agent in a social network also sends messages across the network. The choice of which messages to spread is, ultimately, a problem of decision-making (Hackel et al., 2020.). To connect an agent’s decisions to its memory, we used Gonzalez et al.’s (2003) instance-based learning framework (IBL). In IBL, agents select their next action by generating expectations reflecting previous experiences in memory that match the current context. This framework is particularly appealing because it meshes well with the ACT-R declarative model and has a long history of successful applications in decision-making (e.g., Erev et al, 2010). Furthermore, while rooted in declarative memory models, IBL gives predictions that are largely consistent with reinforcement learning (Chelian et al., 2015), another framework that has been successfully applied to social networks.

In IBL previous memories are aggregated through a mechanism called blending, which combines different outcomes in a weighted average, based on the probability P_i of retrieval of each memory reflecting its activation and similarities between the contents of the memories:

$$V = \operatorname{argmin} \sum P_i \cdot (1 - \operatorname{Sim}(V_i, v_i))^2$$

Our model represented narratives and their associated valence (defined in a [-1, 1] interval) as chunks in memory. At each iteration, the model representing a node in the network would store in memory all the narratives received from its neighbors. It would then compute the node’s attitude by performing a blended retrieval over all narrative chunks in memory, extracting a consensus valence. The model would then generate a narrative to spread to its neighbors by matching the node’s attitude against the valence of the narrative chunks in memory. The resulting output reflects a combination of attitude of the node and popularity of narratives in its ego network.

Information Entropy as a Common Measure of Cognitive and Behavioral Dynamics

Because our goal is to measure changes in social behavior and in agent cognition at the same time, it is useful to have a common metric that applies to both.

To do so, we used Shannon’s information entropy H (Shannon, 1948), which can be defined over the set of beliefs S (pairs of narrative and valence):

$$H = -\sum_{i \in S} P(i) \log P(i)$$

where $P(i)$ is the probability of encountering the i -th belief, Although the definition of H is the same, the interpretation is different within a social context (between agents) and a cognitive context (within a single agent's memory).

Social Entropy. Social entropy is a measure of uncertainty or consensus of the narratives that were propagated by all agents in the network during the final time step of a simulation. The probability $P(i)$ of the i -th belief is defined as the proportion of times it is propagated over the network in a given interval time. Thus, social entropy reflects the order or disorder of each simulation's final state. Roughly speaking, 0 bits of social entropy indicates consensus, i.e. all agents propagating the same narrative with the same valence, whereas fragmentation (a diversity of opinions) is indicated by 2 or more bits of social entropy. One bit of social entropy indicates polarization.

Cognitive Entropy. Within a single agent, entropy is defined by the activation of beliefs in memory. Because the combination of narrative/valence pairs are encoded as chunks in ACT-R, entropy is calculated from the probability that a given chunk in declarative memory (DM) will be retrieved and spread over the network. In turn, the retrieval probability of a chunk i is related to a memory's base-level activation by the function:

$$P(i) = e^{B(i)} / \sum_{j \in DM} e^{B(j)}$$

Thus defined, Shannon's entropy captures the degree of the internal organization of memories in a given agent and captures the agent's need to allocate cognitive resources to the different narratives. In this sense, information entropy has been previously used, for example, to derive predictions about the size of the hippocampus in humans (Smith et al., in press).

Materials and Methods

We conducted a 5 (Memory Decay Rate) x 10 (Network Topology) x 3 (Narrative Valence) simulation-based experiment with 10 replications per cell. The Memory Decay Rate manipulation varied the architectural decay rate parameter to address the general question of whether memory matters for simulations of information diffusion. The Network Topology and Narrative Valence manipulations addressed general questions about the effects of social context and the types of messages exchanged.

During each of the 1500 simulations, a connected social network of 200 cognitive agents exchanged a set of 10 narratives over a period of 100 ticks. During each tick, an agent encoded the narratives conveyed by all alters in its ego-network, decided on a narrative to convey, and then conveyed that narrative to all neighbors in its ego-network at the next tick. Agent behavior thus arose from a combination of neighbors' opinions and ego's evolving attitude in a closed-loop system defined by a simulation's initial state.

Initialization of social structure and cognitive agents proceeded as follows.

Network Topology: Agents were embedded in one of 10 network structures, all of which were based on a classic "caveman" graph (e.g., Watts, 1999). In our caveman networks, agents are divided into 10 "caves" of 20 agents each. All agents within each cave are fully connected with each other, except for two agents, each of which communicates with one other cave (see Figure 1).

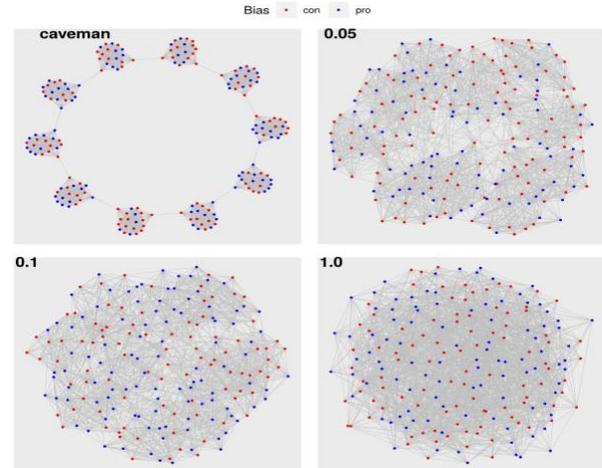


Figure 1: Different network topologies used in this study.

We manipulated the dense clustering of social interactions within caves by randomly replacing in-cave connections with new between-cave connections with probability p_{rewire} . Ten levels of p_{rewire} were used to transition from regularity to randomness: 0, 0.025, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75, and 1.0.

Agent Initialization. As described above, an agent's attitude is a consensus valence produced by blended retrievals of belief chunks from declarative memory. Each belief is a combination of a narrative with an associated *valence*, a number representing the extremity of the belief conveyed by the narrative. For instance, a neutral narrative would have a value close to zero, while extreme narratives would have values close to +1.0 or -1.0 (e.g., Eagly & Chaiken, 1993).

To establish an initial attitude for each agent, we seeded declarative memory with 100 belief chunks. The procedure to generate and allocate these agent belief histories involved three steps. First, we generated a population of notional attitudes by drawing 200 values from a truncated normal distribution (mean = 0, standard deviation = 0.25, minimum = -1, maximum = +1).

Each notional attitude was then combined with the valences associated with 10 narratives to determine the probability that a belief chunk would appear in an agent's history. Specifically, each agent's history was generated by drawing 100 samples from a discrete, truncated normal distribution (mean = notional attitude, standard deviation = 0.8, minimum = -1, maximum = +1).

For our experiment, we created three types of narrative-valence associations: polarized, centrist, and linear.

Polarized narrative valences represent strong, extremist message content (5 narratives with valences of -1 and 5 with valences of +1). Centrist narrative valences represent the use of moderate language, where valences for narratives were drawn from the same truncated normal distribution used to generate the notional attitudes. Linear narrative valences represent well-defined narratives that convey attitudes which span the valence spectrum (-1, -0.8, -0.6, -0.4, -0.2, 0.2, 0.4, 0.6, 0.8, 1).

Finally, we randomly allocated notional attitudes and their histories to agents in the network. The red and blue node colors in Figure 1 illustrate the distribution of negative and positive notional attitudes, respectively.

Results

Social networks have an inherent duality. They can be described by focusing on (a) global properties at a network Level Of Analysis (LOA), or (b) local properties at a node LOA. Our initial analyses, reported below, reflect this duality in separate ANOVAs: one concerned with entropy at a network (i.e., social) LOA, the other concerned with entropy at a node (i.e., cognitive) LOA.

A Memory Decay Rate x Rewiring Probability x Narrative Valence ANOVA of social entropy yielded a 3-way interaction, $F(72, 1350) = 1.37, p < .05$. Figure 2 shows the means and 95% confidence intervals for each experimental condition in the Decay x Rewiring x Valence interaction.

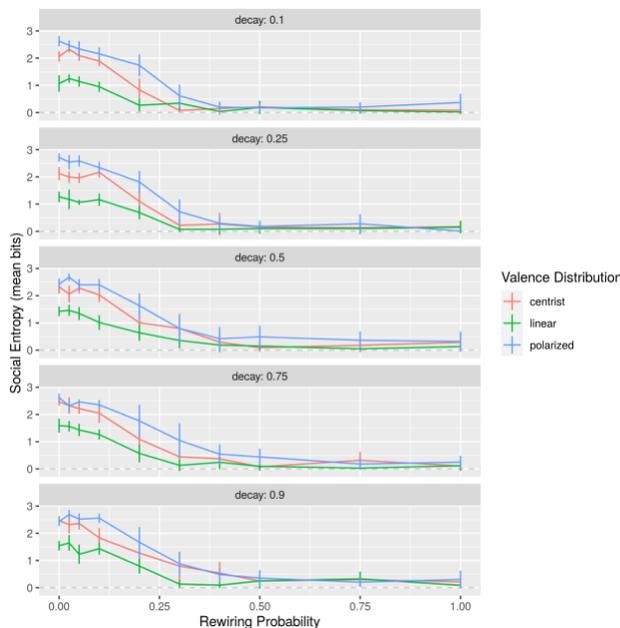


Figure 2: Effects of network topology on social entropy.

The pattern of interaction indicates that networks tend toward consensus (i.e., social entropy near 0) at rewiring probabilities of 0.4 and greater, independently of memory decay rate and narrative valence. Thus, in networks that are poor imitations of real-world social networks (i.e., those lacking local coherence), memory and message content have minimal effects on social entropy. This tendency toward consensus replicates agent-based simulation studies demonstrating assimilation with simplified agent models.

At rewiring probabilities that produce networks more similar to real-world social networks (i.e., those exhibiting dense clusters of peers), the networks tend toward polarization (i.e., entropy near 1¹) or fragmentation (i.e., entropy near 2 or more bits), depending on memory and narrative valence. In the context of messages with a linear valence distribution, polarization is more likely to occur at reasonable values of memory decay (i.e., near the default ACT-R value of 0.25 for the transient activation noise parameter). As decay rate increases toward unrealistic values, social entropy increases and the networks tend toward fragmentation (narrative diversity) as patterns of narratives fluctuate across the network without the damping effect of memory to stabilize them. Messages with polarized and centrist valence distributions tend to produce fragmentation regardless of memory decay rate.

Generally, these results indicate that polarization is a relatively infrequent phenomenon that arises when narratives of a particular type are exchanged in realistic social networks by agents who act in a manner that is congruent with memory (e.g., strong, stable attitudes). The narratives that lead to polarization are distinct from one another (linear). Narratives that are more easily confused with one another (polarized, centrist) lead to a diversity of opinions.

Furthermore, it was puzzling that polarization, when viewed across the complete network, was rare even when we tried to force the issue by using extreme values of belief valence for the initial conditions (e.g., in the polarized condition). This may seem paradoxical, but what may explain it is that, within each cave, there existed low social entropy, due to the strong effect of the polarized initial condition. When aggregating across caves, however, the entropy is naturally larger as each cave has settled on a local set of beliefs that are uncorrelated with other caves and this is reflected as more equal probabilities for each of the 10 narrative beliefs (this would be especially true of the zero-rewire condition). A prediction, for cognitive entropy (which we explore next), is that the differences across the three valence distribution conditions will be much less than we see in social entropy, especially when the rewiring probability is zero or low.

¹ One bit of entropy is an indication that two narratives dominate a network, not necessarily that narratives representing two opposing attitudes dominate the network.

To study cognitive entropy, we conducted a 5 (memory decay rate) x 3 (narrative valence distribution) x 10 (rewiring probability) ANOVA yielded 3 2-way interactions of $p < .05$: Decay x Rewire, $F(36, 1350) = 1.86$; Decay x Valence, $F(8, 1350) = 2.15$; Rewire x Valence, $F(18, 1350) = 15.99$.

Figure 3 shows how cognitive entropy changes as a function of social context and the memory decay rate. Cognitive Entropy (and variance in cognitive entropy) generally increases (up to a certain level) as the local coherence of networks decreases (i.e., as rewiring probability increases). The relatively homogeneous social contexts provided by locally coherent networks minimize the effect of decay rate on entropy. As local coherence decreases, the decay-rate effect grows more pronounced. Thus, networks that tend toward a social consensus produce more cognitive entropy than do networks tend toward polarization or fragmentation. Our memories help reduce the degree of cognitive entropy experienced from social pressures to conform in contexts that lack the redundancy of cliquish peers.

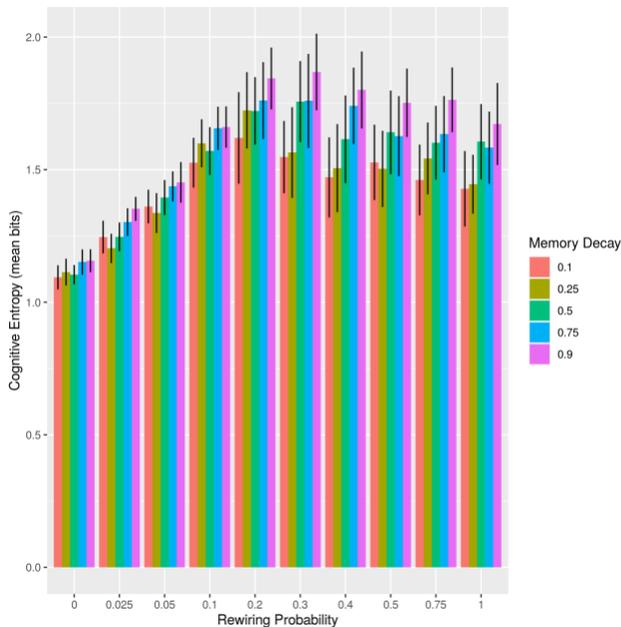


Figure 3: Effects of network topology and decay rate on cognitive entropy.

Our memories also help reduce the entropy experienced from narratives we encounter in social environments -- if the narratives can be distinguished from another. As can be seen in the left panel of Figure 4, the degree of cognitive entropy generally increases as the distinctiveness of narratives decreases. Thus a linear valence distribution (with clearly differentiable narratives) generally produces less entropy than a centrist valence distribution (with narratives that are more similar to one another), and centrist narratives produce less entropy than a polarized distribution in which everyone is using strong language which essentially conveys attitudes for or against some issue.

Interestingly, as shown in the right panel of Figure 4, heterogeneous social environments maintain the general effects of narratives on cognitive entropy: polarized > centrist > linear. In more cliquish environments, the effects of narrative valences on entropy are very similar (especially for zero rewiring probability), with the difference between polarized and centrist narratives being the most similar.

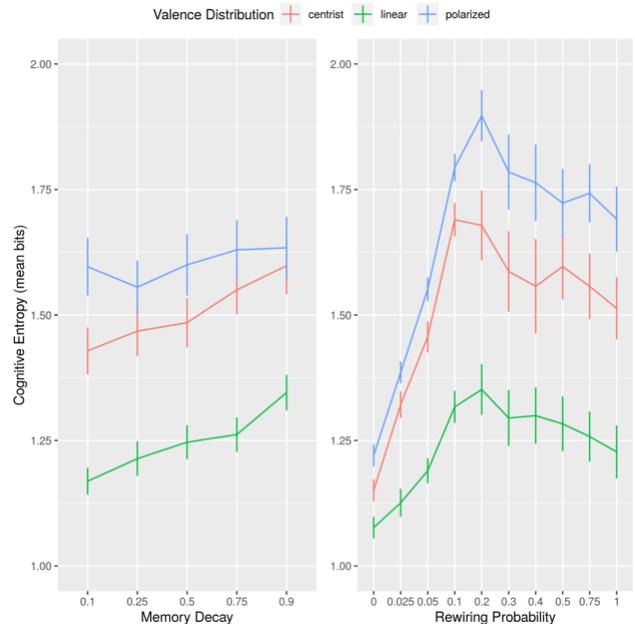


Figure 4: Effects of memory decay (left) and network topology (right) on memory entropy, divided by the distribute of narrative valence.

Discussion

We set out to show the potential value in exploring social kinds of phenomena from the perspective of cognitive architecture. At a high-level, this meant developing an understanding of the relation between internal mental representations and social structure using a single construct, information entropy. We demonstrated the ability to manipulate both cognitive and social entropy via both distribution of valence as an initial condition and network structure as a static condition.

Relations Between Cognitive & Social Entropy

According to our definition, the distribution of activation of the chunks in declarative memory determines cognitive entropy. If we imagine the hippocampus (a wetware component of declarative memory) as a communication channel between stimulus environment and response, the interpretation of cognitive entropy seems straightforward. Cognitive entropy describes the resources (i.e., channel capacity, aka attention) required to encode future events (which is compatible with the biological interpretation of Smith et al., in press).

In our simulations, agents with low cognitive entropy exist in predictable (orderly) social environments. Opinions from the neighbors of such agents provide little information for responding (i.e., propagating particular narratives). Responses of these agents thus may be driven more by expectations (cf. attitudes) than by social environments. When cognitive entropy is less than 1 bit, for example, agents "could" choose to propagate narratives that are "socially appropriate" without bothering to encode narratives received from their neighbors.

High cognitive entropy, on the other hand, indicates that social context requires responses that are more data-driven than conceptually driven. This implies that attitude strength, in some sense, should decrease as cognitive entropy increases. It also implies that agents with high cognitive entropy can be more easily influenced than those with low entropy. Furthermore, it implies that agents with low cognitive entropy may be difficult to influence, not because they harbor strong attitudes, but because experience limits their capacity for effectively encoding more complex messages; they simply do not have the bandwidth required to carry the information in complex messages that is relevant for accurate comprehension. They overgeneralize (and communication fails) because they have learned to attend to a non-discriminating subset of the features of meaning underlying the narratives of their neighbors.

Limitations

A number of limitations need to be acknowledged. This work used a highly-stylized social system to explore polarization. These results were not designed to provide insight into real-world social network dynamics, but to illustrate the approach. Another limitation is that the attitude of each belief was simulated at a purely symbolic level, without any connection to the possible effects of valence in cognition. These effects, instead, are well documented in the literature and have been incorporated into ACT-R agents in the past (Juvina et al., 2018; Smith et al., 2021). Future studies should aim to remove these limitations and test our findings in simulations with a greater degree of realism.

Implication for Polarization

These limitations notwithstanding, we believe that our results entail a number of interesting implications. First, these results might also shed light on a related, but different, problem in the social sciences: the fact that individuals who hold one extreme belief tend to harbor other extreme ones (Wood et al., 2012). In a striking example, individuals who believe in one conspiracy theory (i.e., "Princess Diana faked her own death to escape the Crown") were also found to believe in other, incompatible ones (i.e., "Princess Diana was murdered by the Crown").

It is highly unlikely that multiple radical beliefs spontaneously arise within a single person, as individuals. Instead, extreme beliefs likely spread from person to person

across social networks. This hypothesis is confirmed by the fact that the widespread use of social media in first-world countries, which amplify the reach and exposure of information, has been linked to increased partisanship, radicalization, and the spreading of fake news (Bail et al., 2018). In our results, entropy within a single agent likely tracks entropy in belief systems, and the rise of entropy in proportion to the polarization of narratives is consistent with such findings.

Finally, the finding that the network structure affects the inconsistency of beliefs has important applications for balancing policy and regulation of social media. The cognitive perspective may yield insights that are hard-gotten otherwise--understanding the micro-structure of the dynamics of social change, e.g., information operation campaigns and public health messaging, may provide the levers needed for beneficial social change.

References

- Bhat, D., & Redner, S. (2019). Nonuniversal opinion dynamics driven by opposing external influences. *Physical Review E*, 100, 050301(R).
- Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. F., ... & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 115(37), 9216-9221.
- Chelian, S. E., Paik, J., Pirolli, P., Lebiere, C. & Bhattacharyya, R. (2015). Reinforcement learning and instance-based learning approaches to modeling human decision making in a prognostic foraging task. *Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, pp. 116-122, doi: 10.1109/DEVLRN.2015.7346127.
- Eagly, A. & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace: New York.
- Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of personality and social psychology*, 67(3), 382.
- Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S., Hau, R., Hertwig, R., Stewart, T., West, R., Lebiere, C. (2010). A choice prediction competition, for choices from experience and from description. *Journal of Behavioral Decision Making* 23(1): 15-47.
- Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, 27(4), 591-635.
- Hackel, L. M., Berg, J. J., Lindström, B. R., & Amodio, D. M. (2019). Model-based and model-free social cognition: investigating the role of habit in social attitude formation and choice. *Frontiers in psychology*, 10, 2592.
- Huberman, Bernardo & A., Bernardo & Pirolli, Peter & T., Peter & Pitkow, & E., James & Lukose, & M., Rajan. (1998). Strong Regularities in World Wide Web Surfing. *Science*. 280. 95-. 10.1126/science.280.5360.95.

- Irwin, H. J., Dagnall, N., & Drinkwater, K. (2015). Belief inconsistency in conspiracy theorists. *Comprehensive Psychology*, 4, 17-CP.
- Lindström, B., Bellander, M., Chang, A., Tobler, P. N., & Amodio, D. M. (2019). A computational reinforcement learning account of social media engagement. *PsyArXiv*, (78mh5).
- Orr, M., Thrush, R. & Plaut, D. C. (2013). The Theory of Reasoned Action as Parallel Constraint Satisfaction: Towards a Dynamic Computational Model of Health Behavior. *Plos One*, <https://doi.org/10.1371/journal.pone.0062490>.
- Romero, O., & Lebiere, C. (2014). Simulating Network Behavioral Dynamics by using a Multi-agent approach driven by ACT-R Cognitive Architecture. In *Proceedings of the Behavior Representation in Modeling and Simulation Conference (BRIMS-2014)*. Washington, DC, April 2014.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379-423.
- Schelling, Thomas C. "Dynamic models of segregation." *Journal of Mathematical Sociology* 1.2 (1971): 143-186.
- Smith, B. M., Chiu, M., Yang, Y. C., Sibert, C., & Stocco, A. (in press). When fear shrinks the brain: A computational model of the effects of post-traumatic stress on hippocampal volume. *Topics in Cognitive Science*.
- Stanley, Clayton & Byrne, Michael. (2016). Comparing vector-based and Bayesian memory models using large-scale datasets: User-generated hashtag and tag prediction on Twitter and Stack Overflow. *Psychological Methods*. 21. 542-565. 10.1037/met0000098.
- Thagard, P., & Verbeurgt, K. (1998). Coherence as constraint satisfaction. *Cognitive Science*, 22(1), 1-24.
- Van Prooijen, J. W., Krouwel, A. P., & Pollet, T. V. (2015). Political extremism predicts belief in conspiracy theories. *Social Psychological and Personality Science*, 6(5), 570-578.
- Watts, D. J. (1999). Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology*, 105(2), 493-527.
- Wood, M. J., Douglas, K. M., & Sutton, R. M. (2012). Dead and alive: Beliefs in contradictory conspiracy theories. *Social psychological and personality science*, 3(6), 767-773.

Prediction Advantage as Retrieval Interference: An ACT-R Model of Processing Possessive Pronouns

Umesh Patil (umesh.patil@gmail.com)

University of Cologne
50923 Cologne, Germany

Sol Lago (sollago@em.uni-frankfurt.de)

Goethe University Frankfurt
60629 Frankfurt, Germany

Abstract

We propose a retrieval interference-based explanation of a prediction advantage effect observed in Stone et al. (2021). They reported two dual-task eye-tracking experiments in which participants listened to instructions involving German possessive pronouns, e.g. ‘Click on **his** blue button’, and were asked to select the correct object from a set of objects displayed on screen. Participants’ eye movements showed predictive processing, such that the target object was fixated before its name was heard. Moreover, when the target and the antecedent of the pronoun matched in gender, predictions arose earlier than when the two genders mismatched — a prediction advantage. We propose that the prediction advantage arises due to similarity-based interference during antecedent retrieval, such that the overlap of gender features between the antecedent and possessum boosts the activation level of the latter and helps predict it faster. We report an ACT-R model supporting this hypothesis. Our model also provides a computational implementation of the idea that prediction can be thought of as memory retrieval. In addition, we provide a preliminary ACT-R model of how linguistic processes could drive changes in visual attention.

Keywords: pronoun resolution; prediction; retrieval interference; ACT-R; possessive pronouns

Introduction

In a sentence such as “Peter wanted to go jogging with Paula, but his sneakers were torn out”, finding out the referent of the pronoun *his* involves: (i) using the linguistic knowledge that the referent should prototypically have a masculine gender, (ii) maintaining the memory representation of all the referents encountered so far, i.e. *Peter* and *Paula*, and (iii) retrieving the correct antecedent, *Peter*, and co-referring it with *his*. The task of finding an appropriate antecedent is partly facilitated by the gender feature of the pronoun, at least in languages where gender is reflected in the surface form of words.

Pronoun resolution as cue-based retrieval

The psycholinguistic processes involved in pronoun resolution can be modeled well in the cue-based retrieval (henceforth CBR) theory of sentence processing (Lewis & Vasishth, 2005; Lewis, Vasishth, & Van Dyke, 2006). The CBR theory, implemented in ACT-R (Anderson, Byrne, Douglass, Lebiere, & Qin, 2004), describes sentence processing as a series of activation-based skilled memory retrievals. Lexical knowledge and a current partial representation of the input

(the parse) are maintained in declarative memory (chunks) and psycholinguistic processes are represented in procedural memory (production rules). Incremental sentence processing occurs through the selection of production rules, which retrieve chunks from declarative memory and operate on them to update the representation of the sentence. In a CBR model of pronoun resolution, antecedent retrieval is achieved by using features such as the gender of the pronoun (Patil, Vasishth, & Lewis, 2016; Parker & Phillips, 2017; Engelmann, Jäger, & Vasishth, 2019). The CBR theory has also been used to model retrieval processes with other linguistic dependencies such as subject-verb agreement and negative polarity items (Vasishth, Brüßow, Lewis, & Drenhaus, 2008; Dillon, Mishler, Sloggett, & Phillips, 2013).

Pronoun resolution and prediction

In addition to the backward-looking processes implemented through memory retrieval, some types of pronouns also involve forward-looking processes, i.e., expectations about the identity of upcoming words. Predictions about upcoming material are an integral part of sentence processing (Huettig, 2015; Kuperberg & Jaeger, 2016). Comprehenders can generate these expectations based on the dependencies between the predicted words and previously processed ones. For example, in languages where articles and determiners need to agree in gender with a following noun, an article with masculine gender allows comprehenders to predict an upcoming noun, e.g., in the German phrase “der Knopf” (‘the.MASC button.MASC’). These agreement-based predictions are not restricted to gender, but extend to different kinds of morphosyntactic features, such as number, person and case (Dahan, Swingley, Tanenhaus, & Magnuson, 2000; Kamide, Scheepers, & Altmann, 2003; Lew-Williams & Fernald, 2010; Hopp, 2012; Zhang & Knoeferle, 2012; Grüter, Lau, & Ling, 2020).

In previous work, predictions have mostly been studied in words that solely encode forward-looking agreement, such as articles and determiners. However, there are linguistic categories that simultaneously encode backward- and forward-looking dependencies, such as linking elements (e.g. “however” or “despite of”), verbs and possessive pronouns. The current study focuses on possessive pronouns (e.g. “his” or “her”) because they are useful to investigate the interaction between antecedent retrieval and word predictions.

German possessive pronouns

Our study models the comprehension of German possessive pronouns. A German possessive pronoun such as “seinen” (‘his’) shows a bi-directional pattern of agreement: The stem “sein-” indicates a preceding masculine possessor (like “his” in English) but additionally, the suffix “-en” indicates an upcoming masculine possessum noun. These backward- and forward-looking agreement relationships mean that German comprehenders can use the two gender features of the possessive to retrieve a preceding antecedent and to predict an upcoming possessum. Thus, German possessives provide a good test case to examine whether retrieval and prediction mechanisms interact during sentence processing.

Stone et al. (2021) addressed this question in a visual world eye-tracking study and reported an interaction between these mechanisms: Participants predicted the upcoming possessum noun faster when the possessum and possessor matched in gender than when they mismatched, i.e., a prediction advantage. Here, we provide an explanation of this prediction advantage by modeling the eye-tracking experiments of Stone et al. (2021). Our model uses the sentence processing mechanism in the CBR theory and the principles of ACT-R. By doing this, the model extends the CBR architecture and further proposes that the prediction advantage is due to similarity-based interference during the antecedent retrieval process.

Data: Stone et al. (2021), Experiment 2

Stone et al. (2021) reported two visual world eye tracking studies. We first describe and model the second experiment (Experiment 2), and then extend the model to the first experiment (Experiment 1). We proceed in this order because Experiment 2 had a simpler experimental design and a larger sample size than Experiment 1, which likely yielded more precise estimates. Experiment 2 was performed by seventy-four German native speakers. At the beginning of the experiment, participants were introduced to two characters, *Martin* and *Sarah*, whose faces were displayed on screen. Participants’ task was to help Martin and Sarah tidy up their house by finding their belongings before their parents arrived. They were told that they would see images and hear instructions, and that their task was to click on the object mentioned in the instruction as quickly and accurately as possible.

During the experimental trials, participants heard an auditory instruction and saw a visual display with a target object (e.g. a blue button.MASC) and a color competitor of different gender (e.g. a blue bottle.FEM). Each object had one of four colors: red, green, blue, or yellow. There were 96 items distributed across two conditions. In the MATCH condition, shown in (1a), the possessor and target noun in the auditory instruction had the same gender, i.e. both were masculine or both were feminine. In the MISMATCH condition, shown in (1b), the possessor mismatched in gender with the target object but matched with the competitor.

- (1) a. **MATCH condition**
Klicke auf seinen blauen Knopf!
Click on his.MASC blue.MASC button.MASC
- b. **MISMATCH condition**
Klicke auf ihren blauen Knopf!
Click on her.MASC blue.MASC button.MASC

Stone et al. (2021) used a Bayesian bootstrapping procedure to estimate the earliest point in time at which participants’ fixations to the target object increased compared to those to the color competitor. This point, together with a 95% credible interval was taken as the prediction effect onset. The comparison of the predictive onset in the MATCH vs. MISMATCH condition showed that predictions were 307 [262, 352] ms earlier in the MATCH condition (Figure 1, top row). This difference indicates that predictions arose earlier when the antecedent of the possessive matched in gender with its target object, despite the fact that the antecedent gender was syntactically irrelevant for the target noun prediction.

Prediction advantage as retrieval interference

We propose that the prediction advantage observed in Experiment 2 is a consequence of interference due to a partial-cue match during retrieval, a kind of similarity-based interference (Vasishth et al., 2008). Interference occurs during the antecedent retrieval triggered by the possessive pronoun — the gender cue used in the retrieval of the antecedent in the MATCH condition boosts the activation of the gender-matching target object, but the gender cue in the MISMATCH condition boosts the activation of the competitor object. During the prediction stage, this higher activation of the target object compared to the competitor object in the MATCH condition enables a faster prediction of the target. By contrast, the higher activation of the competitor in the MISMATCH condition delays the prediction of the target object.

Next, we illustrate a model of the two experiments described in Stone et al. (2021). The model is an extension of the sentence processing mechanism from CBR that is tailored to the task of selecting the target object on screen after processing an input sentence. We use the ACT-R architecture to model non-linguistic processes. Note that we do not explicitly model eye movement processes or visual search processes as, for example, in EMMA (Salvucci, 2001) or other aspects of the visual system as, for example, in PAAV (Nyamsuren & Taatgen, 2013). The goal here is to provide an explicit proposal of how top-down psycholinguistic processes, such as antecedent retrieval and prediction, could influence the activation of elements in memory and how these activation levels could impact visual attention and fixation probabilities. We are able to model fixation patterns as they unfold in real-time, thus going beyond previous CBR models on pronoun resolution, which have solely focused on average reading time effects (Patil, Vasishth, & Lewis, 2016; Parker & Phillips, 2017; Engelmann et al., 2019).

Model of Experiment 2 from Stone et al. (2021)

The model combines the cue-based retrieval model of antecedent retrieval (Patil, Vasishth, & Lewis, 2016) and the ACT-R model for predicting the target picture matching the sentence (Patil, Hanne, Burchert, De Bleser, & Vasishth, 2016). To model the dual-task in Experiment 2, we modified the values of three ACT-R parameters (Table 1) and made the following new assumptions.

Model assumptions

- (1) At each input word, the model tries to predict the target object (the possessum) based on the information in the sentence encountered up to that point in time. This configuration seeks to replicate participants’ goal during the experiment, since their task was to click on the target object as quickly as possible. Thus, we assume that they would try to predict the target object with each new bit of linguistic information.
- (2) We assume that the objects on screen are stored as referents in declarative memory. This means that the memory representations of, for example, *Martin*, *Sara*, *button* and *bottle* are referents that are accessible during sentence processing.
- (3) The prediction of the target object is implemented as a retrieval of the memory representation of its referent. This is motivated based on the model of sentence-picture matching task in Patil, Hanne, et al. (2016). Additionally, the prediction steps weight color cues higher than linguistic cues (see Parker, Shvartsman, & Dyke, 2017 for similar cue-weighting proposals). This was done to model the saliency of visual features over linguistic features in a visual world task (Coco & Keller, 2015).
- (4) When processing the possessive pronoun, the antecedent retrieval precedes the target prediction. This reflects the linear order of the two agreement morphemes in the possessive.
- (5) The probability of fixating an object is modeled through the activation of the memory representations of the object — higher activation means higher probability of fixation. This is also based on the model in Patil, Hanne, et al. (2016).

Table 1: List of ACT-R parameter values that were modified during model fitting in Model 1. The parameters were modified to improve model fit. All other parameters had their default values or values used in earlier CBR model.

ACT-R parameter	Default	New
Activation noise (ANS)	0.2	0.15
Maximum associative strength (MAS)	1	3
Match Scale (MP)	1	0.2

Results and discussion

The model predictions for the MATCH and MISMATCH conditions are illustrated in Figure 1 (bottom row). The object with higher activation is predicted to be the object that is fixated. The activation values for objects are sampled after every temporal event, such as production firing or mem-

ory retrieval. This is done because the increment of time and memory retrievals cause the activation to change which influences the decision to fixate an object (see assumption (5) in ‘Model assumptions’). The predicted fixation curves are smooth because they are binned averaged fixation probabilities (bin size = 200 ms) across 10000 simulations. Vertical red bars denote the divergence points between the two curves predicted by the model. The divergence onset was predicted to be 400 ms earlier in the MATCH conditions compared to the MISMATCH.

The predicted fixation probabilities capture the two key effects in the empirical data. First, the prediction of the target object before hearing its name, which in the empirical data emerged as a 66 [64, 68]% target-over-competitor advantage over the entire predictive window. The model captures this effect by using the gender and color features of the possessive and the adjective to retrieve the target object (e.g. “masculine” and “blue”). Second, the model captures the earlier prediction onset in the MATCH than MISMATCH condition, which was on average 307 [262, 352] ms in the empirical data. The model captures this effect through an interaction between retrieval and prediction processes, on the one hand, and similarity-based interference, on the other. Specifically, the gender cue (masculine for the stem “sein-”) in the antecedent retrieval in the MATCH condition boosts the activation of gender-matching objects in memory, which includes the memory representation of the target object (the button, “Knopf.MASC”). Meanwhile, the gender cue in the antecedent retrieval in the MISMATCH condition (feminine for the stem “ihr-”) boosts the activation of the memory representation of the competitor object (the bottle, “Flasche.FEM”) relative to the target object. This difference in the activation of the target and the competitor during the antecedent retrieval process leads to a faster prediction of the target in the MATCH condition.

Model of Experiment 1 from Stone et al. (2021)

The goal of this model is to test the predictions made by the previous model with new data, without making any new assumptions. With this goal, we modeled the data from Experiment 1 in Stone et al. (2021). Experiment 1 was also performed by seventy-four German native speakers and it was similar to Experiment 2, with a couple of exceptions. First, Experiment 1 featured four objects on screen: in addition to the target and color competitor (‘button’ and ‘bottle’, as in Experiment 2) there were two additional objects: a “gender competitor”, which matched the target object in gender but not in color (e.g., ‘the balloon.MASC’), and a “distractor”, which mismatched the target in both color and gender (‘the flower.FEM’). Due to the presence of these differently colored objects on screen, the target object was only predictable after the onset of the color adjective (e.g., ‘blue’), because both color and gender were necessary to identify the target.

Experiment 1 had only 24 experimental trials, a smaller number than the 96 trials of Experiment 2. The experimen-

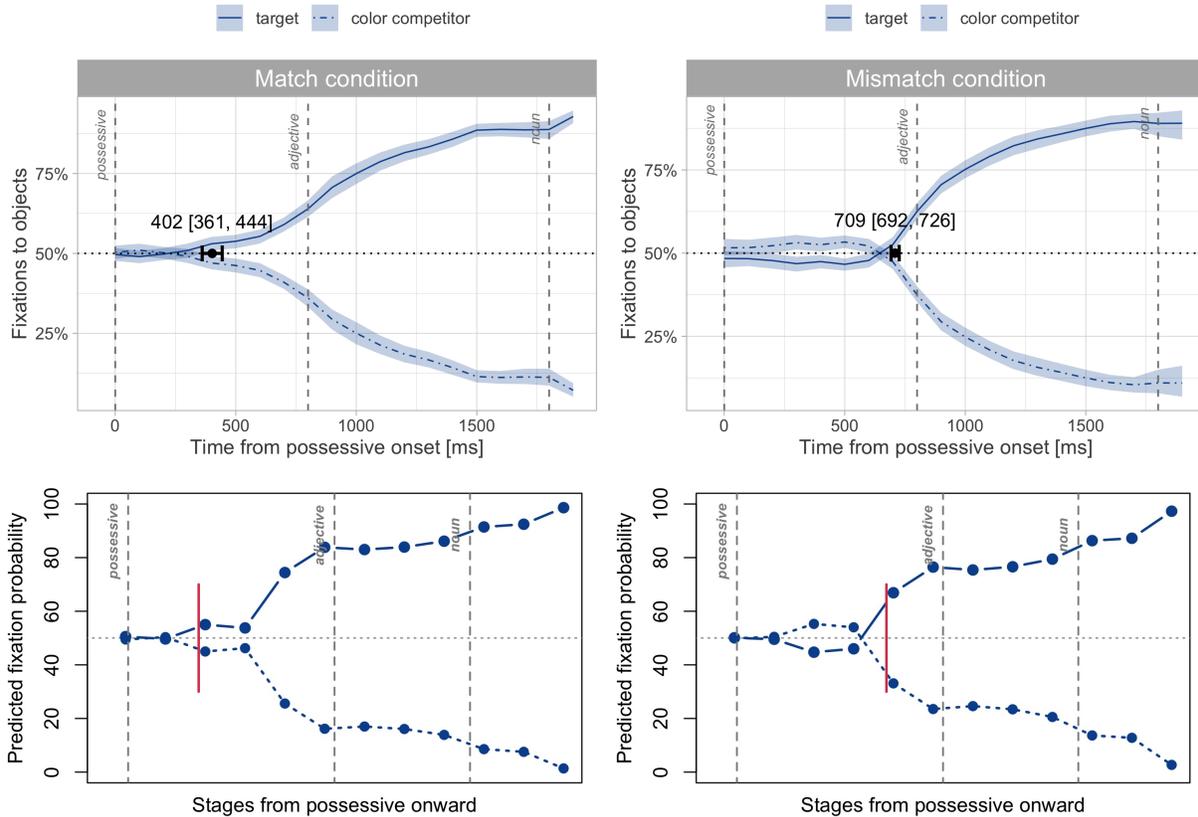


Figure 1: **Top row** (behavioral data): Fixation curves to the two objects averaged across items and participants in Experiment 2. The predictive window extended from the onset of the possessive to the onset of the noun, adding 200 ms for saccade planning (Salverda et al., 2014). The x-axis is time-locked to the possessive. Estimated prediction onsets and their 95% credible intervals are overlaid on the fixation curves in each condition. **Bottom row** (model): Predictions of the model for fixation probabilities to the target and competitor object (the button and bottle, respectively). Red bars denote the predicted divergence points between the two curves.

tal conditions were identical to Experiment 1 and featured a MATCH and a MISMATCH condition. The results were consistent with those of Experiment 1 (Figure 2, top row). First, a 59 [53, 66]% target-over-competitor advantage was observed across the whole predictive window. Second, the onset of the prediction effect was 106 ms [-56, 268] ms earlier in the MATCH than in the MISMATCH conditions. The direction of the effect suggested earlier predictions when the antecedent of the possessive matched in gender with the target object, despite the fact that the antecedent gender was syntactically irrelevant for prediction purposes. However, the magnitude of the between-condition difference was smaller than in Experiment 2 (106 vs. 307 ms on average).

The assumptions of our model were kept constant in terms of modeling the task. We also generate the predicted fixation probabilities in the same manner. The only difference is that declarative memory in the current model includes two extra referents corresponding to the two additional objects shown on screen: one for the gender competitor object (e.g. the balloon) and one for the distractor object (e.g. the flower). We

examined whether the model was able to predict the effects observed in Experiment 1 without additional assumptions.

Results and discussion

The model predictions are illustrated in Figure 2 (bottom row). The predictions are generated using the same procedure as in the previous model, with the only difference being that here the predictions are generated also for the two additional objects. The predicted fixation probabilities for the two conditions show patterns comparable to those in the data. The model partially captures the prediction advantage effect in the data: the earlier onset of the target prediction between the MATCH and MISMATCH conditions only when the prediction of target vs. the color competitor is considered. The model captures the other key effect in the data: the prediction of the target object before hearing its name. However, some predictions of the model do not correspond well to the empirical patterns. First, the model predicts similar fixation proportions to the target and gender competitor after the processing of the pronoun. This was not observed in the em-

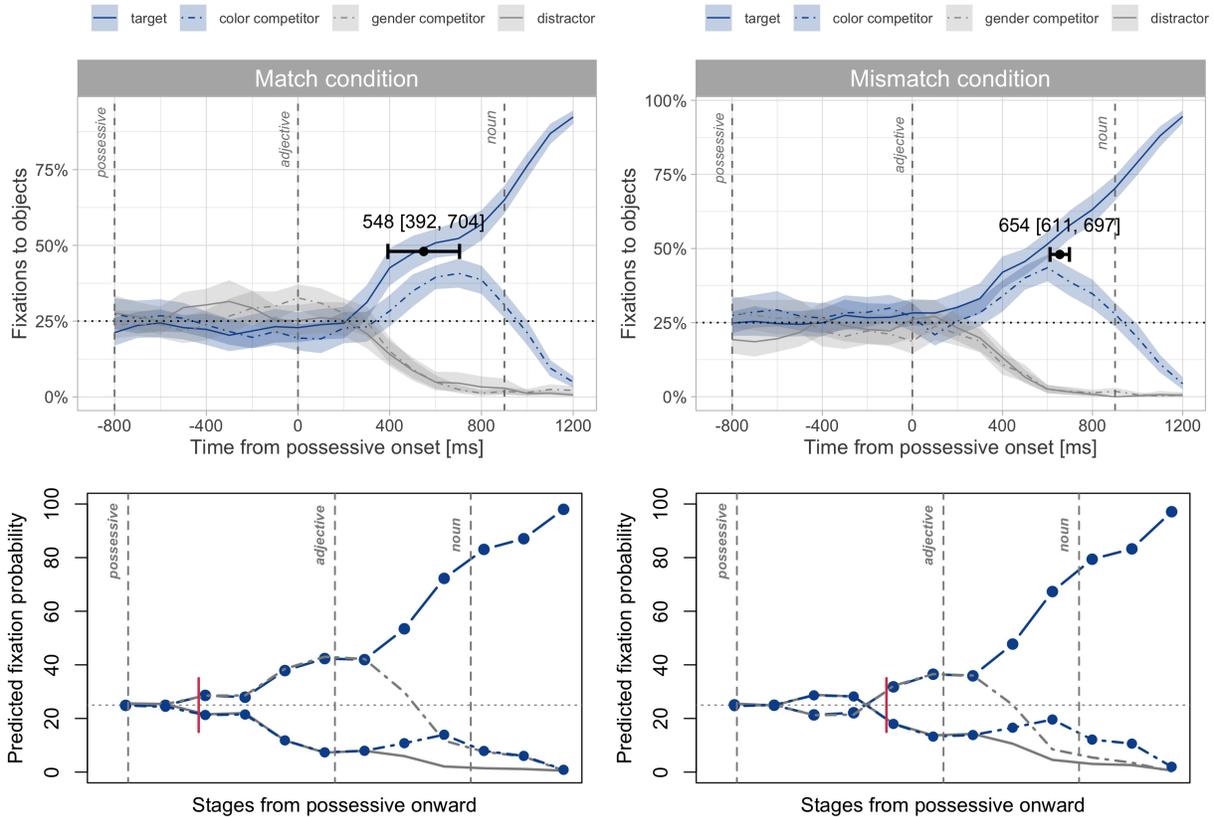


Figure 2: **Top row** (behavioral data): Fixation curves to the four objects averaged across items and participants in Experiment 1. The predictive window extended from the onset of the adjective to the onset of the noun, shifted 200 ms to the right. The x-axis is time-locked to the adjective. Estimated predictive onsets and their 95% credible intervals are overlaid on the fixation curves in each condition. **Bottom row** (model): Predictions of the model for fixation probabilities to the target (the button), the color competitor (the bottle), the gender competitor (the balloon) and the distractor object (the flower). Red bars denote the predicted divergence points between the curve for the target and the color competitor.

pirical data, in which fixations to the gender competitor were very infrequent and patterned with the fixations to the distractor object. Moreover, the magnitude of the prediction advantage (between the target and the color competitor) predicted by the model was higher than in the data. The model predicted a prediction advantage of 400 ms, however, in the data it was only 106 [-56, 268] ms. We discuss these issues in the general discussion.

General discussion

This paper reports two modeling experiments that test the hypothesis that the prediction advantage observed in Stone et al. (2021) is due to similarity-based interference during the antecedent retrieval of a possessive pronoun. Stone et al. (2021) reported two dual-task experiments involving sentence processing in the visual world paradigm. Participants listened to German sentences with possessive pronouns and were asked to select an appropriate object on a screen. German possessive pronouns have a bi-directional pattern of gender agreement: their stem encodes agreement with a previously men-

tioned antecedent but their suffix encodes agreement with a following possessum. Stone et al. (2021) found that participants predicted the target object faster when the possessum and possessor matched in gender (MATCH condition) than when they mismatched (MISMATCH condition).

We hypothesized that the prediction advantage in the MATCH condition was due to the interaction between the antecedent retrieval and the possessum prediction at the pronoun. We tested this hypothesis by modeling the dual-task from Stone et al. (2021). The model is an extension of the cue-based retrieval model of sentence processing in ACT-R. The model captures the key effect of the prediction advantage in MATCH condition in both the experiments. The prediction advantage arises due to retrieval interference during antecedent retrieval at the possessive pronoun — the overlap of the gender feature between the antecedent and the possessum boosts the activation level of the possessum which later helps in predicting it faster.

Stone et al. (2021) also observed that the prediction advantage is smaller and happens at a later stage when the visual

scene contains two extra objects, a gender competitor and a distractor object as in Experiment 1. The model only partially captures this effect — it captures the prediction advantage effect between the target and color competitor, but not between the target and the gender competitor. The model uses all the information present in the input immediately but sequentially to predict the target: first, the gender information encoded in the suffix of the possessive pronoun to rule out the color competitor, and then, the color information encoded in the adjective to rule out the gender competitor. By contrast, in Experiment 1, participants seem to delay the prediction decision until after they have heard the adjective. We acknowledge that this is a limitation of the model and needs to be investigated further. One possible way to improve the predictions of the model for Experiment 1 could be using different combinations of weights for linguistic and visual cues. However, since the size of the data in Experiment 1 was substantially smaller than in Experiment 2, we also consider that the process of adjusting parameters should be deferred until the effects in Experiment 1 are replicated using a larger sample size.

By modeling the prediction advantage in Stone et al. (2021) data, we have also created a preliminary working model of: (1) prediction in terms of retrieval, and, (2) how psycholinguistic processes might influence visual attention. This is supported by the effect that the model captures across the MATCH and MISMATCH conditions: the prediction of the target object before hearing its name. Since this effect emerged due to interactions between linguistic, visual and predictive processing, we suggest that our model is a good starting point for implementing a full-fledged model of sentence processing in the visual world paradigm. A next step towards such a full-fledged model would be to combine our model with a model that can relate higher level cognitive processes with lower-level eye movement processes and visual search, such as EMMA (Salvucci, 2001) which is an extension of ACT-R's vision module. Such an extension should also be useful for studying how the visual system could influence declarative memory and, in effect, language processing.

It has been proposed that prediction can be conceived of as a memory retrieval, but without any implemented model of this proposal (Chow, Momma, Smith, Lau, & Phillips, 2016). The current model fills this gap by demonstrating that prediction and memory retrieval do not have to be thought of as two separate or encapsulated cognitive processes. If this were the case, then interference at retrieval should not have affected prediction speed. Instead, our results show that prediction and retrieval processes interact in comprehension, either because they act on the same memory representations, or because they draw from a shared set of resources. Our model captures the prediction effect without having to posit any special prediction mechanism.

Our model converges with psycholinguistic accounts that view prediction as a memory retrieval problem, in which the linguistic and non-linguistic context are used to access rep-

resentations in working memory, with the goal of inferring which words are likely to come next (Chow et al., 2016). In this account, previously encountered words and predicted words are held in the same working memory space, with their activation levels being modulated by factors such as recency and frequency of use. Within the larger framework of cognitive science, such an account is more compatible with unified models of memory (e.g., Cowan, 1988; McElree, 2000; Oberauer, 2002) than with multistore models (e.g., Baddeley, 2000) or models that posit a specialized memory system for the storage of predictions (“prospective memory”, e.g., Zogg et al., 2012).

Funding information & data

Umesh Patil and Sol Lago were supported by the German Research Foundation (project numbers 248411105 and 317308350, respectively). The model code and supplementary files are available at: <https://osf.io/fsbwz/>

References

- Anderson, J. R., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1050.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. doi: [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2)
- Chow, W.-Y., Momma, S., Smith, C., Lau, E., & Phillips, C. (2016). Prediction as memory retrieval: timing and mechanisms. *Language, Cognition and Neuroscience*, *31*(5), 617–627. doi: [10.1080/23273798.2016.1160135](https://doi.org/10.1080/23273798.2016.1160135)
- Coco, M. I., & Keller, F. (2015). The interaction of visual and linguistic saliency during syntactic ambiguity resolution. *Quarterly Journal of Experimental Psychology*, *68*(1), 46–74. doi: [10.1080/17470218.2014.936475](https://doi.org/10.1080/17470218.2014.936475)
- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*, 163–191.
- Dahan, D., Swingle, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in french. *Journal of Memory and Language*, *42*(4), 465–480. doi: <https://doi.org/10.1006/jmla.1999.2688>
- Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, *69*(2), 85–103. doi: <https://doi.org/10.1016/j.jml.2013.04.003>
- Engelmann, F., Jäger, L. A., & Vasishth, S. (2019). The effect of prominence and cue association on retrieval processes: A computational account. *Cognitive Science*, *43*(12), e12800. doi: <https://doi.org/10.1111/cogs.12800>
- Grüter, T., Lau, E., & Ling, W. (2020). How classifiers facilitate predictive processing in 11 and 12 chinese: the role of semantic and grammatical cues. *Lan-*

- guage, *Cognition and Neuroscience*, 35(2), 221-234. doi: 10.1080/23273798.2019.1648840
- Hopp, H. (2012). The on-line integration of inflection in l2 processing: predictive processing of german gender. In A. K. Biller (Ed.), *Proceedings of the 36th annual boston university conference on language development* (pp. 226–245). Somerville, Mass.: Cascadilla Press.
- Huetig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118-135. (Predictive and Attentive Processing in Perception and Action) doi: <https://doi.org/10.1016/j.brainres.2015.02.014>
- Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003, Jan 01). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from german and english. *Journal of Psycholinguistic Research*, 32(1), 37-55. doi: 10.1023/A:1021933015362
- Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32-59. (PMID: 27135040) doi: 10.1080/23273798.2015.1102299
- Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29, 1-45.
- Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, 10(10), 447–454.
- Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native spanish speakers. *Journal of Memory and Language*, 63(4), 447-464. doi: <https://doi.org/10.1016/j.jml.2010.07.003>
- McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, 29(2), 111–123.
- Nyamsuren, E., & Taatgen, N. A. (2013). Pre-attentive and attentive vision module. *Cognitive Systems Research*, 24, 62-71. (Cognitive Systems Research:Special Issue on ICCM2012) doi: <https://doi.org/10.1016/j.cogsys.2012.12.010>
- Oberauer, K. (2002, May). Access to information in working memory. exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 411 – 421. (Publisher: American Psychological Association) doi: 10.1037//0278-7393.28.3.411
- Parker, D., & Phillips, C. (2017). Reflexive attraction in comprehension is selective. *Journal of Memory and Language*, 94, 272-290. doi: <https://doi.org/10.1016/j.jml.2017.01.002>
- Parker, D., Shvartsman, M., & Dyke, J. A. (2017). The cue-based retrieval theory of sentence comprehension: New findings and new challenges..
- Patil, U., Hanne, S., Burchert, F., De Bleser, R., & Vasishth, S. (2016). A computational evaluation of sentence processing deficits in aphasia. *Cognitive Science*, 40(1), 5-50. doi: <https://doi.org/10.1111/cogs.12250>
- Patil, U., Vasishth, S., & Lewis, R. L. (2016). Retrieval interference in syntactic processing: The case of reflexive binding in english. *Frontiers in Psychology*, 7, 329. doi: 10.3389/fpsyg.2016.00329
- Salverda, A. P., Kleinschmidt, D., & Tanenhaus, M. K. (2014). Immediate effects of anticipatory coarticulation in spoken-word recognition. *Journal of Memory and Language*, 71(1), 145-163. doi: <https://doi.org/10.1016/j.jml.2013.11.002>
- Salvucci, D. D. (2001). An integrated model of eye movements and visual encoding. *Cognitive Systems Research*, 1(4), 201-220. doi: [https://doi.org/10.1016/S1389-0417\(00\)00015-2](https://doi.org/10.1016/S1389-0417(00)00015-2)
- Stone, K., Veríssimo, J., Schad, D. J., Oltrogge, E., Vasishth, S., & Lago, S. (2021). The interaction of grammatically distinct agreement dependencies in predictive processing. *Language, Cognition and Neuroscience*, 0(0), 1-21. doi: 10.1080/23273798.2021.1921816
- Vasishth, S., Brüßow, S., Lewis, R. L., & Drenhaus, H. (2008). Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science*, 32(4), 685–712.
- Zhang, L., & Knoeferle, P. (2012). Visual context effects on thematic role assignment in children versus adults: Evidence from eye tracking in german. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th annual meeting of the cognitive science society* (pp. 2593–2598). Austin, TX: Cognitive Science Society.
- Zogg, J. B., Zogg, J. B., Woods, S. P., Woods, S. P., Saucedo, J. A., Saucedo, J. A., ... Simoni, J. M. (2012). The role of prospective memory in medication adherence: a review of an emerging literature. *Journal of behavioral medicine*, 35(1), 47-62.

Cognitive Modelling of a Mental Rotation Task Using a Generalized Spatial Framework

Kai Preuss (preuss@tu-berlin.de) and Nele Russwinkel (nele.russwinkel@tu-berlin.de)

Department of Psychology and Ergonomics, Marchstraße 23
10587 Berlin, Germany

Abstract

Bespoke cognitive models of mental spatial transformation, like those used in mental rotation tasks, can generate a very close fit to human data. However these models usually lack grounding to a common spatial theory. In turn, this makes it difficult to assess their validity and impedes research insights that go beyond task-specific limitations. We introduce a spatial module for the cognitive architecture ACT-R, serving as a framework offering unified mechanisms for mental spatial transformation to try and alleviate those problems. This module combines symbolic and spatial information processing for three-dimensional objects, while suggesting constraints on this processing to ensure high theoretical validity and cognitive plausibility. A mental rotation model was created to make use of this module, avoiding custom-made mechanisms in favor of a generalizable approach. Results of a mental rotation experiment are reproduced well by the model, including effects of rotation disparity and improvement over time on reaction times. Based on this, the spatial module might serve as a stepping stone towards unified, application-oriented research into mental spatial transformation.

Keywords: spatial cognition; mental spatial transformation; mental rotation; ACT-R

Introduction

The ability to imagine physical interaction with arbitrary objects in a physically existing space and to assess objects and their attributes based on this mental representation is a fundamental aspect of human life. Forming such a mental representation is possible through the interplay of multiple cognitive processes. These processes of mental spatial transformation are governed by common criteria that directly influence the complexity, perceived difficulty and feasibility of altering the representation (Harris, Hirsh-Pasek, & Newcombe, 2013).

Mental rotation research, as a subfield of research into those processes, has established itself as a mainstay paradigm of experimental psychology. Mental rotation refers to the mental examination of real or simulated objects so that statements about their attributes can be made beyond their initial presentation, most often their similarity to other objects. Processes of mental rotation are, on one hand, ubiquitous in everyday life: they contribute greatly to our understanding of environments by helping us assess objects and possible interactions with them. On the other hand, the phenomenon is usually studied with the use of stripped-down, abstract objects; facilitating testing in laboratory conditions but removing real-world meaning.

Performance in mental rotation tasks is heavily influenced by task difficulty and experience (Shepard & Feng, 1972).

Individual traits like reference frame proclivity (Gramann, 2013), additional workload or time pressure additionally influence mental spatial transformation. An understanding of processes underlying mental rotation has great potential, especially for economic applications, like improved product ergonomics (e.g. improving ease of use and perceptibility of features), but also for accessibility (e.g. identifying [dis-]advantages of individual traits). Cognitive modeling bases its predictions of task performance on postulated process models, which are in turn embedded into the framework of a cognitive architecture. This architecture combines multiple theories of mental processing to a holistic system that allows researchers to make general and plausible statements on cognitive processes. In the cognitive architecture ACT-R (Anderson et al., 2004), this is achieved through so-called cognitive modules which represent abstract processing stages set between neurophysiological activity and psychological correlates. Therefore, the action and interaction of these modules correspond to mental processing of declarative information and procedural action knowledge. Process models are quantified and encoded into production rules, i.e. assumptions about mental processes, that are subsequently validated with experimental data and, if necessary, engineered towards a closer fit to these data. While cognitive models encode assumptions about mental processes for specific tasks, more general mental mechanisms are implemented in the form of aforementioned modules. Hence, to model mental spatial transformation validly, the architecture needs to support a plausible implementation of it. This would then allow cognitive models of similar modalities to make use of a common, unified processing framework. By mitigating the reliance on highly task-specific assumptions and tailor-made process models in favor of a general framework, models of spatial cognition could offer higher validity and broader generalizability of their predictions.

Prior Research

Shepard and Metzler (1971) introduced an experimental paradigm for mental rotation. Their work examined the influence of rotation between two same or mirrored objects on the time needed by participants to decide if the presented objects match. They found a linear relationship of rotation discrepancy on reaction times. A follow-up study on mental folding (Shepard & Feng, 1972) showed similar results. Here, a fold-

ing pattern was required to be assembled into a cube shape to decide if it was a copy of a reference cube that was also presented. A linear effect of task difficulty on reaction times was found. Interestingly, the experiment also showed what the researchers perceived to be an upper limit on mental spatial transformation ability –above a certain threshold of required folds, reaction times increased considerably and non-linearly. Consequently, this result could be a pointer towards a general limitation on the amount of transformations that can be applied on an internal spatial representation.

Just and Carpenter (1976) used eye tracking during a mental rotation study to determine the existence of distinct cognitive stages. Based on their results, they proposed three general stages of cognitive processing: initial search, transformation and comparison, and confirmation. These stages can serve as an approximation for spatial cognition in general: a visual encoding phase, a transformation phase and a comparison or matching phase.

Eye tracking was also used during a mental folding experiment to try and find correlates for cognitive stage switching (Preuss, Hilton, & Russwinkel, 2020). Differences in gaze position switches and gaze durations were found that correlated with task difficulty. This was interpreted as signifiers of stage switching and stage duration, respectively.

To further differentiate processes during mental rotation and investigate possible solving strategies, Yuille and Steiger (1982) presented a study on objects with different complexities. While showing that object complexity has a direct influence on solving time, they introduced their theory of two distinct solving strategies: if an object is “familiar” enough, it can be transformed holistically, meaning as a whole; if the object is not recalled, it must be transformed in a piecemeal fashion, meaning it is separated into several parts or features which are then processed in sequence. This distinction proved to be a popular explanation for learning effects in mental rotation and mental spatial transformations in general.

Harris et al. (2013) reviewed differences and similarities between mental rotation and mental folding as the most common paradigms in mental spatial transformation research. While the tasks differentiate in the specific way a stimulus is processed, Harris et al. identified several attributes that underlie both processes, for instance physical analogy, malleability and predictiveness of success in *Science, Technology, Engineering & Mathematics* (STEM) fields. This work points to spatial cognition as a technical, trainable skill. Similar results were obtained by Wright, Thompson, Ganis, Newcombe, and Kosslyn (2008), who also compared skill development in a mental rotation and a mental folding task, in addition to a verbal analogy task. Learning one spatial task improved proficiency in the other tasks, but not as pronounced for the non-spatial task. Notably, the researchers argue that improvement comes mostly from improved encoding and transformation preparation processes, less from transformations per se, implying learning to stem largely from non-spatial mechanisms.

A cognitive model for a mental rotation task was previ-

ously introduced by Peebles (2019a). Peebles implemented both piecemeal and wholesale strategies on a simplified visual representation. Different to the approach presented here, the model was mostly self-contained and relied on default ACT-R mechanisms, with only slight changes to the architecture.

Gunzelmann and Lyon (2007) first proposed the concept of a cognitive module dedicated to spatial transformations. They presented a relatively complex mechanism, making use of several smaller information processing units. Unfortunately this approach has not yet been implemented into a cognitive architecture.

Several other approaches for mental transformations not relying on internal, three-dimensional representations exist: arguments for reliance on mental imagery (Peebles, 2019b; Lovett & Forbus, 2013), purely physical reasoning (Forbus, 1984; De Kleer & Brown, 1984) or syllogistic representations (Barkowsky, Knauff, Ligozat, & Montello, 2007) have been made for mental spatial transformations. The cognitive architecture SOAR offers a mechanism (*Spatial and Visual System*, SVS) that combines symbolic and spatial information (Laird, 2008).

This paper presents a cognitive task model for a mental rotation task that incorporates such a spatial framework for ACT-R, proving the usefulness of an additional module dedicated to spatial processing. This module is proposed as an extension to the cognitive architecture, integrating seamlessly into its existing structures and allowing multiple modalities of mental spatial processing to be simulated in a unified manner. It serves as an interface for the mathematically correct computation of three-dimensional space while processing it in a cognitively plausible way, without having to rely on overly task-specific assumptions about spatial processing. The spatial module presented in this paper shows a similar concept to the one suggested by Gunzelmann and Lyon (2007), but foregoes many of their proposed mechanisms in favor of a seamless integration into ACT-R’s existing architecture. Default ACT-R modules are used for memory retrieval and for comparison purposes. Additionally, by integrating the proposed module into existing methods for simulating module activity and, by extension, brain activity, model predictions can be compared by brain-imaging data of participants in the actual experiment (Prezenski & Russwinkel, 2016).

The module’s validity is pending on further assessment of its ability to predict mental spatial transformation processes for several modalities beyond mental rotation. As multiple design decisions are as of now made intuitively, open questions on structure, function and cortical localization of the module are tended to by current and upcoming research.

Methods

Spatial Module

The mental rotation model uses a dedicated spatial module added to ACT-R’s default architecture, facilitating the processing of mental spatial transformations. Based on work by Gunzelmann and Lyon (2007), the idea is to offer seamless

functionality for three-dimensional data in ACT-R in a cognitively plausible fashion. In contrast to the aforementioned work, the framework presented herein avoids episodic and allocentric buffers and relies instead on standard ACT-R mechanisms.

The spatial module aims to offer better explainability, applicability and validity for cognitive models of spatial cognition by offering a common theoretical ground for frequently shown effects such as differences in spatial strategies or influence of higher task difficulties on task solving. A unified mechanism for simulating mental spatial transformations would offer modelers both the ability and the constraints necessary to do so with high reliability and high validity, respectively. Effectively this would create a general framework spanning multiple paradigms of mental spatial cognition research, such as mental rotation or mental folding. In consequence, to the best of our knowledge it would be the first cognitive modeling approach to explain both paradigms in a satisfactory manner.

Spatial objects are encoded in standard chunks, ACT-R's basic unit of information, extended by information representing the object in 3D space in the form of so-called point clouds. This additional information is predefined for each spatial object, either implicitly by the model's environment or the modeler themselves. Point clouds were chosen for their versatility, scalability and relative ease of computation. They are able to represent objects in arbitrary detail, allowing modelers to focus on features relevant to their model. Extending chunks in this manner allows for full compatibility with all default ACT-R mechanisms such as vision or memory modules, while at the same time allowing algebraic manipulation of objects defined in this manner, i.e. being translated, rotated and rescaled in three-dimensional space. Furthermore, spatial objects can be compared and angles between objects can be measured. In practice, this extends the symbolic capabilities of ACT-R with the ability to perceive and interact with geometric properties. Analogous to how visual information is processed in ACT-R, the transition between geometric and symbolic information of perceived features is handled by the models themselves, in contrast to e.g. SOAR's similar mechanism (Laird, 2008).

Cognitive operations on spatial objects are handled by two buffers: a storage buffer for maintaining mental spatial representations (the *spatial* buffer) and an action buffer for applying transformation intention to said representation (the *spatial action* buffer). Spatial chunks contain point clouds and optionally additional spatial information like separable parts, angles for internal transformation or other features. Transformations on the representation are requested through the action buffer and, if within limits set by architectural and modular constraints, applied to the spatial object. The core function of the module is calculating a time delay for operations conducted through it. It does this through a transformation cost function which draws from currently available information to calculate an appropriate time frame for a transformation pro-

cess to take place. Currently the following simple formula is used:

$$\text{Transformation delay} = F * M * x$$

including a delay factor (F) which can be set as a parameter with a default value of 0.005s, an optional modality factor (M) to assign weights to different transformation modalities, such as mental rotation or mental folding (if required for model adjustments) and the raw input value of the transformation (x). This formula is an attempt to find a common denominator underlying mental spatial transformations. By combining symbolic processing with three-dimensional spatial information, several limitations by aforementioned prior research could be alleviated or overcome. Contrary to task-specific approaches, this framework constrains models to adhere to established mechanisms of cognitive spatial processing which facilitates explainability, validity and generalizability in model creation. Additionally, compared to methods relying on default mechanisms of ACT-R (e.g. using the imaginal module to store and process simplified spatial information as in Peebles, 2019a), the presented module enables symbolic calculations with true three-dimensional data. Finally, this module serves as a solid foundation for more complex models orientated away from lab conditions and towards real-world applications.

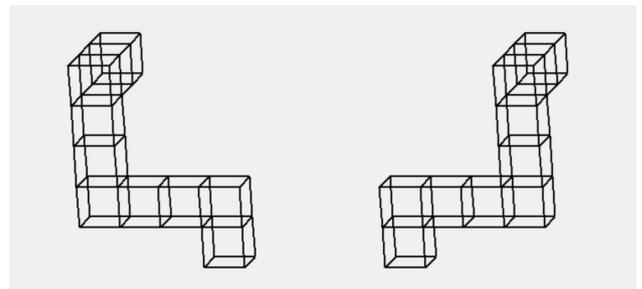


Figure 1: An example picture of 3D mental rotation stimuli as presented to the model. For simplicity, each cube is drawn around a single 3D coordinate. Multiple coordinates make up the point clouds of the whole figure and its features (i.e. straight sections orthogonal to each other), respectively.

Experiment

Participant data was collected during a mental rotation experiment as part of a Bachelor's thesis (Raddatz, 2014). The experiment was based on the classic mental rotation paradigm by Shepard and Metzler (1971). In a trial, one out of 16 figures is presented to the participant without any rotation. After 1 second, either the original figure or a mirrored version of it is presented and rotated by either 0, 50, 100 or 150 degrees on the picture plane. The participant must decide whether the presented objects are equal or mirrored variants of each other. To this end, the participants are instructed to mentally rotate one of the objects clockwise until an informed decision can be made if the objects match or mismatch. 6 Blocks of each

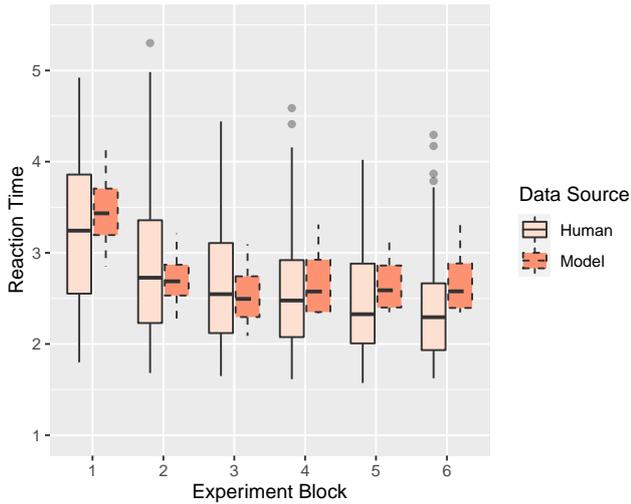


Figure 2: Aggregated human (leftmost, solid outline) and model reaction times (rightmost, dashed outline) for each experiment block.

possible trial combination (16 figures * 4 degrees of rotation * 2 types of mirroring = 128 combinations) take place, resulting in 768 trials overall. The cognitive model was designed to solve a simulated version of this experiment.

Mental Rotation Model

By making use of the spatial module’s ability for both symbolic and spatial information, the mental rotation task model implements a cognitively plausible approach for human-like solving. The cognitive model follows the process model originally proposed in Just and Carpenter (1976), and follows their proposal of three rough stages –initial search, transformation and comparison, and confirmation. The model offers two strategies, first differentiated by Yuille and Steiger (1982) as “holistic” comparison (also referred to as “wholesale”) and “piecemeal” comparison: if the presented figure is “known”, meaning the object is sufficiently familiar, the object can be transformed and compared as a whole. On the other hand, if the presented figure is unknown to the solver, meaning it was not seen before or forgotten, it has to be sequentially transformed and compared by its individual features or *pieces*. In the case of mental rotation stimuli, pieces are the respective straight sections formed by multiple cubes, of which each figure has either 3 or 4. Thus, use of a piecemeal strategy explains longer reaction times for “unknown” figures in human trials.

At the start of each trial, the current reference stimulus is presented: both its individual features and the complete object are placed in the environment as visual features visible to the model. First, the model encodes the whole object and attempts a declarative memory query, testing for object familiarity –if successful, the wholesale strategy is initiated. If the presented object can not be remembered, the model waits

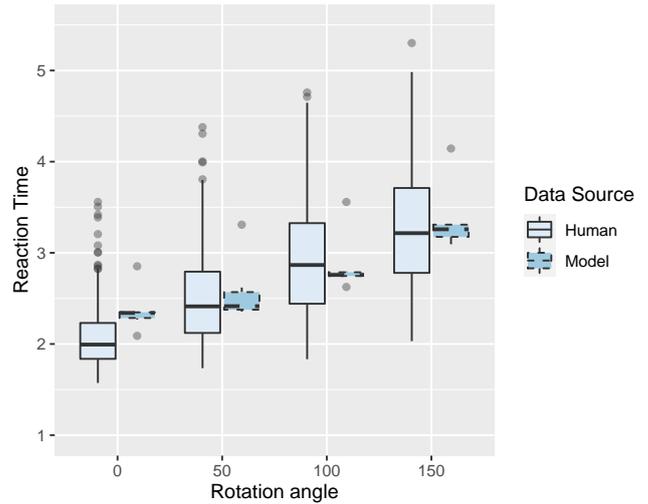


Figure 3: Aggregated human (leftmost, solid outline) and model reaction times (rightmost, dashed outline) for each rotation angle.

for the appearance of the target stimulus, which happens one second after the reference stimulus appears. Then, either the whole target object is visually encoded and prepared for the wholesale strategy, or its separate features are visually encoded and considered for the piecemeal strategy.

While solving the mental rotation task, the model rotates the object or parts of the object –depending on the strategy – by a fixed amount of 45 degrees, chosen to be close but avoid equality to the experiment’s rotation conditions. After each rotation, a comparison process measures the mean euclidean distance between paired points of the point clouds of the target object with its reference counterpart, resulting in a similarity value. If this comparison results in a similarity higher than a preset threshold, but lower than the last value computed (or is the first comparison for this trial), an additional rotation is planned and executed. If the comparison yields a value higher than the threshold but also a value higher than the last similarity value, the model assumes that a low enough similarity value cannot be reached and gives a “mismatch” answer. If the similarity value is lower than the threshold, a match of objects is assumed. In the wholesale strategy, the trial is then directly confirmed as a “match”. In the piecemeal strategy, the degrees of rotation necessary to reach this similarity are remembered and applied to subsequent pieces. If all pieces yield similarity values under the threshold, the object is considered a match and a “match” answer is given. For this experiment, a threshold value of 20 and 45 degrees of rotation per transformation yielded the best results. Additionally, the following parameters were adjusted as follows:

- Latency factor: 0.3 (default: 1.0)
- Retrieval threshold: -1.0 (default: 0.0)
- Activation noise: 0.5 (default: none)

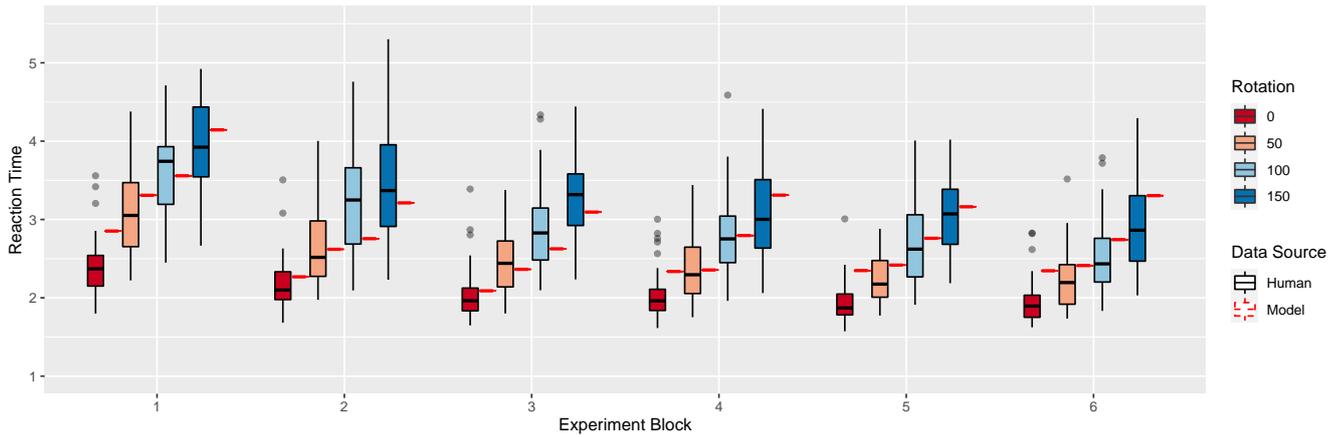


Figure 4: Human reaction times (leftmost per color, solid black outline) and model-predicted reaction times (rightmost per color, dashed red outline), grouped by rotation condition and experiment block.

- Utility noise: 2 (default: none)
- Spatial delay: 0.005 (default: n/a)

The aforementioned strategy choice is integrated in the form of a memory retrieval to mimic object familiarity –if the reference stimulus was presented often and recently enough, the model can proceed with the wholesale strategy, otherwise the piecemeal strategy is chosen. Reinforcement learning is implemented in the form of utility learning: model decisions that consistently result in fast and correct task solving will be reinforced and chosen more frequently in subsequent trials.

Results

Both datasets were prepared by Median Absolute Deviation outlier correction. Figure 4 shows reaction times from participant data and model predictions.

Model Data

The model predictions correlate nicely with behavioral data. A strong overall correlation with little deviation to human data was achieved ($r(22) = .92, p < .001$; $RMSE = .23$). More specifically, the influence of rotation disparity on model and human reaction times aggregated over all blocks reached a very strong correlation ($r(2) = .97, p < .05$, $RMSE = .139$) (see also Figure 3), while comparing the influence of experiment block aggregated over all rotations shows a strong correlation ($r(4) = .84, p < .05$, $RMSE = .22$) (see also Figure 2). Overall standard deviation of model-predicted reaction times is close to the original data, but slightly lower ($SD_H = 1.953, SD_M = 1.421$).

Regression Analysis

A linear model was created, gauging the influence of experiment block, rotation disparity and data source (human or model) on reaction times. The three predictors explained 53.9% of the variance ($R^2 = 0.539$, $F_{47,840} = 23.04$, $p <$

.001). Rotation angle significantly predicted reaction times ($\beta = 0.66, p < 0.001$), as did the interaction between experiment block and angle ($\beta = -0.29, p < 0.05$). Data source has no influence on reaction times ($\beta = 0.02, p = 0.43$), implying no significant differences between human and model results. As shown in Figure 4, a linear effect is visible, with increased rotations leading to increased reaction times. Over blocks, reaction times are generally lowered, with a more pronounced effect for higher rotations.

Discussion

Interpretation of Results

The behavioral data collected shows a linear effect of difficulty typically reported in mental rotation studies (Shepard & Metzler, 1971). A decrease of reaction time over the experiment blocks suggests a learning effect that is more pronounced for higher task difficulty, which mirrors results previously reported for a mental folding task (Preuss, Raddatz, & Russwinkel, 2019). The model results show a promising fit to the behavioral experiment data. Aside from a strong general correlation, it accurately models learning over experiment blocks, which validates the implemented strategy choice mechanism based on object familiarity.

Correlation between the two datasets is comparable to results from similar modeling approaches to mental rotation (e.g. Peebles, 2019a). Of note is that our results stem from the reliance on generalized spatial processes instead of mechanisms tailored to the task at hand, giving strong support for the validity of a unified approach.

Open Questions and Known Issues

The spatial module for ACT-R enqueues itself into a line of similar theoretical approaches and implementations. Mental imagery (Peebles, 2019b; Lovett & Forbus, 2013), qualitative reasoning (Forbus, 1984; De Kleer & Brown, 1984), syllogistic representations (Barkowsky et al., 2007), or spatial-

visual integration (Laird, 2008) offer alternatives to tackle open questions in spatial research. As of now, our common spatial framework does not challenge these theories, as insight into the nature of the cognitive mechanisms underlying spatial processes is still vague. Further research could increase support for our approach, or dismiss it altogether.

Most design decisions for the spatial module are made under consideration of prior research as outlined above. Still, many of its mechanisms are currently in need of verification. For now, the mental rotation model is the only cognitive model fully realized using this framework. While this model proved successful, additional work on cognitive models for other spatial paradigms is necessary to validate the framework further.

The underlying experimental data was originally collected for an EEG study –therefore, the experimental design was kept simple to reduce unwanted artifacts (i.e. stimuli were only rotated on the picture plane, low overall task difficulty). This restricts the use of these data for several interesting questions in the modeling domain: does mental spatial transformation happen statically and stepwise, or is it dynamic? Is there a number of maximal transformations applicable on a mental spatial object? These issues will be addressed in future study designs.

Outlook

Since ACT-R simulates cognitive functions in a modular fashion, it lends itself to modeling effects beyond behavioral data: a method proposed by Prezenski and Russwinkel (2016) would allow a comparison of ACT-R module activity to EEG data of experiment participants. To this end, components are calculated from EEG data, i.e. clusters of neurons that are frequently active in parallel. In the case of independent component analysis (ICA), components with the highest degree of independence from one another are generated, meaning that in theory, cortex areas fulfilling distinct functions are mapped for each participant during task solving. These independent components can then each be associated to ACT-R's modules by correlating brain activity with predicted module activity. This could help verify or falsify the existence and/or location of one or several dedicated spatial area(s). Another promising approach lies in computing principal components of EEG signals for comparison with module activity produced by the cognitive models. A principal component analysis (PCA) ranks components by variance explained which then can be associated with activity of specific modules during specific times during task solving (Borst & Anderson, 2015; Tenison, Fincham, & Anderson, 2016). While ICA matching helps localizing specific brain activity, PCA matching allows for temporal correlation. Both methods are currently being tested on data sets created by the mental rotation model.

[Mention of related project omitted for anonymity] In addition to a study on mental rotation, an experiment on a mental folding paradigm (Shepard & Feng, 1972) was conducted and simulated in a cognitive model using the spatial module (Preuss et al., 2019). Applying the spatial module to a re-

lated mental spatial paradigm allows for further verification or falsification of its validity and should lead to adjustments necessary for its further generalization. To arrive at a module representing universal spatial cognition, it will be important to follow the constraints dictated by both cognitive architecture and neurobiological plausibility to avoid parameter overfitting.

As the effects of several factors on spatial processing time are yet to be gauged and additional spatial paradigms yet to be implemented on the basis of the spatial module, it currently computes the time necessary for mental spatial transformations on the basis of an admittedly simple multiplication. Other variables influencing the outcome are in consideration to be included in later versions of the spatial module, for instance added noise to reduce the formula's deterministic behavior, increased processing time depending on the number of transformations already applied to the object in the spatial buffer or an upper limit to the transformations applicable in a row.

A follow-up experiment will combine both mental rotation and mental folding into one experimental paradigm. By forcing the use of cognitive folding and rotation processes at the same time, this study and further upcoming work will rely less on lab conditions and move towards real-world applications. Requiring both spatial modalities for problem solving will allow further evaluation of the proposed module's validity.

Acknowledgments

The authors would like to thank Leonie Raddatz and the chair of Biological Psychology and Neuroergonomy for data acquisition, Dan Bothell for technical support & Linda Heimisch and Severin Reuter for additional data analysis and model feedback. This research is financed through the German Research Foundation (DFG), as part of project #396560184.

References

- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060.
- Barkowsky, T., Knauff, M., Ligozat, G., & Montello, D. R. (2007). *Spatial Cognition V: Reasoning, Action, Interaction*. Springer Science & Business Media.
- Borst, J. P., & Anderson, J. R. (2015). Using the ACT-R cognitive architecture in combination with fmri data. In *An Introduction to Model-Based Cognitive Neuroscience* (pp. 339–352). New York, NY: Springer New York.
- De Kleer, J., & Brown, J. S. (1984). A qualitative physics based on confluences. *Artificial Intelligence*, *24*(1-3), 7–83.
- Forbus, K. D. (1984). Qualitative process theory. *Artificial Intelligence*, *24*(1-3), 85–168.
- Gramann, K. (2013). Embodiment of spatial reference frames and individual differences in reference frame proclivity. *Spatial Cognition & Computation*, *13*(1), 1–25.

- Gunzelmann, G., & Lyon, D. R. (2007). Mechanisms for Human Spatial Competence. In T. Barkowsky, M. Knauff, G. Ligozat, & D. R. Montello (Eds.), *Spatial Cognition V: Reasoning, Action, Interaction* (pp. 288–307). Springer.
- Harris, J., Hirsh-Pasek, K., & Newcombe, N. S. (2013). Understanding spatial transformations: similarities and differences between mental rotation and mental folding. *Cognitive Processing*, 14(2), 105–115.
- Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8(4), 441–480.
- Laird, J. E. (2008). Extending the soar cognitive architecture. *Frontiers in Artificial Intelligence and Applications*, 171, 224.
- Lovett, A., & Forbus, K. (2013). Modeling spatial ability in mental rotation and paper-folding. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 930–935). Berlin, Germany: Cognitive Science Society.
- Peebles, D. (2019a). Modelling alternative strategies for mental rotation. In T. Stewart (Ed.), *Proceedings of the 17th International Conference on Cognitive Modelling* (pp. 138–143). Waterloo, Canada: University of Waterloo.
- Peebles, D. (2019b). Modelling mental imagery in the ACT-R cognitive architecture. In A. Goel, C. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st Annual Conference of the Cognitive Science Society* (pp. 2550–2556). Montreal, Canada: Cognitive Science Society.
- Preuss, K., Hilton, C., & Russwinkel, N. (2020). Cognitive processing stages during mental folding are reflected in eye movements. In *Symposium on Eye Tracking Research and Applications (ETRA'20 Adjunct)*. New York, USA: ACM.
- Preuss, K., Raddatz, L., & Russwinkel, N. (2019). An implementation of universal spatial transformative cognition in ACT-R. In T. Stewart (Ed.), *Proceedings of the 17th International Conference on Cognitive Modelling* (pp. 144–150). Waterloo, Canada: University of Waterloo.
- Prezenski, S., & Russwinkel, N. (2016). A proposed method of matching ACT-R and EEG-data. In D. Reitter & F. Ritter (Eds.), *Proceedings of the 14th International Conference on Cognitive Modeling* (pp. 249–251). PA: Penn State: University Park.
- Raddatz, L. (2014). *EEG correlates of sex differences during mental rotation*. Lüneburg, Germany. (Unpublished Bachelor's thesis)
- Shepard, R., & Feng, C. (1972). A chronometric study of mental paper folding. *Cognitive Psychology*, 3(2), 228–243.
- Shepard, R., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701–703.
- Tenison, C., Fincham, J. M., & Anderson, J. R. (2016). Phases of learning: How skill acquisition impacts cognitive processing. *Cognitive Psychology*, 87, 1–28.
- Wright, R., Thompson, W. L., Ganis, G., Newcombe, N. S., & Kosslyn, S. M. (2008). Training generalized spatial skills. *Psychonomic Bulletin & Review*, 15(4), 763–771.
- Yuille, J. C., & Steiger, J. H. (1982). Nonholistic processing in mental rotation: Some suggestive evidence. *Perception & Psychophysics*, 31(3), 201–209.

Towards Precise Measures of Individual Performance in Complex Tasks

Roussel Rahman (roussel.rahman@gmail.com) &

Wayne D. Gray (grayw@rpi.edu)

Cognitive Science Department, Rensselaer Polytechnic Institute
Troy, NY 12180 USA

Abstract

Simple laboratory tasks typically allow one or a few methods of task performance. In contrast, moderately complex tasks, such as video games, provide many methods of task performance which, in essence, provide many ways of completing the task without necessarily completing all possible components. Although performance on complex tasks improves with practice, the improvements do not represent the simple effects of power-law learning but, rather, they tend to reflect the discovery and practice of a diverse set of methods. Understanding what we see during complex task learning, requires us to evaluate individual performance against benchmarks of optimality. In this report, we use the game of Space Fortress (SF) as a complex experimental paradigm in which we demonstrate two alternative measures that reveal scopes of individual differences in the discovery and implementation of an optimal method that would be missed by traditional measures of the game.

Keywords: Complex Task Learning; Individual Learning; Plateaus; Dips; Leaps; PDL; SpotLight

Introduction

General laws which explain human learning as a function of practice (e.g., the power law or the exponential law) implicitly assume that practice alone is sufficient to reach the asymptote of performance. Although such assumptions may be reasonable for simple tasks that afford few alternative methods, they do not hold for more complex tasks, such as video games, which afford many alternative methods. A growing body of work (e.g., Siegler, 1987; Rickard, 1997; Delaney, Reder, Staszewski, & Ritter, 1998; Towne, Boot, & Ericsson, 2016; Thompson, McColeman, Blair, & Henrey, 2019; Rahman & Gray, 2020; van der Mijl & van Rijn, 2021) shows that individuals demonstrate both inter- and intra-individual differences of task execution methods during learning and also that the practice benefits are largely localized to the specific methods practiced. Indeed, even for seemingly simple video games (e.g., Pacman or Tetris), it may be difficult to identify the optimal method from amongst its numerous alternative possibilities.

The difficulty in finding the best or even an appropriate method can be observed in many real-world tasks; for example, finding the fastest route in traffic, finding a sure-win formula for Chess or Football, solving mathematical problems, even choosing the tasks to learn in a lifetime. How do humans search for and find the optimal method(s) in such tasks? To reach the asymptote, optimal methods must be discovered or invented. Therefore, theories of complex task learning must

include an account of how the individuals' task execution methods evolve with learning to reach the optimal one(s) at the asymptote of performance.

Until now, we portrayed the complexity of complex tasks from a performer's perspective. But similar difficulties also persist for the researchers of complex skill learning in deciding where to look for measurable changes and which measures to use (Gray & Lindstedt, 2017). Looking at the wrong or imprecise measures can easily lead to false negatives of learning or training benefits, as underlying improvements may remain undiscovered (Gray, 2017). Moreover, if the asymptote(s) of performance and the corresponding optimal method(s) are both unknown, it is difficult to ensure that increments in performance measures are indeed steps towards the asymptote. The reason is that individuals may be using suboptimal methods that would lead to plateaus instead of the asymptote (Gray, 2017; Rahman & Gray, 2020).

An approach that has been useful in evaluating complex task performance is comparing performance against benchmarks of optimality. For example, Anderson, Kleinberg, and Mullainathan (2017) recently investigated the predictors of blunders in chess endgames, by comparing each move against known optimal moves. Relevantly, they found that the players are more likely to err in positions with fewer optimal or near-optimal moves within very large pools of possible moves. This relationship was consistent across all skill levels, even for the best human players with ELO ratings above 2300. In cases where optimal performance is not known, expert performance may serve as a substitute. For example, van Meeuwen et al. (2014) compared performance of novice air-traffic controllers against experts' performance to investigate how effective strategies are formed in solving complex visual problems (e.g., finding the optimal landing order for incoming planes).

In this work, we explore the benefits of evaluating individual performance against benchmarks of optimality in a historic experimental paradigm – the complex game of SF (Mané & Donchin, 1989). Since its development, SF has been used in many studies of complex skill learning to enrich our understanding of human learning process. However, several studies observed that two very important measures of SF – Velocity and Control – that represent the most fundamental skill needed in the game (flying in the game universe), are prone to ceiling effects; consequently, the measures asymptote before humans do (Boot et al., 2010; Destefano, 2010;

Gray, 2017). Here, we use two alternative measures – (1) angular velocity of player ship and (2) approximations of π (π) from ship paths – both tailored to capture progress towards optimal flight strategy of moving in slow, small circles around the enemy (i.e., the Fortress). These measures depict a much clearer picture of individuals’ route to optimality and reveal scopes of changes in individuals’ performance that would be missed by the traditional measures of SF.

The Game of Space Fortress

The game of SF was developed by Mané and Donchin (1989) as a common complex task for different research groups to study complex skill acquisition. The goal was to create an experimental task representative of real-world complex tasks incorporating dimensions of complexity based on existing research. Complexity of SF stems from both the multiplicity of tasks to be performed and the specificity of the ways they need to be performed. In each game of SF, the player flies a ship (yellow plane in Figure 1) equipped with a limited number of missiles to engage in a five-minute battle against the Fortress (located at the center of the screen). The Fortress needs to be destroyed in two steps: (1) make it vulnerable by 10 or more hits (at intervals > 400 ms), then (2) a double shot with an interval within 250-400ms to destroy it – any deviation results in instant recovery of the Fortress. The Fortress fights back by shooting shells at the ship; in addition, its minions (the mines, Figure 1) spawn periodically at random locations to chase the ship. There are two types of mines, each of which requires identification by letter-codes shown at start screen and specific handling. The player must also protect the ship from getting hit by enemies, as four hits would result in ship destruction. Both the ship and the Fortress respawn upon destruction and the battle resumes. Finally, the player needs to manage the ship’s arsenal. Each game starts with a full ar-

senal and the player receives several bonus opportunities to replenish the arsenal. The player can still shoot missiles with depleted arsenal, but sacrificing points per missile.

Game-generated Scores as Performance Measures

The objective of the game is to maximize the Total score, which is the sum of four subscores – Points, Speed, Control and Velocity – capturing performance in different sub-tasks. The Points score serves as a measure of several skills together; such as, skills in fighting the Fortress and the mines, defending own ship, managing resources. The Speed score rewards speed of killing mines and penalizes if mines escape.

The rest of the two scores are both measures of ship maneuvering skills. The Control score measures the performers’ control over OS’ spatial location; the player is rewarded at a higher rate for staying within the large hexagon than outside (Figure 1). The Velocity score measures the performers’ control over OS’ velocity; the player is rewarded for flying the ship within a speed limit and penalized for any violations. As mentioned previously, these two measures are prone to ceiling effects and do not consistently reflect improvements in associated skills. In the next section, we briefly review findings of optimal/expert flight behavior, before discussing alternative measures.

Review of Optimal (Flight) Strategies in SF

As mentioned earlier, SF was developed as a common paradigm to compare different training regimens (Mané & Donchin, 1989). The original game included only the Points score. The other three scores were added by Gopher, Weil, and Siegel (1989) for their *Emphasis Change* study. In their experiment, players practiced in the whole task, but were instructed to prioritize different parts at different points during practice. In contrast, Frederiksen and White (1989) adopted a part-task training approach by discretizing the gameplay into sub-tasks and trained the players by building up from small to more integrated subtasks. The purpose was to develop a better understanding of the dynamics of the SF universe. The researchers first identified the gameplay variables that affect task execution methods and the high-level goals of the game, after verbal protocol analysis of expert players’ methods; then, decided on an optimal method as the foundation of a hierarchical training regimen.

Frederiksen and White observed that the optimal methods in low-level subtasks of Space Fortress are regulated by three high-level goals: (1) Hit Fortress without getting hit (2) Handle mines with the least possible disruption to the first goal (3) Allocate resources to maximize the Points score. For the two first goals, they suggested that players should fly around the Fortress in circles constructed by a series of pre-planned, linear trajectories at low speeds, and when a mine appears, players should wait until mines move to locations which require minimum deviation from the circles.

Recent works confirm that the flight paths of expert players indeed converge to circles around the Fortress (Destefano, 2010; Towne et al., 2016; Rahman & Gray, 2020). The need

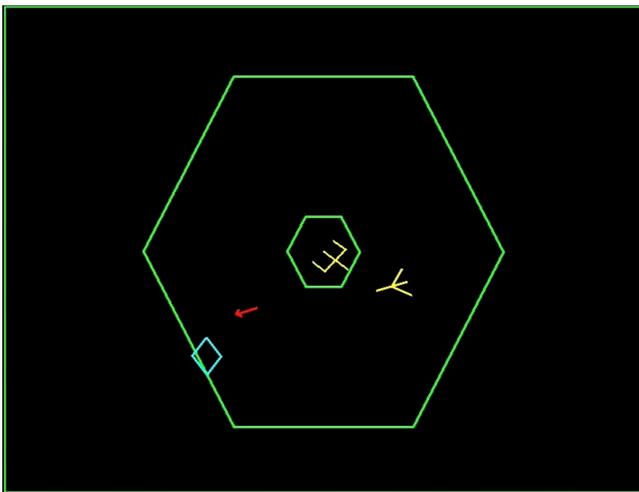


Figure 1: Screenshot of Pygame Space Fortress 4 (Destefano, 2010). The Fortress is at the center; the player’s ship (yellow) have recently fired a missile (red) at a mine (blue diamond).

to construct circles with small lines stems from the constraints imposed in SF's input system: a player can either move the ship along straight lines (using *Thrust* key) or rotate (using *Rotate* key) to change ship direction, but cannot simultaneously use both to move at angles. To obtain a circular path (and to attack the enemies), a player needs to periodically repeat a sequence of keypresses throughout the game: *Thrust-Rotate* -(*Shoot*). This way, movement constraints force players to construct the circles with numerous straight lines. Similar constraints also exist for joystick-based input systems. The need to precisely synchronize actions indicates that even if a high-level description of the optimal method is known, it cannot be implemented without understanding the mechanics of the SF universe and mastering the low-level action components. Confirming this view, Rahman and Gray (2020) found that players demonstrate both inter- and intra-individual differences of flight paths during learning, even when explicitly instructed on the optimal choice of slow circles.

Recent works also fine-tune Frederiksen and White's suggestions for optimal flight control. For example, Destefano (2010) observed that although low velocities are optimal for the Velocity score, players would benefit in defending against the Fortress by flying the ship faster. The reason is: the Fortress only fires at the ship if it can locate the ship for more than 1 second, in one of 36 equally divided segments around the Fortress (Destefano, 2010, pp. 34, Figure 14). Hence, by moving at an angular velocity faster than $360/36 = 10$ degrees/second, a player can prevent the fortress from shooting at the ship. Therefore, the upper limit of velocity is the speed limit for Velocity score, whereas the lower limit of velocity is determined by the minimum angular velocity. In the next section, we knit these pieces of information together for a more precise description of optimal flight control in SF.

Methodology

Dataset used

We use the dataset from Destefano (2010). This dataset is publicly available (osf.io/v5mzx/) and has been used in several previous studies (Destefano & Gray, 2016; Gray, 2017; Rahman & Gray, 2020). We chose this dataset as it contains millisecond-level performance records of nine individuals over 31 hours of gameplay. Each individual played 8 games in each 1-hr session (one session per day), resulting in total 248 games per player. As the players needed time to familiarize themselves with the complex rules of SF, we exclude the data from the first day. Therefore, the final dataset contains 240 games for each player.

To provide a glimpse of the richness of information, the dataset contains about 40 game-aggregated measures to capture performance in different subtasks (e.g., number of Fortress kills, ship deaths, missiles bought). More importantly for our work, 9000 datapoints were collected at 30 Hz frequency from each 5-minute game, documenting each keypress by the player, each tiny movement by the Fortress or the mines and many more. The detailed records at the lowest-

level performance mean that a researcher can develop performance measures tailored to answer specific research questions, at any level of the complex task.

Description of Optimal Flight Performance

To maximize the Control score, a player must fly the ship in the area between the large and small hexagons; using specifications of the hexagons, this statement can be written as:

$$50 < \text{Ship distance from Fortress (in pixels)} < 182$$

To maximize the Velocity score, a player only needs to maintain *linear velocities* below a specified limit. But moving at a fast enough *angular velocity* in circles would prevent the Fortress from firing at the ship. Therefore:

$$\begin{aligned} \text{limit from ang. vel.} < \text{Linear velocity (pixels/sec)} < 120 \\ 10 < \text{Angular velocity (degrees/sec)} < \text{limit from lin. vel.} \end{aligned}$$

Measures of Flight Control

To demonstrate the reasons for ceiling effects of Control and Velocity scores, we examine their main constituents – respectively, distance from the Fortress and ship velocity. As examples of alternative measures, we demonstrate two measures: (1) angular velocity of the OS and (2) approximations of π using Archimedes' method.

We chose angular velocity because, in circular motions, the radius of the circle (i.e., the distance from Fortress) and the linear velocities on the circles follow this mechanistic relation: *linear velocity* = *angular velocity* * *radius* (Beer et al., 1972). Previously, it has been noted that the Control and Velocity scores are correlated (Boot et al., 2010; Gray, 2017); the aforementioned relation indicates that this correlation would exist only when the optimal method of moving in circles is adopted.

Finally, as a measure of goodness of the circles, we use approximations of π from the individuals' flight paths using Archimedes' method. The method of constructing circles with many lines, closely resembles Archimedes' method of calculating π by approximating circumferences of circles from perimeters of polygons. In both cases, with increasing number of sides, the polygons converge to the circles, resulting in better approximations of π .

Results and Discussions

Discrepancies between the Stories Revealed by the Control Score and its Constituents

As mentioned earlier, several previous works noted that the Control and the Velocity scores asymptote before humans reach the limit of performance and thereby hide underlying improvements in low-level constituents of the scores. In Figure 2, we show the game-averages of the Control score's main low-level constituent – ship distance from the Fortress – which confirms that the players improved in flying close to the Fortress deep into practice. However, the Control score would stop showing these improvements beyond the red-dashed line. The reason is, to max out the score, a player only

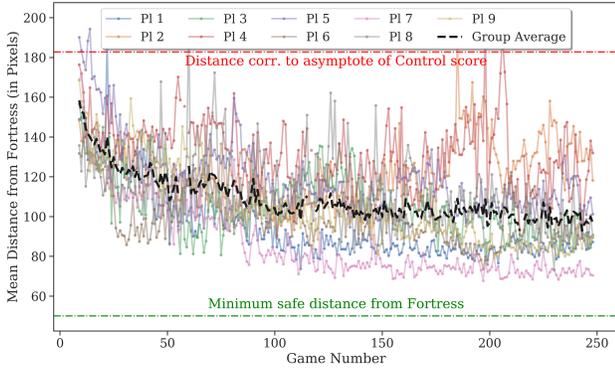


Figure 2: Mean distance for all players, almost none of whom seem to have reached the limit of performance even at the end of practice.

needs to stay within the large hexagon which corresponds to the ship distance denoted by the horizontal red-dashed line. However, as can be observed, players continue to improve beyond this threshold and gradually approach the minimum safe distance from the Fortress (green-dashed line).

The asymptote of the Control score is clearly false, but what about the score’s ability to depict changes in individuals’ performance with learning? Figure 3 demonstrates the Control score (red line) along with the game averages of the distance from Fortress (blue line), for our best player (Player 7) alone. Previously, it has been demonstrated that Player 7 went through a period (games 50-80) of extensive explorations of optimal flight paths before permanently adopting the optimal flight paths of circles around the Fortress (Rahman & Gray, 2020, samples of within-game trajectories in Figure 3, pp. 981). But the player’s Control score in Figure 3 shows hardly any signs of these explorations. Rather, the player seemed to have suddenly leapt from a plateau to the asymptote. The mean distance (blue line) shows the improvement to be much more gradual than the Control score does, with no obvious plateau in the preceding period.

Although we discuss only the Control score here, the same discrepancies were also observed for the Velocity score and its only low-level constituent, ship velocity. To find the reasons behind these discrepancies, we next look at how these scores are constructed.

Discontinuous Reward Functions ⇒ Disproportionate Rewards with Performance

The step functions for rewards in Equations 1 and 2 explain why we see a stepwise progression of the scores despite continuous improvement of players in associated low-level performance, as step functions convert continuous input to stepwise, discontinuous outputs. To elaborate, these two scores are largely insensitive to any intermediate improvements in flying skills apart from right at the transition point of the function. For example, Velocity score rewards would be the same for flying OS at 10x, 2x and 1.01x of the speed limit, but dif-

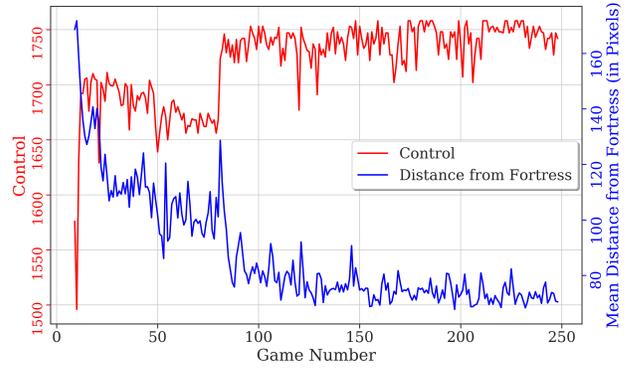


Figure 3: Control score vs its main constituent (mean distance from Fortress) for Player 7.

ferent at 0.99x. The situation is analogous to having a digital watch showing only the hours of time; as the changes of minutes or seconds would not be observable, progression of time would seem to follow a step function to uninformed eyes.

$$\text{Control score reward} = \begin{cases} +6 \text{ per second; } & \text{Inside large hexagon} \\ +3 \text{ per second; } & \text{Out of large hexagon} \end{cases} \quad (1)$$

$$\text{Velocity score reward} = \begin{cases} +7 \text{ per second; } & \text{within speed limit} \\ -7 \text{ per second; } & \text{above speed limit} \end{cases} \quad (2)$$

In summary, the disproportionate relation between reward and performance level leads to (i) the asymptotes of the game-generated scores not being equal to the asymptotes of players skills, and may lead to (ii) false plateaus in individual performance hiding underlying improvement and (iii) false leaps by rewarding long-term improvement in one burst.

Exploiting Knowledge of Optimal Methods to Capture Progress towards Optimality

Spurious plateaus and asymptotes are likely to lead researchers to false conclusions about learning patterns and training effects, especially when the performance records are not as detailed as ours and are studied only at a high level. Therefore, the step-wise reward functions of the Control and the Velocity scores should be replaced with more continuous functions to develop measures sensitive to fine-grained changes in performance. One option is to adopt a reductionist view and deconstruct elements of performance to investigate progress towards optimality in each element. However, as all subtasks within a whole complex task are not independent, the true asymptote of performance in the whole task would inevitably be lower than the one estimated from parts. A more practical approach is to use the knowledge of optimal methods from previous works to identify or develop measures that would unambiguously reflect progress towards optimality.

For example, as for flight paths in SF, we know that the optimal method is to move in circles around the Fortress. To investigate progress towards this optimal method, we may

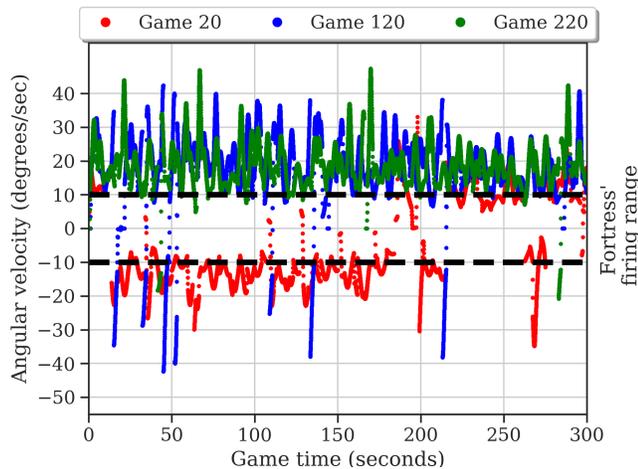


Figure 4: Instantaneous angular velocity (positive in the clockwise direction) in three sample games for Player 7

simply investigate how these circles improve with learning. The circles are implemented by controlling the ship’s angular velocity, which we use as our first measure. Next, to investigate the goodness of end results (i.e., the circles), we use approximations of π from ship’s paths.

Angular Velocity: The Velocity score aims to capture players’ control over the ship based on its linear velocity (i.e., the rate of change of linear position) of OS; but for circular motions, angular velocity (i.e., the rate of change of angular position) is a more appropriate determinant of control. To create a perfect circle, the player needs to maintain a constant angular velocity throughout. This requirement, due to the input constraints of SF, needs to be approximated by consistently oscillating about steady reference values. In addition, the Fortress can be prevented from shooting by maintaining an angular velocity greater than 10 degrees/second within the circles. Therefore, instantaneous angular velocity provides an excellent measure to investigate within-game flight control.

Figure 4 shows the instantaneous angular velocities in three example games (at 100-game intervals) played by Player 7. As can be seen, from as early as the 20th game (shown in red), the player demonstrates remarkable consistency in maintaining a wave-like angular velocity about steady reference values. In the 120th game (blue), the player shows marked improvements in controlling the angular velocity and in staying away from the Fortress’ firing range (marked in figure 4), and then shows comparatively smaller improvements in the 220th game (green). As mentioned earlier, Player 7 extensively explored and practiced different flight paths within games 50-80 before permanently adopting the circles, providing a specific explanation for the diminished returns from practice. Finally, based on the patterns observed in angular velocity, we can safely conclude that the player indeed progressed towards optimal flight performance with practice.

Approximations of π using Archimedes’ Method: Archimedes had observed that, with increasing number of

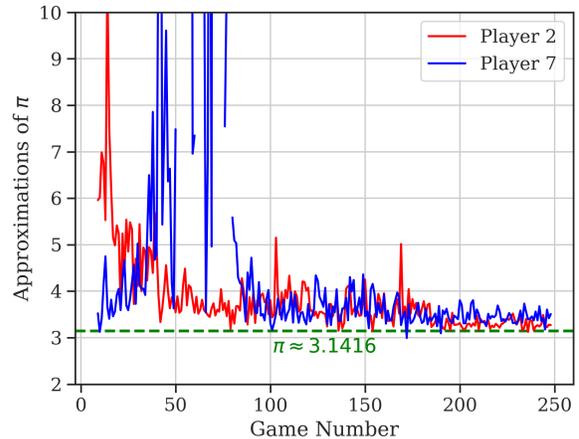


Figure 5: Approximations of π using Archimedes’ method, for our best player (Player 7) and the worst (Player 2).

sides (n), regular n -sided polygons become increasingly better approximations of circles. This simple observation led him to develop one of the earliest methods to calculate π as the ratio between the perimeter of the n -polygon and its largest diagonal. Although the flight paths taken by our players are not regular polygons, this ratio can still be used to approximate π for each full circle around the Fortress.

Figure 5 demonstrates the game averages of these π values for two players: Players 7 and 2, respectively the best and the worst performing players according to the Total scores achieved in the last 50 games. As mentioned earlier, Player 7 experimented with different flight paths (e.g., moving along lines or half-circles) within games 50-80. The impact of these experimentations are clearly observable, as either π could not be calculated (13 games) or were very inaccurate during this period. To facilitate comparison between the players, we limit the y-axis to show only values below 10, which occludes 11 games for Player 7 and one for Player 2.

Although these players demonstrate opposing trends early in practice, both players can be observed to be approaching the asymptote (i.e, the true value of π), yielding increasingly better approximations of π with more practice. To illustrate the level of accuracy reached at the end of practice, average π in the last 50 games is 3.3 (SD = 0.09) for Player 2 and 3.4 (SD = 0.14) for Player 7.

To note, the approximations of π from each circle around the Fortress can also be used as a within-game measure of performance in maintaining the circles. We skip this demonstration due to space constraints, but the within-game approximations fluctuate a lot more than the game-averages do, indicating substantial detours from the circles. Therefore, even though the game-averages suggest that the players are approaching the asymptote of performance, ample room for improvement may still remain.

Conclusions

In this work, we highlight the need to evaluate individual performance in complex tasks against benchmarks of optimality. Individuals demonstrate ample differences of task execution methods in complex tasks, therefore, looking in the wrong scopes of improvement may lead to false negatives regarding individuals' training or practice benefits. In such cases, measures tailored to capture performance within scopes of optimality, provides a common ground to compare different individuals and search for general patterns underneath the individual differences.

For our demonstrations, we use the complex game of SF and investigate individuals' acquisition of one fundamental skill – flying the ship – using two measures tailored to capture progress towards the optimal flight strategy. We chose angular velocity as our first measure, as the optimal path of circles needs to be implemented by controlling the angular velocity. Second, we use approximations of π as a measure of goodness of the circles created.

These measures – directed to capture performance within scopes of optimality – are able to reveal scopes of consistency and changes in individuals' performance that would be missed by undirected measures. For example, our results indicate that the individuals did realize that the circular paths need to be achieved by maintaining a consistent (optimally, constant) angular velocity and improved in doing so with practice. Excellent approximations of π towards the end of practice show that these players attained near-asymptotic skill levels in executing the optimal flight strategy. Importantly, as the asymptotes are known for both measures (i.e., constant angular velocity and the true value of π), improvements in these measures can be unambiguously interpreted as progress towards optimal performance. The known asymptotes also allow us to reliably investigate within-game performance of individuals with the same measures and identify both the scopes of current expertise and for further improvements.

Although our demonstrations are in one game only, the game of SF represents real-world complex tasks that present performers with the general difficulty to identify optimal methods among many alternatives. Evaluating performance against benchmarks of optimality would help us find general explanations for how individuals' different routes converge towards the same optimal methods and when do they diverge towards plateaus of stable, suboptimal performance. This way, by helping to uncover the evolution of individuals' task execution methods, precise measurement and evaluation of individuals' performance can help us progress towards the general laws of individual learning.

Acknowledgments

Correspondence should be sent to Roussel Rahman, Cognitive Science Department, Rensselaer Polytechnic Institute, Troy, NY 12180. Email: rousssel.rahman@gmail.com . The work was supported, in part, by grant N00014-16-1-2796 to Wayne Gray from the Office of Naval Research, Dr. Ray

Perez, Project Officer.

References

- Anderson, A., Kleinberg, J., & Mullainathan, S. (2017). Assessing human error against a benchmark of perfection. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 11(4), 1–25.
- Beer, F. P., Johnston Jr, E. R., Mazurek, D. F., Cornwell, P. J., Eisenberg, E. R., & Sanghi, S. (1972). *Vector mechanics for engineers* (Vol. 1). Tata McGraw-Hill Education.
- Boot, W. R., Basak, C., Erickson, K. I., Neider, M., Simons, D. J., Fabiani, M., ... others (2010). Transfer of skill engendered by complex task training under conditions of variable priority. *Acta Psychologica*, 135(3), 349–357.
- Delaney, P. F., Reder, L. M., Staszewski, J. J., & Ritter, F. E. (1998). The strategy-specific nature of improvement: The power law applies by strategy within task. *Psychological Science*, 9(1), 1–7.
- Destefano, M. (2010). *The mechanics of multitasking: The choreography of perception, action, and cognition over 7.05 orders of magnitude*. Unpublished doctoral dissertation, Rensselaer Polytechnic Institute.
- Destefano, M., & Gray, W. D. (2016). Where should researchers look for strategy discoveries during the acquisition of complex task performance? The case of Space Fortress. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 668–673).
- Frederiksen, J. R., & White, B. Y. (1989). An approach to training based upon principled task decomposition. *Acta Psychologica*, 71(1-3), 89–146.
- Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica*, 71(1-3), 147–177.
- Gray, W. D. (2017). Plateaus and asymptotes: spurious and real limits in human performance. *Current Directions in Psychological Science*, 26(1), 59–67.
- Gray, W. D., & Lindstedt, J. K. (2017). Plateaus, dips, and leaps: Where to look for inventions and discoveries during skilled performance. *Cognitive Science*, 41(7), 1838–1870.
- Mané, A., & Donchin, E. (1989). The space fortress game. *Acta Psychologica*, 71(1-3), 17–22.
- Rahman, R., & Gray, W. D. (2020). Spotlight on dynamics of individual learning. *Topics in Cognitive Science*, 12(3), 975–991.
- Rickard, T. C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General*, 126(3), 288.
- Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, 116(3), 250–264.
- Thompson, J. J., McColeman, C. M., Blair, M. R., & Henry, A. J. (2019). Classic motor chunking theory fails to

- account for behavioural diversity and speed in a complex naturalistic task. *PLOS ONE*, 14(6), 1-24.
- Towne, T. J., Boot, W. R., & Ericsson, K. A. (2016). Understanding the structure of skill through a detailed analysis of individuals' performance on the space fortress game. *Acta Psychologica*, 169, 27–37.
- van der Mijl, R., & van Rijn, H. (2021). Attention does not affect the speed of subjective time, but whether temporal information guides performance: A large-scale study of intrinsically motivated timers in a real-time strategy game. *Cognitive Science*, 45(3), e12939.
- van Meeuwen, L. W., Jarodzka, H., Brand-Gruwel, S., Kirschner, P. A., de Bock, J. J., & van Merriënboer, J. J. (2014). Identification of effective visual problem solving strategies in a complex visual domain. *Learning and Instruction*, 32, 10–21.

The Algebra of Cognitive States: Towards modelling the Serial Position Curve

Stefan Reimann (stefan.reimann@protonmail.com)

Department of Psychology, Cognitive Psychology
University of Zurich, Switzerland

Abstract

A computational framework for modelling storage and retrieval of information in human working memory is proposed. The aim is to analyse the corresponding algebra alone, especially with regard to its congruence with empirical findings including the serial position curve. That algebra builds on the high-dimensional holographic representation of information together with two operations for computation: multiplication for binding and addition for bundling. Unlike other models, the bundling operation defined is not associative and preserves serial order information in terms of activation gradients. Consequently, the cognitive states representing a memorised list exhibit a primacy as well as a recency effect generically. The typical concave-up and asymmetrically shaped serial position curve is derived as a linear combination of those gradients. The serial position curve for cued recall, including similar items, is derived within this formalism. Quantitative implications of the algebra are shown to agree well with empirical data from basic cognitive tasks.

Keywords: human working memory; activation gradients; serial position curve; holographic representation; high-dimensional computing

Introduction

Human Working Memory is commonly regarded as a functional subsystem of memory, whose goal is to hold and to organise information for some short period of time in order to make it available for higher cognitive processes (Cowan, 2017). Experiments in this field rely on the subtle construction of input data such as memory lists and produce output data such as recall probabilities or response times (Murdock, 1974; Kahana, 2012; Oberauer et al., 2018). Among these, the most prominent finding is the serial-position curve, which shows the accuracy of item retrieval varying as a function of serial position in a memory list, averaged over a sample of participants. As observed across (probably all) immediate memory tasks, it has a concave-up shape and is asymmetric. Its particular shape depends on the particular cognitive task. For example in recognition and in cued (probed) recall the serial position curve shows a strong recency effect, while the primacy effect is weak. Strong primacy effects are seen in forward recall, while recency effects are strong in backward or in free recall.

To describe particular aspects of the functioning of the human working memory, models with different characteristics have been used, differing both in terms of the medium in which the information is stored and the storage operations used. Models include local code models such as REM

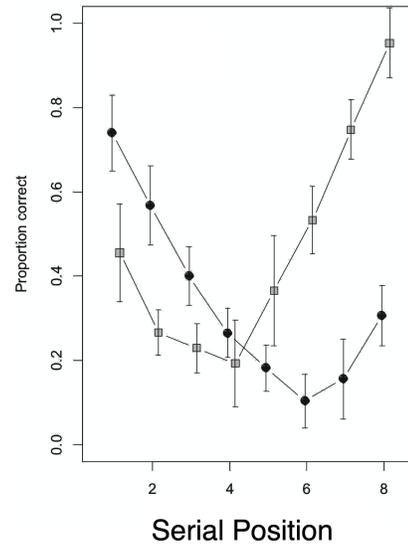


Figure 1: **Primacy effect and recency effects:** Data are for immediate forward and backward serial recall (Oberauer et al., 2018), but are similarly in other immediate tasks (Murdock, 1974; Kahana, 2012)

(Shiffrin & Steyvers, 1997), distributed models of memory such as SOB (Farrell & Lewandowsky, 2002) and TCM (Howard & Kahana, 2002), as well as holographic models such as TODAM (Murdock, 1982, 1993) which uses high-dimensional probabilistic encoding for the holographic representation of information (Plate, 1991). Holographic models gain from the properties, which are implied by high dimensionality together with randomness, see (Kanerva, 2009) for an overview about the framework of high-dimensional computing.

The holographic approach appears as a natural candidate to model the functioning of cognitive processes. Input items evoke activity patterns in the respective neural field; The fact that these representations are sparse and the consequences thereof are not explicitly considered in this note. Computation consists in transforming those patterns according to two elementary operations: The additive-like superposition realises the bundling of item information, while multiplication realises binding of items. The high-dimensional space

of binary patterns together with these two operations form a high-dimensional algebra governing storage and computation in this system.

Before giving an outline of the paper, a remark seems worthwhile: The aim is not to provide a full-blown model rather than to propose an elementary computational structure, an algebra, on top of which a model could be constructed. The main question is, how much of experimental findings can already be described on the basis of that algebra alone.

The outline of this paper is briefly as follows: Firstly, the state-space is defined as a high-dimensional Hamming space (eq 1) equipped with some distance on it. A similarity measure is proposed which is derived from that distance. It allows both, to judge about the familiarity of two states as well as about their distinctiveness (eq 2). Computing is by manipulating states according to two operations on that space: multiplication for binding and a not associative addition for bundling. This completes the definition of the algebra (eq 4) to be considered. Non-associativity is an essential feature of that bundling since it implies that the sum of components depends on their sequential ordering (eq 5). As a consequence, information about the order of sequentially presented list items is conserved. The corresponding left-associative sum and the right-associative sum of list items correspond to states exhibiting a recency and a primacy gradient, respectively (Fig 4). As applications basic cognitive tasks such as item recognition and probed recall are considered. The typical concave-up and asymmetrical shape of the serial-position curve is derived as a mixture of these two activation gradients (Fig 9).

The algebra of cognitive states $(\mathbb{X}, +_p, *)$

The state-space In the course of perceiving a physical item, the corresponding sensory input invokes an activity pattern in the neuronal field it is projected to. That way, each physical item can be represented by a binary pattern, in which 1's indicate active neurons, while 0's indicate inactive ones. Due to the size and structural complexity of the neuronal correlate, patterns are described by high-dimensional random binary vectors. These patterns are the states of the cognitive system. The state-space therefore is

$$\mathbb{X} = (\mathbb{X}_q^N, d). \quad (1)$$

$N > 100$ is its dimension, q is the degree of sparseness, i.e. the mean activity of a state, and d is some metric on \mathbb{X}_q^N .

The state-space is a (metric) Hamming space allowing for some similarity measure derived from the distance d . This measure should respect both: the closeness of two states as well as their distinctiveness as points in the state-space. A cosine-similarity only reveals information about closeness since it is locally defined. In a probabilistic setting, two points are the more difficult to distinguish, the less likely it is to find another state at random which is 'in between' the two. To capture this, the definition of similarity must contain global information about the state space.

Definition (Similarity). *The similarity of two states having distance d from each other is*

$$S(d) := e^{-\kappa F_{\mathbb{X}}(d)}, \quad \kappa > 1 \quad (2)$$

where $F_{\mathbb{X}}(d) = \mathbb{P}_{\mathbb{X}}[D \leq d]$ is the distribution function for distances on \mathbb{X} .

Different items are represented by uncorrelated states, while similar items will be represented by similar states. $\kappa > 1$ is chosen to have highest sensitivity with respect to almost identical or near-by states.

The operations The two operations to be defined on the state space correspond to binding and bundling. Two items are (associatively) bound to each other, if one can be retrieved by cueing with the other item. The corresponding formal operation is *multiplication* $*$, which is defined in eq 3. Binding of items happens by simultaneously activated components in the neural pattern. This similarity measure directly relates to a recall probability or accuracy of retrieval.

Bundling means collecting items by adding their respective states. Assume that two neurons X and Y converge on a third neuron Z . If both are inactive, i.e. $x = y = 0$, neuron Z will also be, $z = 0$, while if both are active, Z will be active, i.e. $1 + 1 = 1$. If only X or Y is active, it depends on some threshold, whether Z is active. If the activation threshold is low, $p \approx 0$, Z is likely to be active, while if the activation threshold is high, $p \approx 1$, Z will remain inactive. Addition x_p is defined in eq 3.

$$\begin{array}{c|cc} * & 0 & 1 \\ \hline 0 & 1 & 0 \\ 1 & 0 & 1 \end{array} \quad \begin{array}{c|cc} +_p & 0 & 1 \\ \hline 0 & 0 & \zeta \\ 1 & \zeta & 1 \end{array} \quad (3)$$

where $\zeta \in \{0, +1\}$ is random with $\mathbb{P}[\zeta = 0] = p$. This completes the definition of the algebra used to calculating with cognitive states.

$$(\mathbb{X}, +_p, *) \quad (4)$$

In the following, its elementary properties are further investigated. What properties are already implied by this elementary algebra and how much of empirical findings can be already described by those?

Bundling preserves sequential information in the memory list

Usually, bundling is realised by vector-addition (Schlegel, Neubert, & Protzel, 2020), which is commutative and associative, so that $x + (y + z) = (x + y) + z = z + (y + z)$, i.e. the order of components doesn't matter. That is: If addition is associative, sequential order information is lost!

Observation. *For $0 < p < 1$, addition $+_p$ is not associative.*

$$x +_p (y +_p z) \neq z +_p (y +_p x) \quad (5)$$

Note that, if $p = 1$, addition equals component-wise *AND*, while for $p = 0$, addition is component-wise *OR*. These operations are associative.

In the following, the state resulting from left-associative addition is denoted by \mathbf{L} , i.e. $\mathbf{L} = (x +_p y) + z$, while the state resulting from right-associative addition is denoted by $\mathbf{R} = x +_p (y +_p z)$. For the sake of readability, I will write $+ = +_p$ in the following, while assuming that $p = \frac{1}{2}$.

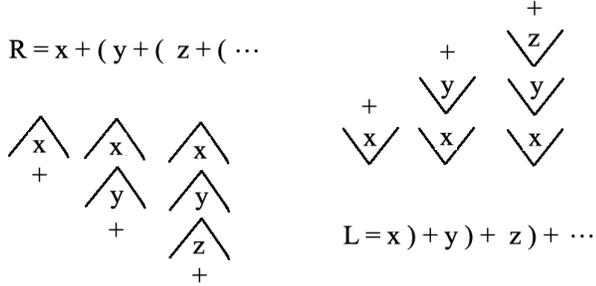


Figure 2: In right-associative \mathbf{R} addition, early items are kept prominent, while in left-associative \mathbf{L} addition, later items are superposed on earlier ones.

The states representing a memory list \mathbf{L} and \mathbf{R} states can be constructed for a list of any length. Construction starts from a pre-experimental state η and proceeds by iteratively adding items to the memory states \mathbf{L} and \mathbf{R} according to left-associative addition and right-associative addition to the respective branch as follows: For the \mathbf{L} -state

$$\begin{aligned}
 \mathbf{L}^0 &= \eta \\
 \mathbf{L}^a &= \eta + a \\
 \mathbf{L}^b &= \eta + a + b \\
 &\vdots \\
 \mathbf{L}^\Lambda &= \left(\left(\left(\left(\eta + a \right) + b \right) + c \right) + \dots \right) + f + g,
 \end{aligned}$$

while for the \mathbf{R} -state

$$\begin{aligned}
 \mathbf{R}^0 &= \eta \\
 \mathbf{R}^a &= \eta + a \\
 \mathbf{R}^b &= \eta + (a + b) \\
 &\vdots \\
 \mathbf{R}^\Lambda &= \left(\eta + \left(a + \left(b + \left(c + \left(\dots + \left(f + g \right) \right) \right) \right) \right) \right)
 \end{aligned}$$

After its sequential presentation, the memory list $\Lambda = (A, B, C, \dots)$ is thus represented by the two states

$$\mathbf{L} = \left(\left(\left(\left(\eta + a \right) + b \right) + c \right) + d \right) + f + g \quad (6)$$

$$\mathbf{R} = \left(\eta + \left(a + \left(b + \left(c + \left(d + \left(f + g \right) \right) \right) \right) \right) \right), \quad (7)$$

In (Murdock, 1982) η is assumed to be empty, while in (Franklin & Mewhort, 2015) it comprises a holographic collection of items and item-item associations. a, b, c, \dots are the cognitive states representing the physical list items A, B, C, \dots . These states preserve the serial order of items in

the memory list in that distances change monotonously along subsequent items, see Fig. 3

$$d(\eta, \mathbf{R}) < d(\eta, \mathbf{L}) \quad (8)$$

$$d(a, \mathbf{L}) > d(b, \mathbf{L}) > d(c, \mathbf{L}) > \dots \quad (9)$$

$$d(a, \mathbf{R}) < d(b, \mathbf{R}) < d(c, \mathbf{R}) < \dots \quad (10)$$

Correspondingly, both states inherit serial order in that item distances increase along \mathbf{R} , while they decrease along \mathbf{L} , see Fig 3. These distance gradients directly translate into activation gradients.

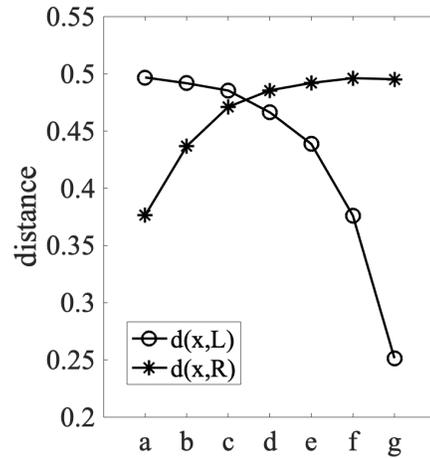


Figure 3: **Distance profiles** of the two states \mathbf{L} and \mathbf{R} as in eq 8 ff. \mathbf{L} has smallest distances to the most recent items, while \mathbf{R} is closest to the early list items.

Implied activity gradients

From the concept of similarity, two other concepts can be immediately derived: activation and memory strength. The intuition is closely related to the idea of a projection. Given that the memory state \mathbf{M} represents a memorised list, and that a cue item is presented. The cue item activates the memory state more, the more similar it is to that memory state (Hintzman, 1984). Conversely, the more the corresponding memory element is engraved in the memory state, the more the memory state is activated by the cue state.

Definition (Activation). Let \mathbf{M} be a memory state constructed during representing some memory list. A cue state x activates the memory state \mathbf{M} according to their similarity, see eq 2

$$\alpha_{\mathbf{M}}(x) := S(d(x, \mathbf{M})). \quad (11)$$

The activation gradient of \mathbf{M} is the vector $\alpha_{\mathbf{M}}$ with components $\alpha_{\mathbf{M}}(x)$, where x is a state representing a list item.

In terms of strength theory, $\alpha_{\mathbf{M}}(x)$ is the strength by which x is memorised in \mathbf{M} . One might also call $\alpha_{\mathbf{M}}(x)$ the familiarity of x given \mathbf{M} .

Consequently, the distance gradients in eq 8 ff directly translate into activity gradients, see Fig. 4. Since activation

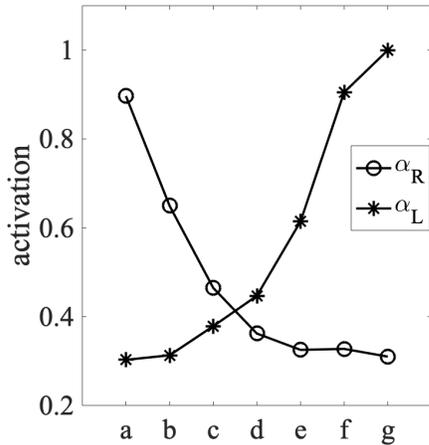


Figure 4: **Primacy and recency gradients** implied by the two states **L** and **R** are α_R and α_L .

as well as strength are increasing functions of similarity and hence decreasing functions of distance, **L** implies a *recency gradient* α_L , while **R** implies a *primacy gradient* α_R .

Activation gradients are nowadays widely accepted to play an important role in working memory. Various mechanisms have been discussed as sources of these gradients, see (Oberauer, 2003). In many models including TODAM, TCM and SOB, these gradients are separately modelled and superimposed on top of the model. In contrast, these gradients directly result from the bundling operation defined in eq 4 and its non-associativity: While non-associativity preserves information about serial order, right-associative addition and left-associative addition imply the primacy and the recency gradient, respectively.

The response function for recognition and recall

After presentation of a memory list, the participant has to fulfil some task. Most cognitive tasks involve cues such as cued item recognition or cued recall, associative or serial. The answer the participant gives is the result of a decision process which depends on both, the memory state as well as the cue. The response function in recognition only depends on familiarity, while the response function in recall additionally depends on distinctiveness (Murdock, 1982). Thus it is reasonable to make the response function a function of activation as defined in eq 11.

Definition (Response function). *The response function given a cue x facing the memory state \mathbf{M} is an increasing function of induced similarity, e.g.*

$$\Phi(x|\mathbf{M}) = \alpha_M(x) \quad (12)$$

Accordingly an activation gradient directly translates into a serial position curve. Particularly, the recency effect refers to the activation gradient of the **L**-state, while the primacy effect corresponds to the activation gradient of the **R**-state.

Experimental data indicate that the recency effect does not depend on list length and shows a slightly sigmoid curve shape, see Fig 5 (left). Both empirical observations are well captured by the modelling algebra proposed, see Fig 5 (right).

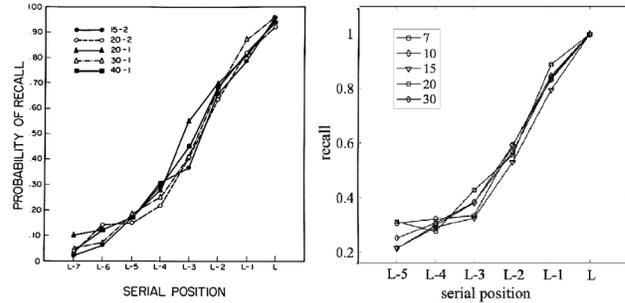


Figure 5: The recency effect does not depend on list length. Left: Experimental data from Murdock (Murdock, 1982), Right: Simulated data from the model for various list length'.

Application to some basic cognitive tasks

In this section some examples are presented to demonstrate how the formalism works, i.e. how to describe tasks such as cued recall in this formalism. Results are direct consequences of the algebra defined, i.e. no further assumptions are made. In the following only the **R**-state is concerned, i.e. states are bundled according to right-associative addition, while corresponding brackets are skipped for the sake of readability.

Repetition increases strength

It is intuitively expected that a repeated occurrence of an element in a list will increase its coding strength. This effect is indeed observed in the model. As a benchmark, consider the list $\Lambda = (A, B, C, D, \dots)$, in which all items are different. In $\Lambda^{(1)}$ a neighbouring pair is similar, e.g., $B \sim C$. In $\Lambda^{(2)}$, $B \sim D$ and so forth. k can be regarded as the lag from B until the similar item. Fig 6 shows the serial position curves for lists Λ , $\Lambda^{(1)}$, and $\Lambda^{(2)}$. Note that the coding strength of B is increased by any other item which is similar to B , while the strengthening is greater, the smaller the lag is, i.e. the effect of $C \sim B$ on the coding strength of B is larger than the effect of $D \sim B$.

Cued recall

Cued associative recall In this task, the participant is presented a paired memory list $(A - X, B - Y, C - Z, \dots)$. After memorizing this list, a memory item, i.e. a member of some pair, is presented as a cue, and the participant is asked to identify the memory item, which was bound to that cue item. The memory state corresponding to the paired list is

$$\mathbf{R} = \eta + a * x + b * y + c * z + \dots, \quad (13)$$

where $a * x$ is the state representing the binding between items X and A in the list.

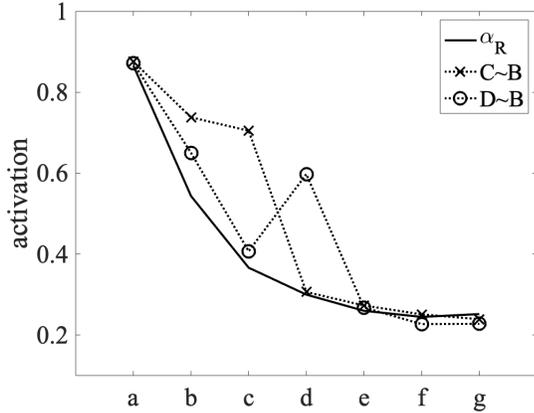


Figure 6: **The effect of similar items on activation** The solid black curve is the activation profile of state \mathbf{R} for the list Λ in which all items are different. Dotted lines are the profiles if that list contains one item, e.g. C or D , which is similar to item B .

When a memory item X is presented as a cue and the task is to retrieve the item which is bound to X in the list, consider the activation of

$$x * \mathbf{R} = x * \eta + a + a * b * y + a * c * z + \dots \quad (14)$$

The activation $\alpha_{x*\mathbf{R}}$ attains its maximal value for $\alpha_{x*\mathbf{R}}(a)$, see Fig. 7. Thus the cue X activates the A component most, so that the participant will answer " X is bound to A ", with some probability. Analogously, if the cue is Y , the activation $\alpha_{y*\mathbf{R}}$ attains its maximum in B , so that B is retrieved, and so forth. These maximal points form a curve, which is identical to the activation gradient $\alpha_{\mathbf{R}}$.

Retrieval from similar contexts Assume that the paired list $(A - X, B - Y, C - \tilde{X}, D - Z, \dots)$ is given, in which items A and C are bound to similar contexts X and \tilde{X} . The corresponding state yields

$$\mathbf{R} = \eta + a * x + b * y + c * \tilde{x} + \dots \quad (15)$$

Cueing with X will not only retrieve A but also C , just to a lesser extent. The effect of cueing with x is displayed when considering the activation gradient $\alpha_{x*\mathbf{R}}$, see Fig. 8. The gradient has two peaks, one at a and a weaker one at c , saying that cueing with X reveals two items, A and C . Cueing with Y uncovers only one, which is B .

In the recall task, the participant has to make a choice between the two alternative items bound to X . Thus invoking Luce's choice axiom, the probability to recall X yields

$$P(a|x) = \frac{\alpha_{x*\mathbf{R}}(a)}{\alpha_{x*\mathbf{R}}(a) + \alpha_{x*\mathbf{R}}(c)} \quad (16)$$

which is less than the probability to recall a without an alternative. The existence of an item similar to the cue impairs the corresponding recall.

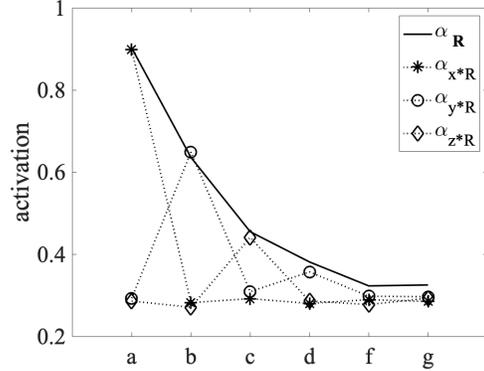


Figure 7: **Cued Recall**: Given cues such as x, y, z , the corresponding activations $\alpha_{x*\mathbf{R}}, \alpha_{y*\mathbf{R}}, \alpha_{z*\mathbf{R}}$ are plotted, see eq 14. The cue x causes the activation $\alpha_{x*\mathbf{R}}$ to have a peak at the corresponding item, which is a . $\alpha_{\mathbf{R}}$ is the activation profile of the \mathbf{R} -state.

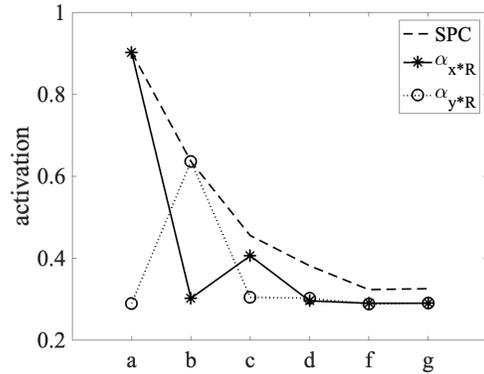


Figure 8: **Recall of items**: Activation profiles of $\alpha_{x*\mathbf{R}}$ (*), and $\alpha_{y*\mathbf{R}}$ (o). The activation profile $\alpha_{\mathbf{R}}$ of the \mathbf{R} state by distinct list items is shown as a reference.

Putting things together: The serial position curve

During memorizing a list, the two states \mathbf{R} and \mathbf{L} are constructed. Since there is no a priori reason to favour one over the other, I assume that both cognitive states \mathbf{L} and \mathbf{R} coexist and are the components of a *memory state* \mathbf{M} ,

$$\mathbf{M} = \begin{pmatrix} \mathbf{L} \\ \mathbf{R} \end{pmatrix}. \quad (17)$$

A single cue thus activates both components. The total activation of the memory state \mathbf{M} is a linear combination of the activation gradients of its two components.

$$\alpha_{\mathbf{M}} = \rho \mathbf{R} + \ell \mathbf{L} \quad (18)$$

where ρ and ℓ are non-negative parameters governing the mixture of respective activations. The response function to a cue is $\Phi(x|\alpha_{\mathbf{M}})$, so that the serial position curve is the graph $\Phi(x|\alpha_{\mathbf{M}})$, where x is a state representing a list item, see Fig. 9.

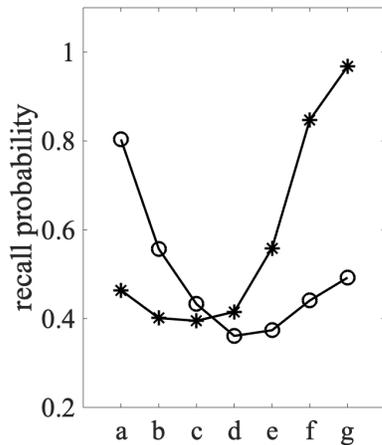


Figure 9: **The Serial Position Curve** is simulated for different pairs of parameters according to eq. 18. It shows a strong recency effect and a weak primacy effect for $\rho = 0.4, \lambda = 0.9$, while for $\rho = 0.9, \lambda = 0.4$ there is a strong primacy effect and a weak recency effect.

The serial position curve thus results from the linear combination of the primacy gradient α_R and the recency gradient α_L . As seen in Fig 9, a large ρ together with a small ℓ makes the recency effect, while a small ρ together with a large ℓ leads to a prominent primacy effect. The relative strength of the primacy and the recency effect will generally depend on the experimental set-up, including the task to be performed. For example in recognition and in cued (probed) recall the serial position curve shows a strong recency effect, while the primacy effect is weak. Strong primacy effects are seen in forward recall, while recency effects are strong in backward or in free recall.

Conclusion and out-look

In the previous sections, an elementary algebra (eq. 4) for storage and retrieval of information in basic cognitive tasks was proposed. The aim was not to present a full-blown model but to investigate how far one can get with the algebra alone.

Item information and associative information are represented by two operations, bundling and binding, respectively. If bundling is realised by an associative operation such as ordinary (vector-) addition, information about sequential order is lost. On the other hand, tasks such as serial recall require that order information. Consequently in corresponding models order information has to be implemented separately. This can be achieved by postulating serial position markers, chaining by associative mechanisms between consecutive items, or weight functions varying over serial position governing the recency and the primacy effect.

This is different in the approach presented: Information about sequential ordering is preserved. This is due to the non-associativity of the addition operation by which item in-

formation is bundled into a memory state. Reading from that state thus reveals order information necessary to related tasks, which is represented by corresponding gradients. Activation gradients are implied rather than postulated separately. The serial position curve comes as a linear combination of both. Its shape is concave-up and asymmetric as observed as a typical experimental finding, see Fig 1 for experimental data and Fig 9 for simulations of our model.

As already mentioned, the aim was not to present a full-blown model but to investigate how far one can get with the algebra alone. So it does not come as a surprise that several experimental observations were not captured. For example, while the recency effect does not depend on list length, the primacy effect does. This robust finding cannot be explained by our algebra alone but needs an additional assumption about attention, which then imposes an additional constraint on den attention gradient. Furthermore, serial recall can not be described by our algebra alone but needs an additional assumption such as output-suppression, as supposed in many models, or an other feedback mechanism, see (Franklin & Mewhort, 2015).

The cognitive algebra proposed appears to provide a reasonable basis for modelling since it generically implies several features that fit empirical observations quite well, in a qualitative sense in that no attempt was made to fit data. Modelling then could consist in carefully adding assumptions on top of the cognitive algebra such as discussed above.

Acknowledgement

I would like to thank the anonymous referee for value input and comments.

References

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic bulletin & review*, 24(4), 1158–1170.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic bulletin & review*, 9(1), 59–79.

Franklin, D. R., & Mewhort, D. (2015). Memory as a hologram: An analysis of learning and recall. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 69(1), 115.

Hintzman, D. L. (1984). Minerva 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, 16(2), 96–101.

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, 46(3), 269–299.

Kahana, M. J. (2012). *Foundations of human memory*. OUP USA.

Kanerva, P. (2009). Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors. *Cognitive computation*, 1(2), 139–159.

- Murdock, B. B. (1974). *Human memory: Theory and data*. Lawrence Erlbaum.
- Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, 89(6), 609.
- Murdock, B. B. (1993). Todam2: a model for the storage and retrieval of item, associative, and serial-order information. *Psychological review*, 100(2), 183.
- Oberauer, K. (2003). Understanding serial position curves in short-term recognition and recall. *Journal of Memory and Language*, 49(4), 469–483.
- Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D., Conway, A., Cowan, N., ... others (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, 144(9), 885.
- Plate, T. (1991). Holographic reduced representations: Convolution algebra for compositional distributed representations. In *Ijcai* (pp. 30–35).
- Schlegel, K., Neubert, P., & Protzel, P. (2020). A comparison of vector symbolic architectures. *arXiv preprint arXiv:2001.11797*.
- Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: Rem—retrieving effectively from memory. *Psychonomic bulletin & review*, 4(2), 145–166.

Estimating Individual Differences in Working Memory through ACT-R Modeling and Resting State Connectivity

Patrick Rice (pjrice@uw.edu)

University of Washington Department of Psychology, Campus Box 351525 Seattle, WA 98195 USA

Andrea Stocco (stocco@uw.edu)

University of Washington Department of Psychology & Institute for Learning and Brain Sciences, Campus Box 351525 Seattle, WA 98195 USA

Abstract

A complete and holistic understanding of human cognition should be able to relate idiographic parameters representing cognitive functioning to interactions between connected brain networks identified by neuroimaging methods. Here, using the ACT-R cognitive architecture, we examine the possibility of producing idiographic parameterizations of cognitive functioning in a task environment and show that these parameterizations produce reasonable predictions of individual behavior. We then demonstrate a method of determining a subset of parameters that are adequate for prediction of behavior before confirming that the most critical of these task-based parameters is related to functional connectivity measures in individual resting-state fMRI data.

Keywords: Cognitive Modeling; Long-Term Memory; Resting-state fMRI; Functional Network

Introduction

One of the advantages of the utilization of computational models in the study of cognition is the possibility to estimate parameters that characterize behavior and/or cognitive performance on a per-individual, or *idiographic*, level. For example, reinforcement learning (RL) models can be fit to behavioral data, and the resulting parameter estimates can be used to make inferences about individual differences in dopamine function or to distinguish between healthy and pathological groups (Frank et al. 2004). Similar work has been performed with drift diffusion models (DDM), which model decision-making through a noisy information accumulation process that “drift” towards one of two decision boundaries. In addition to being more clearly interpretable than raw behavioral data, parameters inferred through DDM are often more reliable in detecting individual differences than behavioral metrics (White et al. 2016). In the past, ACT-R models have been used to make such inferences as well. For example, Daily et al. (2001) estimated goal spreading activation from behavioral data, used it as a proxy for working memory, and successfully predicted performance on a different task.

This individual-difference approach, however, has not been applied consistently - instead, the majority of modeling efforts have focused on fitting parameter values that are descriptive at the group level. Furthermore, ACT-R is a far more complex computational framework than RL or DDM, and it encompasses dozens of parameters. While this complexity makes it possible to capture complex tasks

that lay outside the scope of RL or DDM models, it also poses some significant challenges: is it possible to identify idiographic parameter values that reliably characterize the behavior of a given individual? How many parameters are needed to characterize individual differences within a group? How can each parameter’s contribution to predicting these differences between individuals be determined?

Here, we provide an empirical answer to this question. We created a model of the zero-back condition of the standard n-back working memory task, and then fit the model to behavioral data from ~150 participants. We show it is possible to use convex optimization techniques to identify points in multidimensional parameter space that accurately capture an individual’s performance. We then provide a method to determine which estimated parameters contribute most meaningfully to the prediction of individual performance. Furthermore, we demonstrate that these idiographic parameterizations are predicted by the individual’s resting-state functional connectivity, indicating that the parameterization captures fundamental aspects of individual cognitive function.

Materials and Methods

The study presented herein consists of an analysis of $N=178$ individuals from the Human Connectome Project, the largest existing repository of young adult neuroimaging data. The analysis was restricted to the resting fMRI subset in conjunction with the zero-back condition of the “Working Memory” (WM) task component. The resting fMRI data collection consisted of two 30-min recording sessions, performed 24 hours apart; the task fMRI data collection consisted of two 30-min task sessions performed directly after each resting-state acquisition session. During each task session, participants performed six other tasks in addition to the WM component, per the HCP protocol. All subject recruitment procedures and informed consent forms were approved by the Washington University in St. Louis’ Institutional Review Board. The present study met criteria for exemption at the University of Washington’s Institutional Review Board.

Task Data

Each working memory task session consisted of four 0-back blocks, with each block containing 10 trials. Each block begins with a 2.5 s cue that informs the participant of the target stimulus for the proceeding block of trials. Each trial presents a single image centered on the screen, and participants are required to indicate if the trial’s stimulus is identical to the cue stimulus by pressing one of two buttons. The stimuli belong to one of four possible categories: faces, places, tools, and body parts. These categories were presented in a block-wise fashion such that two of the eight blocks presented a given category. Each trial stimulus is presented for 2 s with a 500 ms ITI, for a total duration of 27.5 s per block. Additionally, 15 s fixation blocks were presented after the second and fourth task blocks within a session. This paradigm produces stimuli of three conditions: targets (match to the block cue); lures (non-targets that have been presented at least once before within the block); and non-target, non-lures (non-targets that are presented for the first time within a block).

fMRI Image Acquisition and Preprocessing

Functional neuroimages were acquired with a 32-channel head coil on a 3T Siemens Skyra with TR = 720 ms, TE = 33.1 ms, FA = 52°, FOV = 208 × 180 mm. Each image consisted of 72 2.0 mm oblique slices with 0-mm gap in-between. Each slice had an in-plane resolution of 2.0 × 2.0 mm. Images were acquired with a multi-band acceleration factor of 8X.

Images were acquired in the “minimally preprocessed” format (Van Essen et al., 2013), which includes unwarping to correct for magnetic field distortion, motion realignment, and normalization to the MNI template. The images were then smoothed with an isotropic 8.0 mm FWHM Gaussian kernel.

ACT-R Modeling of WM Task Data

An ACT-R task device and model were developed in order to characterize individual behavior in the zero-back task. The task device implements the zero-back task by updating the ACT-R visicon with a representation of the task elements in the form of three strings: one that identifies the category of the stimulus, a second representing the stimulus itself, and a third indicating the “kind” of the stimulus - either a block-cue or a trial-stimulus. The model automatically attends to this information before transferring the chunk representation of the display to the imaginal buffer. In the case of a cue, the model updates the goal buffer to represent that the target of future retrieval requests is the block-cue “kind”, and then waits until a new visual display - automatic buffer harvesting ensures that the chunk representing the cue is entered into declarative memory. In the case of a stimulus, after the chunk representation is loaded into the imaginal buffer, the model attempts to retrieve a chunk to compare against the stimulus by making a retrieval request

specifying the category of the stimulus and the block-cue “kind”. If retrieval is successful, the model proceeds to determine if the stimulus identity represented by the chunks in the imaginal and retrieval buffers are matched. If so, it responds that the current stimulus is a target; otherwise, nontarget. In some cases, the retrieval process may not complete before the trial ends. If so, the model detects the presentation of the ITI and interrupts the ongoing retrieval attempt through a secondary retrieval request. A flowchart depicting the strategy of the model can be found in Figure 1.

The behavior of the model in Figure 1 ultimately depends on the parameters that influence memory retrieval. In ACT-R, retrieval is affected by a memory’s *activation*, $A(m)$, which is the sum of a base-level term $B(m)$ and a contextual spreading activation $S(m)$. $B(m)$ is the log sum of the decaying traces of previous uses of m :

$$B(m) = \log \sum_i t_i^{-d}$$

where t_i is the time elapsed from the i -th time m was used and d is the decay rate. The spreading activation is defined as an additional boost coming from the information stored in a buffer:

$$S(m) = \sum_b \sum_j (W_b/N) s_{i,m}$$

where W_b is the amount of activation spreading from buffer b and $s_{i,m}$ is the association between slot i in buffer b and memory m . In our model, two such sources of activation exist, one for the goal buffer W_g and one for the imaginal buffer W_i . The strength of association $s_{i,m}$ is computed through a function which returns a scalar integer value equal to the number of source chunks j contained in chunk i ; 1 if chunks j and i are identical; and 0 otherwise. Task accuracy depends on both the availability of a memory and the probability of unintentionally retrieving a wrong item; the latter is controlled by a partial matching similarity parameter c that determines the penalty between two slots. Thus, chunks that do not match the retrieval specification are penalized, but can still be retrieved.

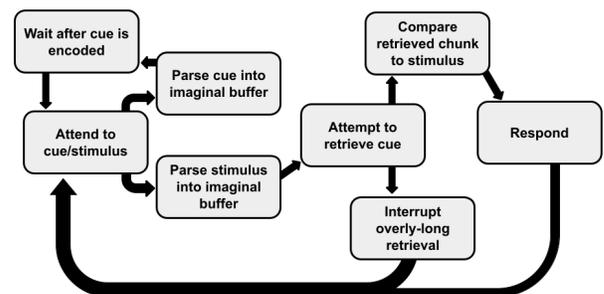


Figure 1. Flowchart of the ACT-R model strategy for performing the zero-back condition of the n-back task.

Finally, the relationship between the activation of a memory and the time RT it takes to retrieve is given by the equation:

$$RT(m) = Fe^{A(m)}$$

In summary, the dynamics of the model depend on five parameters: d , W_i , W_g , F , and c .

Individual-specific Estimation of ACT-R Parameters

Five model parameters were fit to individual participant behavior on an idiographic basis: d , W_i , W_g , F , and c . These parameters were chosen as they each have a strong effect on the model's response time or accuracy, the two participant measures that the model was fit against. The parameters d , W_i , W_g , and c all influence the likelihood that the correct cue-chunk is retrieved to be compared against (with additional minor influence on the RT due to changes in the activation level of the retrieved chunk), while the F parameter largely affects the retrieval time of the retrieved chunk (and therefore, the response time of the model). Parameters were only estimated for participants who demonstrated greater-than-chance performance on the target, lure, and non-target/non-lure conditions. To perform the fitting process, the *optimize.minimize()* method of the Python *scipy* package was utilized to minimize the RMSE between a set of participant measures and the commensurate model measures through the minimization function's *Powell* method. Bounds were placed on the five parameters (c : (-1,1); W_g : (0,2); W_i : (0,2); F : (1,3.5); d : (0.2,0.8)) to ensure that the minimization function remained within either reasonable or required ranges for these parameters. To compute a single RMSE across both RT and accuracy, these measures were placed on the same scale by dividing the trial-by-trial model and participant response times by 2 (as the maximum allowable RT by the task was 2 s). Missing RTs for both model and participant were replaced with the corresponding nan-measured RT. Additionally, as the binary trial-by-trial accuracy outcomes had the potential to be exceedingly punishing to the model-fitting process, the aggregate block-wise and condition-wise (target/lure/non-target, non-lure) accuracies were used instead. Model and participant trial-by-trial scaled RTs and block-wise/condition-wise accuracies were then vectorized in order to compute the RMSE. Once the minimization algorithm converged to parameter estimates for each participant, model predictions were produced by running the model 100 times for each set of participant parameters, and then first taking the trial-by-trial average of the predicted RTs and accuracies over these runs before determining the average RT and accuracy for each participant.

Evaluation of Parameter Estimates

To evaluate the relative importance of each of the estimated parameters to the predictive efficacy of the model, a "decremental leave-one-out" (dLOO) procedure

was applied. In this procedure, a set of models utilizing a subset of the estimated parameters are first produced from the full parameter set n by applying n choose k , where $k = n-1$. For each participant and each model in this set, the k chosen parameters are set to the participant's estimated values, while the "left out" parameter is set to the mean of that parameter's estimates (across participants). Model predictions are produced for each of the models in this set (as described above), and the R^2 between model predictions and participant measures are determined for both RTs and accuracies. The model with the largest mean R^2 (across RTs and accuracies) is determined to be the "best-fitting" model in this set, and the parameter that was "left out" of this model is "decremented" from the set of parameters. This procedure is then repeated for the remaining parameters, with both the "left-out" parameters and the "decremented" parameters set to the mean of that parameter's estimates, until only a single parameter remains.

Brain Parcellation

To calculate functional connectivity, each participant's brain was divided into discrete regions using a parcellation proposed by Power et al (2011). Although other parcellations have been proposed, this parcellation is notable for including both cortical and subcortical regions (see also Cole et al., 2016).

Statistical Learning Model

To identify the optimal combination of functional connectivity measures that reliably predicts individual parameters, resting-state functional connectivity was analyzed using a Lasso regression, a statistical learning method that combines feature selection and parameter fitting (Tibishirani, 1996). As a variant of linear regression, Lasso results remain interpretable in terms of beta weights that linearly scale a set of regressors. Unlike linear regression, Lasso reduces the complexity of the model by adding a penalty term that reduces to zero the weight of unnecessary variables, dramatically reducing the number of regressors provided. This feature is crucial for high-dimensional data such as the set of connectomes associated with a group of participants.

While in canonical linear regression the weights β are obtained by minimizing the quantity $\|\mathbf{y} - \beta\mathbf{X}\|_2$ (where the notation $\|\mathbf{v}\|_n$ represents the $L(n)$ norm of a vector \mathbf{v}), in Lasso the quantity to minimize includes a penalty term:

$$\beta = \operatorname{argmin}(\|\mathbf{y} - \beta\mathbf{X}\|_2 + \lambda\|\beta\|_1)$$

The value of λ represents the tradeoff between model simplicity (captured by the first-order $\|\beta\|_1$ penalty) and accuracy (captured by ordinary least squares minimization term $\|\mathbf{y} - \beta\mathbf{X}\|_2$). When $\lambda = 0$, the model reduces to canonical linear regression. As λ grows, however, more and more regressors are eliminated to satisfy the constraints.

Results

Participant Task Performance

Participants who did not achieve greater-than-chance performance (binomial test) on the target, lure, or non-target/non-lure condition were not included in the analysis. This resulted in the exclusion of 36 out of 178 participants. Mean response times and accuracies were calculated for each participant for whom parameters were estimated. Mean RT was 0.78 ± 0.10 s, while mean accuracy was 0.94 ± 0.06 . The distribution of RTs and accuracies across participants can be seen in Figure 2.

Idiographic Parameter Estimation and Prediction

For each participant, the set of parameter values that minimized the RMSE between trial-by-trial RTs and block-wise/condition-wise accuracies were estimated. The estimated cue-stimulus similarities c had a mean value of -0.43 ± 0.22 , with a range of $(-0.88, 0.19)$. While the majority of participants were found to have a negative c value, the c value for 2% of participants was estimated as slightly positive. The goal buffer spreading activation value W_g estimates had a mean of 0.93 ± 0.33 and a range of $(0.18, 1.74)$, while the imaginal buffer spreading activation value W_i had a mean of 0.68 ± 0.34 and a range of $(0.06, 1.71)$. As W_g and W_i provide a complementary but opposing influence on the retrieval process in this model (except in the case of target stimuli, for which they both promote the retrieval of the correct cue chunk), the difference between these two parameter estimates ($W_g - W_i$) was examined. The mean difference was 0.24 ± 0.51 , with a range of $(-0.85, 1.38)$. Over 70% of participants were estimated to have a W_g value greater than their W_i value, indicating that overall, information in the goal buffer drove the retrieval process. The mean of the latency factor F estimates was 2.49 ± 0.42 , with a range of $(1.31, 3.29)$.

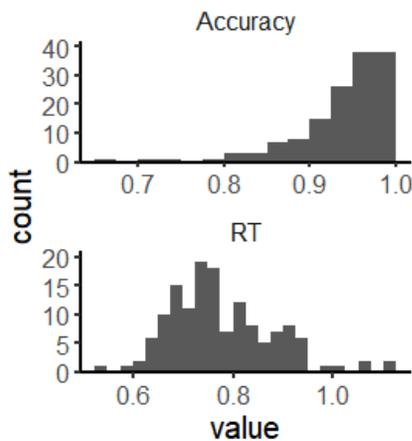


Figure 2. Histograms of accuracies and response times across participants in the zero-back condition of the HCP n-back task.

While F linearly affects retrieval times (and, by extension, response time) and the mean of the estimates was greater than the maximum allowable response time, the magnitude of this parameter compensates for retrieval time speeding caused by the influence of spreading activation and partial matching. Finally, for the decay-rate parameter d , the mean estimated value was 0.53 ± 0.10 , and the range was $(0.29, 0.71)$.

Once parameters for each participant were estimated, model predictions of participant performance were produced. Across predicted participants, the model's mean RT was 0.64 ± 0.09 , and the model's mean accuracy was 0.89 ± 0.08 . Individual participant RTs and predicted RTs were strongly correlated ($r = 0.56, p < 0.001$), while participant accuracies and predicted accuracies were moderately correlated ($r = 0.23, p < 0.01$). Scatterplots of participant measures versus predicted measures can be seen in Figure 3.

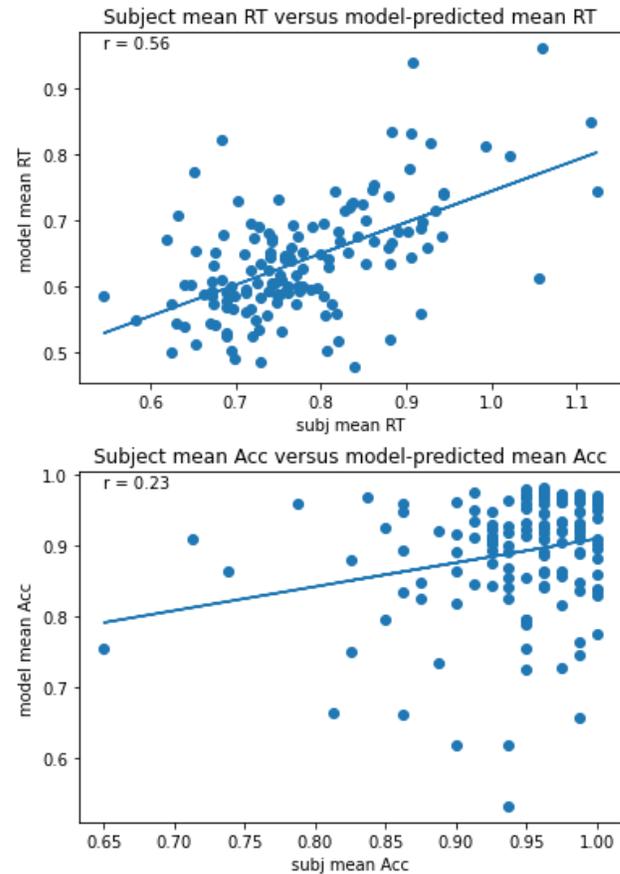


Figure 3. Scatterplots of participant mean RTs/accuracies versus model-predicted mean RTs/accuracies. Pearson's r between participant measures and model predictions shown in the upper left.

Decremental Leave-one-out Procedure

To determine which parameters contributed most strongly to the model’s ability to predict individual participant RTs and accuracies, the parameter set was subjected to a “decremental leave-one-out” procedure. In the first round (five sets of four out of five parameters included, one parameter in each set assigned to the mean estimated value), it was found that the model containing the individual predictions of W_g , W_i , F , and d parameters had the largest mean R^2 (mean $R^2 = 0.21$; RT $R^2 = 0.32$; accuracy $R^2 = 0.11$); consequently, the c parameter was “decremented”. In the second round, the model containing the W_g , W_i , and F parameters was the strongest predictor of participant behavior (mean $R^2 = 0.21$; RT $R^2 = 0.33$; accuracy $R^2 = 0.09$); the d parameter was dropped. In the third round, the model that included the W_g and W_i parameters was the most successful (mean $R^2 = 0.18$; RT $R^2 = 0.27$; accuracy $R^2 = 0.09$), and in the final round, the model including only the W_g parameter was the most predictive (mean $R^2 = 0.18$; RT $R^2 = 0.24$; accuracy $R^2 = 0.11$).

rs-fMRI Prediction of Individual-specific W_g

For each individual, we extracted a matrix of functional connectivity by calculating the Pearson correlation coefficient of each pair of the 264 x 264 regions in the Power (2011) parcellation. The group-level average of the individual correlation matrices, known as the *connectome*, was then visually inspected for comparison with similar functional connectivity studies. The connectome was found to be consistent with previous findings using the same parcellation scheme (compare, for example, to Cole et al. 2016). Because correlations between pairs of regions tend to be partially driven by common, unobserved factors (such as motion and physiological noise), the matrices were re-calculated using partial correlations (Cole et al., 2016), so that correlations between each region in the pair and the remaining 262 regions were partialled out. The resulting mean connectome is a much more sparse matrix (Figure 4B) and includes both negative and positive correlations (as expected from the spontaneous dynamics of brain activity: Fox et al., 2005).

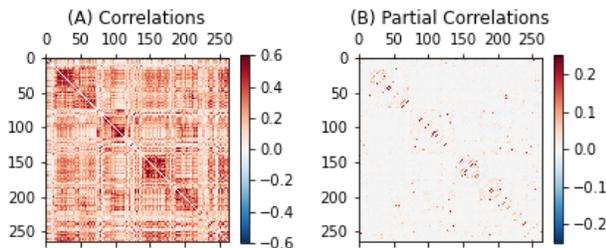


Figure 4: (A) Raw correlations between each of the 264 regions ; (B) Partial correlations between the same regions. In each matrix, rows and columns are ordered by network.

Each participant’s sparse correlation matrix was then reshaped into a row vector of $(264 \times 263) / 2 - 264 = 34,452$ elements. The number of possible regressors was further reduced by excluding connectivity measures related to three irrelevant networks (the Auditory, Cerebellar, and “Uncertain” networks in Power et al., 2011).

Lasso Fit and Cross-Validation

A cross-validation procedure was used to find the optimal value of λ . A sequence of possible λ values was generated, and, for each value, the performance of the Lasso algorithm in predicting the parameter W_g on a per-participant basis was measured using leave-one-out validation (LOOV). In LOOV, the algorithm is run 142 times, each time leaving out a different participant as the test set while the β values are fit to the remaining 141 participants as the training set. The mean error in predicting the parameter W_g for the left-out participant was then measured for all values of λ , and the value of λ that produced the smallest cross-validation error across all participants was chosen.

Resulting Connectivity

At the optimal level of λ , only 19 functional connections were left with a $\beta > 0$, involving a total of 36 brain regions from eight different functional networks. These connections and their regions are shown in Figure 5.

Notably, this list of regions includes the four ROIs in the Power parcellation that span the anterior cingulate cortex (ACC), corresponding to ACT-R’s goal buffer (Anderson, et al. 2008). The list also includes five regions in the salience network, a set of regions involved, like the ACC, in the top-down control of attention. The functional connectivity values that best predict individual values of the W_g parameter include connections between the salience network and the default mode network, which is known to correlate with long-term memory function, and the sensorimotor network, including motor regions corresponding to the right hand.

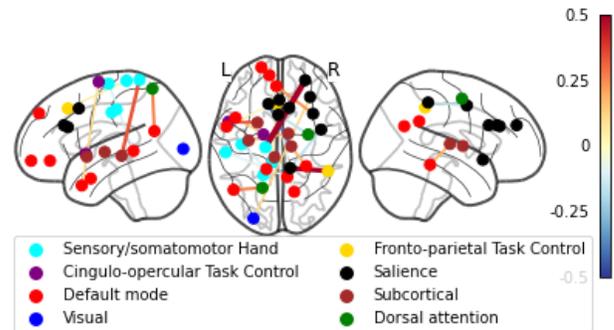


Figure 5: Functional brain connections predictive of individual rates of forgetting. Colored edges between nodes represent functional connectivity between the connected regions, while colors of nodes represent the network each region belongs to, using the Power et al (2011) scheme.

Together, connectivity between a group of networks including the salience network, the default mode network, and the sensorimotor network comprised a majority of the identified connections (16 out of 19, $\chi^2(1) = 15.474$, $p < 0.001$) and regions (24 out of 36, $\chi^2(1) = 10.667$, $p < 0.005$), significantly greater than what could be expected by chance.

As a final examination of this connectivity, we determined how well the parameter W_g can be recovered from functional connectivity alone. To do so, we multiplied each individual-specific set of functional connectivity values by the beta weights produced by Lasso, and compared them to the values inferred from the behavioral data by the ACT-R model. The predicted and observed values had a correlation of $r(142) = 0.775$, $p < 0.001$ (Figure 6).

Discussion

In the present study, idiographic parameterization of working memory function was investigated through the application of an ACT-R model. Values of five different parameters capturing various aspects of cognitive functioning were estimated for each participant through minimization of the RMSE between participant behavior and parameterized model predictions. A rank-ordering of the importance of these five parameters to the predictive efficacy of the model was determined through a “decremental leave-one-out” procedure, demonstrating that the goal-buffer spreading activation parameter W_g was critical to the model’s predictive ability. Furthermore, it was shown that this essential parameter is predicted by an individual’s resting-state functional connectivity.

This work makes it clear that ACT-R parameter estimates are capable of producing quality predictions regarding individual-level behavior. The correlation between participant RTs and model-predicted RTs was strong; while the predictions of accuracy were somewhat weaker, this can be partially attributed to the fact that the model was fit to the block-wise and condition-wise accuracies, instead of trial-by-trial accuracies (as RTs were). While this approach avoids the overly-punishing nature of RMSEs computed on binary outcomes, it reduces the amount of information the minimization algorithm has to fit the participant accuracy, relative to the participant RT. This effect was also apparent in the “decremental leave-one-out” procedure; the RT R^2 measures were overall larger than the accuracy R^2 measures across the iterations of the procedure. This procedure rank-ordered the “importance” of the five parameters to the model’s predictive efficacy ($c < d < F < W_i < W_g$), and made it clear that W_g was by far the most valuable parameter for this model to predict individual differences in behavior, as both RT and accuracy R^2 measures changed negligibly as the other parameters were set to the mean values. As W_g was indicated to be the most crucial for individual prediction, it was chosen to be examined in relation to the participant’s resting-state fMRI data.

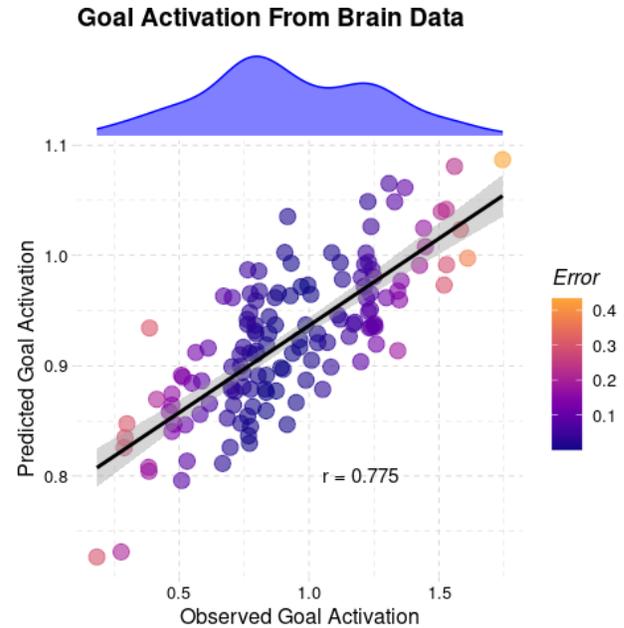


Figure 6: Correlation between observed values of the W_g parameter and the values of W_g predicted from functional connectivity (Figure 5).

A functional connectome for each participant was generated and subsequently used as the set of predictors for W_g in a Lasso regression. This resulted in the identification of a set of functional connections between regions inclusive of the salience, default mode, and sensorimotor networks as being maximally predictive of individual values of W_g . Moreover, this particular set of functional connections is entirely compatible with the putative role played by the goal buffer’s spreading activation in the model, where it is used to assist in the retrieval of the correct cue from *long-term memory* (compatible with the salience-default mode connections). The result of this work would allow for the prediction of individual-specific W_g parameter values on individuals for whom resting-state measures exist, and through the ACT-R model, prediction of their behavior in a task environment.

In conclusion, this work exemplifies the potential of utilizing ACT-R modeling in conjunction with neuroimaging measures for the identification and prediction of signatures of cognitive functioning on an individual basis. Potential future efforts in this area of work include identification of a maximally predictive subset of parameters for each individual, as well as determination of resting-state nodes and functional connections that allow for the prediction of these parameters.

Acknowledgements

This research was supported by grant FA9550-19-1-0299 from the Air Force Office of Scientific Research to AS.

References

- Anderson, J. R., Fincham, J. M., Qin, Y., & Stocco, A. (2008). A central circuit of the mind. *Trends in Cognitive Sciences*, *12*(4), 136-143.
- Cole, M. W., Ito, T., Bassett, D. S., & Schultz, D. H. (2016). Activity flow over resting-state networks shapes cognitive task activations. *Nature Neuroscience*, *19*(12), 1718-1726.
- Daily, L. Z., Lovett, M. C., & Reder, L. M. (2001). Modeling individual differences in working memory performance: A source activation account. *Cognitive Science*, *25*(3), 315-353.
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, *102*(27), 9673-9678.
- Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science*, *306*(5703), 1940-1943.
- Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., ... & Petersen, S. E. (2011). Functional network organization of the human brain. *Neuron*, *72*(4), 665-678.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267-288.
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., & Wu-Minn HCP Consortium. (2013). The WU-Minn human connectome project: an overview. *Neuroimage*, *80*, 62-79.
- White, C. N., Curl, R. A., & Sloane, J. F. (2016). Using decision models to enhance investigations of individual differences in cognitive neuroscience. *Frontiers in Psychology*, *7*, 81.

Lessons Learned From Modelling Situated Cognitive Agents Interacting With a Dynamic Environment

Lukas Seiling*, Maximilian Plitt*, Paul Schweidler*, Oliver Klückmann*, Lilian Befort, Robin Konczir, Severin Reuter, Kai Preuss, Sebastian Wiese and Nele Russwinkel

Department of Psychology and Ergonomics, Marchstraße 23
10587 Berlin, Germany

* Authors contributed equally

Abstract

The study of knowledge representations and reasoning problems faced by a cognitive agent interacting with a dynamic and incompletely known world is relevant to cognitive robotics and understanding complex cognition and related fields. The paper introduces four cognitive agents that were modeled in a student project with specific requirements. The cognitive architecture ACT-R was used to model flexible agents that interact with objects in a grid field with only a limited field of view. Long-term planning is not possible here: the meaning of objects needs to be discovered and the field explored to find the goal as quickly as possible. The project demonstrates how the four agents learn from interactions and what information needs to be kept available to flexibly decide in unpredictably occurring situations. All four agents are shortly described in more detail. The project covers on a small scale some aspects that are crucial for autonomous agents in a simple game environment. The four agents are faced with 15 challenge environments that need to be explored and managed. The challenge performance results show that a higher number of productions does not necessarily lead to better performance.

Keywords: mental representation; situated cognition; embodied cognition; situated agent; interactive learning

Introduction

Everyday we cope with new challenges, experience time pressure when we try to complete our tasks and need to flexibly react to changes and new information. Normally we do not have time to evaluate different options, but have to instantly find a good solution and act accordingly. Therefore, in such situations our cognitive system not just relies on independent, well-elaborated rational processes, but often depends on the situation and context in which cognition occurs. Embodied cognition holds that cognitive processes are deeply rooted in the body's interactions with the world (e.g. Wilson, 2002). The research field of embodied cognition is still widely heterogeneous, but there are some distinct characteristics of embodied cognition most researchers agree upon. Embodied cognition is situated, cognition is time-constrained, and whenever possible we offload cognitive work onto the environment. According to Levesque and Reiter (1998), cognitive robotics is the study of knowledge representation and reasoning problems faced by an autonomous robot (or agent) in a dynamic and incompletely known world. This leads to the research question what cognitive mechanisms are central to build a cognitive agent that is able to cope with such an environment and is still to accomplish its goals. Situated agents in dynamic environments are good test beds to investigate different implementations of such mechanisms because

of their required abilities to flexibly manage changes in the environment, explore unknown objects and to handle novel challenges. In the long run this kind of research is relevant to learn more about complex cognition. According to (Funke, 2010) complex cognition deals with all mental processes that are used by an individual for deriving new out of given information, with the intention to make decisions, solve problems, and plan actions. This assumes an active and goal-directed information processing by an agent that is able to perceive its environment and to use its memory. In a complex situation, the result is more than the sum of perceptual, learning, and memory processes. In this sense perception can be seen as part of a higher structure. The context delivers the meaning which is not only given by itself but in combination with other events and objects. In addition this kind of research has the potential to develop good solutions for cognitive robotics or human-robot collaboration.

The aim of this paper is to explore these questions within a simple task environment. We want to show four different realizations of such an agent for the same requirements and task environments within the cognitive architecture ACT-R (Anderson, 2000). These examples can support other researchers faced with similar task requirements to reason about ways to build such an agent. Usually each modeler starts with their own idea, therefore such model challenges are useful to explore a wider field of possible implementations and an evaluation thereof. Furthermore, we hope to contribute information to the question of how situated cognition can be realized with a cognitive architecture and what possible architectural developments are promising in order to address such research fields. In a student project of about 2 months, the given task was to develop a situated agent that should use mainly cognitive plausible mechanisms to deal with several challenges.

Agent requirements

The game (a grid field with several colored object on its tiles, see Figure 1) required the interactive agent to

1. find out what object on the grid represents the agent and the color it has
2. search for the goal that has to be approached. The goal is not visible at the beginning

3. find out what object with what color shows what kind of effect on contact (obstacle, add points, deduct points), in other words to explore the environment and infer the best way towards the goal
4. cope with constantly appearing and disappearing objects (due to a *fog of war* mechanism, only objects in immediate surroundings are shown) and make decisions based on a mental representation of the environment.

The goal is to move onto the goal tile as fast as possible, preferably with a high score. At the end of the course, 15 novel challenge fields were provided and the agents were tested in order to explore and evaluate their respective performance and flexibility in unknown environments. Prior knowledge for all agents was:

- the color of the goal is green,
- the agent starts on the top row (color is unknown),
- the goal is below the agent, in the lower half of the grid,
- possible movements are left, right, up, down; these are restrained by a yellow bounding box ,
- movements towards a colored tile can have three different consequences (blocked movement, pass and win points, and lose points),
- the effect of object colors are randomized each trial, except for the green goal object.

The main cognitive skills required are therefore to learn about the environment through interactions, to explore the grid in order to find the goal, to draw inferences, to gather information and hold this information in mental representations and to make decisions based on available information. In the following sections, the different cognitive agents will be introduced, and for each agent it will be explained how and where different aspects of information are gathered and represented (for instance by chunk representations in some buffer, e.g. goal or imaginal; or production rules) and how this knowledge is used in specific situations. Then it will be sketched out how the four requirements mentioned above are realized for the different agents. Lastly the main benefits and weaknesses of each agent are discussed and performance in some situations is described.

Some requirements are realized similarly across the agents, such as identifying itself through an action and checking for the object that moved. Since it is known that the agent starts in the top row, all visible objects are encoded and movement is initiated. When a change is registered by the visual module, the color of the moved object is stored as a mental representation in either the goal or imaginal buffer. Most agents also initiate object tracking via the visual module to keep their agent representations in focus. To facilitate self-localization and an understanding of the grid's dimensions, the agents make use

of geometric data of the objects and borders for simple heuristics, such as moving towards the center of the grid. This data is also stored as part of the agent's mental representation of the task.

Models

Speedy

Orients itself according to sub-goals. Bonus points are collected when close. Acts pre-attentively and therefore quickly.

Relevant information for the agent is stored in the imaginal buffer including its current position, movement intention as well as its specific color. At a later point, the colors of additional tiles are stored in the imaginal buffer according to their meaning. After successful self-identification the agent starts searching for the goal. To reach the goal tile as quickly as possible, a strategy of subgoals is pursued. Subgoals represent specific waypoints the agent tries to reach. Information regarding the agent's current subgoal is stored in the goal buffer. The goal buffer also contains the minimal and maximal x and y coordinates of the grid field, representing the borders and the distance to the subgoal. Since the goal object is located in the lower part of the grid field, the agents tries to reach its first subgoal, which is directly in the middle of the grid field in order to explore the space where the goal could be located.

Searching the environment First, the agent routine searches its visible field for the green goal and follows its subgoal. The adjacent tile in the agents movement direction is checked. In case of an object of unknown color, the object gets evaluated. Otherwise a movement according to the object's meaning is executed. An unknown object is tested by the agent moving onto it. Object meanings are evaluated by using the visual-location module that searches for an appearing text (score). In case a red text appears, it is examined for "+" and "-" signs and accordingly the bonus or malus chunk stored in the imaginal buffer is filled. In case no text appears, the color is inferred to be an obstacle. For obstacles or malus objects, movement direction is changed. Bonus tiles are collected whenever in the vicinity.

Locating the goal As soon as the agent gets close to a subgoal location, the subgoal chunk gets updated. Further subgoals in order to find the goal are pursued, namely reaching the bottom-left corner and reaching the right grid border at three quarters of the grid's height. The green goal object is detected with a pop-out effect due to the high utility of the search production. As soon as the goal is recognized, its location is stored as the new subgoal.

Decision making and problem solving Even if the agent's heuristic is to move directly towards its subgoal, the walking path is not implemented as a straight line. Steps are chosen randomly, whereas the movement in subgoal direction is prioritized. If there are obstacles or malus objects on the path, they are avoided. In case the movement direction is changed

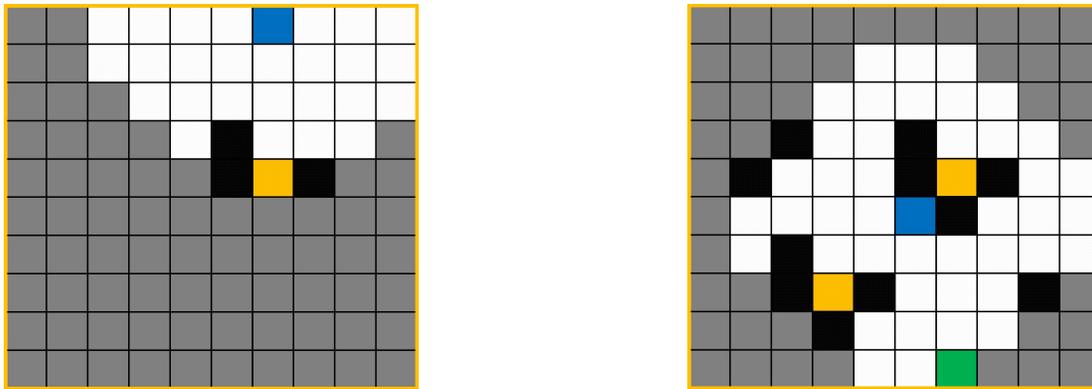


Figure 1: Examples of two situations in the same environment. The area in the proximity of the agent is "visible" to it, anything beyond is invisible, i.e. covered by a *fog of war*. On the left, the agent is in a starting position, with 4 objects in 2 colors, but not the goal directly visible. Black tiles are obstacles while yellow tiles are bonus fields, although in the beginning the meaning of these colors is unknown to the agent. On the right, the agent advanced to a point where the goal tile is now visible. Note that exploration revealed "traps" to the agent, enticing it to enter a dead end of obstacles.

the blocked location is stored in the imaginal to prevent the agent approaching the same object again. This mechanism is helpful to free the agent from triangular "traps" or to walk around obstacles blocking the path.

Strengths and weaknesses Since the agent perceives all objects only pre-attentively, the speed to reach the goal tile is maximized. This is supported by the fact that it does not take bonus points into consideration unless they are located directly on the path. Therefore, the resulting score is low compared to the other agents. If the goal tile is spotted but not reachable, the agent will be stuck at an obstacle and cannot free itself due to its goal-directed heuristics (with a total of 41 productions). During testing, the agent was able to complete 12 out of 15 challenges. This represents an overall satisfying performance with special regard to the enormous speed with which Speedy solves the challenges. However there is still room for improvement, including priority collection of bonus objects for higher scores as well as the ability to recover from entrapment. As soon as the agent spotted the goal object with its path blocked, the agent gets stuck. Therefore, a possible option would be to change the subgoal in case the distance to this subgoal is not reduced within several steps.

Forest

Represents the current goal by separate state, intention and searching slots in the goal chunk. Unknown tiles are sought to be identified.

This agent attends a random tile with an object in the top row of the grid. When an object is located pre-attentively the agent will try to exercise a movement. After the movement the agent is attending the same location again to check whether the color of the tile has changed. In case of a color change - the agent was found. In the other case a different color will be perceived and the same process is repeated. This color information is stored in the goal-chunk in order to have a sustained awareness of itself.

Searching the environment Before executing a move the agent checks whether the next targeted tile is blocked by another object. If the object is unknown the agent will try to move on that tile. By that the classification will start. If the color of the desired field changes to the color of the agent the object will be classified as a bonus or malus tile. If the color of the desired fields does not match - the object is classified as an obstacle. When the agent encounters an object, it will retrieve a chunk from the declarative memory. In case of an obstacle or malus tile it will try to avoid that tile and in case of a bonus tile it will try to move onto it.

Locating the goal With the analyzed and calculated grid the agent has a good starting point and orientation to use its first heuristic - make your way to the middle of the grid. Since the agent is aware that the main goal will be in the lower half of the grid. After reaching it desired location the agent will use another heuristic. This heuristic is based on a waypoint system by exploring the left and the right side, making its way to the bottom of the grid. In the routine a high utility production ensures that a visible green tile will be prioritized as the new main goal.

Decision making and problem solving In general the agent's processes are organized in routines as visualised in Figure 2A. Inside of the main routine a hierarchic structure is used to ensure that distinct routines are available at certain times. For example the agent checks before every movement whether the green main-goal is visible or the adjacent field is an unknown or known object. A wide variety of 53 productions were used to solve other problems on its way to the green tile. That included several escape mechanisms inside the decide-action routine to get around the obstacles. A specific goal-chunk is used to hold slots of the actual state, an intention and a searching slot. Each slot had a distinct assignment: the state-slot was used to guide the agent through the heuristics, the intention-slot was used to remember the loca-

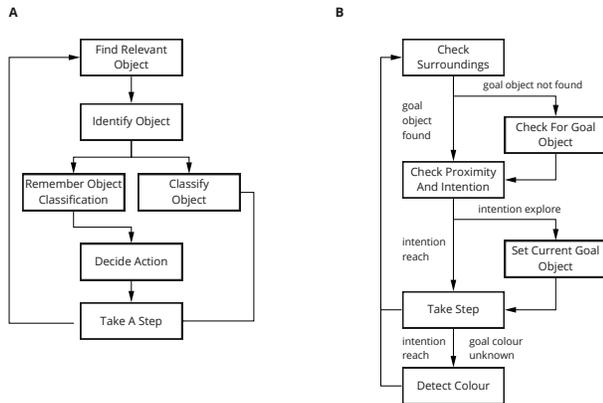


Figure 2: Flowcharts of the goal-finding routines of the agents named “Forest” (A) and “Intell-agent” (B) after self- and board-detection.

tion the agent was moving to and the searching-slot gave it the actual state of the routine.

Strengths and weaknesses Our agent was comparatively fast when moving through the presented environments. The approach was to keep it simple and to process everything pre-attentively. This meant less costs in terms of time. On the other hand the agent was not able to solve neither complex nor medium difficulty levels, because avoiding movements were not context sensitive. A key aspect for improvement would be the agent’s capacity to remember where it is coming from.

Ms. Captain Curious

Explores the environment randomly. Actively seeks out novel tile colors. Identifying all encountered tiles is prioritized over reaching the goal.

Ms. Captain Curious was developed to be a lightweight model employing only a few heuristics that are generally useful, instead of specialized strategies for different cases. The model does not build up a lot of permanent knowledge. Most of the relevant information is retrieved from the visual module when needed and then temporarily stored in the goal or imaginal buffer. In the following, some of the heuristics, or routines, are described followed by a brief look at the overall performance of the model.

Searching the environment The localization of the agent relative to the grid was necessary to determine a suitable direction for the next moves - instead of steering the agent against the border, an unexplored direction was preferred. The heuristic was to find the borders of the field and move in the direction of those borders that are away the farthest. Since only the eight fields adjacent to the spaceship were visible at any given time, movements were necessary to explore the grid. The agent then started moving from the current quadrant to the one diametrically opposite. Only when touching borders or unknown objects, the direction was reversed or re-

evaluated. On an empty grid, the movement pattern would thus resemble a billiard ball rolling unchecked from edge to edge across a billiard table.

Locating the goal In order to find out the meaning of the different colored objects, the agent had to be steered on unknown object tiles to then memorize the consequence (point gain, point deduction or obstacle) and subsequently deal with the tile types differently (seek, avoid). This curious, minimal learning behavior can be described as “Whenever you see a novel object, check it out and remember the consequence”. This primacy of curiosity was eponymous for the agent. In case the colors of obstacles or malus fields have already been determined they were avoided. Whenever there was a known object on a field in the intended direction, a new direction of movement was selected randomly. This is no long term planning strategy, only a consideration of the next step.

Decision making and problem solving After each single movement, the agent checks whether the goal or any known or unknown object is in sight, and if there is a contact with the border. The checks have different priorities: Identifying the color of bonus tiles has the highest priority. Once the color of the bonus tile has been identified, approaching the target is set above the tile identification. Finally it is evaluated whether the spaceship is at a border, if necessary the direction of movement is reversed. Additionally, further abstract procedural patterns were identified and made explicit, which, for example, governed the handling of obstacles and malus fields or determined the hierarchy of routines.

Strengths and weaknesses Overall, the agent performed comparatively well, solving most of the challenges within the given time constraints with a total of 83 productions. Yet, since an initial decision was to design an agent that moves around a lot rather than thoughtfully weighing each of its decisions, wall-like rows of obstacles posed serious problems: When an unknown object or the goal was situated behind a row of obstacles, the agent would move towards it, resulting sometimes in the agent getting irresolvably stuck. It was serendipity, provoked by the random selection of movement directions, that sometimes helped the agent to circumvent those problem situations nonetheless.

Intell-Agent

Uses a 12 tile diamond-shaped visual representation for reasoning. Intentions determine goal pursuit behaviour. Preference to collect visible bonus points followed by reaching the goal.

The Intell-agent represents its environment as a diamond-shaped field consisting of 12 tiles (see Figure 3). This information, the agent’s current x- and y-position as well as its last move, are saved in a “vision”-chunk in the imaginal buffer consisting of 15 slots. Consequently it is able to detect and avoid immediately adjacent triangular traps, which are token (malus and obstacles) in an disadvantaged formation.

The goal buffer is relevant for representing the agents’

states and intentions. States describe the various stages the agent passes through as part of its goal-finding routine as detailed below. Intentions specify the overall tactic of approaching the set goal. The minimal and maximal x and y coordinates of the grid field are also contained in the goal buffer and represent the borders of the grid field. Lastly, the goal is responsible for upholding information about kind and position of the current goal. The Intell-agents goal-finding routine is structured in five stages shown in Figure 2B.

Searching the environment The visual representation is updated based on the agents current position and its last motion. The visual-location module is used to check for tiles on which the agent does not have any information yet. Tiles outside of the grid field are marked so that they won't be moved on.

Locating the goal The agent checks if the green goal object is visible. If so, its location is stored in the declarative memory. In case a chunk with the goal's location can be retrieved, the agent directly proceeds to the next stage. The agent's current position and the goal position are then compared to determine if the agent is next to the current goal. Based on its intention at that stage, it will then either skip the next stage and move directly towards the current goal (intention: reach) or enter the goal setting stage (intention: explore).

During goal setting, the agent sequentially goes through a list of priorities, only moving onto the next one in case the previous one does not apply.

1. The agent's top priority is to find out which color represents the bonus field. Therefore, if a tile with an unknown color is detected and the bonus color is unknown its position is set as the current goal and its intention is set to "reach".
2. If however, the bonus color is known, reaching bonus tiles becomes the highest priority. The agent tries to find the visual-location of the nearest tile with the bonus color. If successful, the bonus tile becomes the current goal and its intention is set to "reach".
3. If the bonus color is known but no bonus object is visible, the agent looks for the green goal object. In case the green goal object is not visible, the agent tries to retrieve the goal position from the declarative memory. If successful, the green object's position will be set as the current goal and the intention is set to "explore".
4. If the green tile is neither visible nor retrievable, the agent applies a searching heuristic. Depending on current position of the agent relative to the grid field boundaries, the agent will set its current goal to the center of the playing field (if in the top half of the playing field), the lower left corner (after the middle was successfully reached) and eventually the lower right corner (after the left corner was successfully reached).

Decision making and problem solving The agent tries to reduce either its vertical or horizontal distance to the current goal. If this is not possible, it will move onto an object that it has not yet visited. This behaviour is facilitated using 'breadcrumbs' which are placed on tiles' representations within the imaginal buffer that were previously visited. Such tiles are avoided by the agent. All breadcrumbs are removed whenever any goal is reached.

When the agent seeks to identify the meaning of different colors, it does so by moving on to a tile with unknown color and checking for a visual location with red text. If none is detected the tile color is saved as the obstacle color. If red text is perceived, the agent's movement is evaluated. Given the agent has not moved, the color of the current goal is saved as malus color. Otherwise the agent has successfully found out the bonus color.

Strengths and weaknesses The strategy resulting from this cycle is to maximize the score during level completion. This can lead to prolonged run times and sometimes even unexpected behaviour, for example when the agent moves away from the goal just to collect another bonus. Overall the performance of the agent was satisfactory, but due to its tendency to collect every bonus and the fairly high complexity of the model (with a total of 226 productions) the agent is sometimes quite slow and exceeds the time limit of some challenges.

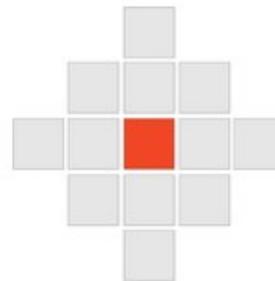


Figure 3: Intell-Agent's twelve tile representation of the agent's environment. Each tile was indexed and could carry the values "empty", "breadcrumb", "out of bounds" and the various object colors.

Performance Review

All four agents tried to cope with a partly unknown environment and tried to apply complex cognition as defined by (Funke, 2010). Thus the agents used mental processes for deriving new information out of given information, with the intention to make decisions, solve problems, and plan actions. They had to perceive their environment and use their memory. Not all mechanisms are cognitively plausible and are sometimes oversimplified, yet the agents were able to flexibly handle new environments and settings and find their way to the goal. Although the task was the same and required the same cognitive skills for fulfilling the four requirements, the implementation of the individual agents varied a lot. The number of productions is a first indication for this, with a range from 41 to 226 productions - with the lowest production agent still performing fairly well.

Table 1: Aggregated Agent Model Results

Agent	# of productions	Completed (out of 15)	Avg. # of Moves	Avg. Time (s)	Avg. Score
Speedy	41	12	26.75	14.184	1145
Forest	53	4	24.10	29.50	80
Ms. Cap. Curious	83	13	39.46	39.78	1180
Intell-agent	226	13	26.61	62.04	1177.2

In Table 1 the different performances are listed and it becomes apparent that most agents were able to manage most of the challenges - some of which were quite complex and difficult to solve. Most importantly lessons were learned on what was helpful to model such an agent having to cope with a dynamic and incompletely known world, i.e. what kind of knowledge representations were crucial and how reasoning problems were solved.

The following insights were gained from modelling situated cognitive agents interacting with the introduced dynamic environment. First, relevant information about the agent (self representation such as colour, once identified) and non-changing information about the environment such as its size are held available, usually in the goal buffer. Further relevant information about the current context, acquired information about the meaning of objects and information about the goal location are also stored, usually in the imaginal buffer. Second, sub-goals or identifying different phases of the task are helpful to find flexible ways to solve specific problems or to have access to specific productions. Third, as speed is important, it is essential to identify what information is relevant and would take too much time to retrieve often. This information should be available in one of the buffers. Information that is only necessary infrequently can be retrieved from memory. Also pre-attentive visual processes were used whenever possible and cognitively plausible, such as for self-localization or when searching for specific information, consequently saving time. Forth, strategies that are too rigid usually lead to situations where agents get stuck and have difficulties to free themselves. Furthermore, keeping information about past movement of the agent or what areas have already been searched are greatly helpful.

Discussion

The main lessons learned by the students were (1) to develop a better understanding of what cognitive plausible mechanisms really are, where difficulties lie and how to change the usual approach to this kind of task. Still, parts of the agents show computational aspects rather than cognitive, but time was restricted for the project. (2) The second lesson learned was to realize how important it is to use detailed task analysis and visualizations of model structure for group communication while modelling. (3) The third point was that pre-attentive visual processes are sometimes sufficient for simple localization and checking purposes of the agent.

Lessons learned regarding the architecture used

ACT-R offers a lot of structures that are helpful to model flexible and learning agents in task environment such as these. Debugging and handling its output was a challenge sometimes. Visual grouping or perception of a "field" and identifying borders was difficult to realize in a cognitively plausible way. Additional visual support would be a very helpful component for research on self-sufficient agents.

The project nicely showed the aspects and requirements that Kurup and Lebiere (2012) listed for high-level cognition in robotics. (1) Represent, integrate and use large amounts of knowledge: it needs to be carefully considered what information really needs to be stored and where. Storing the whole grid field in our example would have slowed down the agents, so this was not done. Rather, the students tried to find ways to solve the most important problems with as little stored information as possible, since human participants would also not store all tiles in the grid. (2) Learning patterns: in this project, the agents learned the meaning of objects by interacting with them and adjusted their planning accordingly. There was not enough time to learn from difficult situations and obstacle patterns, which would be highly interesting and potentially address (3) Problem solving and reasoning. (4) Flexible, adaptive, dynamic, and real-time behavior was shown by the agents. Explicit encoding of visual objects was prevented as much as possible in order to not lose valuable time. The agents were also able to flexibly cope with newly appearing objects, an unseen goal and different environments - therefore long-term planning was not possible. The last requirement, (5) Interact with humans in a natural way, would require a more refined approach.

This type of challenges for cognitive agents, as mentioned earlier, offers intriguing test beds to explore how flexible models based on a cognitive architecture really are and how much such approaches could add to existing agent approaches. Especially the topic of mental representations (e.g. Clark & Grush, 1999) is crucial in such unpredictable environments and for adaptive and complex behavior. This potential should be explored in more detail.

Acknowledgments

The authors would like to thank their respective project partners and all other participants of the 2020/21 winter semester course "Applied Cognitive Modeling" at the Technische Universität Berlin.

References

- Anderson, J. R. (2000). *Learning and memory: An integrated approach* (2nd ed.). Wiley.
- Clark, A., & Grush, R. (1999). Towards a cognitive robotics. *Adaptive Behavior*, 7(1), 5-16.
- Funke, J. (2010). Complex problem solving: A case for complex cognition? *Cognitive processing*, 11(2), 133–142.
- Kurup, U., & Lebiere, C. (2012). What can cognitive architectures do for robotics? *Biologically Inspired Cognitive Architectures*, 2, 88–99.
- Levesque, H., & Reiter, R. (1998, March). High-level robotic control: Beyond planning. A position paper. In *AIII 1998 Spring Symposium: Integrating Robotics Research: Taking the Next Big Leap*.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625–636.

People are Insensitive to Within-Category Feature Correlations in Categorization

Florian I. Seitz (florian.seitz@unibas.ch)
 Department of Psychology, Missionsstrasse 62A
 Basel, 4055 Switzerland

Jana B. Jarecki (jana.jarecki@unibas.ch)
 Department of Psychology, Missionsstrasse 62A
 Basel, 4055 Switzerland

Jörg Rieskamp (joerg.rieskamp@unibas.ch)
 Department of Psychology, Missionsstrasse 62A
 Basel, 4055 Switzerland

Keywords: categorization; similarity; cognitive modeling; correlational sensitivity; within-category information

Introduction

This work investigates people’s sensitivity to within-category feature correlations in perceptual categorization by testing two types of psychological similarity against each other.

Similarity is a central component of many psychological categorization theories. Exemplar theories, for instance, assume that people categorize new objects based on their similarity to previously seen category members (Nosofsky, 1986, 1989). Traditionally, the underlying psychological similarity between a pair of objects is modeled as the sum of the objects’ squared feature value differences (*Euclidean similarity*; e.g., Goldstone, 1994). The Euclidean similarity, however, ignores the distribution of objects by assuming uncorrelated features within categories. In turn, the *Mahalanobis similarity* extends the Euclidean similarity by accounting for within-category feature correlations. Results from machine learning have shown that the Mahalanobis similarity can outperform the Euclidean similarity in categorization problems involving correlated features within categories (Mao & Jain, 1996; Weinberger & Saul, 2009). Yet, in the psychological categorization literature, there are mixed results regarding the extent to which people take within-category feature correlations into account (Chin-Parker & Ross, 2002; Ell, Smith, Peralta, & Hélie, 2017; Lancaster, Shelhamer, & Homa, 2013).

Therefore, the present work investigated if people use within-category feature correlations for categorization. Our work rigorously compared the correlation-insensitive Euclidean similarity against the correlation-sensitive Mahalanobis similarity by means of mathematical modeling on data from an optimized category learning task.

Methods

We designed a standard trial-by-trial supervised, binary category learning task (e.g., Nosofsky, 1989) with two strongly correlated features within each category ($r = .98$). In the task, participants learned to categorize a set of stimuli from feedback and then categorized new test stimuli

without feedback. The category structure was selected using simulation-based optimal experimental design (Myung & Pitt, 2009). Our formal modeling framework were two versions of the exemplar model of Nosofsky (1986, i.e., the *generalized context model*); one version used the Euclidean similarity, the other used the Mahalanobis similarity.

To optimize the experimental design, we searched for a category structure that both model versions can learn accurately and that maximizes the classification prediction differences between the two model versions for the test stimuli. Figure 1 shows that in the resulting optimal design the Euclidean similarity assigns the test stimuli into the category with lower straight-line distances, whereas the Mahalanobis similarity assigns them into the category with the matching correlational structure.

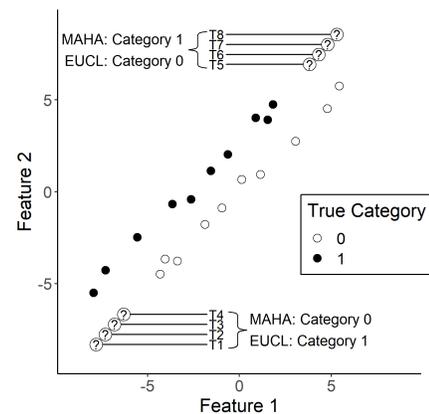


Figure 1: *Category structure of the task involving a high within-category feature correlation.* Shown are the 20 learning stimuli with their true category (black and white circles) and the 8 test stimuli presented after learning without feedback (circles with question mark). EUCL = predictions of the model using the Euclidean similarity; MAHA = predictions of the model using the Mahalanobis similarity.

In the experiment, participants ($N = 43$; 14 females, $M_{age} = 25.56$ years, $SD_{age} = 6.81$ years) learned to classify 20 stimuli with feedback until they were more than 90% accurate

across the last 100 trials; then, they classified 8 new test stimuli without feedback. The stimuli were geometric figures consisting of a circle of varying size and a line of varying orientation (as in Nosofsky, 1989); category labels and visual feature assignments were randomized across participants.

Results

The free parameters of both model versions were estimated for each participant based on their classification learning data using maximum likelihood. The resulting individual parameter values were used to predict the respective participant’s classification behavior for the test stimuli.

The results clearly show that most participants ignored the within-category correlations. Specifically, participants tended to classify stimuli ”T1-4” into category 1 and stimuli ”T5-8” into category 0, in line with the predictions of the Euclidean similarity, see Figure 2. Mathematical modeling showed that, in the aggregate, the Euclidean similarity predicted participants’ categorizations better than the Mahalanobis similarity, median log-likelihood across participants: Euclidean similarity model = -7.21, Mahalanobis similarity model = -21.63, random-choice model (predicting category probabilities of .50) = -16.64.

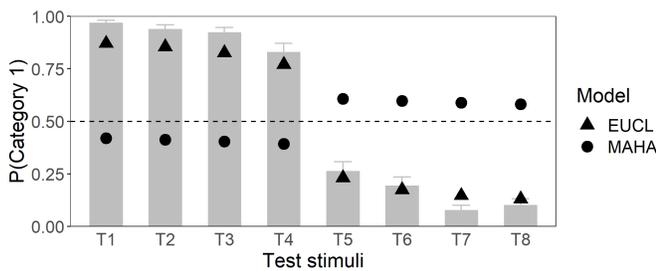


Figure 2: Responses and predictions for the test stimuli. Bars and whiskers show the mean and the standard deviation, respectively, across participants’ mean category responses for a given test stimulus. The model predictions are aggregated over participants with the mean. EUCL = Euclidean similarity model; MAHA = Mahalanobis similarity model.

At the individual level, a participant was assigned to a model if the model’s Akaike weight exceeded .67 for this participant. The results show most participants were best described by the Euclidean similarity (n = 33 of 43) with strong evidence, see Figure 3. The remaining participants were described by the Mahalanobis similarity (n = 4), the random-choice model (n = 2), or unclassifiable (n = 4).

Conclusion

Our findings indicate that people do not include within-category feature correlations in their representation of similarity during categorization. Instead, people treat every object as stemming from a category with uncorrelated features, even if the true classification environment has strong within-category feature correlations.

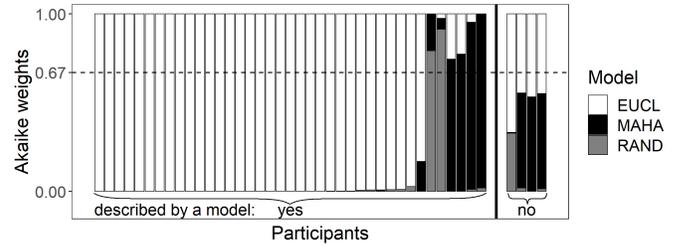


Figure 3: Model evidence strengths for each participant. For each participant, the Akaike weights of the different models are stacked upon each other and sum up to 1. EUCL = Euclidean similarity model; MAHA = Mahalanobis similarity model; RAND = random-choice model.

Acknowledgments

This work was supported by the Swiss National Science Foundation [grant number P0BSP1_195389 to F.I.S.].

References

Chin-Parker, S., & Ross, B. H. (2002). The effect of category learning on sensitivity to within-category correlations. *Memory & Cognition, 30*(3), 353–362. doi: 10.3758/bf03194936

Ell, S. W., Smith, D. B., Peralta, G., & Hélie, S. (2017). The impact of category structure and training methodology on learning and generalizing within-category representations. *Attention, Perception, & Psychophysics, 79*(6), 1777–1794. doi: 10.3758/s13414-017-1345-2

Goldstone, R. L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition, 52*(2), 125–157. doi: 10.1016/0010-0277(94)90065-5

Lancaster, M. E., Shelhamer, R., & Homa, D. (2013). Category inference as a function of correlational structure, category discriminability, and number of available cues. *Memory & cognition, 41*(3), 339–353. doi: 10.3758/s13421-012-0271-8

Mao, J., & Jain, A. K. (1996). A self-organizing network for hyperellipsoidal clustering (HEC). *Ieee transactions on neural networks, 7*(1), 16–29. doi: 10.1109/72.478389

Myung, J. I., & Pitt, M. A. (2009). Optimal experimental design for model discrimination. *Psychological Review, 116*(3), 499–518. doi: 10.1037/a0016104

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115*(1), 39–57. doi: 10.1037/0096-3445.115.1.39

Nosofsky, R. M. (1989). Further tests of an exemplar-similarity approach to relating identification and categorization. *Perception & Psychophysics, 45*(4), 279–290. doi: 10.3758/BF03204942

Weinberger, K. Q., & Saul, L. K. (2009). Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning Research, 10*(2). doi: 10.5555/1577069.1577078

The Structured Mind at Rest: Evidence for the “Common Model of Cognition” in Resting State fMRI

Catherine Sibert (sibert@uw.edu)

Department of Psychology, University of Washington, Seattle, WA 98195 USA

Holly Hake (hakehs@uw.edu)

Department of Psychology, University of Washington, Seattle, WA 98195 USA

Andrea Stocco (stocco@uw.edu)

Department of Psychology, University of Washington, Seattle, WA 98195 USA

Abstract

The Common Model of Cognition (CMC) has been proposed as a high level framework through which functional neuroimaging data can be predicted and interpreted. Previous work has found the CMC is capable of predicting brain activity across a variety of tasks, but it has not been tested on resting state data. This paper adapts a previously used method for comparing theoretical models of brain structure, Dynamic Causal Modeling (DCM), for the task-free environment of resting state, and compares the CMC against six alternate architectural frameworks. For a large sample of subjects from the Human Connectome Project (HCP), the CMC provides the best account of resting state brain activity, suggesting the presence of a general purpose structure of connections in the brain that drives activity when at rest and when performing directed task behavior.

Keywords: Brain architecture, Cognitive Architecture, Computational models, Dynamic Causal Modeling, fMRI, Resting state.

Introduction

Despite a shared goal of understanding the underlying mechanisms of the brain, research that focuses on high-level structural models of cognition remains largely isolated from efforts to interpret direct measurements of brain activity. Many neuroscientists are reluctant to rely on the results and conclusions from cognitive architectures because, while the behavior of the models often closely matches observed human data, the mechanisms driving that behavior are rooted in the principles of computer science and information theory. Efforts have been made to connect components of cognitive architectures to corresponding brain regions, but direct biological brain functions are rarely well captured by the more conceptual architecture modules, and architectures often make incompatible assumptions about the basic functional components that are needed to support cognition.

The Common Model of Cognition

One successful attempt to achieve consensus is represented by the so-called Common Model of Cognition (CMC; Laird et al., 2017). The CMC is a computational framework that can serve as a blueprint to understand the organization of a human-like mind.

Abstract computations are categorized into five functional components (long-term memory, working memory, procedural memory, perception systems, and action systems) with specific directional relationships (Fig. 1A) between them.

Although it was not proposed specifically as a *brain* architecture, a number of studies have found that the CMC is surprisingly effective at modeling brain activity across tasks and individuals (Steine-Hanson et al., 2018; Stocco et al., 2018, 2021). In this interpretation, the CMC’s functional components are mapped onto large-scale brain regions (Fig. 1B) and their relations are translated into predicted patterns of functional connectivity. In other words, the neural counterparts of the functional components and their connections serve as a simplified architecture for the human brain, not only the human mind.

Resting State Brain Activity

A secondary problem with cognitive architecture models is their focus on the brain *at work*. Virtually all mappings between cognitive architectures and brain activity have been carried out based on neural responses to specific tasks (Anderson et al., 2008; Eliasmith et al., 2012). This bias was inherited from the brain imaging analyses carried out to test the CMC, which, so far, have similarly focused on task-based activity.

In contrast, while many analyses of fMRI data compare differences in activity while subjects perform a variety of tasks, a lot of recent work has instead focused on the connectivity of the brain at rest. This line of research was spawned by the observation that even spontaneous brain activity shows a high degree of structure (Fox et al., 2005; Sherzhad et al., 2008), which is revealed in terms of correlations between the time courses of the activity of different brain regions. These patterns of correlations are fairly stable across individuals (Gratton et al., 2018), to the point that variations in the patterns of correlations can be used to reliably predict abnormal neurological conditions (Hohenfeld, Werner, & Reetz, 2018) and can even be used to successfully predict the patterns of brain activity during tasks (Cole et al., 2016; Yeo et al., 2011). These findings suggest the possibility of an underlying

structure to the brain that can be adapted to tasks as needed, but is still present even when resting.

This paper extends the work of Stocco et al. (2021) by testing the Common Model of Cognition on brain activity *at rest* using a pre-defined network of brain regions. Specifically, this paper adapts the framework of Dynamic Causal Modeling (DCM) and compares the Common Model of Cognition (CMC), against six other exemplar network structures that could capture the underlying structure of the mind.

Dynamic Causal Modeling

The DCM framework aims to identify the causal influences of neuronal systems by quantifying the dynamic fluctuations in brain activity (Friston et al., 2003).

$$dy/dt = \mathbf{A}y + \mathbf{C}x \quad (1)$$

In this equation, hemodynamic brain activity, represented by vector y , is multiplied by matrix A , which contains a set of parameters constituting a proposed structure of connectivity between regions. Thus, the structure of matrix A can be adapted to test alternative connectivity architectures. C is a matrix of the parameters that specify how external or driving inputs elicit changes in brain activity, and x defines the matrix of task inputs. Since there are not any external inputs driving activity in the resting state data, the C matrix was adapted to model low frequency fluctuations seen in this state using deterministic inputs as task conditions (see Materials and Methods, below).

Alternate Model Architectures

As pointed out in Stocco et al. (2021), DCM is a strictly top-down, theory-driven method, and cannot be used to infer an architecture from the data. Instead, to evaluate the CMC as an architecture, its predictions were compared against a collection of alternative networks that consist of the same components, but different connection patterns (Stocco et al., 2021). These alternate models are not exact implementations of other cognitive architecture systems, like ACT-R or SPAUN, but instead represent the space of possible theoretical neural architectures.

The alternate architectures fall into two broad categories, or families. In the “Hub-and-Spoke” family (Fig. 1C), a single ROI is designated as the central “Hub”, and is bidirectionally connected to all other ROIs. However, none of the “Spoke” ROIs are connected to any other - all activity must travel through the “Hub”. Three different Hub-and-Spoke models are considered, based on whether the role of the hub is played by the Prefrontal Cortex, mapped to Working Memory (as proposed by Cole et al., 2012), the basal ganglia, mapped to Procedural Memory (as proposed by Anderson, 2007), or the temporal lobe, mapped to Long Term Memory (as proposed by Visser et al., 2012).

The “Hierarchical” family of models proposes an alternate structure, wherein brain connectivity implements hierarchical levels of processing that initiate with Perception and culminate with Action (Fig. 1D). Networks in this family conceptualize the brain as a feedforward neural network model in which different regions perform progressively greater levels of representational abstraction (Huntenburg et al., 2018). Three different hierarchical architectures are generated based on the relative position of the basal ganglia (mapped to Procedural Memory) in the hierarchy. Specifically, the basal ganglia can be placed between perception and long-term memory (as in models of procedural categorization: Kotz et al., 2009; Seger et al., 2008), between long-term memory and working memory (as in models of memory retrieval: Scimeca & Badre, 2012), or between working memory and action (as in models of action selection: Houk et al., 2007).

Broadly speaking, the CMC can be considered as a “Hub-and-Spoke” structure, using Working Memory (mapped to the Prefrontal Cortex) as the “Hub” ROI, with an additional direct connection between Perception and Action.

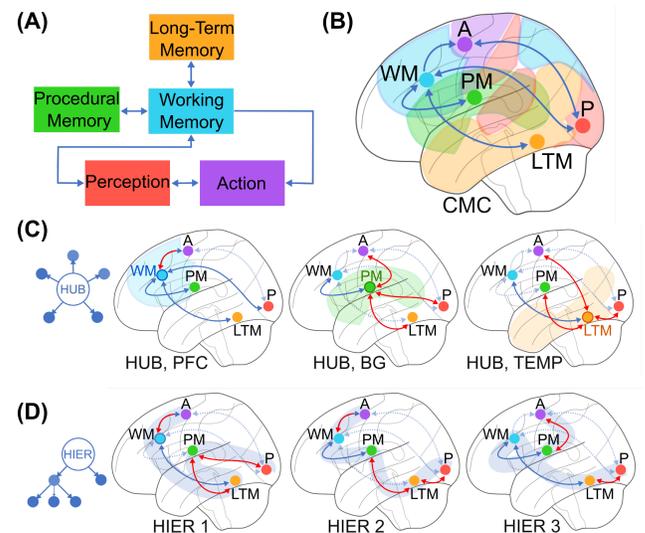


Figure 1: (A) The Common Model of Cognition (CMC); (B) Proposed associations between components and anatomical brain regions. (C) Three variations of Hub-and-Spoke (HUB) models, and of (D) Hierarchical (HIER) models. Arrows: dark blue, connections present in both CMC and candidate models; red, connections unique to candidate models; and dotted, connections present in CMC and absent in candidate models.

Materials and Methods

The Human Connectome Project Dataset

The data used in this analysis was drawn from the Human Connectome Project (HCP), a large scale effort to collect neuroimaging data from healthy young adults. This study

in particular analyzed a subset ($N=168$) of rsfMRI data exclusively. For each subject, 14 minutes of rest data (eyes open with fixation) were recorded prior to a run of task data collection. A second rest run was recorded after the task battery, and was not included in this analysis. Between the two collection days, each subject had a total of 28 minutes of data. Each day's data was modeled separately, and then combined in the final analysis.

Data Processing and Analysis

Image Acquisition and Preprocessing. MRI images were acquired and minimally preprocessed according to HCP guidelines (Barch et al., 2013; Van Essen et al., 2013). Scans were taken on a 3T Siemens Skyra using a 32-channel head coil with acquisition parameters set at TR = 720 ms, TE = 33.1 ms, FA = 52°, FOV = 208 × 180 mm. Each image contained 72 2.0mm oblique slices with an in-plane 2.0 × 2.0 mm resolution. Images were acquired with a multi-band acceleration factor of 8X. These raw images then underwent minimal preprocessing including unwarping, motion realignment, and normalization to the standard MNI template. From there, the images were then smoothed with an isotropic 8.0 mm full-width half maximum Gaussian kernel.

Simulated Task Events Both general linear modeling (GLM) and DCM analysis require a design matrix that specifies the timing of external events that drive brain activity. Traditionally, these events are task related; the onset or absence of some stimuli. Rest data, by contrast, is collected without any specific task structure, and the recorded activity must be driven by internal and unobservable patterns. Following Di and Biswal's method (2014), a series of slow oscillatory waves of different frequencies were created as input "events" that simulate background brain activity (Fig. 2A). Specifically, eight different driving waves were generated as sine and cosine waves with frequencies of 0.01, 0.02, 0.04, and 0.08 Hz, respectively. The frequencies of these oscillations capture the canonical frequency range (0.1 - 0.01 Hz) of spontaneous fluctuations in brain activity (Fox et al., 2005). An event is considered to be occurring during the positive cycle of the wave. A second assumption concerns how different events affect the different regions. In task-based DCM analysis, it is possible to make reasonable assumptions about which regions are affected by which events, such as the presentation of visual stimuli affecting a perceptual region. Di and Biswal (2014) explored a subset of possible regressor-by-region combinations to determine the most appropriate. Here, we followed the procedure of Ketola et al (2020) and let each region be potentially affected by each oscillatory regressor (Fig. 2B). Note that, while being the most general approach, this method goes *against* our hypotheses that spontaneous brain activity would follow a structured architecture, as it gives every region the greatest opportunity to have its time

series modeled by external inputs rather than by the network effects of other regions.

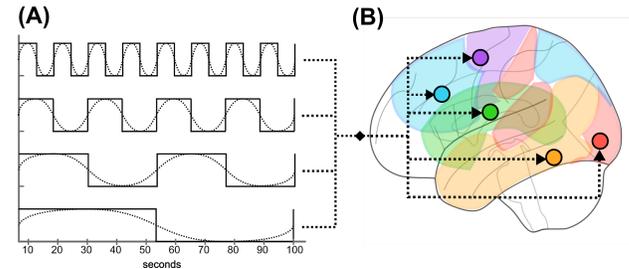


Figure 2: (A) Oscillatory waves (dotted lines) translated into "box-car" plots of events (solid lines). (B) Each event used as drivers for activity in all ROIs.

GLM. A GLM analysis was carried out to define the event matrix x that is used in the DCM equation (Eq. 1). In task based DCM, events were differentiated by type and served as input to specific regions of interest. Since the resting state does not have any tasks, and the artificial events were used to capture background activity patterns, all "events" were used as direct inputs to all regions of interest (Fig. 2).

Regions of Interest Definition. Previous DCM analyses relied on task-based activity to define specific regions for each model component, but in the absence of a task structure for rest data, an alternate method was needed to determine regions of expected activity. Initial region masks were created using NeuroSynth (www.neurosynth.org), a platform that combines the results of thousands of published fMRI results and produces meta-analysis images of activity associated with various higher level conceptual category terms. For each of the five model components of the CMC model, a corresponding term was chosen from NeuroSynth's database, and a summary statistical mask was produced for each term, with each voxel having an associated Z value representing the probability that the voxel would show up a study associated with the term. These individual masks, however, were large and produced significant overlap when combined, meaning that activity in a particular voxel could belong to more than one region. To solve this problem, two thresholds were applied to the original masks, one height threshold applied to each individual voxel statistic and a minimal extent threshold applied to each cluster size. Both thresholds were calculated *proportionally* for each region, i.e. as a proportion of the highest Z-score and of the largest cluster within an image, respectively. The proportional adjustment was done to prevent regions with large clusters and high statistics, like perception, from overtaking regions with comparatively low Z score levels, like procedural memory. The Nelder-Mead (1965) optimization algorithm was then applied to find thresholds in the two-parameter space that would produce the largest possible regions without any overlapping voxels. The final values identified by the Nelder-Mead

algorithm were a proportional height threshold of 0.5359, and a proportional extent threshold of 0.4164. The final masks are shown in Figure 3.

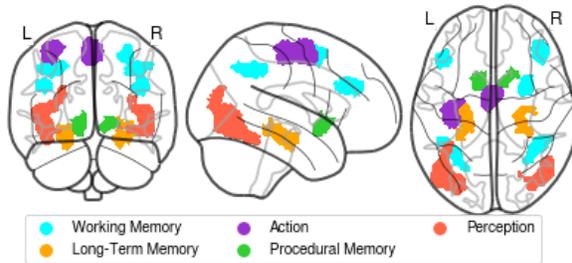


Figure 3: Final regions of interest derived from Neurosynth activity masks. Individual ROIs were selected from most active voxels within these areas for each subject.

Model Fitting. Once the time-series for each ROI was extracted, different networks were created by connecting all of the individually-defined ROIs according to the specifications of each model (Fig. 1). The predicted neural activity for each model was then calculated using Equation 1, and the predicted time course of BOLD signal was then generated by applying a biologically-plausible model of neurovascular coupling to the simulated neural activity of each region. All of the model parameters were estimated through an expectation-maximization procedure (Friston et al., 2003) to reduce the difference between the predicted and observed time course of the BOLD signal in each ROI.

Model Comparison. The models were compared on the basis of their likelihood function $L(m | x)$. A model's likelihood is the probability of it producing the observed data x ; that is, $L(m | x) = P(x | m)$. Group-level likelihood values for a model m can then be expressed as the product of the likelihood of that model fitting each participant p , i.e., $\prod_p L(m | x_p)$. The log-likelihood is the sum of all of the individual log-likelihoods: $\sum_p \log L(m | x_p)$. Although more sophisticated model comparison procedures have been proposed (e.g., Stephan et al., 2009), the log-likelihood based metric used here is not only the most easily interpretable, but also the most relevant, as it specifically applies to cases in which it is assumed that the model is constant or architectural across individuals (Kasess et al., 2010).

Results

Regressor Quality Analysis

We first conducted a GLM analysis to ensure that our oscillatory regressors successfully captured brain activity. To do so, we calculated an omnibus ANOVA across all oscillatory regressors at the participant level. This test captures any variance that can be accounted for by any of the oscillatory regressors. The resulting F -statistic map was then log-transformed, yielding a measure of the

difference between the variance explained by regressors and the residual variance (i.e., noise). Finally, a group-level T -test was performed on the individual-specific log-transformed F -maps. The result of this analysis is a statistical test of whether the variance captured by the regressors was significantly greater than the variance of the residuals. The results are shown in Figure 4, thresholded at a value of $t(160) > 5.212$, which corresponds to $p < 0.05$ when corrected for multiple comparisons through the Family Wise Error correction procedure.

As Figure 4 shows, most of the grey matter voxels exhibit oscillatory activity that was captured by our regressors. Importantly, the significant voxels encompass regions in all of our predefined ROIs, including the medial temporal lobes (long-term memory ROI in Figure 3, visible in the coronal section of Figure 4) and the subcortical basal ganglia (procedural memory ROI in Figure 3, visible in the axial and sagittal sections of Figure 4), which are notoriously affected by lower signal-to-noise ratios in high-density neuroimaging protocols.

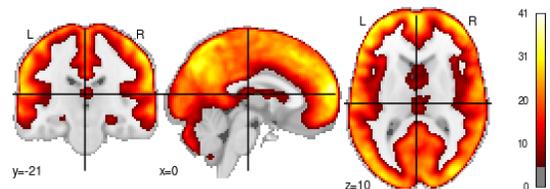


Figure 4: T -test showing voxels whose brain activity was significantly captured by the oscillatory regressors.

Comparison of Architectures

Each subject had two sessions of rsfMRI data, collected on two separate days. Each session was modeled individually, and then both sessions were combined on a subject level for the comparison analysis. Figure 5 illustrates the group-level log-likelihoods of the models in the rest condition. The figure presents *relative* log-likelihoods: the lowest log-likelihood is subtracted from all the others. As a result, the worst-fitting model always has a relative log-likelihood value of zero, with the best fitting model having the highest positive value.

Across both sessions, the CMC provides the best account of resting state brain activity, when compared against each of the six alternate structures. Because log-likelihood is not sensitive to model complexity, it is common to compute log likelihood in some penalized form. For example, the common Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) penalize likelihood by the number of parameters. Both measures assume, however, that parameter values are independently distributed, which is not the case for DCM models (for example, connectivity values for the same node tend to be correlated). For this reason, it is common to use a different, penalized form of likelihood known as Free Energy (Penny et al., 2012), which

accounts for non-independent parameters. The values reported in Figure 5 depict this penalized form of likelihood, and thus already account for varying model complexity.

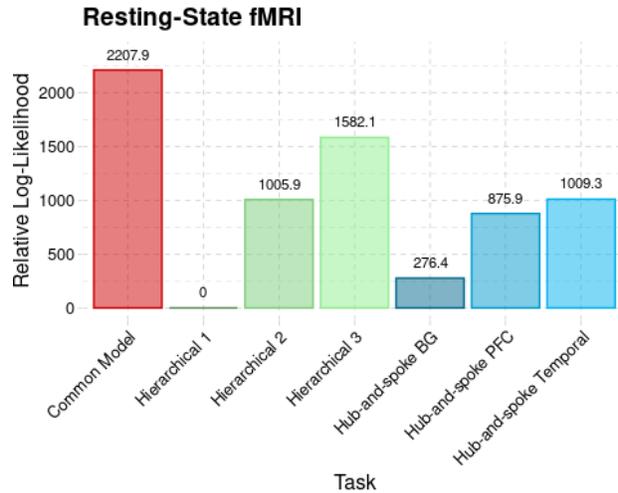


Figure 5: The log-likelihood of the CMC architecture compared to six alternate architectures across both sessions of rsfMRI data.

Analysis of Bayes Factors

Although the evidence in favor of the CMC is apparent, one might wonder exactly how significant the difference in log-likelihood is. To express log-likelihood in an interpretable form, we will use Bayes Factors (BF). The $BF_{1,2}$ between two models m_1 and m_2 is defined as:

$$BF_{1,2} = P(m_1 | \mathbf{x}) / P(m_2 | \mathbf{x})$$

In other words, the value of $BF_{1,2}$ represents the odds of model 1 fitting the data better than model 2. Given the definition of likelihood as $L(m|\mathbf{x}) = P(\mathbf{x}|m)$, $BF_{1,2}$ can be expressed as

$$BF_{1,2} = e^{\Delta L}$$

where $\Delta L = \log L(m_1 | \mathbf{x}) - \log L(m_2 | \mathbf{x})$ is the difference in log-likelihoods between model 1 and model 2. As a guideline, Kass and Raftery (1995) suggest that values of $BF > 20$ correspond to a value of $p < .05$ in a canonical null-hypothesis test and provide “strong” evidence in favor of model 1 over model 2, while values of $BF > 150$ provide “very strong” evidence. All of the BF values for the comparisons of the CMC against all the other models exceeded 10^{250} , indicating that the evidence in favor of the CMC is, in fact, overwhelming.

Random-Effects Analysis

Although the results provide strong evidence in favor of the CMC, it should be noted that they are not directly

comparable with the model comparison approach reported by Stocco et al. (2021). In the original paper, the authors compared the different architectures by measuring the relative probabilities that each architecture would fit any given participant (Stephan et al., 2009). This approach is conceptually different from the log-likelihood approach because it is based on relative, rather than absolute, fit to the data and because participants are considered as a random factor, thus giving different architectures the opportunity to fit different subgroups of participants.

To provide a better comparison to the original findings, we replicate the analysis method of Stocco et al. (2021) with the current resting-state data. The results are reported in Figure 6. In the figure, the curves represent the densities of the relative probabilities that each architecture would fit a participant. The superiority of the CMC is shown by the fact that its probability density function lies to the right of all other architectures. Architectures can be quantitatively compared in terms of exceedance probabilities, i.e. the probability that a point randomly sampled from their density distributions would have a higher probability than any other architectures. In this case, the Common Model had an exceedance probability of 96.4%, further confirming its superiority.

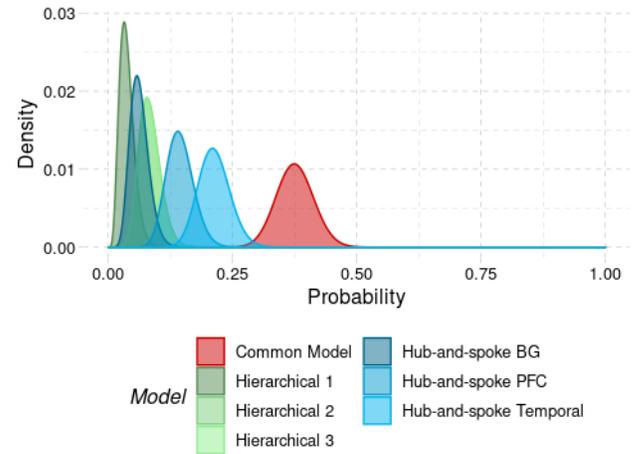


Figure 6. Probability densities that each architecture would best fit the data from a participant in our sample.

Discussion

The major finding of this paper is the apparent presence of an underlying structure of brain connectivity that predicts activity even during undirected and task free behavior. The implications of these results are broad.

First, they demonstrate the success in adapting a traditional DCM analysis to resting state through the use of simulated task events and externally generated ROIs, paving the way for future explorations of resting state data. In particular, the use of summarized fMRI data to determine ROIs presents the opportunity to explore increasingly complex model structures involving more specific brain areas. While the CMC provides the best

account of underlying connectivity, it remains only the best model of those that we have tested so far, and is deliberately composed of a few, high level components. The use of more localized ROIs opens the door to examining each component in greater detail; separating visual perception from auditory perception, for example, or decomposing the long term memory component into semantic and episodic memory. The DCM framework also allows models to account for modulatory connections between regions, which, while not used in this paper, provide further opportunities to define and specify a general purpose framework of cognition.

While specific ROIs will always differ slightly across subjects and tasks, the localized ROIs used in the present study represent a much smaller search space than the broad parcellation used to define ROIs in the original CMC study. The data from the original study should be reanalyzed using the more specific maps to ensure that the findings still hold, with the ultimate aim of defining more exact regions that can be used in future analyses.

The goal behind this study was to test large-scale architectures in a task-free paradigm, using only signals originating from spontaneous neural activity that would capture the intrinsic organization of the brain (Fox et al., 2005). Although this procedure has become the accepted standard in neuroimaging research, the extent to which resting state activity is truly spontaneous remains debated: even at rest, participants do typically engage in some form of thought, such as daydreaming or mind wandering. A recent computational model of mind wandering (Taatgen et al., 2021), for instance, argues that mind-wandering is generated by the occasional intrusion of task-unrelated goals and that, when activated, it triggers a cascade of mental processes, such as memory retrieval, that are similar to those required by canonical tasks. The fact that the same architecture that was found to best capture brain activity across multiple cognitive tasks (Stocco et al., 2021) also explains brain activity at rest supports the assumptions of this model and the idea that spontaneous thought follows the same patterns as task-directed thought.

The implications of a general framework for cognition that remains persistent in the resting state will significantly increase its applicability to other domains, such as computational psychiatry and neurology.

Acknowledgements

This research was supported by grant FA9550-19-1-0299 from the Air Force Office of Scientific Research to AS.

References

Anderson, J. R. (2007). How can the human mind occur in the physical universe? *Oxford University Press*.

Cole, M. W., Yarkoni, T., Repovš, G., Anticevic, A., & Braver, T. S. (2012). Global connectivity of prefrontal cortex predicts cognitive control and intelligence.

Journal of Neuroscience, 32(26), 8988-8999.

Cole, M. W., Ito, T., Bassett, D. S., & Schultz, D. H. (2016). Activity flow over resting-state networks shapes cognitive task activations. *Nature neuroscience*, 19(12), 1718-1726.

Di, X., & Biswal, B. B. (2014). Identifying the default mode network structure using dynamic causal modeling on resting-state functional magnetic resonance imaging. *Neuroimage*, 86, 53-59.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, 102(27), 9673-9678.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, 19(4), 1273-1302.

Gratton, C., Laumann, T. O., Nielsen, A. N., Greene, D. J., Gordon, E. M., Gilmore, A. W., ... & Petersen, S. E. (2018). Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation. *Neuron*, 98(2), 439-452.

Hohenfeld, C., Werner, C. J., & Reetz, K. (2018). Resting state connectivity in neurodegenerative disorders: Is there potential for an imaging biomarker? *NeuroImage: Clinical*, 18, 849-870.

Houk, J. C., Bastianen, C., Fansler, D., Fishbach, A., Fraser, D., Reber, P. J., ... & Simo, L. S. (2007). Action selection and refinement in subcortical loops through basal ganglia and cerebellum. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1485), 1573-1583.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430), 773-795.

Ketola, M., Thompson, S., Madhyastha, T., Grabowski, T., & Stocco, A. (2020). Applying the Common Model of Cognition to resting-state fMRI Leads to the identification of abnormal functional connectivity in Parkinson's Disease. *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society*.

Kotz, S. A., Schwartze, M., & Schmidt-Kassow, M. (2009). Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. *Cortex*, 45(8), 982-990.

Laird, J. E., Lebiere, C., & Rosenbloom, P. S. (2017). A Standard Model of the Mind: Toward a Common Computational Framework across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics. *AI Magazine*, 38(4), 13-26.

Nelder, J. A. & Mead, R. (1965). A simplex method for function minimization. *Comput. J.*, 308-313.

Penny, W. D. (2012). Comparing dynamic causal models using AIC, BIC and free energy. *Neuroimage*, 59(1), 319-330.

Scimeca, J. M., & Badre, D. (2012). Striatal contributions to declarative memory retrieval. *Neuron*, 75(3),

380-392.

- Seger, C. A. (2008). How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. *Neuroscience & Biobehavioral Reviews*, 32(2), 265-278.
- Shehzad, Z., Kelly, A. C., Reiss, P. T., Gee, D. G., Gotimer, K., Uddin, L. Q., ... & Milham, M. P. (2009). The resting brain: unconstrained yet reliable. *Cerebral cortex*, 19(10), 2209-2229.
- Steine-Hanson, Z., Koh, N., & Stocco, A. (2018). Refining the Common Model of Cognition Through Large Neuroscience Data. *Procedia Computer Science*, 145, 813–820.
- Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. *Neuroimage*, 46(4), 1004-1017.
- Stocco, A., Sibert, C., Steine-Hanson, Z., Koh, N., Laird, J. E., Lebiere, C. J., & Rosenbloom, P. (2021). Analysis of the Human Connectome Data Supports the Notion of A “Common Model of Cognition” for Human and Human-Like Intelligence Across Domains. *Neuroimage*, 235, 118035
- Taatgen, N. A., van Vugt, M. K., Daamen, J., Katidioti, I., Huijser, S., & Borst, J. P. (2021). The Resource-availability Model of Distraction and Mind-wandering. *Cognitive Systems Research*.
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., & WU-Minn HCP Consortium. (2013). The Human Connectome Project: An overview. *NeuroImage*, 80, 62–79.
- Visser, M., Jefferies, E., & Lambon Ralph, M. A. (2010). Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *Journal of cognitive neuroscience*, 22(6), 1083-1094.
- Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology*, 106(3), 1125–1165.

Collective Intelligence in Latent Imagination

Amit Kumar Singh (amitkumarsingh5@acm.org)

Cognition Researcher, Active Inference Lab
California, USA

Abstract

Intelligence is fundamentally the ability for an agent to infer causal dependencies in its environment. However, the precise conceptualization across systems and scales is a polemical question. The concept of “Intelligence” may as well refer to a quantitative measure of formal cognitive ability than to a qualitative property of skilled agency. This difficulty in understanding the concept compounds when we try to scale to descriptive and predictive models of collective behavior. While it is self-evident that groups may leverage pairwise interactions or their collective resources to tackle complex problems, is that process only the sum of individual intelligence or is the group intelligent in its own right? If the latter, what does it mean for the classical internalist conception of intelligence and agency? If the former, then what is the proper scale of analysis in systems of nested organization, such as human societies? This question can be approached rigorously through a non-reductive account of the physical processes underlying intelligence. Here I propose that the latent model framework (with active inference as intrinsic reward mechanism) framework is a promising approach that could live up to the multiple dimensions of adeptness required by any framework that would attempt to generalize cognition across scales. A statistical state model for mathematical state transitions can be built and can be used to further define cognitive model.

Keywords: Collective cognition; Mathematical modeling; Active Inference; State spaces; Latent spaces

Introduction

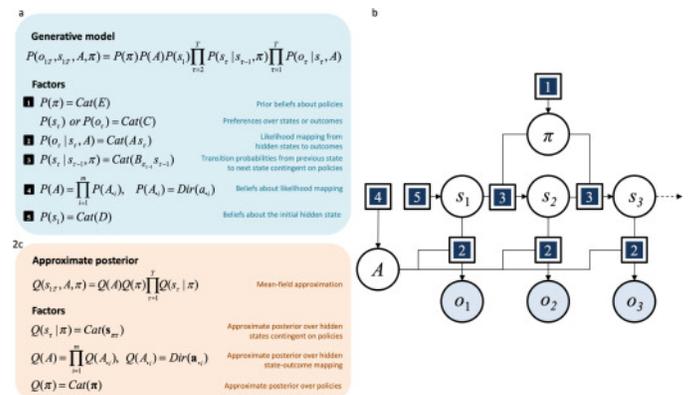
The view of mind as an experience generating machine or a generative model goes further back than just recent Machine learning breakthroughs (like Variational Autoencoders). This paper takes a similar route. Starting out with explaining possible state-space configurations, a world-model (World Models, Ha and Schmidhuber) is mapped in the latent space. This could be enough framework to explain individual actions in an environment, just like it does with many model-based approaches to Reinforcement Learning. But, in contrast to most RL approaches where reward is extrinsic and task structure changes with it which works well in specific RL environments, where rewards are intrinsic in the complex environment itself (Reward is enough, Silver et. al) It fails to explain intelligence at the collective level where the agents apart from the environment, have formed a dynamic between themselves too. This dynamic is represented through a collective latent which can be traversed in an abstract space by the agents of the collective for inference to eventually reach cognition as a collective.

Towards a Free Energy Agent Cognitive Model

Agent’s configuration at an instant is defined by its state with parameters interacting to form state variables. To establish a stable ground I invoke the Free-Energy principle. While we can argue about all derivations of intelligence, we can come to

standstill that the system exists. This is the basic formulation of the Free Energy principle and everything is deduced from this assumption with agent and environment in the frame. The Active inference principle can be framed as the minimisation of surprise (Friston, 2009) by perception and action. Here, in discrete state models - agents select from different possible courses of action (i.e., policies and their gradient of preferences) in order to realise the preferences and thus minimise the surprise that they expect to encounter in the future. This enables a Bayesian formulation of the perception–action cycle (Fuster, 1990): agents perceive the world by minimising variational free energy, ensuring their model is consistent with past observations, and act by minimising expected free energy, to make future sensations consistent with their model.

Active inference describes the dynamics of systems that persist (i.e., do not dissipate), and that can be statistically segregated from their environment—conditions which are satisfied by biological systems. Mathematically, the first condition means that the system is at non-equilibrium steady-state. This implies the existence of a steady-state probability density to which the system self-organises and returns to after perturbation (i.e., the agent’s preferences). The statistical segregation condition is the presence of a Markov blanket, where a set of variables through which states internal and external to the system interact (e.g., the skin is a Markov blanket for the human body).



Above is an example of a discrete state-space generative model which is how the agent represents the world. The generative model is a joint probability distribution over (hidden) states, outcomes and other variables that cause outcomes. In this representation, states unfold in time causing an observation at each time-step. The likelihood matrix [A] encodes the probabilities of state–outcome pairs. The policy (Pi) specifies which action to perform at each time-step. Note that the agent’s preferences may be specified either in terms of states or outcomes. It is important to distinguish between

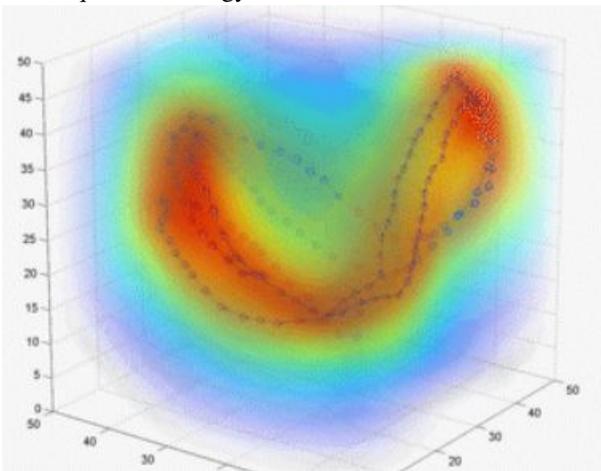
states (resp. outcomes) that are random variables, and the possible values that they can take in S (resp. in O), which we refer to as possible states (resp. possible outcomes). Note that this type of representation comprises a finite number of timesteps, actions, policies, states, outcomes, possible states and possible outcomes. The arrows represent causal relationships (i.e., conditional probability distributions). The variables highlighted in grey can be observed by the agent, while the remaining variables are inferred through approximate Bayesian inference and called hidden or latent variables. The Markov blanket of a random variable in a probabilistic graphical model are those variables that share a common factor.

Concept of an Unsupervised Loss function and Intrinsic Motivation

Since, most behaviour in individuals is directed or supervised by a goal, agents seem conformity and don't actually build on cognitive structures. Here, the agents is set to traverse in the abstract space without any prior goal or anything such, it forms a geometrical projection on its own manifold, as the process repeats, we can see what the function is being optimised for intrinsically over the timesteps.

Convergence on the Latent

Letting the agents interact with the environment unsupervised and intrinsically, they map out common latent abstract geometric manifolds (shown below). This is the moment of **Cognitive Convergence** on abstract space. The energy-based modeling view would be how a collective converges to a manifold of equivalent energy.



Results and further research directions

The Convergent latent can later manifest itself at common playground of lingual abstractions through language, implicit demographic knowledge through culture or any common cognitive structure that developed intrinsically within the collective. The framework can also be used to describe any process where goal(s) is(are) not explicit and system is set to evolve with random initial configurations.

References

Friston and Kiebel (Predictive coding under the free-energy principle), Philosophical Transactions of the Royal Society, Series B (Biological Sciences), 364
Fuster J.M (Prefrontal cortex and the bridging of temporal gaps in the perception-action cycle), Annals of the New York Academy of Sciences, 608
David Silver et. al (Reward is Enough), Artificial Intelligence - Volume 299.
Ha and Schmidhuber (World Models), NeurIPS 2018.
La Costa et. al (Active inference on discrete state-spaces: A synthesis), Journal of Mathematical Psychology - Volume 99

A cognitive computational model of collective search with social information

Sabina J. Sloman, Robert L. Goldstone and Cleotilde Gonzalez

For many if not most of the decisions we make in day-to-day life, such as picking a restaurant for dinner when traveling in a new city or choosing a university to attend or a job offer to accept, our knowledge about the valuations of possible outcomes is incomplete. However, many current decision theories specify how to make decisions when the possible outcomes and their valuations are fully specified.

When information is incomplete, the quality of that information and how we use it to inform our decisions is especially important. *Social information* is a critical source of information that we use to assess the costs and benefits of our actions: Usually, we extrapolate from not only our own experiences, but also from the experiences of others, e.g., anecdotes told by friends or reviews read online. The structure of the social environments in which we receive information from others—our social *networks*—can vary in important ways. While a friend’s anecdote may only be told to a few in a tight-knit social network, online reviews are available to everyone who cares to look. It turns out the structure of our social networks can have dramatic effects on our ability to identify and make the best decision possible.

In a 2008 experiment, Mason, Jones, and Goldstone explored the effect of participants’ social network structure on their abilities to identify the best decision given incomplete information. Mason, Jones, and Goldstone (2008) designed a spatial search task where participants selected integers on a number line and received points as a function of the number they guessed. Participants had access to not only their own outcome information, but also to the outcome information of their network neighbors. Mason et al. (2008) found that when the payoff function was smooth with an obvious global maximum—when extrapolating from a relatively small amount of information was usually sufficient to identify the best decision—members of more interconnected networks more frequently guessed within the global maximum. By contrast, for jagged problems with more than one local maximum—complex problems for which identifying the best decision usually required exploration—members of more dispersed networks more frequently guessed within the global maximum.¹

Our work builds on this result. We synthesize work from various areas of cognitive science into a computational cognitive model of search in a social context: the Social Interpolation Model (SIM). We then explore the implications of our model by running simulations of interacting agents whose behavior is determined by the SIM. By embedding these agents in the same task structure as the one designed by Mason et al. (2008), we explore how these dynamics are affected by the structure of the agents’ social networks.

All of our simulations were run with groups of 15 agents, arranged in the same four network structures as the participants in Mason et al. (2008)’s studies. On each of 15 consecutive rounds, these agents “guessed” integers between 0 and 100. On a given round, an agent’s guess was informed by the outcome information generated by their own previous guesses and the guesses of their network neighbors.²

The SIM posits that agents rely on similarity-based generalization (Shepard, 1987) to integrate the outcome information they’ve already seen in order to infer the number of points they are likely to receive from decisions whose payoffs are as-yet unknown. An agent’s probability of selecting a particular number is proportional to the number of points they think that decision is most likely to confer. The SIM has three free parameters, or avenues for individual difference: 1) the breadth of the agent’s generalization gradient, 2) the quality of the agent’s uninformed prior about unseen options, and 3) the degree to which the agent weights their own experiences more heavily than the experiences of others.

Like Mason et al. (2008), we find that network structure matters: Interconnected networks perform well when the best decision is easy to identify, but comparatively worse when the payoff function is more complex. We find that the most effective parameter settings also depend on the complexity of the problem: Agents who generalize broadly or attach a high value to unobserved options do well on more complex landscapes—and

¹The most interconnected network configuration was a fully-connected network, in which every agent had access to the outcome information for every other agent. The most dispersed network configuration was a regular lattice, in which group members were arranged in a circle and had access to the outcome information of only their two nearest neighbors (with some members also connected to others two steps away). Two other network structures were characterized by average path lengths that fall between the extremes defined by the fully-connected and regular lattice networks. See Mason et al. (2008) for complete details.

²While the full version of the SIM allows for retention and aggregation of outcome information from multiple rounds, agents in the simulations reported here retain and aggregate information from only the most recent round.

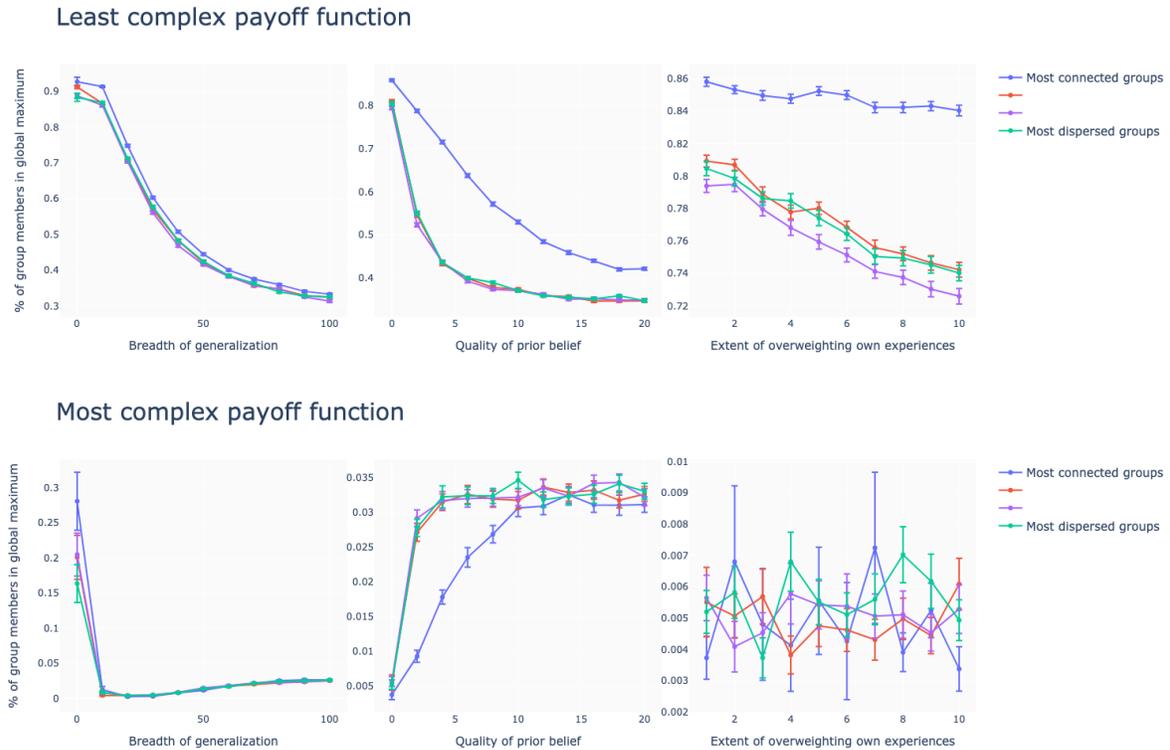


Figure 1: Percentage of simulated agents' guesses that were within one unit of variance of the payoff functions' global maximum.

worse on simpler landscapes (see Figure 1).³ Taken together, our results show that the effect of reducing the flow of social information by putting agents in a more dispersed social network can have similar effects to inducing the same agents to generalize more broadly or to be more optimistic about unseen options.

In summary, our main contributions are 1) the development of a computational cognitive model of search in a social context, 2) an exploration of the effects of the values of the SIM's free parameters, 3) the deployment of our theory in an agent-based model, and 4) an exploration of the effect of different social and reward environments on the SIM's dynamics. Our work has important practical and theoretical implications. Practically, our agent-based framework can allow exploration of the effects of different interventions in different contexts. Theoretically, we synthesize various areas of cognitive science into a single model that can make predictions about individual- and group-level behavior in decision-making environments characterized by incomplete information and the availability of social information.

References

- Mason, W. A., Jones, A., & Goldstone, R. L. (2008). Propagation of innovations in networked groups. *Journal of Experimental Psychology: General*, 137(3), 422–433. doi: 10.1037/a0012798
- Shepard, R. (1987, September). Toward a universal law of generalization for psychological science. *Science*, 237(4820), 1317–1323. doi: 10.1126/science.3629243

³Figure 1 shows that the effect of generalization on the complex payoff function is actually U-shaped: Agents in the most connected groups who generalize extremely narrowly perform the best. We speculate that this exception to the general trend reveals a tension between two components of success in this environment: On the one hand, exploration helps agents escape local maxima; on the other, narrow generalization more accurately reflects the ruggedness of the payoff function.

Modeling Aperture Passage Affordances in ACT-R 3D

Sterling Somers (sterling@sterlingsomers.com)

Department of Cognitive Science, Carleton University
Ottawa, ON(, Canada

Abstract

In this paper I present a model of aperture passage judgment (judgment of whether an agent can walk through an aperture, rotating shoulders as needed) and performance (initiation and termination of shoulder rotation while walking through an aperture) in ACT-R 3D. The model is adapted from Somers (2016) and represents a first attempt to unify findings across multiple experiments with a single model. The cognitive model is embodied in a robotics simulator, with motor control operated directly by the ACT-R model. The model exhibits an improved fit as compared to Somers (2016), in the same experiment, and a reasonable fit in an additional experiment, in exaggerated conditions.

Keywords: ACT-R; embodied cognition; motor control; cognitive modeling; affordances;

Introduction

Walking through narrow apertures, such as a narrow doorway, may require a shoulder rotation in order to reduce the frontal width of the body to afford passage. From an ecological psychology perspective, an ‘affordance’ is a property or set of properties (or relations, depending on author) in the environment that specify to an agent what actions are available (Chemero & Turvey, 2007; Chemero, 2003; Stoffregen, 2000; Şahin, Cakmak, Doğar, Uğur, & Üçoluk, 2007; Turvey, 1992). Research in support of affordances has come from a range of domains including stair-climbing (Warren, 1984), aperture passage (Warren & Whang, 1987), reaching (Carello, Groszofsky, Reichel, Solomon, & Turvey, 1989), grasping (Tucker & Ellis, 1998), and a number of sports abilities (Fajen, Riley, & Turvey, 2008).

A common theme in affordance research is to identify π -numbers that relate some dimension of the environment (E) with some dimension of body (A) as a ratio: $\pi = E/A$. These π -numbers are typically presented as support for Gibson’s notion of *direct perception* (Gibson, 1986). Direct perception is the claim that our actions are not mediated by strong, internal, semantically-laden representations of the environment. Affordances are, instead, presented to us when the properties of the environment match the action capabilities of the agent.

Aperture passage was first studied by Warren and Whang (1987). In their series of experiments they attempted to identify the π -number that modulates shoulder rotation when walking through apertures. Since Warren and Whang’s classical paper, there have been a number of follow-up experiments that either support or extend their findings. Fath and

Fajen (2011), for example, modify visual properties in a virtual environment, in aperture passage experiments, to identify a set of visual properties that contribute to the aperture passage affordance. A number of studies have investigated the aperture passage affordance for participants carrying objects (Wagman & Taylor, 2005; Wagman & Malek, 2007; Higuchi, Cinelli, Greig, & Patla, 2006; Higuchi, Seya, & Imanaka, 2012). Higuchi, Takada, Matsuura, and Imanaka (2004) studied passability judgments and aperture passage performance for novel wheelchair users. Finally, Chang, Wade, and Stoffregen (2009) studied passability judgments of people grouped in a dyad. In most cases these authors subscribe, to varying degrees of commitment, to Gibson’s theory of direct perception, and therefore offer very little with respect to an information processing description.

In recent work by Somers (2016, 2017), a processing description and accompanying computational cognitive model of the first experiment in Warren and Whang (1987) is provided. Introduced as proof-of-concept for the simulation environment, ACT-R 3D, the aperture-passage model proposes that aperture-passage judgments and aperture-passage performance rely on a comparison of the geometric properties of body schema and the geometric properties of the environment (Somers, 2017). While their model has a reasonable fit to the data in Warren and Whang (1987), given the results in Higuchi et al. (2012) (discussed below), one can anticipate that their model cannot account for aperture passage performance in exaggerated conditions. In this work we adapt their model to account for experiments by both Warren and Whang (1987) and Higuchi et al. (2012).

Aperture Passage Research

Warren and Whang (1987) performed a series of experiments aimed at showing that aperture passage is directly perceived. In their first experiment they had participants walk through apertures of various sizes, rotating their shoulders as needed. Participants were grouped according to size: *large* or *small*. Larger participants rotated their shoulders more than smaller participants when passing through apertures of equal width. When expressed as an aperture-width to shoulder-width ratio, however, group differences were eliminated, suggesting that shoulder rotation is modulated by the ratio between aperture width and shoulder width. This experiment established a critical ratio (π -number) of 1.3 at which participants, regard-

less of their size, would change from a forward posture to a posture that included a shoulder rotation. This π -ratio, they maintain, is a constant, used by an agent to determine when shoulder rotation is required.

The second and third experiments in Warren and Whang (1987) are aimed at establishing the source of optical information contributing to the passability affordance judgment. In these experiments the authors modify binocular/monocular vision, movement and non-movement conditions, as well as introduce an Ames-room-like illusion. These experiments are meant to establish that the perception of passability is scaled to body units as opposed to absolute size judgments (in some extrinsic dimensions). While these experiments are out of the scope of the models developed for this work, it is worth noting that the conclusions of these two experiments are not entirely incommensurate with the model as the model is agnostic with respect to the source of optical information contributing to the geometric comparison process.

Higuchi et al. (2012) had participants walk through apertures while carrying bars of varying lengths in order to exaggerate the frontal width of participants. While the authors align themselves theoretically with Warren and Whang (1987), with respect to direct perceptions, they also somewhat diverge, offering some insight about the control of rotation. They propose that the central nervous system controls rotation by maintaining a constant safety margin between the agent and the edges of the aperture. By exaggerating the length of the bar they are able to test whether rotation is extremely exaggerated (as would be the case if the π -ratio of 1.3 was used). Their reasoning is as follows:

Consider an agent, 40 cm in width. From Warren and Whang (1987) we know that an agent would rotate their shoulders at a π -ratio of 1.3, leaving a 6 cm safety margin. If that same person was carrying a bar 100 cm in length and rotated based upon the same ratio, then they would create a 15 cm safety margin $((100 * 1.3 - 100) / 2 = 15)$. This over-rotation would be markedly inefficient (Higuchi et al., 2012)

Instead, Higuchi et al. (2012) propose that the central nervous system controls rotation to maintain a consistent safety margin. Assuming a safety margin of 6cm, a π -ratio of 1.12 $((100 * 1.12 - 100)/2 = 6)$ should only be required for safe passage. That is, participants should only begin rotation their shoulders when the ratio between the aperture and themselves (including a bar) is 1.12, and should only continue rotation until they have established a 6cm gap between themselves (or the bar) and the edge of the aperture. Their hypothesis predicts: 1) that the amplitude of rotation should become smaller as width increases and π -ratio is maintained; and 2) that spatial margins should remain constant regardless of absolute size or π -ratio.

In their experiment they manipulated aperture ratio and agent widths by having the participants carry bars that modify their frontal width by either a factor of 1.5 or 2.5 (as well as a control condition, bar length 30cm). Aperture widths are set to create ratios of 0.9, 1.0, and 1.1 to encourage large ro-

tation. Authors found a main effect of bar length, such that the angle of rotation was smaller as bar length increased (addressing 1). With respect to spatial margin, they found a main effect of bar length.

ACT-R 3D

ACT-R 3D (Somers, 2016) is a time-synchronized simulation environment for the Python variant of ACT-R (Stewart & West, 2006) that consists of a middleware, a camera class, vision module, motor module, as well as a humanoid robot. The ACT-R 3D middleware is time-synchronized with the Mobile OpenRobots Simulation Engine (MORSE) (Echeverria, Lassabe, Degroote, & Lemaignan, 2011; Echeverria et al., 2012).

Vision ACT-R 3D adds a new camera class to MORSE, Geometric Camera, that provides a single, structured, retinotopic geometric description of the scene from the perspective of the agent. On the ACT-R side, an updated vision module, inspired by the SOS Vision System (West & Emond, 2002) in Python ACT-R, which makes accessible ‘features’ of the environment to the agent (the agent has no access to object labels). Currently the vision module has algorithms for detecting *obstacles* and *openings*. Requests to the vision module from the ACT-R production system are parameterized in order to filter information top-down. For example, when given a request for an *opening*, chunks that describe the minimum size of the opening are used as parameters for the request. If multiple features match the request (e.g. there are multiple openings), the returned chunk is selected based on a weighted random choice, weighted by a *salience* factor, as described by West and Emond (2002).

Motor Control The motor module in ACT-R 3D maintains a hierarchical, symbolic and numerical representation of body parts (currently only the ones being modelled). Each body part has degrees of freedom represented by minimum and maximum values on axes of rotation. As the agent moves its body, the minimum and maximum values achieved are stored in declarative memory, functioning as body schema. (Schwoebel, Branch Coslett, & Buxbaum, 2001; Schwoebel & Coslett, 2005; Coslett, Buxbaum, & Schwoebel, 2008). The motor module also includes functionality to provide proprioceptive feedback to estimate 3D body dimensions in a given posture. How the representations are achieved are currently beyond the scope of the module and is implemented with a measure of the agent’s bounding box. The bounding box values are stored with the body schema in declarative memory.

Further details about the time synchronous middleware, the Geometric Camera, vision module, motor module, as well as the simulated robot are available in (Somers, 2016).

Geometry-Based Affordances Theory

Presumably due to the theoretical commitments of ecological psychology, research into aperture passage (Warren & Whang, 1987; Wagman & Taylor, 2005; Higuchi et al.,

2012; Fath & Fajen, 2011) is generally sparse with respect to an overall information processing description. Warren and Whang (1987), for example, discuss optical information that might contribute to a passability judgment but miss critical details about cognitive control of the overall amplitude of rotation. Alternatively, Higuchi et al. (2012), suggest that the central nervous system maintains a safety margin between the shoulders and the edges of the aperture (which is a theory explored here), but do not offer a theory about the processes or representation involved.

The model presented in this work represents an instantiation of geometry-based affordances presented in (Somers, 2017). At a functional level, the theory proposes that a certain class of affordances are realized by an agent as a result of a comparison process that compares the geometric properties (width, depth, height) of some feature in the environment and the geometric properties of a current or stored body schema. For more details regarding geometry-based affordances and evidence of the role of body schema, see the work by (Somers, 2017).

The theory/model proposes four phases: 1) a body schema encoding phase, 2) passability judgment phase, 3) rotation initiation, and 4) rotation completion. The four phases are described below.

Body Schema Encoding Phase The body schema encoding phase occurs pre-experiment as part of the agents life. As instantiated in ACT-R 3D, body schemas are stored when a rotated joint reaches its minimum or maximum rotation along the principle axes of rotation. In the simulations, once a simulated robot is generated (according to the size constraints for the experiment), the robot performs shoulders rotations in each direction multiple times to encode the body schema in declarative memory.

Judgment Phase One of the processes not discussed in aperture passage literature is that passability judgments rely not only on the current frontal width but also, potentially some future frontal width, after the shoulders have been rotated. A π -ratio simply cannot account for passability judgments without also introducing either a representation, association process, or simulation process. The judgment phase in this theory/model results from two possible cases. In the first case, body geometry is estimated from a body schema of the current posture, and is used top-down in a visual search for an aperture of appropriate size. If the vision system is able to return a feature in the environment that meets those constraints, the returned aperture is deemed ‘passable’. If no environment feature is returned by the vision system, the second case proceeds.

In the second case, a (potential) series of memory requests are made for stored body schema that closely match the current body posture (e.g. standing, no shoulder rotation) and the current action capabilities (e.g. walking) but relaxed in

an increasing number of postural details. In the case of aperture passage, memory requests would be for a posture that affords walking, allowing for variation in torso posture (such as shoulder rotation). If a suitable schema is returned, the geometric properties of that schema are used top-down to filter the visual results in the manner described above.

Rotation Initiation Another aspect of performing aperture passage not discussed in the literature is how the rotation is initiated. In this phase, the agent is already walking towards the aperture, and in the model, rotation is initiated when the bottom-up vision system is triggered by the proximity to the aperture. When the edges of the aperture are within a multiple of the agent’s rotation radius, the vision system pushes information into the visual buffer, and the agent responds by carrying out a motor plan. The body schema retrieved in the *judgment* phase is maintained in working memory, and used at this point as the motor plan.

Rotation Completion The theory proposes that rotation completion is the result of a moment-to-moment comparison between body schema and optical information about the aperture. This is, to some degree, similar to the theory in Higuchi et al. (2012). The moment-to-moment comparison continues until frontal width of the agent is less than the width of the aperture. Although the body schema retrieved in the *judgment* phase, of fully rotated shoulders, was used as a goal state for the motor module, the agent need not always rotate the shoulders maximally. In other words, the goal state of the motor system was to fully rotate the shoulders, but a moment-to-moment visual update limits the rotation as a result of the comparison process. It is in this process that the current model differs from that of Somers (2016). In particular, the model presented by Somers, inspired by (Warren & Whang, 1987), multiplies the current body schema by a constant to overestimate body width. The model presented in this work favors a comparison process that maintains a safety margin, following the findings and theory of Higuchi et al. (2012).

Model and Experiments

In Somers (2016), the author used the same metric for passability judgment as for rotation completion. That is, their model ended rotation when it was determined that the agent’s frontal width, multiplied by a constant (1.139), was less than the width of the aperture. Given the experimental findings in (Higuchi et al., 2012), however, one can fully expect that the model would over rotate in exaggerated agent-width conditions, especially considering their model exhibited a mild over-rotation in large aperture conditions. In the following we present changes to the model in Somers (2016) and run experiments to for both Warren and Whang (1987) and Higuchi et al. (2012).

Model

The model described in this section goes through the four steps described above: body schema encoding, judgment, ro-

tation initiation, and rotation completion. One of the main factors in producing measurable behavior (rotation degree) is the temporal dynamics of the model. The temporal constraints imposed on the model due to the production system, motor module, and the vision module affect, in particular, when the model will initiate or terminate rotation, creating a source of variability. That said, the kinematics of the simulated-robot agents also has a major affect on degree of rotation.

One of the main assumptions across all models is that they rotate with a constant, instantaneous velocity of 120° per second. The only known aperture-passage study to report on the kinematics of shoulder rotation is from Fath and Fajen (2011), where participants are immersed in virtual environments. Fath and Fajen reported participants initiating rotation between 0.5 and 0.7s before reaching the aperture with rotation degree varying from approximately 20° and 60° . A parameter search was conducted with approximate values (from literature) and a rotation rate of 120° per second (the upper bound as described in Fath and Fajen) was used in all models. This is the same rotation rate used in the model by Somers (2016).

The other kinematic assumption in the model that has a major impact on the rotation prediction is walking rate. Warren and Whang (1987) provide a set of average walking rates in the four condition of their first experiment of: 1.29 m/s and 1.28 m/s (small vs. large) *normal* speed conditions and 1.61 m/s and 1.77 m/s (small vs large) in the *fast* speed conditions. The simulated robots moved at the average speeds reported in Warren and Whang (1987), according to size and speed, for all experiments.

There are three main parameters that affect the rotation in the model. *RadiusMultiplier* is used by the model to affect when to initiate rotation, bottom-up. The *RadiusMultiplier* parameter was set the same value as in the model in Somers (2016) (3.0). A new parameter was introduced for the purposes of this study: *VisionConstant*. The *VisionConstant* parameter represents the safety margin in Higuchi et al. (2012) and is set to 3cm accordingly ($3 \times 2 = 6\text{cm}$). Given those parameters as constants, a parameter search for the parameter *VisionMultiplier* was run. In the model by Somers (2016) the *VisionMultiplier* parameter was used both in the judgment phase and in the rotation completion phase, as a means of over-estimating body width. In this model, the parameter is only used in the judgment phase (to detect apertures) and after a coarse parameter search, for apertures of 40cm and 55cm , *VisionMultiplier* parameter was set to (1.36). Note, this value is similar to the π -ratio of 1.3 found by Warren and Whang (1987).

Experiment 1

We re-ran the experimental conditions from Somers (2016), a simulated version of the first experiment in Warren and Whang (1987). Warren and Whang had participants (group: small vs. large) walk through aperture of different sizes in two speed condition (normal vs fast). They found that par-

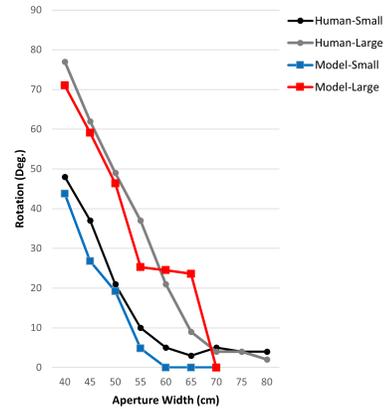


Figure 1: Human vs. Model, rotation angle by aperture width, normal speed. Black and gray line represent small and large human rotation (respectively). Blue and red lines represent small and large model (respectively).

ticipants rotated more in response to smaller apertures, that larger participants rotated more than smaller participants, and faster speed resulted in higher degrees of shoulder rotation.

Because there is a floor effect in human data, the models were only run through apertures up to maximum width of 70cm . All other experimental conditions in Warren and Whang (1987) were re-created as accurately as possible within the simulation environment. In the simulation there were 5 agents per group condition and agent sizes were chosen from a normal distribution centered around the mean human sizes for each group (40.4cm for small and 48.4cm for large) with a standard deviation as reported ($SD = 2.0\text{cm}$ for small and $SD = 0.7\text{cm}$ for large). Agents walked at the average speeds per group reported in Warren and Whang, as described above. There were a total of 20 agents, 10 per size group (large and small). Each agent walked through the apertures 15 times for a total of $15 * 20$ (agents) * 2 (speed) * 5 (apertures) = 3000 simulation runs.

Results (Ex 1) Because the original data from Warren and Whang (1987) was not available, limited analysis of fit is provided. A visual comparison between the results in Warren and Whang and the simulation runs are presented in Figure 1 for the slow condition and Figure 2 in the fast condition. A Pearson’s correlation on the means (as all data was not available) indicate a fit of 0.98 and 0.91 for the small and large agents in the normal speed condition; and 0.98 and 0.92 for small and large agents in the fast speed condition. A combined RMSE for large and small agents was 8.78° in the normal speed condition and 8.27° in the fast speed condition. In addition to comparative statistics, an ANOVA was run on the model data to see if the same main effects were present in the model as in the human data. Large participants had larger degrees of rotation than smaller participants. Participants rotated more

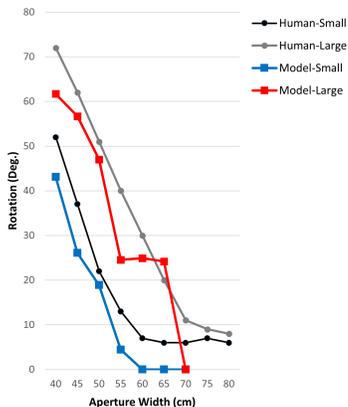


Figure 2: Human vs. Model, rotation angle by aperture width, fast speed. Black and gray line represent small and large human rotation (respectively). Blue and red lines represent small and large model (respectively).

for narrower apertures. These results are similar to Warren and Whang (1987). Unlike the human data, the model rotated less in fast speed conditions than in slow conditions.

Discussion (Ex 1) Visually the model has a reasonable fit to human data in both the normal and fast conditions. There is evident an over-rotation for large agents (red) at apertures 55cm, 60cm, and 65cm before no longer rotating at an aperture of 70cm. Two potential factors (and their combination) could account for this over rotation. First, delays caused by the constraints of the productions system can very easily lead quick rotation inaccuracy. Second, the rotation rate (120°per second) is the high-end of that reported by Fath and Fajen (2011). A more thorough fit for rotation rate could have been done but would not have been informative and would have been a parameter fitting exercise without rotation rate data. This model exhibits a better fit to the data than the model presented in Somers (2016).

Experiment 2

The model for the second experiment is the exact same model, with the exact same parameters as in experiment one. The only differences between them are the differences in the simulated robot which reflect the size of participants in Higuchi et al. (2012) and, as described above, the addition of a bar at agent-width ratios of 1.5, 2.5, as well as a control condition (30cm). There were a total of 10 agents, who each performed 15 trials of each aperture * bar combination.

Results (Ex 2) Figures 3 and 4 illustrate the mean angle of rotation and mean safety margins for both human and model data. Agents rotated less with larger bars, and rotated less at higher aperture ratios. The effect of bar and aperture ratio are both significant for the model ($ps < 0.01$). A Pearson’s correlation indicates a fit of 0.80 for rotation angle and 0.21 for safety margin. Note, however, that the model exhibits a

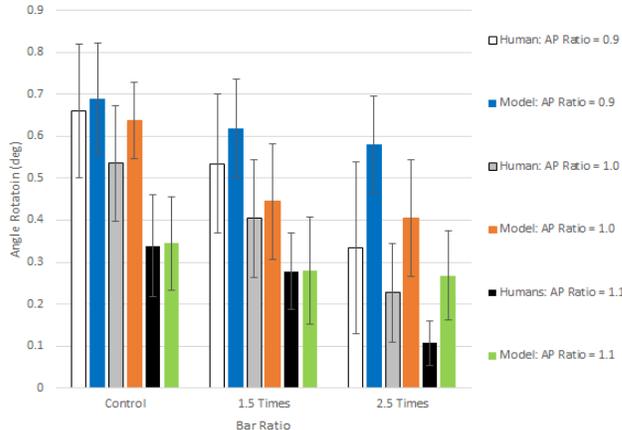


Figure 3: Human (grayscale) vs Model (color) rotation angle for bar ratios: control, 1.5, 2.5 and aperture ratios: 0.9, 1.0, and 1.1.

large over rotation at the 2.5 times bar condition. Excluding that condition, the Pearson’s correlation is 0.84 for absolute rotation and 0.89 with respect to safety margin. For the control and the 1.5 bar ratio condition, mean absolute rotation is comparable to human participants.

As shown in Figure 4, the model has a reasonable fit for mean spatial margin in both the control and 1.5 times condition. Agents leave greater spatial margins when carrying larger bars. All effects are significant ($ps < 0.01$).

RMSE, excluding the 2.5 bar condition was approximately 9 degrees of absolute rotation and approximately 2cm with respect to safety margin.

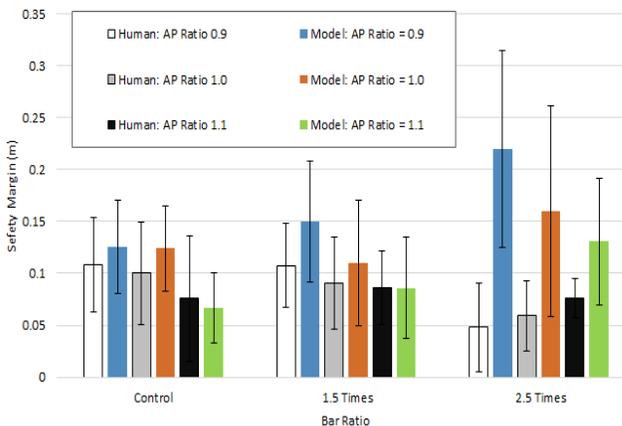


Figure 4: Human (grayscale) vs Model (color) safety margin for bar ratios: control, 1.5, 2.5 and aperture ratios: 0.9, 1.0, and 1.1.

Overall Discussion

Importantly, by implementing this research in a cognitive model, interesting questions are raised about the overall in-

formation processing involved in the task. Previous literature has largely overlooked the need to explain how apertures are judged as passable in some future posture. While body schema is one possible answer, and the one explored here, there could be other explanations worth investigating. Proposing body schema also requires a means of storage, a means of retrieval, detail on the representational content, and requisite processes provided in the present work.

Over-rotation is evident in the model across both experiments. It is very likely that the high rotation rate of 120° per second is a large contributor to the over rotation. For example, in the 2.5 times bar condition of experiment 2, human participants may be exhibiting more caution by either rotating or walking more slowly, and attending to the rotation more thoroughly. The qualitative change in pattern in the human data in Figure 4, at the 2.5 condition is at least suggestive that there may be an alternate strategy as compared to the other conditions. It is perfectly plausible that a more thorough parameter search could have resulted in a better model fit, however, to do so would not be well motivated, as the model is constrained by the physical and kinematic properties involved in the experiment. Alternatively, there may be low-level implementation details in the processing for the camera, due to calculations at such an obtuse angle, which could account for both the increase in variance and the higher means in Figure 4. Finally, of course, it could be that the theory in the model is wrong, and an alternate theory and set of processes is required for a unified explanation of the experiments.

The purpose of the research presented here is not to present an absolutely correct model but, rather, to motivate empirical research that could falsify it and, in turn, lead to refinements, or alternatives. Given the reliance on the temporal dynamics of the model, and the relationship to the physical and kinematic properties of rotation, this model motivates a more thorough account of the physical responses of participants, particularly rotation rate, as discussed above. There are, further, more qualitative observations from the model such as rotation initiation that could benefit from empirical measures. From a cognitive perspective, an alternate account of the bottom-up process for initiating rotation as proposed here, could include a more thorough motor plan, or some form of simulation that allows the agent to program the rotation initiation, rotation speed, and, possibly, the rotation termination, without moment-to-moment monitoring.

Finally the role of body schema in a cognitive model presents an interesting research direction. The implementation of body schema for this project, as body configurations stored in memory, is undoubtedly crude, however, it would be interesting to see the development of stronger motor control mechanisms in an architecture such as ACT-R. This is especially true with respect to modeling complex tasks, in complex environments, where processes such as aperture passage judgments enhance agents with capacities to make autonomous action decisions without requiring pre-labeling or apriori knowledge of simulation environments.

References

- Carello, C., Groszofsky, A., Reichel, F. D., Solomon, H. Y., & Turvey, M. T. (1989). Visually perceiving what is reachable. *Ecological psychology*, 1(1), 27–54.
- Chang, C.-h., Wade, M. G., & Stoffregen, T. A. (2009). Perceiving affordances for aperture passage in an environment–person–person system. *Journal of Motor Behavior*, 41(6), 495–500.
- Chemero, A. (2003). An outline of a theory of affordances. *Ecological psychology*, 15(2), 181–195.
- Chemero, A., & Turvey, M. T. (2007). Gibsonian affordances for roboticists. *Adaptive Behavior*, 15(4), 473–480.
- Coslett, H., Buxbaum, L. J., & Schwoebel, J. (2008). Accurate reaching after active but not passive movements of the hand: Evidence for forward modeling. *Behavioural Neurology*, 19(3), 117–125.
- Echeverria, G., Lassabe, N., Degroote, A., & Lemaignan, S. (2011). Modular open robots simulation engine: Morse. In *2011 IEEE International Conference on Robotics and Automation* (pp. 46–51).
- Echeverria, G., Lemaignan, S., Degroote, A., Lacroix, S., Karg, M., Koch, P., ... Stinckwich, S. (2012). Simulating complex robotic scenarios with Morse. In *International conference on simulation, modeling, and programming for autonomous robots* (pp. 197–208).
- Fajen, B. R., Riley, M. A., & Turvey, M. T. (2008). Information, affordances, and the control of action in sport. *International Journal of Sport Psychology*, 40(1), 79–107.
- Fath, A. J., & Fajen, B. R. (2011). Static and dynamic visual information about the size and passability of an aperture. *Perception*, 40(8), 887–904.
- Gibson, J. J. (1986). *The ecological approach to visual perception*. Psychology Press.
- Higuchi, T., Cinelli, M. E., Greig, M. A., & Patla, A. E. (2006). Locomotion through apertures when wider space for locomotion is necessary: adaptation to artificially altered bodily states. *Experimental Brain Research*, 175(1), 50–59.
- Higuchi, T., Seya, Y., & Imanaka, K. (2012). Rule for scaling shoulder rotation angles while walking through apertures. *PloS one*, 7(10), e48123.
- Higuchi, T., Takada, H., Matsuura, Y., & Imanaka, K. (2004). Visual estimation of spatial requirements for locomotion in novice wheelchair users. *Journal of Experimental Psychology: Applied*, 10(1), 55.
- Şahin, E., Cakmak, M., Doğar, M. R., Uğur, E., & Üçoluk, G. (2007). To afford or not to afford: A new formalization of affordances toward affordance-based robot control. *Adaptive Behavior*, 15(4), 447–472.
- Schwoebel, J., Branch Coslett, H., & Buxbaum, L. J. (2001). Compensatory coding of body part location in autotopagnosia: Evidence for extrinsic egocentric coding. *Cognitive Neuropsychology*, 18(4), 363–381.
- Schwoebel, J., & Coslett, H. B. (2005). Evidence for multiple, distinct representations of the human body. *Journal of*

- cognitive neuroscience*, 17(4), 543–553.
- Somers, S. (2016). Act-r 3d: A 3d simulation environment for python act-r. In *14th international conference on cognitive modeling* (pp. 107–112).
- Somers, S. (2017). Geometry-based affordances. In *Proceedings of the 39th annual conference of the cognitive science society* (p. 3221-3226).
- Stewart, T. C., & West, R. L. (2006). Deconstructing act-r. In *Proceedings of the seventh international conference on cognitive modeling* (pp. 298–303).
- Stoffregen, T. A. (2000). Affordances and events. *Ecological psychology*, 12(1), 1–28.
- Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human perception and performance*, 24(3), 830.
- Turvey, M. T. (1992). Affordances and prospective control: An outline of the ontology. *Ecological psychology*, 4(3), 173–187.
- Wagman, J. B., & Malek, E. A. (2007). Perception of whether an object can be carried through an aperture depends on anticipated speed. *Experimental Psychology*, 54(1), 54.
- Wagman, J. B., & Taylor, K. R. (2005). Perceiving affordances for aperture crossing for the person-plus-object system. *Ecological Psychology*, 17(2), 105–130.
- Warren, W. H. (1984). Perceiving affordances: visual guidance of stair climbing. *Journal of experimental psychology: Human perception and performance*, 10(5), 683.
- Warren, W. H., & Whang, S. (1987). Visual guidance of walking through apertures: body-scaled information for affordances. *Journal of experimental psychology: human perception and performance*, 13(3), 371.
- West, R., & Emond, B. (2002). Sos: A simple operating system for modeling hci with act-r. In *Seventh annual act-r workshop*. pittsburg, pa.

How Good Can an Individual’s Conclusion Endorsement be Predicted?

Sara Todorovikj (todorovs@cs.uni-freiburg.de)

Marco Ragni (ragni@cs.uni-freiburg.de)

Cognitive Computation Lab, Technical Faculty, University of Freiburg
 Danish Institute for Advanced Study, University of Southern Denmark

Abstract

Reasoning about conditional statements is relevant in science, culture, and our everyday life. It has been shown that humans do deviate from a classical logical interpretation of conditionals. Consequently, in the past years a number of cognitive models based on Bayesian or mental model approaches have been developed, whose performance is normally judged based on their ability to fit aggregate data of participants. Here, we diverge by focusing on the *individual* instead. Moreover, we propose a different model testing paradigm by analyzing on an existing large data set, how good current models are in *predicting* an endorsement of an individual reasoner on a scale from 0 to 100%. Towards this goal we reanalyze the data by rigorously distinguishing between test and training data set, by making existing models for conditional reasoning predictable such as the Dual Source Model (Singmann, Klauer, & Beller, 2016) and a model by Oaksford, Chater, and Larkin (2000). We also implement a modeling idea of Pearl based on possible worlds. We can show that all three models perform equally good in predicting an individual reasoner’s endorsement and that they meet an empirical baseline (the median of the most frequent answer). A discussion on the gained insights in understanding conditional reasoning concludes the paper.

Keywords: Predictive modeling; cognitive modeling; conditional reasoning

Introduction

In order to understand how human cognition works, a variety of cognitive models have been developed throughout the years and fitted to various experimental data. For example, consider the following reasoning task about a conditional statement (c.f., Singmann et al., 2016):

If a balloon is pricked with a needle, then it will pop.
 A balloon is pricked with a needle.

How likely is it that it will pop?

Your task would be to provide an answer between 0 and 100%. Now, imagine that a cognitive model is provided with the same task and makes a *prediction* of your response. Given experimental data, we propose that cognitive models are applied in such a predictive setting to each *individual*, as illustrated in Fig. 1. Comparing the true response and the prediction for all participants leads to a novel approach of cognitive model performance evaluation.

Motivated by the idea of Feynman that in order to fully understand something, one needs to be able to re-create it, Riesterer, Brand, and Ragni (2020) introduce a predictive modeling task in the syllogistic reasoning domain. They evaluated the predictive performance of syllogistic theories using

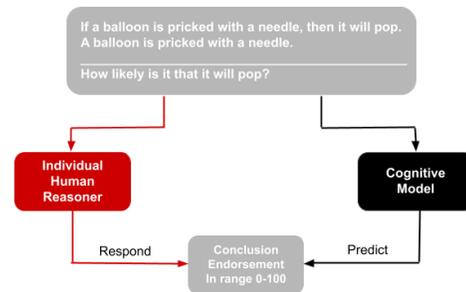


Figure 1: Predictive modeling task for endorsement rates

a modeling framework¹ called CCOBRA. Riesterer et al.’s (2020) focus is on the syllogistic domain, where a reasoner has only 9 answer options, meaning that a model either predicts the correct answer out of the possible 9 or not. This is where the scenario differs in our case. Here, we are dealing with a more complicated predictive task – one for *endorsements* that can be *any* value in the range 0-100. To understand the complexity of this task, consider the previously given reasoning task example once again. Since you are provided with the rule that if a balloon is pricked, it will pop, you would most likely gravitate towards answering with a 100%. But what happens, if:

The balloon is without air, i.e., empty.

Then the balloon would *not* pop and you might give an answer that is less than 100%. Such aspects are called *disablers*. If an individual is aware of many disablers, their conclusion *endorsement* might be lower. On the other hand, there can be additional cases, called *alternatives*:

A balloon can also pop, when it is pricked with something else than a needle.

Hence, depending on the cases different reasoners *have* in their minds, the given responses might differ. This introduces challenges when trying to predict how much a specific reasoner endorses a possible conclusion.

Existing models and their comparison

In this paper we focus on the *conditional* reasoning domain. Conditionals are statements of the form “If X then Y” (also written as $X \rightarrow Y$, where X is called the *antecedent* and Y the

¹<https://orca.informatik.uni-freiburg.de/ccobra/>

consequent), often used to describe a causal relationship between two propositions X and Y. Given a conditional (called *major premise*) and a current state of a proposition (called *minor premise*), a *conclusion* can be inferred about the state of the other proposition. There are four inference forms: *modus ponens* (MP), *modus tollens* (MT), *affirming the consequent* (AC) and *denying the antecedent* (DA), as shown in Table 1.

Table 1: Inference Forms

Name	MP	AC	DA	MT
Premise 1	$X \rightarrow Y$	$X \rightarrow Y$	$X \rightarrow Y$	$X \rightarrow Y$
Premise 2	X	Y	$\neg X$	$\neg Y$
Conclusion	Y	X	$\neg Y$	$\neg X$

Singmann et al. (2016) studied the endorsements of the respective conclusions for the four inference forms in four experiments. Three of them focus on contents with varying amounts of disablers and alternatives. The fourth experiment introduces speaker expertise. We want uniform data, so we do not consider it. The authors also present a performance comparison of *Bayesian* modeling approaches for conditional reasoning. They are built upon the idea that a conditional “If X then Y” is understood by a conditional probability $P(Y|X)$.

Oaksford et al. (2000) proposed that reasoning about a conditional rule can be modeled by the three parameters $P(X)$, $P(Y)$ and $P(\neg Y|X)$, the last one allowing for exceptions. Two extended versions using one and two additional exception parameters (Oaksford & Chater, 2007) and a model based on the Kullback-Leibler-Distance, have been statistically compared to a newly developed model – the Dual-Source-Model (DSM) – that assumes that individuals integrate two different kinds of processes: A knowledge-based component where they take Oaksford et al.’s (2000) approach and extend it with an additional form-based component, integrating both with a weight λ . We will present and explain the technicalities of these models in a following section.

While explicitly stating that model comparison should take model fit and model flexibility into account, due to the lack of Maximum Likelihood Estimation abilities for AIC/BIC, only a model fit using R^2 has been computed (Singmann et al., 2016). The R^2 goodness-of-fit values for the four models were used in a Linear Mixed Model (LMM) with random effects. Overall the DSM had the highest R^2 , meaning that it was able to account for the highest percentage of variance, i.e., it had the best performance. So far the models have been evaluated on a statistical level given the analysis approach based on the R^2 and the LMM. In the following, however, we will focus on process aspects and – we will analyze if a model queried for a yet untested person is even able to predict an endorsement of a conclusion from 0 to a 100%.

Our goal: Evaluating the predictive power of models

The current state of analysis does not convey yet, *if* the described models are predictive. When provided with observa-

tions on other participants’ endorsement answers to a set of reasoning problems (= *training data set*), a cognitive model is called *predictive* for a (untested) reasoner, if it can correctly predict the inference endorsements (between 0 and 100%) for those problems (= *test data set*). This is rather easy for a yes/no question, as we only have two answer options for the model’s prediction. However, it is *much more* challenging to develop a predictive setting for endorsement rates that range in the interval 0 - 100. Hence, this paper’s first research question is: How can we *develop a predictive task setting and evaluate the predictions* and how can we adapt and evaluate the existing models to provide this prediction?

Our second research question – as current models are probabilistic – is it possible to have a cognitive model based on mental models? This is often questioned, as endorsement problems are usually considered *new paradigm*. Pearl has suggested approaches that combine a model structure with probabilities, which we will implement and compare too.

The paper is structured as follows: First, we present existing experimental data and Bayesian cognitive models for conditional reasoning. Second, we present an idea of Pearl, which we adapt to represent inference form endorsements. Third, we elaborate on how the benchmark was implemented. To conclude the paper, we present its predictive results, followed by a discussion.

Data and Cognitive Models for Conditionals

We consider the experimental data provided in Singmann et al. (2016)², specifically the Experiments 1, 3a and 3b with 199 participants. In Exp. 3a and 3b, participants are divided in three groups. In two groups, participants are provided additional information in the form of alternatives and disablers, whereas the participants in the last group are provided only with the conditional task. All three experiments use the same four contents that have a varying amount of disablers and alternatives, both quantified with ‘Few’ and ‘Many’, shown in Table 2. The participants’ task was to provide endorsement rates for the four inference forms as a probability in the range 0 - 100%. Each content is presented as a full conditional inference and as a reduced inference, i.e., no major premise, e.g., MP:

A balloon is pricked with a needle.

How likely is it that it will pop?

Bayesian Cognitive Models

In the 60s a *deductive* path of cognitive modeling was followed, based on the assumption that logic is the basis for reasoning (Evans & Over, 2004). However, with time it has been shown that humans deviate from logic when given deductive reasoning tasks, and therefore, their responses are deemed false. That motivated the development of a new, Bayesian paradigm, where the models are based on probabilities and allow for background knowledge to be integrated when reasoning (Oaksford & Chater, 2020).

²The data can be found at <https://osf.io/zcdfq>.

Table 2: Contents used in Singmann et al. (2016) experiments.

Keyword	Content	Disablers	Alternatives
Predator	If a predator is hungry, then it will search for prey.	Few	Few
Balloon	If a balloon is pricked with a needle, then it will pop.	Few	Many
Girl	If a girl has sexual intercourse, then she will be pregnant.	Many	Few
Coke	If a person drinks a lot of coke, then the person will gain weight.	Many	Many

Oaksford et al.’s (2000) model (OC) Oaksford et al. (2000); Oaksford and Chater (2020) propose a probabilistic model for conditional reasoning. By using a 2×2 contingency table, as in Table 3, they represent conditional rules, where $a = P(X)$ and $b = P(Y)$, probabilities of the antecedent and consequent, respectively and $\varepsilon = P(\neg Y|X)$ is the exception parameter.

Table 3: Contingency table for a conditional rule “If X then Y” Oaksford et al. (2000). There are three parameters: the probability of the antecedent $P(X)$ denoted by a ; the probability of the consequent $P(Y)$ denoted by b ; and a third parameter ε for the probability of the exception $P(\neg Y|X)$.

	Y	$\neg Y$
X	$a(1 - \varepsilon)$	$a\varepsilon$
$\neg X$	$b - a(1 - \varepsilon)$	$(1 - b) - a\varepsilon$

Derived from Table 3, this model uses the following equations for inference endorsement:

$$\text{MP: } P(Y|X) = 1 - \varepsilon \quad \text{DA: } P(\neg Y|\neg X) = \frac{1 - b - a \cdot \varepsilon}{1 - a}$$

$$\text{AC: } P(X|Y) = \frac{a(1 - \varepsilon)}{b} \quad \text{MT: } P(\neg X|\neg Y) = \frac{1 - b - a \cdot \varepsilon}{1 - b}$$

As already mentioned, Oaksford and Chater (2007) present a more sophisticated version of this model. We decide to still take the original 2000 variant into consideration as the DSM builds up on it, as explained in the following.

Dual-Source Model (DSM) The DSM (Singmann et al., 2016) is an extension of Oaksford et al.’s (2000) model. It assumes that individuals integrate two different kinds of information: background knowledge about the content and information related to the logical form of the inference. It uses three types of parameters:

$\xi(C, x)$ – knowledge-based component, depending on the content C and inference x , i.e. how much does an individual endorse an inference solely based on their background knowledge about the content

$\tau(x)$ – form-based component, reflecting the subjective probability of the inference form x , i.e. how much does an individual believe in the validity of an inference regardless of the content

λ – a weight given to the form-based component (integrating $\xi(C, x)$ and $\tau(x)$ using Bayesian model averaging)

Applying the DSM to experimental data requires that participants have given endorsements to both a *reduced inference*

and a *full conditional inference*. The model expresses the reduced inference endorsement through its knowledge-based component for content C and inference x :

$$E_r(C, x) = \xi(C, x)$$

The $\xi(C, x)$ parameters are obtained by using Oaksford et al.’s (2000) equations, as shown above. Then, the endorsement of the full inference x with content C is given by:

$$E_f(C, x) = \lambda \cdot \{\tau(x) + (1 - \tau(x)) \cdot \xi(C, x)\} + (1 - \lambda) \cdot \xi(C, x)$$

The λ parameter determines how much do individuals rely on form validity versus their background knowledge. $\tau(x)$ is the degree of belief in the full inference form. In case of uncertainties concerning the inference, the individual falls back to their background knowledge, through the weight $(1 - \tau(x))$ given to the knowledge-based component.

Models and Probabilities: Applying an Idea of Pearl

ε -semantics Pearl (1991) introduced ε -semantics, a ‘formal framework for belief revision’, where belief statements are interpreted as statements of high probability and belief revision shapes current beliefs on newly available evidence. This approach seems to be most fruitful in our case, because disablers or alternatives can be such ‘updates’. The idea of Pearl is based on the idea of possible worlds (or models) that can be assigned a probabilistic assignment (Pearl, 1991, p. 5):

“Let L be the language of propositional formulas, and let a *truth-valuation* for L be a function t , that maps the sentences in L to the set $\{1, 0\}$, (1 for ‘true’, 0 for ‘false’). To define a probability assignment over the sentences in L , we regard each truth valuation t as a world w and define $P(w)$ such that $\sum_w P(w) = 1$. This assigns a probability measure to each sentence l of L .”

Before diving into our application of Pearl’s idea, we will briefly touch upon mental models. A mental model consists of the truth states of the premise’s propositions. Given a conditional premise “If X then Y”, the initial mental model that an individual would construct is the one where both propositions are true, i.e. XY.

The Mental Model Theory (MMT) (Johnson-Laird & Byrne, 1991, 2002; Johnson-Laird, Khemlani, & Goodwin, 2015) assumes that once the initial model is created it triggers the recollection of relevant facts and knowledge. Those facts can either serve as evidence that the initial model is correct or will stimulate a search for alternatives leading to a second process where an extended mental model representation is obtained, also called a fleshed-out representation. It

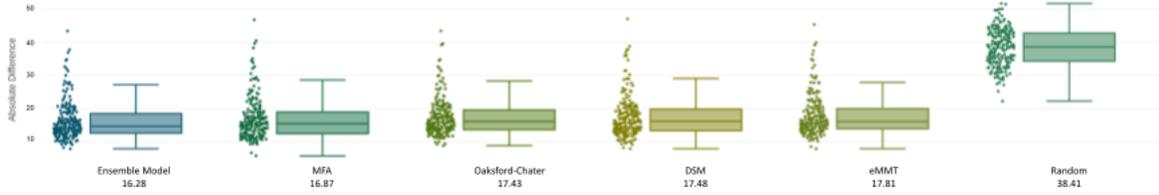


Figure 2: Boxplot depicting individual performance through the absolute difference between the predicted value and the true response. Overall mean absolute difference beneath the model’s name. The comparable performance between OC, DSM and ϵ -MMT points to (a partial) functional equivalence.

contains models where X is false ($\neg X$), as shown in Table 4a. This representation consists of all possible combinations of truth-values for X and Y for which the conditional “If X then Y” is true, which Johnson-Laird and Byrne (2002) call the *principle of truth*. This coincides with the material implication definition which is the leading interpretation of conditionals in the deductive paradigm.

The ϵ -MMT

ϵ -MMT takes the mental model representation of *all* the conditional’s propositions’ truth state combinations, which we will refer to as *possible worlds*. In contrast to MMT, it also allows for the world $X\neg Y$ to exist, thus abandoning the material implication interpretation. Given a premise containing two propositions, X and Y, all possible worlds described by the premise along with the corresponding probability values are shown in Table 4b. Given a conditional “If a balloon is pricked with a needle then it pops”, the probability of the world ω_2 , where the balloon is not pricked with a needle ($X = 0$) and it pops ($Y = 1$) is p_2 .

Table 4: Representations of a conditional premise “If X then Y” with mental models and as possible worlds.

Mental M.		Fleshed-out M.		World	X	Y	P
X	Y	X	Y	ω_1	0	0	p_1
...		$\neg X$	$\neg Y$	ω_2	0	1	p_2
		$\neg X$	Y	ω_3	1	0	p_3
				ω_4	1	1	p_4

(a) Johnson-Laird and Byrne (2002)

(b) Possible worlds, probability distribution P and values $p_i, i \in (1, 2, 3, 4)$

ϵ -MMT follows the same approach of previous accounts in the Bayesian paradigm, e.g. Oaksford et al. (2000), assuming that an individual’s inference form endorsement can be expressed as a conditional probability of the conclusion given the minor premise.

$$P(\beta|\alpha) = \frac{P(\alpha \wedge \beta)}{P(\alpha)} \tag{1}$$

Following the definition of conditional probability, as shown in Eq.1 the four expressions shown below are ob-

tained. They describe the endorsement of the four inference forms through the probability distribution P of the conditional’s worlds (Table 4b):

$$\begin{aligned} \text{MP: } P(Y|X) &= \frac{p_4}{p_3 + p_4} & \text{DA: } P(\neg Y|\neg X) &= \frac{p_1}{p_1 + p_2} \\ \text{AC: } P(X|Y) &= \frac{p_4}{p_2 + p_4} & \text{MT: } P(\neg X|\neg Y) &= \frac{p_1}{p_1 + p_3} \end{aligned}$$

The parameters are bound by their sum, $\sum_i p_i = 1$, meaning that the number of free parameters for modeling one task is three. Total number of parameters to model an *individual* hence depends on the number of tasks they have to complete.

Benchmark

In order to evaluate the three presented cognitive models, we implemented a benchmark within the framework CCOBRA, following Riesterer et al.’s (2020) approach. As already mentioned, their focus is on the syllogistic domain, where a model either predicts the correct answer out of the possible 9 or not, whereas our goal is to predict a value in the range 0-100. This poses a difficulty in adopting the same idea of judging a model based on whether it predicted the exactly correct answer or not. Instead, we are interested in *how close* the model’s prediction is to the true reasoner’s answer. The framework was extended to calculate the *absolute difference* between answers and predictions, rather than check for accuracy. In their benchmark, a theory is considered to have a good performance if it has a high accuracy rate. In our case, a cognitive model aims for a low absolute difference.

Generally, in order for the cognitive model to be able to predict a response as accurately as possible, it needs to be exposed to already existing data, i.e. a training set, from which it can learn. Here, we used Singmann et al.’s (2016) experimental data which we presented earlier. Since all three experiments have the same contents, we combined their data into one set, as the authors did in their original study. When provided with the same data for both training and testing, as in our case, the CCOBRA framework uses a leave-one-out cross-validation method – models are fitted on every participant, except the one whose answers are to be predicted. The same process is repeated for each participant.

In the training phase, we fit the models to the participants’

Table 5: Medians of the models’ parameters per task and conditional presentation form. Values discussed below are in bold.

Form	Task	ϵ -MMT ¹				Oaksford-Chater ²				Dual-Source Model ³		
		p_1	p_2	p_3	p_4	a	b	ϵ	$\xi(C,MP)$	$\xi(C,AC)$	$\xi(C,DA)$	$\xi(C,MT)$
Reduced Inference	Predator	.44	.06	.03	.59	.56	.60	.10	.90	.85	.80	.86
	Balloon	.48	.12	.05	.38	.38	.48	.12	.88	.70	.77	.91
	Girl	.23	.06	.42	.13	.63	.23	.68	.33	.87	.92	.45
	Coke	.27	.20	.14	.29	.47	.53	.37	.63	.56	.55	.63
Conditional Inference	Predator	.46	.05	.02	.62	.56	.59	.08				
	Balloon	.49	.07	.02	.53	.46	.53	.08				
	Girl	.32	.06	.22	.35	.60	.41	.39				
	Coke	.33	.15	.06	.40	.47	.55	.23				

¹ Probabilities of possible worlds ω_i , see Table 4b, we have $p_1 = P(\omega_1)$, $p_2 = P(\omega_2)$, $p_3 = P(\omega_3)$, $p_4 = P(\omega_4)$, note that only 3 parameters are necessary because of $\sum_i p_i = 1$; ² The three parameter values are: $a = P(X)$, $b = P(Y)$, $\epsilon = P(\neg Y|X)$; ³ Knowledge-based parameters $\xi(C,x)$ for content C and inference x . The same values are used in the conditional case.

answers by optimizing the models’ parameter values such that the absolute difference between the predicted answer and the reasoner’s response is minimized. In order to do that we used Python’s `scipy.optimize.minimize`³ with the method Sequential Least Squares Programming (SLSQP). This method was chosen because it allows for constrained minimization.

Following Riesterer et al. (2020), we included a Random model as a lower bound, which in our case gives a random value in the range 0-100 as a prediction. Our models do not adapt to the individual, so we also included a Most Frequent Answer (MFA) model as an upper bound. In old paradigm experiments such a model would count the number of times an inference has been accepted or rejected and would predict the outcome that was most frequent. However, now we have a far more complex situation, dealing with a big range of values, to which we had to adapt this idea by having the MFA model give the *median* of the responses as a prediction.

Predictive Modeling Results

We judge a model’s performance by the mean of the absolute differences between the model’s predictions and the individuals’ answers. A lower absolute difference indicates more accurate predictions and therefore, better performance.

Figure 2 illustrates the model performance for each individual. The probabilistic models have similar results, all three greatly outperforming the Random model, while being comparable to the MFA model, but not better. OC and the DSM, both established models in the current Bayesian paradigm give an impressive performance. But, now we can also see that ϵ -MMT, a *model-based approach* is a valuable competitor in this probabilistic paradigm.

Having a predictive performance that is as good as an empirical model is an accomplishment for the probabilistic cognitive models. However, if we compare only the three of them – their performance is not very different. So, we proceed with the analysis by investigating the models’ parameter values and how they aid in explaining the individuals’ conditional interpretations. The median values of the models’

parameters are shown in Table 5. In the reduced inference case participants are not provided with a rule, so their background knowledge is more prominent and that is reflected in the parameter values. We use now X for the antecedent of a conditional, and Y for its consequent (“If X then Y”). In the case of ϵ -MMT, the parameter p_2 describes the probability of the world ω_2 where Y happens even if X does not and its values are higher for tasks with ‘Many’ alternatives, in contrast to ‘Few’. The parameter p_3 , on the other hand, is the probability of the world ω_3 where X is true, however Y is not and through higher values shows the presence of ‘Many’ disablers. It can be seen how when a conditional has been provided, the belief in these two worlds diminishes. For OC, the most noticeable impact is on the ϵ parameter which is the probability of the exception $P(\neg Y|X)$. Its values are exceptionally higher for tasks with ‘Many’ disablers. A lower value for $a = P(X)$ is present in the case of ‘Many’ alternatives, showing that X does not need to be true for Y to happen. Likewise, $b = P(Y)$ reflects the presence of ‘Many’ disablers which would prevent Y from occurring. The influence of alternatives and disablers is reflected in the conditional case as well, though at a smaller scale due to the conditional rule restricting the integration of background knowledge, similarly to ϵ -MMT. For the DSM, we have the four knowledge-based parameters $\xi(C,x)$ for each content C and inference form x . Their values correspond to the inference form endorsements in the reduced inference case. Alternatives suppress the logically invalid forms, AC and DA, which is shown through ξ ’s values for the tasks with ‘Many’ alternatives. Similarly, as disablers suppress the logically valid MP and MT, the corresponding ξ values for tasks with ‘Many’ disablers are noticeably lower. The other parameters have the following median values: $\tau(MP) = 1.00$, $\tau(AC) = 0.40$, $\tau(DA) = 0.49$, $\tau(MT) = 0.88$ and $\lambda = 0.78$. Larger values of τ for MP and MT show higher beliefs in the logically valid forms MP, MT.

Considering each individual from 199 participants, 81 were best predicted by OC, another 81 by the DSM and 37 by ϵ -MMT. That lead us to the conclusion that among these three models, there is not a single one that “dominates” the others. Therefore, in order to support the idea that one single model can *not* capture *every* individual, we combined all

³<https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.minimize.html>

three models into what we call an ensemble model, using for each individual the model that made the best prediction. This model consists of the best that these cognitive models can offer and it outperformed the MFA Model by reaching a mean deviation of ca. 16%. The purpose of this ensemble approach is to show that integrating strategies captures individuals best.

Discussion and Conclusion

Our first research question was if it is possible to predict a conclusion endorsement varying in the range between 0 and 100, and not a dichotomous “yes” or “no” response. Yes – our results show that two Bayesian models exposed to a training data set can generate predictions on unseen data with a mean deviation of 17%. With this, we are establishing a new testing paradigm by not asking how good are the models in explaining existing data, but rather how good can they *predict* a reasoner’s answers? By doing that we can elicit new insights. E.g., Singmann et al. (2016) showed that the DSM outperforms other probabilistic models when comparing their fits. However, evaluating the predictive power, the DSM does not perform better than the model it is built upon, the Oaksford et al.’s (2000) original probabilistic model, meaning that in this task only that one source is enough. We posed the question whether it would be possible that a model-based approach could compete with Bayesian models. Elqayam and Over (2013) discuss how old paradigm theories, like the MMT, focus on truth preservation from assumptions and cannot account for irrationality in human decision making. Here we took MMT’s conditional representation and adapted it such that it does not follow the old paradigm’s material implication interpretation and extended it with probabilities based on Pearl’s (1991) ϵ -semantics. With that, we showed that – yes, a model-based approach can indeed compete with established Bayesian models. The comparable performance of the 3 cognitive models indicate a functional equivalence and similar processes, but, they do differ in their representation. None of the single models predictive performance was better than the MFA. This has been regarded in other domains such as syllogistic reasoning as an empirical upper bound for static models (Riesterer et al., 2020). By combining them into an ensemble model and introducing a better representation flexibility we showed that this performance upper bound can be surpassed while still having the tools to give insight into individuals’ conditional reasoning, capturing individual differences. By looking into the models’ parameter values we learn how disablers and alternatives influence the reasoners’ representation of the conditional from different perspectives. Consider a task with ‘Many’ disablers, through ϵ -MMT’s p_3 parameter we understand that the individual’s belief in the world ω_3 , where the antecedent has happened but the consequent has not, is stronger. OC shows us that individuals assign a high probability to the conditional’s exception, $P(\neg Y|X)$. The DSM shows through its ξ parameters how disablers suppress the logically valid MP and MT, which is a reasoning effect that has been long recognized in this field (Byrne, 1989). We

took into consideration experiments that deal with meaningful contents. Data is (still) quite scarce, as the focus in experiments has largely been on reasoning about abstract material. Our interest is in how humans reason in their *everyday life*, where most of our reasoning takes place. Hence we use such material. Nonetheless, the methods can be applied to abstract problems too.

This work opens future research lines in comparing how parts of models can be translated into each other. It not only allows to ground some of the *functional equivalence* we have already identified, but it would additionally help recognize where models deviate and what reasoning strategies might be missing when modeling an individual. With that, predictions of the reasoner’s conclusion endorsement would improve, which would lead to a better understanding of the reasoning processes, making this path of not only fitting models, but also challenging their predictive capabilities an exciting one, opening many doors to a new way of adaptive modeling.

Acknowledgements

This work was supported by DIAS and grants MR 1934/4-1 and MR 1934/10-1.

References

- Byrne, R. M. (1989). Suppressing valid inferences with conditionals. *Cognition*, *31*(1), 61.
- Elqayam, S., & Over, D. E. (2013). New paradigm psychology of reasoning: An introduction to the special issue edited by Elqayam, Bonnefon, and Over. *Thinking & Reasoning*, *19*(3-4), 249.
- Evans, J., & Over, D. E. (2004). *If*. Oxford University Press.
- Johnson-Laird, P. N., & Byrne, R. M. (1991). *Deduction*. Lawrence Erlbaum Associates, Inc.
- Johnson-Laird, P. N., & Byrne, R. M. (2002). Conditionals: a theory of meaning, pragmatics and inference. *Psychological review*, *109*(4), 646.
- Johnson-Laird, P. N., Khemlani, S. S., & Goodwin, G. P. (2015). Logic, probability and human reasoning. *Trends in cognitive sciences*, *19*(4), 201.
- Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press.
- Oaksford, M., & Chater, N. (2020). New paradigms in the psychology of reasoning. *Annual Review of Psychology*, *71*, 305.
- Oaksford, M., Chater, N., & Larkin, J. (2000). Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *26*(4), 883.
- Pearl, J. (1991). *Epsilon-semantics* (Tech. Rep.). Computer Science Department, University of California.
- Riesterer, N., Brand, D., & Ragni, M. (2020). Predictive modeling of individual human cognition: Upper bounds and a new perspective on performance. *Topics in cognitive science*, *12*(3), 960.

Singmann, H., Klauer, K. C., & Beller, S. (2016). Probabilistic conditional reasoning: Disentangling form and content with the dual-source model. *Cognitive Psychology*, 88, 61.

Learning Reference Biases from Language Input: A Cognitive Modelling Approach

Abigail Toth (a.g.toth@rug.nl)

Department of Artificial Intelligence, University of Groningen
Nijenborgh 9, 9747 AG Groningen, The Netherlands

Niels Taatgen (n.a.taatgen@rug.nl)

Department of Artificial Intelligence, University of Groningen
Nijenborgh 9, 9747 AG Groningen, The Netherlands

Petra Hendriks (p.hendriks@rug.nl)

Department of Linguistics, University of Groningen
Jatstraat 26, 9712 EK Groningen, The Netherlands

Jacolien van Rij (j.c.van.rij@rug.nl)

Department of Artificial Intelligence, University of Groningen
Nijenborgh 9, 9747 AG Groningen, The Netherlands

Abstract

In order to gain insight into how people acquire certain reference biases in language and how those biases influence online language processing, we constructed a cognitive model and presented it with a dataset containing reference asymmetries. Via prediction and reinforcement learning, the model was able to pick up on the asymmetries in the input. The model predictions have implications for various accounts of reference processing and demonstrate that seemingly complex behavior can be explained by simple learning mechanisms.

Keywords: implicit causality; reference; cognitive modelling

Introduction

If you congratulate someone, most often it is because of something they did, whereas if you apologize to someone, most often it is because of something you did. This preference to attribute the cause of an event to a particular entity is known as the *implicit causality (IC) bias* and as such, verbs like *apologize* and *congratulate* are known as IC verbs. These verbs can be further separated depending on whether causality is attributed to the grammatical subject or object. For example, when asked, participants consistently attribute the cause in (1a) to Kaitlyn and the cause in (1b) to Marie (e.g., Brown & Fish, 1983; Rudolph & Forsterling, 1997).

- (1) a. Kaitlyn angered Marie.
- b. Kaitlyn comforted Marie.

Implicit causality has mainly been applied to investigate reference processing, in particular pronoun resolution (e.g., Garnham, Traxler, Oakhill, & Gernsbacher, 1996; Järvikivi, Van Gompel, Hyönä, & Bertram, 2005; Koornneef & Van Berkum, 2006). In a self-paced reading study Koornneef and Van Berkum (2006) had participants read sentences like those in (2) and found significantly slower reading times for sentences like (2b), where the pronoun was inconsistent with the bias set-up by the verb, compared to sentences like (2a), where the pronoun was consistent with the bias set-up

by the verb. Furthermore, the effect was significant immediately following the pronoun, suggesting that IC information is used proactively, influencing comprehenders' expectations about subsequent reference.

- (2) a. Linda praised David because he had been able to complete the difficult assignment with very little help.
- b. David praised Linda because he had been able to complete the difficult assignment with her help only.

Other studies have investigated implicit causality in cases where the pronoun cannot ultimately be disambiguated by gender information. When participants are presented with sentences like those in (3), continuations for (3a) are predominantly about Molly, the subject, whereas continuations for (3b) are more evenly distributed between the subject and object referents (e.g., Kehler, Kertz, Rohde, & Elman, 2008; Stevenson, Crawley, & Kleinman, 1994). This has been taken as evidence that IC can modulate well established structural biases, such as the *first mention* and/or *subject bias*, which reflect the typical pattern of people interpreting ambiguous pronouns as referring back to first-mentioned and/or subject referent, which in English are most often confounded (e.g., Gernsbacher, 1989; Järvikivi et al., 2005). Eye-tracking studies have also shown that implicit causality affects the online processing of pronouns (e.g., Järvikivi, Van Gompel, & Hyönä, 2017).

- (3) a. Molly apologized to Sophie. She _____.
- b. Molly congratulated Sophie. She _____.

It is not exactly clear how knowledge about certain interpersonal exchanges, like *congratulating* and *apologizing*, ends up influencing language processing. One possibility is that when language users encounter an IC verb (and are not

already privy to the causal information), they naturally generate an expectation about which referent will be referred to subsequently. In the case of subject-biased IC verbs, this expectation manifests as a continuation, such that the listener expects to continue to hear about the subject referent. In the case of object-biased IC verbs, the expectation instead manifests as a shift, such that the listener expects attention to shift from the (grammatically prominent) subject to the (less grammatically prominent) object. These expectations, about which referent will be referred to subsequently, can in turn interact with expectations about the form of the reference (i.e., whether a name or a pronoun will be used). For instance, there is evidence that subject pronouns like *she* are more likely to be used to refer back to preceding subjects, whereas names are more likely to be used to refer back to preceding objects (e.g., Arnold, 1998; Kehler et al., 2008; Stevenson et al., 1994). Thus, in the case of subject-biased IC verbs, listeners may not only expect to hear about the subject referent, but also that a pronoun will be used.

Not only is it unclear how knowledge of implicit causality ends up influencing language processing, it is also unclear how such knowledge is acquired. One possibility is that over time language learners pick up on asymmetries present in the input (and then make use of this knowledge when processing language). Unfortunately it is quite difficult to assess the relative frequency of the different reference possibilities in one's actual language input (see Sukthanker, Poria, Cambria, & Thirunavukarasu, 2020). To our knowledge no corpus study has been conducted to investigate the frequency of referring to the subject versus the object following implicit causality verbs. Similarly, it is difficult to assess how often pronouns like *she* actually refer back to the preceding subject. In a corpus of children's books, Arnold (1998) found that third person subject pronouns co-referred with the previous sentence's subject in 64% of cases. However, it is unclear if this would hold in a larger and more diverse corpus that also includes natural spoken language. Most of what we know actually comes from sentence completion studies (e.g., Ferstl, Garnham, & Manouilidou, 2011), which are limited in the sense that participants have to switch from being the comprehender to the producer. This can result in task demands that do not reflect natural language processing.

Present Study

The aim of the present study was to gain insight into how certain reference biases come to exert their influence on language processing. Specifically we explored whether simple learning mechanisms, such as prediction and reinforcement learning, could help explain why people display certain reference biases. We had a naive cognitive model learn reference biases from an input dataset containing reference asymmetries. The model was presented with simple transitive sentences and had to predict the subsequent referent (subject vs. object), as well as the form of reference (name vs. pronoun). When the model predicted correctly, it was issued a reward.

We wanted to determine if the model could pick up on the asymmetries present in the input, as well as investigate how predictions changed as the model was presented with more input.

Methods

Input Data

Our input dataset consisted of 1000 unique items. All items consisted of a simple sentence, containing a transitive verb with its subject and object arguments, followed by a critical referring expression. An example of the four possibilities for a single verb can be seen in (4) below. The information in brackets indicates the referent of the referring expression.

- (4) a. Kaitlyn angered Marie. Kaitlyn (Kaitlyn)
 b. Kaitlyn angered Marie. Marie (Marie)
 c. Kaitlyn angered Marie. She (Kaitlyn)
 d. Kaitlyn angered Marie. She (Marie)

All items were created by sampling a verb from a list of 10 verbs that differed with respect to their associated implicit causality: 5 subject-biased verbs (*apologized, repulsed, angered, fascinated, disappointed*), 3 object-biased verbs (*congratulated, feared, comforted*), and 2 non-IC verbs (*filmed, interrupted*). We used an unequal number of each verb type because of our critical assumption that real-world asymmetries in the input are what cause people to display biases. The subject and object referents of each item were randomly sampled from a list of 40 unique female names. The second sentence was determined based on two unique probabilities. The first probability determined which *referent* would be referred to subsequently (i.e., subject or object), for which we used probabilities from Ferstl et al. (2011)'s implicit causality sentence completion corpus. For example, in their study, participants' continuations following an *anger* sentence were about the subject 85% of the time. Therefore, for all of our anger items the sampling probability of the referent being the subject versus object was 0.85/0.15. The second probability determined the *form of reference* (i.e., name or pronoun). Across all verb types, we opted for a general pronoun bias when referring to subjects (with a pronoun sampling probability of 0.75) and a general name bias when referring to objects (with a name sampling probability of 0.75). Given the lack of corpus data, these values were inspired by sentence completion literature.

PRIMs Cognitive Model

Our model was implemented using the cognitive architecture PRIMs (*Primitive Information Processing Elements*, Taatgen, 2013, 2014), which evolved from the ACT-R cognitive architecture (Anderson, Bothell, Lebiere, & Matessa, 1998; Anderson, 2007). Like other cognitive architectures, PRIMs serves as a unified theory of cognition, as well as an interface for implementing models. Like ACT-R, PRIMs assumes that

Table 1: Operators responsible for processing input.

Operator	PRIMs	Description
retrieve-V1	V1<>nil RT1=nil WM1=nil ==> lexical-entry->RT1 V1->RT2	slot 1 of the input buffer (V) is not empty slot 1 of the retrieval buffer (RT) is empty slot 1 of the working memory buffer (WM) is empty retrieve a ‘lexical-entry’ chunk from the declarative- where slot 2 of the chunk matches the information currently in V1
store-V1	V1=RT2 WM1=nil ==> RT2->WM1	slot 1 of the input buffer is the same as slot 2 of the retrieval buffer slot 1 of working memory is empty store the information in slot 2 of the retrieval buffer in slot 1 of working memory

the cognitive system is modular and thus has different modules for specific cognitive functions (e.g., vision, motor control, working memory, declarative memory, etc.). The different modules communicate with each other through their respective buffers. Each buffer has a number of slots that can each hold a single piece of information. Together all the buffers comprise the global workspace of the system. Instead of production rules (as in ACT-R), the exchange of information within the workspace is achieved through the use of operators, which reside as chunks in declarative memory. All operators consist of condition ‘PRIMs’ and action ‘PRIMs’ and are retrieved on the basis of their activation. If the conditions of the retrieved operator are met, the actions are carried out; if not, the operator with the next highest activation is retrieved. We will begin by describing how a single trial unfolds within the model. For the current model it is useful to distinguish between operators responsible for processing input (i.e., the sentences the model is presented with) and operators responsible for making predictions.

The model is first presented with a complete sentence (e.g., ‘Kaitlyn angered Marie’) and then processes each word by retrieving an associated lexical chunk from declarative memory and storing the results in working memory. The two operators responsible for processing the first word are presented with descriptive detail in Table 1 above. The ==> arrow separates condition ‘PRIMs’ from action ‘PRIMs’. The process-

ing of the entire first sentence ultimately results in a completed event representation being held in working memory. It should be noted that because the buffer slots in PRIMs do not have names, the order matters. For example, in our model WM1 is always used to store information about the subject and WM3 is always used to store information about the object.

Next the model predicts the subsequent referent (subject vs. object). The two operators responsible for this are presented in Table 2 below. These operators have the exact same conditional PRIMs and thus in a completely naive model have an equal chance of firing. The crucial difference is which information gets copied into WM5 (the slot reserved for holding information about the subsequent referent). Next the model makes a prediction about the form of reference (name vs. pronoun). The operators responsible for this are presented in Table 3 on the next page. The first three operators again have equal conditional PRIMs and thus an equal chance of firing. The final operator in Table 3 (retrieve-PRO) fires in cases where the model predicts a pronoun in order to account for the fact that different pronouns would be needed to refer to referents depending on their gender and number. However, in the current model the correct pronoun is always *she*. After the model predicts the form of reference, a ‘read-next’ action fires and the model is presented with the critical referring expression and information about the referent. In cases where

Table 2: Operators responsible for predicting referent.

Operator	PRIMs	Description
predict-subj	WM3<>nil WM4=nil WM5=nil ==> WM1->WM5	slot 3 of working memory is not empty slot 4 of working memory is empty slot 5 of working memory is empty copy the information in slot 1 of working memory into slot 5 of working memory (Note: information about the subject is stored in WM1)
predict-obj	WM3<>nil WM4=nil WM5=nil ==> WM3->WM5	slot 3 of working memory is not empty slot 4 of working memory is empty slot 5 of working memory is empty copy the information in slot 3 of working memory into slot 5 of working memory (Note: information about the object is stored in WM3)

Table 3: Operators responsible for predicting reference form.

Operator	PRIMs	Description
predict-subj-name	WM4=nil WM1=WM5 RT1=nil ==> WM5->WM4 read-next->AC1	slot 4 of working memory is empty slot 1 and slot 5 of working memory are the same the retrieval buffer is empty copy the information in slot 5 of working memory into slot 4 of working memory perform 'read-next' action
predict-obj-name	WM4=nil WM3=WM5 RT1=nil ==> WM5->WM4 read-next->AC1	slot 4 of working memory is empty slot 3 and slot 5 of working memory are the same the retrieval buffer is empty copy the information in slot 5 of working memory into slot 4 of working memory perform 'read-next' action
predict-PRO	WM4=nil WM5<>nil RT1=nil ==> lexical-entry->RT1 pronoun->RT3	slot 4 of working memory is empty slot 5 of working memory is not empty the retrieval buffer is empty retrieve a 'lexical-entry' chunk from the declarative- where slot 3 of the chunk is 'pronoun'
retrieve-PRO	WM4=nil WM5<>nil RT1<>nil ==> RT2->WM4 read-next -> AC1	slot 4 of working memory is empty slot 5 of working memory is not empty the retrieval buffer is not empty store the information in slot 2 of the retrieval into slot 4 of working memory perform 'read-next' action

this information matches the model’s predictions, a reward is issued.

During the initial trials, when the model is naive, various operators are just as likely to fire. For example, predicting an subject versus object continuation is just likely even for items containing a subject-biased IC verb. However, by utilizing PRIMs’ context-operator learning, the model is able to learn which combination of operators is most likely to lead to a reward given the current context. As mentioned, operators are retrieved based on their activation, which is primarily influenced by spreading activation. The associated strengths for spreading activation to operators are learned via reinforcement learning. Thus, whenever the model is issued a reward, it increases the association between the current context and all of the operators that lead to the reward being issued. In PRIMs ‘context’ can be used to refer to the entire global workspace (i.e., all the buffers), however, for this particular model we were only interested in spreading activation from the WM buffer. Thus, when certain information is held in working memory (e.g., ‘angered’), specific operators (e.g., predict-subj) are more likely to fire given their increased activation. We also utilized PRIMs’ operator compilation, which allows the model to create new operators by combining pairs of operators that resulted in a reward being issued. In our model predicting the 1) referent and 2) form of reference is initially a two-step process. However, we expect the model to compile the operators responsible for predicting the referent (e.g., predict-subj-continuation) and form of reference

(e.g., predict-PRO), given that in the input data subjects are most often referred to using a pronoun and objects are most often referred to using a name, which the model should pick up on. We ran the model 100 times, where a single run consisted of the model being presented all 1000 items from the input dataset in a completely randomized order. This eliminates any order effects and allows for us to analyze ‘average’ behavior.

Results

With respect to the continued referent asymmetry, Figure 1 shows the proportion of subject (dark gray) versus object (light gray) continuations across the three different verb

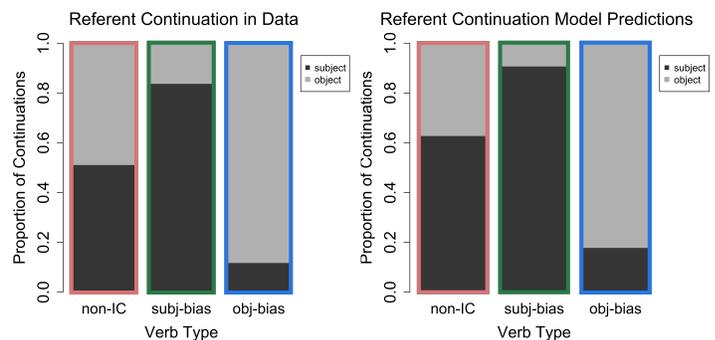


Figure 1: Referent continuations (subject vs. object) by verb type.

types. The left panel is the actual input data and the right panel is the model predictions. As can be seen, the model predictions mirror the input data, primarily predicting subject continuations for items with a subject-biased IC verb and object continuations for items with an object-biased IC verb. With respect to the form of reference asymmetry, Figure 2 shows the proportion of using a name (dark gray) versus pronoun (light gray) when referring to subjects (top panels) and objects (bottom panels) across the three different verb types. The left panels are again the input data and the right panels are the model predictions. When the continued referent was predicted to be the subject, the model largely predicted that the reference would be in the form of a pronoun. When the continued referent was predicted to be the object, the model was more likely to predict the reference would be in the form of a name (except in the case of subject-biased verbs, where it was 50/50 for names and pronouns). These form of reference predictions are in line with the input data (i.e., subjects primarily get referred to with pronouns and objects primarily get referred to with names), however, the model overpredicted pronouns for both subjects and objects.

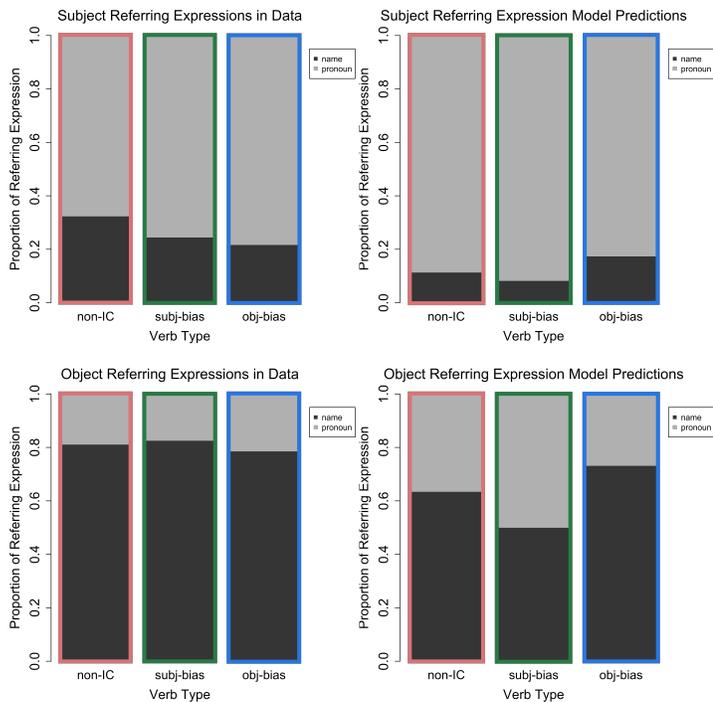


Figure 2: Form of reference continuations (name vs. pronoun) by referent and verb type.

We were also interested in examining learning over trials. Figure 3 illustrates how referent predictions developed across trials (averaged over the 100 model runs). The y-axis represents the proportion of predicting a subject continuation for each of the three verb types. During the initial trials the model predicted subject continuations at chance level across all three verb types. However, as trials unfolded the proportion of predicting a subject continuation increased for items containing

a subject-biased IC verb and decreased for items containing an object-biased IC verb. For the non-IC verbs a gradual increase in subject predictions is seen.

Figure 4 illustrates how form of reference predictions developed across trials. The y-axis represents the proportion of pronoun predictions, for both predicted subject continuations (dark gray) and predicted object continuations (light gray). Here we collapsed over verb type, as the verb itself does not influence the form of reference. During the initial trials the model predicted a pronoun at chance level for both subject and object continuations. However, as trials unfolded the proportion of predicting a pronoun increased in cases where the model predicted a subject continuation and decreased in cases where the model predicted an object continuation. The proportion of predicting a name can be calculated by subtracting the pronoun proportion from 1.

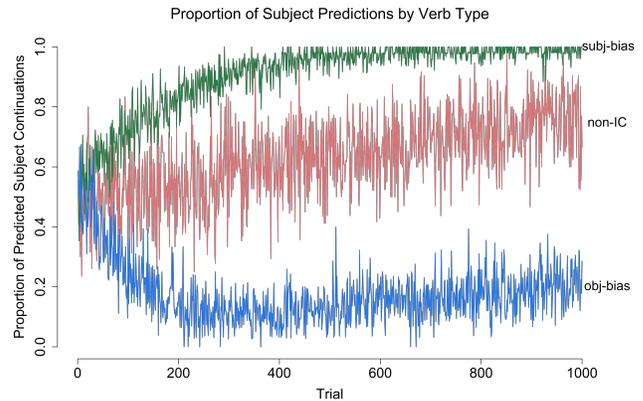


Figure 3: Proportion of subject continuation predictions across trials by verb type.

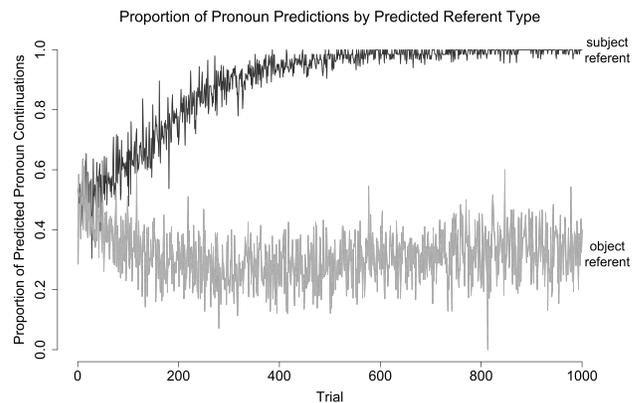


Figure 4: Proportion of pronoun continuation predictions across trials by predicted referent type.

With respect to operator compilation, we expected the model to compile the operators responsible for predicting the referent (e.g., predict-subj) and form of reference (e.g., predict-PRO). However, the model instead combined the operators responsible for predicting the referent (predict-subj

and predict-object), with the previous operator that stores lexical information about the object in working memory (i.e., store-v3). So rather than processing the object of the first sentence (i.e., the Marie in ‘Kaitlyn angered Marie’) and then predicting the subsequent referent, the model combined the two steps. Although this is not what we expected, it also makes sense given that as soon as the verb is stored in working memory, the model can already predict which referent will be referred to next.

Discussion

We constructed a model in the cognitive architecture PRIMs (Taatgen, 2013, 2014) and presented it with input data that contained reference asymmetries. More specifically, in the input dataset we manipulated the proportion of subject versus object continuations by using verbs differing in their associated implicit causality (subject-biased, object-biased and non-IC). Furthermore, we manipulated the likelihood that names versus pronouns would be used to refer to the different referents, such that there was a greater likelihood of using a pronoun when referring back to subjects and a name when referring back to objects. Initially the model was unaware of the asymmetries present in the input. However, by utilizing PRIMs’ context-operator learning (reinforcement learning) the model was able to pick up on the asymmetries present in the data, as reflected by its predictions about subsequent reference.

Our model processed sentences like ‘Kaitlyn angered Marie’ and then made predictions about the subsequent referent, as well as the form of reference. With respect to referent predictions, during the initial trials the model equally predicted subject and object continuations across the three verb conditions. However, as the model was presented with more input this pattern uniquely changed across the three verb conditions, as to mirror the input data. With respect to form of reference predictions, the model initially predicted an equal number of names and pronoun for both subjects and objects. However, as the model was presented with more input, the proportion of predicting a pronoun increased for subjects and decreased for objects (mirroring the input data). Assuming that humans make such predictions of course is not trivial. However, predictive processing is widely assumed to be a core aspect of cognition, especially when it comes to the processing of serial order information, such as language. Furthermore, it is assumed that the reason humans make predictions is to save future processing costs (see Bubic, Von Cramon, & Schubotz, 2010, for a review of prediction). The primary evidence of predictive language processing comes from ERP and visual world eye-tracking studies, which show that people anticipate upcoming arguments following specific verbs, for example, anticipating to hear about something edible, following the verb *ate* (e.g., Altmann & Kamide, 1999; Nieuwland & Van Berkum, 2006).

Our model’s predictions can help explain Koornneef and Van Berkum (2006)’s finding that reading times are slower

when a pronoun is inconsistent with the bias setup by the verb preceding it. For example, following an object-biased IC verb our model most often predicted an object continuation. However, when the continuation ended up being about the subject (i.e., inconsistent with the prediction), the model had to revise the contents of working memory to accurately represent the second sentence, which took additional time. It is difficult to say how exactly the reward issued to our model relates to a reward in the real world. However, one possibility is that the reward in the real world is a successful saving of processing time (or effort).

Similarly, our findings provide insight into why visual world eye-tracking studies find an immediate effect of implicit causality on pronoun resolution (e.g., Järvikivi et al., 2017). One possibility is that following a sentence containing an IC verb, listeners expect to hear about a specific referent. In cases where listeners expected to hear about the subject, they may have also expected a pronoun, whereas in cases they expected to hear about the object, expecting a pronoun would be less likely. Nevertheless, in both cases listeners are presented with an ambiguous pronoun and have no choice but to incorporate it into their discourse representation, which leads to the different gaze patterns. Unfortunately, studies specifically interested in pronoun resolution only report on time windows starting at the pronoun onset. Looking at earlier times (e.g., starting at the verb onset) may actually be crucial for understanding previously reported ‘pronoun’ effects. An empirical prediction of our model, is that that the previous findings are not necessarily about pronouns and how to interpret them, but about how verb biases affect predictions about next referents and their forms.

With respect to future directions, our model was always given enough time to make predictions about upcoming referents and their forms. This is not ideal given that in the real world there is a continuous stream of input that cannot be controlled by the comprehender. This will be addressed in future studies so that the model will only make predictions when it has enough time to do so. Furthermore, we determined our input frequencies based on reasonable assumptions given the current lack of an annotated corpus. In the future it would be informative to explore various relative frequencies and see what effect it has on the model’s predictions. One way to address this would be to include a wide range of verbs, as evidence suggests that implicit causality actually lies on a continuum. Investigating how a continuous measure of IC influences the type of predictions may be fruitful for understanding people’s behavior in previous experimental studies. Finally, although our model predictions can provide insight into some of the previous experimental findings, in none of those studies were participants asked to make explicit predictions. It would be informative to carry out a prediction study so that we could directly compare the model to human prediction data. Although our approach was quite simple, it highlights the fact that seemingly complex behavior can often be explained by simple learning mechanisms.

References

- Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, 73(3), 247-264.
- Anderson, J. (2007). *How can the human mind occur in the physical universe?* 198 Madison Avenue, New York, New York: Oxford University Press.
- Anderson, J., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language*, 38, 341-380.
- Arnold, J. E. (1998). *Reference form and discourse patterns*. Doctoral dissertation, Stanford University.
- Brown, R., & Fish, D. (1983). The psychological causality implicit in language. *Cognition*, 14(3), 237-273.
- Bubic, A., Von Cramon, D. Y., & Schubotz, R. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, 4(25), 1-15.
- Ferstl, E. C., Garnham, A., & Manouilidou, C. (2011). Implicit causality bias in english: a corpus of 300 verbs. *Behavior Research Methods*, 43, 124-135.
- Garnham, A., Traxler, M., Oakhill, J., & Gernsbacher, M. A. (1996). The locus of implicit causality effects in comprehension. *Journal of Memory and Language*, 35(4), 517-543.
- Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition*, 32(2), 99-156.
- Järvikivi, J., Van Gompel, R. P. G., & Hyönä, J. (2017). The interplay of implicit causality, structural heuristics, and anaphor type in ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, 46, 525-550.
- Järvikivi, J., Van Gompel, R. P. G., Hyönä, J., & Bertram, R. (2005). Ambiguous pronoun resolution: contrasting the first-mention and subject-preference accounts. *Psychological Science*, 16, 260-264.
- Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and coreference revisited. *Journal of Semantics*, 25(1), 1-44.
- Koornneef, A. W., & Van Berkum, J. J. A. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language*, 54(4), 445-465.
- Nieuwland, M. S., & Van Berkum, J. J. A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of cognitive neuroscience*, 18(7), 1098-1111.
- Rohde, H., & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience*, 29(8), 912-927.
- Rudolph, U., & Forsterling, F. (1997). The psychological causality implicit in verbs: A review. *Psychological Bulletin*, 121, 192-218.
- Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, 9(4), 519-548.
- Sukthanker, R., Poria, S., Cambria, E., & Thirunavukarasu, R. (2020). Anaphora and coreference resolution: A review. *Information Fusion*, 59, 139-162.
- Taatgen, N. A. (2013). The nature and transfer of cognitive skill. *Psychological Review*, 120(3), 439-471.
- Taatgen, N. A. (2014). Between architecture and model: Strategies for cognitive control. *Biologically Inspired Cognitive Architectures*, 8, 132-139.

A Drift-diffusion Model to Explain Vehicle Deceleration Detection of Vulnerable Road Users

Daniel Trommler (daniel.trommler@psychologie.tu-chemnitz.de)

Department of Psychology, Chemnitz University of Technology
09107 Chemnitz, Germany

Claudia Ackermann (claudia.ackermann@psychologie.tu-chemnitz.de)

Department of Psychology, Chemnitz University of Technology
09107 Chemnitz, Germany

Josef F. Krems (josef.krems@psychologie.tu-chemnitz.de)

Department of Psychology, Chemnitz University of Technology
09107 Chemnitz, Germany

Abstract

The development of automated vehicles is accompanied by the question of how this technology will interact with vulnerable road users (VRUs; e.g. pedestrians, cyclists). Especially in shared spaces, implicit communication signals, such as vehicle deceleration, proved to be crucial. However, previous studies on the parameterization of vehicle deceleration indicated that human detection of vehicle deceleration may depend on various situational and individual factors. This research has two aims: (1) We want to investigate how the detection and perceptual decision-making on vehicle deceleration can be formally described using a computational model. For this, we discuss the applicability of a drift-diffusion model (DDM). (2) Further, we will follow up on previous research regarding the influence of different situational and individual factors on the detection performance and examine how these factors could be related to the DDM parameters. With this research, we would like to contribute to a better understanding and a consistent, formal description of different factors influencing the detection of vehicle deceleration. This could be associated with improved interaction between automated vehicles and VRUs.

Keywords: Automated vehicles; Vulnerable road users; Implicit communication; Deceleration detection; Drift-Diffusion Model

Introduction

The development of automated vehicles is accompanied by challenging issues in the field of human factors (Kyriakidis et al., 2019). One of these issues concerns the communication between automated vehicles (AVs) and vulnerable road users (VRUs, e.g. pedestrians, cyclists; Rasouli et al., 2017). There is a lot of research effort to design an adequate implicit (e.g., vehicle speed adaption) and explicit (e.g., light signals) communication (Markkula et al., 2020). Due to findings which show that the majority of communication in road traffic is realized in an implicit way, especially in the low-speed area (e.g. parking spaces), we take a closer look on this communication approach (Lee et al., 2021).

Deceleration maneuvers are a common implicit communication signal, for example, to indicate the intention of car drivers to give priority to a VRU. However, the

implementation in automated vehicles seems to be non-trivial. On the one hand, the deceleration rate must be strong enough to be perceived by VRUs (Markkula et al., 2018). On the other hand, it should not be too strong to avoid a discomfort for vehicle passengers or a congestion of the road (Markkula et al., 2018).

In this paper, we aim to further investigate the pedestrians detection of vehicle deceleration. First, we summarize empirical results on factors influencing the detection performance of vehicle deceleration. Then, we describe the drift-diffusion model and its previous applications in the transportation context. Next, we present assumptions on how determinants of detection performance may be related to the DDM parameters. Finally, we show preliminary results from an analysis of empirical data.

Background

Factors influencing the detection performance of vehicle deceleration

Ackermann et al. (2019) investigated the relationship of different variables with the detection performance of a vehicle deceleration. In video studies, participants saw the approaching vehicle from the perspective of a pedestrian at the curb. The independent variables included the deceleration rate, vehicle size, different daylight conditions, initial vehicle speed and the onset of the deceleration (early or late onset). The reaction time between the start of the deceleration and the response of the participants was measured as dependent variable.

The results showed no significant effect of daylight conditions on reaction times, in contrast to another study regarding gap acceptance (Beggiato et al., 2017). For the low speed conditions (i.e. 20 km/h), the authors found significant effects for deceleration rate and onset of deceleration. For higher deceleration rate and later onset, participants showed faster reaction times. In addition, the authors found a significant interaction between these main factors. In

particular, the onset of deceleration influenced the detection of the lowest deceleration rate. For the lower and faster speed conditions, the authors found a significant effect of vehicle speed and deceleration rate, as well as a significant interaction of both factors. This means that the higher the vehicle speed and the lower the deceleration rate, the higher the reaction time. However, it was found that there were a high number of missing reaction times (i.e. deceleration was not detected) for conditions with higher speed (40 km/h), low deceleration rate and later onset of deceleration. The influence of vehicle size remained ambiguous. At early onset of deceleration, participants tended to react faster for vehicles with medium size. At late onset of deceleration, participants tended to respond more faster for vehicles with large size. However, the results also varied depending on the deceleration rate.

In their discussion, the authors assume several ways the different variables could influence the detection of vehicle deceleration. They consider the changes in the retinal image size of the approaching, decelerating vehicle as a bottom-up process of information processing. Furthermore, top-down processes such as expectations are discussed.

This view is, among others, consistent with research on collision perception, which also assumes that different sources of information are used for these perceptual decisions (DeLucia, 2015). A possible further influencing factor could be varying risk behaviour under different conditions (e.g. different vehicle sizes). Finally, it might be useful to look at the whole process in which a pedestrian observes a vehicle and not just the time from the onset of deceleration.

Drift-diffusion models in transportation and traffic research

An established model for perceptual decision-making in signal detection tasks or two-alternative forced-choice tasks is represented by the drift-diffusion model (DDM; Ratcliff & McKoon, 2008). The most famous of the evidence accumulation models decomposes behavioral data (i.e., response times and response accuracies) into the underlying cognitive processes and their characteristics. The DDM assumes that humans accumulate (noisy) evidence (information) in the direction of one of two boundaries. This process can be described with a few parameters: The most relevant parameters for our research are drift rate (v), bound height (a), starting point (z) and non-decision time (NDT). The drift rate describes the rate of evidence accumulation which is influenced by the quality of evidence. Thus, the drift rate is associated with the stimulus difficulty. The lower the quality, the lower the drift rate and the higher the difficulty. Evidence is accumulated until it reaches the upper or lower bound representing the two choice alternatives (criteria). The bound height influences the required amount of evidence which is necessary for a decision. A larger bound height is associated with more response caution and more accuracy in decision-making. The starting point defines the position where the accumulation starts. This point can be influenced by expectations or prior knowledge. In this case, the

accumulation starts closer to one of the two boundaries. The non-decision time summarizes the duration for all non-decisional components of response time (i.e. all components except of evidence accumulation), such as stimulus encoding or motor execution (Ratcliff & McKoon, 2008).

While the DDM became increasingly established in the cognitive psychology and cognitive neurosciences (Ratcliff et al., 2016), it was initially unclear to what extent the model could be transferred to the transportation and traffic domain. However, recent studies provide very encouraging indications that the model is also suitable in this context and thus can make valuable contributions to the further development of automated driving. For example, the willingness of pedestrians to cross the road (Giles et al., 2019; Markkula et al., 2018; Tian et al., 2020), car driver reactions to a braking lead vehicle (Engstrom et al., 2017; Xue et al., 2018) or the decision-making of car drivers during unprotected left turns (Zgonnikov et al., 2020) have been successfully modeled so far using the drift-diffusion model.

However, there are also some open questions. Among others, there is limited knowledge about the influence of various situational and individual variables on the DDM parameters in the context of traffic and transportation.

Present work

It seems obvious that the DDM is applicable to the scenario described in Ackermann et al. (2019). A pedestrian at the curb has to make the perceptual decision on a signal detection task, i.e. whether an approaching vehicle is decelerating or not. Furthermore, the DDM seems to be particularly well suitable for our use case because it takes into account both bottom-up (e.g., visual information) and top-down processes (e.g., expectations, cautiousness) of information processing. Therefore, we would like to follow up on this research and investigate how the detection of a vehicle deceleration can be described using a DDM. In particular, we aim investigate the influence of different variables investigated in Ackermann et al. (2019) on the DDM parameters. As a result, we would like to contribute to a better understanding of the cognitive processes involved in the detection of a vehicle deceleration.

Drift-diffusion model for the detection of a vehicle deceleration

In this section, we will discuss our assumptions on the relationship between the variables investigated in Ackermann et al. (2019) and the DDM parameters. We take a closer look on four main parameters of the DDM: Drift rate, bound height, starting point and non-decision time.

Drift rate The drift rate describes the rate of evidence accumulation and is influenced by the quality of evidence (Ratcliff & McKoon, 2008).

We assume that the drift rate results from the pedestrians' speed (change) perception of the vehicle. However, it is questionable which visual information are used. Current research suggests that different cues like looming or distance/duration cues can be used for this task (Lee et al.,

2020). In previous research (e.g. Xue et al., 2018), the use of looming proved to be successful. Therefore, we focus on this visual information.

Looming refers to the change rate of the retinal image size and visual angle related to an object (Lee, 1976). The retinal image and the visual angle becomes larger as a vehicle approaches. The faster the vehicle approaches, the greater the looming. The looming becomes smaller while decelerating. So far, the general looming theory (Lee, 1976) was defined only for frontally approaching objects with constant speed and a small visual angle. Therefore, it is questionable to what extent the looming can be applied to our use case. In an important article by Tian et al. (2020), the looming theory was adapted to the perspective of a pedestrian in a crossing scenario. This shows that looming is a time dependent function depending on speed, distance between vehicle and pedestrian, vehicle size and the pedestrian's distance from the lane. A deceleration would influence the looming via a change in speed.

In addition, we assume that lightness influences the perception of the vehicle. With better daylight, a better vision is possible and a vehicle can be observed more easily. Therefore, we assume that the daylight conditions influence the drift rate with lower drift rates for dusk or in the evening.

Bound height The bound height describes the amount of evidence which is necessary for a decision. This parameter is influenced by the response caution (Ratcliff & McKoon, 2008).

We assume that the bound height is related to the vehicle size. Here, we consider the findings that pedestrians tend to accept a larger gap for larger vehicles (Yannis, Papadimitriou, & Theofilatos, 2013). We assume that pedestrians may be more cautious in decision-making when facing with a larger vehicle which can be associated with a larger bound height.

The same is assumed for daylight conditions. We assume that pedestrians might behave more cautiously in poor light conditions, which could be associated with a larger bound height.

In addition, gender and age might be related to the bound height. Findings indicated that women behave more cautiously and less risky in road traffic than men (Yannis, Papadimitriou, & Theofilatos, 2013). Further, studies indicated that older people are in general more conservative in signal detection tasks than younger people (Ratcliff et al., 2001). Therefore, we assume a larger bound height among women and older people.

Starting point The starting point describes a bias in the evidence accumulation toward one of the two boundaries, for example, due to expectations or prior knowledge (Ratcliff & McKoon, 2008).

We assume that pedestrians tend to not expect a deceleration for faster vehicles. Thus, the starting point would be closer to the corresponding boundary. It is possible that the opposite

effect occurs for slower vehicles, i.e. that pedestrians expect a deceleration.

Non-decision time The non-decision time summarizes the duration of nondecisional components of the response time, such as time for stimulus encoding and motor execution pressing a response button (Ratcliff & McKoon, 2008).

Previous studies showed longer non-decision times for older participants (Ratcliff et al., 2001). Consequently, we hypothesize that non-decision time is related to the age of participants.

Preliminary results

To begin examining our assumptions for the first variables, we conducted an online study using jsPsych (de Leeuw, 2015) following the experiments by Ackermann et al. (2019). $N = 62$ participants ($n = 19$ male, $n = 43$ female) saw videos of approaching vehicles. The initial speed (20 and 40 km/h) and deceleration rate (no deceleration; slight deceleration, i.e. -1.5 m/s²; strong deceleration, i.e. -3.5 m/s²) were varied as independent variables. Participants were instructed to press a button when they decided whether the vehicle decelerated or not. If there was a deceleration, it started immediately after the video's onset. Reaction times and responses were recorded.

Figure 1 and 2 show the mean reaction times and the response accuracy depending on the deceleration rate and the initial vehicle speed.

This shows that the reaction times were always higher for vehicles with higher than for lower initial speed. Furthermore, there are differences in the response accuracy. While the detection of no deceleration was more accurate for vehicles with higher speed, the accuracy for slight and strong decelerations was higher for vehicles with lower speed. The poor detection performance of slight deceleration of vehicles with higher speed confirms the findings of Ackermann et al. (2019)

We conducted a preliminary parameter estimation for a drift-diffusion model using PyDDM (Shinn et al., 2020). Table 1 shows the results for the drift rate (depending on time and deceleration rate), the bound height (depending on gender and age), the non-decision time (depending on age) and the starting points for vehicles with lower and higher speed. To investigate age effects, we divided the participants into two age groups. Participants with an age under 30 years were classified to "young participants". Participants with an age of 30 years and older were classified to "middle-aged participants". This classification was chosen in order to investigate two groups of approximately equal size.

The results show that the drift rate varied with deceleration rate. Here, the values for the slight deceleration were lowest for vehicles with lower as well as higher speed.

The bound height varied slightly depending on gender and age. However, no consistent pattern can be observed. For vehicles with lower speed, the bound height for men were slightly higher than those for women. The opposite direction was observed for vehicles with higher speed. Similarly, the

results for the age groups were contrary. Younger participants showed a lower bound height for vehicles with lower speed and a higher bound height for vehicles with higher speed compared to middle-aged participants. Furthermore, it can be seen that the bound height is generally higher for vehicles with higher speed.

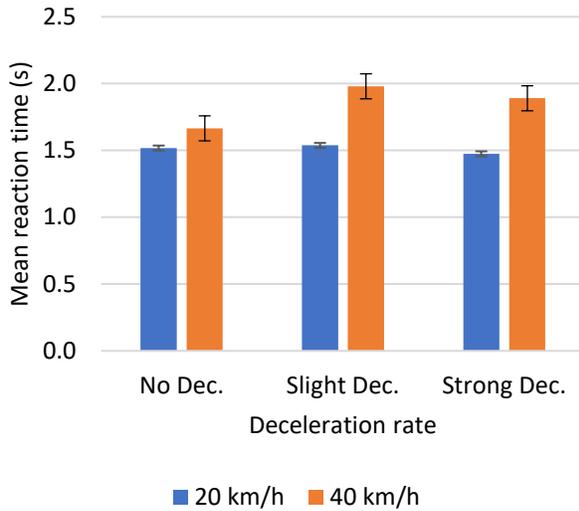


Figure 1: Mean reaction times depending on deceleration rate and initial speed.

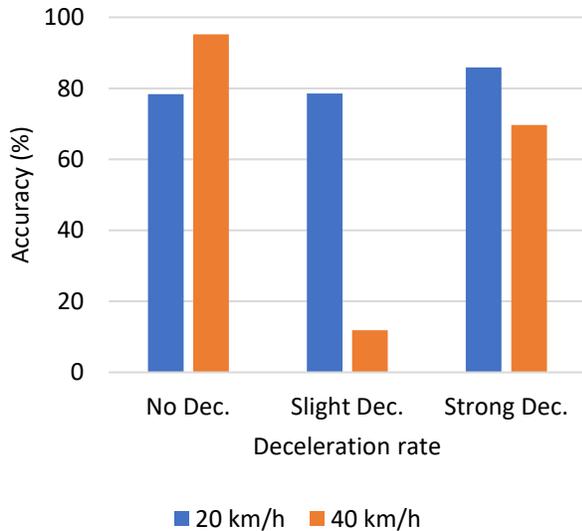


Figure 2: Response accuracy depending on deceleration rate and initial speed.

Furthermore, younger participants showed a lower non-decision time for both vehicles with lower and higher speed compared to middle-aged participants.

Finally, a slightly negative value for the starting point for vehicles with higher speed can be observed.

Table 1: Results from the parameter estimation for a drift-diffusion model.

		Vehicle speed	
		20 km/h	40 km/h
Drift rate	No Dec.	2.303	3.310
	Slight Dec.	1.708	-1.433
	Strong Dec.	2.775	1.804
Bound height	Male	1.999	2.029
	Female	1.898	2.101
	Young participants	1.855	2.389
	Middle-aged participants	1.917	2.227
NDT	Young participants	0.575	0.611
	Middle-aged participants	0.616	0.711
Starting point		0.026	-0.100

Discussion and further work

In this paper, we presented a model for the detection of a vehicle deceleration from the perspective of pedestrians. For this purpose, we used a drift-diffusion model. Previous research showed that several variables can affect the detection performance of a vehicle deceleration. We proposed assumptions on how these variables might be related to the DDM parameters with the aim to formally describe and better understand this cognitive process. Further, we presented a study to examine first variables and their relation with the DDM parameters. The results revealed a strong relation between the deceleration rate and the drift rate and thus the process of decision-making. No and strong decelerations were related to higher drift rates, while slight decelerations were related to lower drift rates. This shows that the deceleration rate represents an important variable regarding the quality of evidence.

Gender and age only slightly affected the bound height and non-decision time. However, it can be emphasized that the bound height were slightly larger for vehicles with higher speed compared to vehicles with lower speed. Here, participants were more cautious in their decision-making.

Finally, a negative value for the starting point for vehicles with higher speed indicates a slight bias, i.e. participants rather expected no deceleration in this speed condition. For vehicles with lower speed, there was no bias observable.

This study, with preliminary estimation of DDM parameters for different conditions represents a first step. In further studies, it seems important to confirm the results within standardized laboratory settings and to investigate further influencing factors on the DDM parameters. These include, among others, the effect of vehicle size, different

daylight conditions and the onset of deceleration on the drift rate and the bound height.

Furthermore, it is necessary to investigate the importance of individual characteristics (gender, age) in the decision making process in more detail using a more balanced sample and a broader range of participants' age.

In addition, a validity study is crucial to check the fit between the model and empirically observed reaction times and response accuracies.

The results extend our understanding of VRUs' perception of vehicle deceleration and, in particular, the effects of bottom-up (evidence) as well as top-down processes (e.g., expectation, cautiousness) in information processing due to various situational (e.g., time of day) and individual (e.g., age) variables. This has several advantages: First, vehicle deceleration can be designed more context-sensitive, which could lead to higher acceptance and user-friendliness both for VRUs and vehicle passengers. Second, a more detailed understanding of human perception and decision-making can be used to derive implications on how to use the communication signals appropriately, for example, through a specific enhancement of implicit communication signals by explicit signals (i.e., external HMI) in case of a low drift rate (Markkula et al., 2018). And third, the findings can be used to improve the feasibility of driving simulations. Currently, developers are focusing on a highly realistic physics for vehicle simulations. Another important focus could be a more realistic behavior of virtual VRUs by modeling their perceptions (Markkula et al., 2018). This would allow a more precise investigation of interactions between (human) drivers and VRUs.

Acknowledgments

This project was funded within the Priority Program 1835 "Cooperative Interacting Automobiles" of the German Science Foundation DFG. The authors appreciate the fruitful collaboration with the project partners.

References

- Ackermann, C., Beggiato, M., Bluhm, L.-F., Löw, A., & Krems, J. F. (2019). Deceleration parameters and their applicability as informal communication signal between pedestrians and automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, 62, 757–768. <https://doi.org/10.1016/j.trf.2019.03.006>
- Beggiato, M., Witzlack, C., Springer, S., & Krems, J. (2018). The Right Moment for Braking as Informal Communication Signal Between Automated Vehicles and Pedestrians in Crossing Situations. In N. A. Stanton (Ed.), *Advances in Human Aspects of Transportation*. Springer International Publishing. https://doi.org/10.1007/978-3-319-60441-1_101
- de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, 47(1), 1–12. <https://doi.org/10.3758/s13428-014-0458-y>
- DeLucia, P. R. (2015). Perception of Collision. In R. R. Hoffman, P. A. Hancock, M. W. Scerbo, R. Parasuraman, & J. L. Szalma (Eds.), *The Cambridge Handbook of Applied Perception Research*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511973017.035>
- Engstrom, J., Markkula, G., & Merat, N. (2017). Modelling the effect of cognitive load on driver reactions to a braking lead vehicle: A computational account of the cognitive control hypothesis. In *Proceedings of the Fifth International Conference on Driver Distraction and Inattention, Paris*. Leeds.
- Giles, O., Markkula, G., Pekkanen, J., Yokota, N., Matsunaga, N., Merat, N., & Daimon, T. (2019). At the Zebra Crossing: Modelling Complex Decision Processes with Variable-Drift Diffusion Models. *Proceedings of the 41st Annual Meeting of the Cognitive Science Society*, 366–372.
- Hawkins, G. E., Forstmann, B. U., Wagenmakers, E.-J., Ratcliff, R., & Brown, S. D. (2015). Revisiting the Evidence for Collapsing Boundaries and Urgency Signals in Perceptual Decision-Making. *Journal of Neuroscience*, 35(6), 2476–2484. <https://doi.org/10.1523/JNEUROSCI.2410-14.2015>
- Kyriakidis, M., de Winter, J. C., Stanton, N., Bellet, T., van Arem, B., Brookhuis, K., Martens, M. H., Bengler, K., Andersson, J., & Merat, N. (2019). A human factors perspective on automated driving. *Theoretical Issues in Ergonomics Science*, 20(3), 223–249.
- Lee, D. N. (1976). A Theory of Visual Control of Braking Based on Information about Time-to-Collision. *Perception*, 5(4), 437–459. <https://doi.org/10.1068/p050437>
- Lee, Y. M., Madigan, R., Giles, O., Garach-Morcillo, L., Markkula, G., Fox, C., Camara, F., Rothmueller, M., Vendelbo-Larsen, S. A., Rasmussen, P. H., Dietrich, A., Nathanael, D., Portouli, V., Schieben, A., & Merat, N. (2021). Road users rarely use explicit communication when interacting in today's traffic: Implications for automated vehicles. *Cognition, Technology & Work*, 23(2), 367–380. <https://doi.org/10.1007/s10111-020-00635-y>
- Markkula, G., Madigan, R., Nathanael, D., Portouli, E., Lee, Y. M., Dietrich, A., Billington, J., Schieben, A., & Merat, N. (2020). Defining interactions: A conceptual framework for understanding interactive behaviour in human and automated road traffic. *Theoretical Issues in Ergonomics Science*, 21(6), 728–752. Business Source Complete.
- Markkula, G., Romano, R., Madigan, R., Fox, C. W., Giles, O. T., & Merat, N. (2018). Models of human decision-making as tools for estimating and optimizing impacts of vehicle automation. *Transportation research record*, 2672(37), 153–163.
- Pekkanen, J., Giles, O., Lee, Y. M., Madigan, R., Daimon, T., Merat, N., & Markkula, G. (2021). *Variable-drift diffusion models of pedestrian road-crossing decisions*. PsyArXiv. <https://doi.org/10.31234/osf.io/f2wsa>
- Rasouli, A., Kotseruba, I., & Tsotsos, J. K. (2017). Agreeing to cross: How drivers and pedestrians communicate. 2017

- IEEE Intelligent Vehicles Symposium (IV)*, 264–269.
<https://doi.org/10.1109/IVS.2017.7995730>
- Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, 20(4), 873–922.
<https://doi.org/10.1162/neco.2008.12-06-420>
- Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging*, 16(2), 323–341.
<https://doi.org/10.1037/0882-7974.16.2.323>
- Shinn, M., Lam, N. H., & Murray, J. D. (2020). A flexible framework for simulating and fitting generalized drift-diffusion models. *ELife*, 9, e56938.
<https://doi.org/10.7554/eLife.56938>
- Tian, K., Markkula, G., Wei, C., & Romano, R. (2020). Creating Kinematics-dependent Pedestrian Crossing Willingness Model When Interacting with Approaching Vehicle. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, 1–6.
<https://doi.org/10.1109/ITSC45102.2020.9294430>
- Xue, Q., Markkula, G., Yan, X., & Merat, N. (2018). Using perceptual cues for brake response to a lead vehicle: Comparing threshold and accumulator models of visual looming. *Accident Analysis & Prevention*, 118, 114–124.
<https://doi.org/10.1016/j.aap.2018.06.006>
- Yan, X., Radwan, E., & Guo, D. (2007). Effects of major-road vehicle speed and driver age and gender on left-turn gap acceptance. *Accident Analysis & Prevention*, 39(4), 843–852. <https://doi.org/10.1016/j.aap.2006.12.006>
- Yannis, G., Papadimitriou, E., & Theofilatos, A. (2013). Pedestrian gap acceptance for mid-block street crossing. *Transportation Planning and Technology*, 36(5), 450–462.
<https://doi.org/10.1080/03081060.2013.818274>
- Zgonnikov, A., Abbink, D., & Markkula, G. (2020). *Should I stay or should I go? Evidence accumulation drives decision making in human drivers.* PsyArXiv.
<https://doi.org/10.31234/osf.io/p8dxn>

Capturing Dynamic Performance in a Cognitive Model: Estimating ACT-R Memory Parameters with the Linear Ballistic Accumulator

Maarten van der Velde¹ (m.a.van.der.velde@rug.nl), Florian Sense¹ (f.sense@rug.nl),
Jelmer Borst² (j.p.borst@rug.nl), & Hedderik van Rijn¹ (d.h.van.rijn@rug.nl)

¹Dept. of Experimental Psychology & Behavioural and Cognitive Neuroscience, University of Groningen, the Netherlands

²Bernoulli Institute, Dept. of Artificial Intelligence, University of Groningen, the Netherlands

Abstract

The parameters governing our behaviour are in constant flux. Accurately capturing these dynamics in cognitive models poses a challenge to modellers. Here, we demonstrate a mapping of ACT-R's declarative memory onto the linear ballistic accumulator, a mathematical model describing a competition between evidence accumulation processes. We show that this mapping provides a method for inferring individual ACT-R parameters without requiring the modeller to build and fit an entire ACT-R model. We conduct a parameter recovery study to confirm that the LBA can recover ACT-R parameters from simulated data. Then, as a proof of concept, we use the LBA to estimate ACT-R parameters from an empirical data set. The resulting parameter estimates provide a cognitively meaningful explanation for observed differences in behaviour over time and between individuals.

Keywords: Memory; dynamic performance; individual differences; ACT-R; linear ballistic accumulator.

Introduction

Cognitive architectures such as ACT-R (Anderson, 2007) provide a framework for developing models of cognition. A challenge commonly faced by modellers is to accurately capture changes in cognitive performance over time, as well as individual differences between people, in the parameters of such models. Current approaches tend to rely on computationally expensive and statistically sub-optimal methods like parameter sweeps to identify the best-fitting parameter values. Mathematical modelling methods can serve as a more efficient and rigorous alternative (Fisher, Houpt & Gunzelmann, 2020). In this paper, we contribute to previous efforts to connect cognitive architectures and mathematical modelling by using the linear ballistic accumulator (Brown & Heathcote, 2008) to infer ACT-R parameters governing memory retrieval.

Retrieval of information from memory can be viewed as a process of evidence accumulation, in which internal and external cues contribute evidence to candidates in memory that are competing for retrieval (Ratcliff, 1978; Anderson, 2007). The first candidate to accumulate enough evidence to cross a boundary wins the race and is retrieved. The dynamics of this process are determined by the amount of evidence each candidate needs to accumulate to cross the boundary, and the rate at which this evidence accumulates.

While such evidence accumulation models have seen most use in the domain of decision making (e.g., Ratcliff, Smith, Brown & McKoon, 2016; Smith & Ratcliff, 2004; Usher & McClelland, 2001; Brown & Heathcote, 2008), there have

been some applications in the domain of memory retrieval. Van Maanen et al. showed that a leaky competing accumulator model could explain performance in picture-word interference tasks (van Maanen & van Rijn, 2007; van Maanen, van Rijn & Taatgen, 2012). In this model, memory chunks accumulate activation by receiving positive and negative spreading activation from other chunks. More recently, Nicenboim and Vasishth (2018) and Fisher et al. (2020) implemented the ACT-R model of declarative memory in a lognormal race model (LNR; Rouder, Province, Morey, Gomez & Heathcote, 2015), in which the rate at which evidence for a chunk accumulates depends on its activation.

Here, we extend this formalisation of ACT-R memory retrieval as an LNR to a more flexible linear ballistic accumulator model (LBA; Brown & Heathcote, 2008). Unlike the LNR, the LBA is able to estimate the rate of accumulation separately from the distance accumulators need to travel to reach the decision boundary. This is useful, because both accumulation rate and distance to the boundary have natural counterparts in ACT-R: the accumulation rate corresponds to the activation of the chunk, while the distance can be linked to the latency factor (F) parameter. As such, the LBA provides a cognitively meaningful interpretation of ACT-R's F parameter as a measure of response caution—the larger the distance, the more evidence needs to be collected before a response is made—and offers a method by which it can be estimated from response data.

In the following sections, we first describe the formal link between ACT-R and the LBA. We then demonstrate how the LBA can be used to recover ACT-R parameters in a simulation study. Finally, we fitted the LBA to an empirical data set, showing how it can offer insight in the mechanisms underlying changes in retrieval performance over time.

Casting ACT-R's Declarative Memory as a Linear Ballistic Accumulator

The linear ballistic accumulator model (Brown & Heathcote, 2008) proposes that response behaviour can be explained through a race between accumulators. Each accumulator has a certain amount of starting evidence k that increments linearly at a drift rate v until it reaches a decision boundary d . The first accumulator to reach the boundary determines the response choice and latency. A constant non-decision time t_0 is also added, representing the time required for other components of

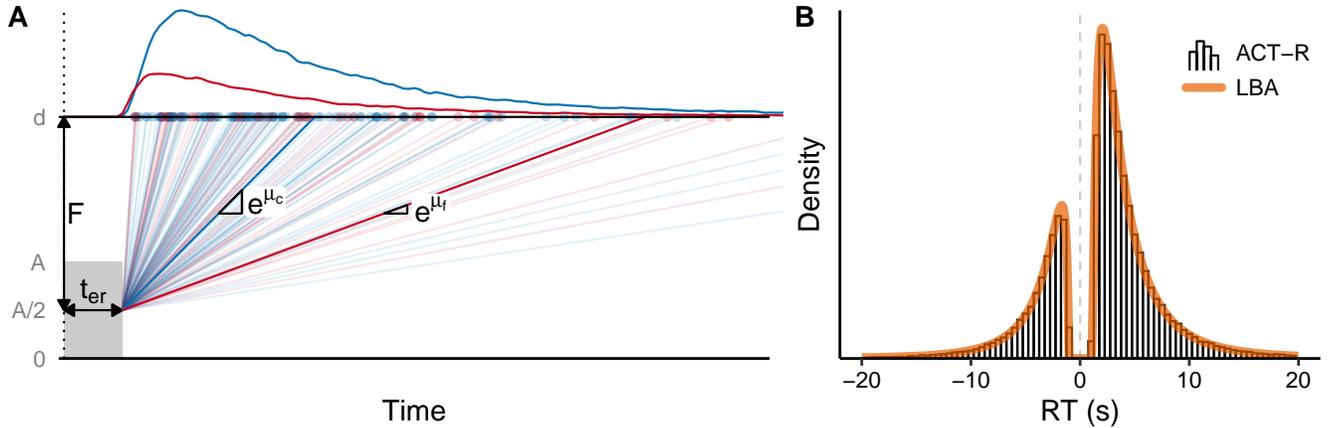


Figure 1: Casting ACT-R memory retrieval as a linear ballistic accumulator. **A**: ACT-R retrieval with two competing chunks visualised as an LBA, with marginal RT distributions shown at the top. See the main text for details. **B**: RT distributions of an ACT-R model (histogram) and the equivalent LBA model (orange curve). Error responses are shown as negative RTs.

the response process, such as perceptual and motor functions. There are two sources of variability between trials: the starting point $k \sim U(0, A)$, and drift rate $v_i \sim N(\mu_i, \sigma_i)$ for each response option i . The LBA assumes a constant rate of evidence accumulation over a trial, so the time required for an accumulator to reach its boundary on a trial j is the distance $d - k_j$ divided by the drift rate, plus non-decision time:

$$RT_j = \frac{d - k_j}{v_j} + t_0 \quad (1)$$

Across trials, the average starting point is $A/2$ and the average drift rate is μ_i , so the expected finishing time for an accumulator is:

$$\mathbb{E}(RT) = \frac{d - A/2}{\mu_i} + t_0 \quad (2)$$

We can map the LBA parameters onto ACT-R memory parameters. ACT-R models declarative memory as a set of symbolic chunks, each with a sub-symbolic activation that decays over time and is subject to noise (Anderson, 2007). The time required to retrieve a chunk depends on its activation: the more active the chunk, the faster its retrieval will be, but like the LBA, the time course of memory retrieval is deterministic once the starting values are known. If multiple chunks match a retrieval request, the chunk with the highest activation—and therefore the lowest retrieval time—at the time of the request wins. A full response also involves non-memory operations, such as stimulus encoding and response execution, which can be captured by adding a term t_{er} to the retrieval time.

ACT-R defines the full time required to retrieve a chunk i with an activation A and respond accordingly by the following equation, in which F is the latency factor, a positive scaling

parameter¹:

$$RT_i = F * e^{-A_i} + t_{er} \quad (3)$$

We can rewrite this equation in a similar form to (2):

$$RT_i = \frac{F}{e^{A_i}} + t_{er} \quad (4)$$

The mapping between ACT-R's parameters (left) and those of the LBA (right) then becomes straightforward:

$$F = d - A/2 \quad (5)$$

$$A_i = \ln(\mu_i) \quad (6)$$

$$t_{er} = t_0 \quad (7)$$

With this mapping, ACT-R's latency factor (F) is equivalent to the average distance between starting point and boundary in the accumulator model, often conceptualised as the response caution: given a constant activation, a higher value of F means that more evidence is required to complete a retrieval. The mapping relates the activation A_i of a chunk to its drift rate μ_i , meaning that a highly activated chunk can be seen as accumulating evidence more rapidly than one with a lower activation. Put differently, the drift rate μ_i is equivalent to e^{A_i} , the odds of needing the chunk. Finally, there is a direct equivalency between the non-retrieval time (t_{er} and t_0) in both models.

Figure 1A visualises the ACT-R retrieval process in the style of an accumulator model. It shows two chunks, c (blue) and f (red), competing for retrieval over multiple trials. In each trial, both accumulators race to cover the vertical distance F to the boundary. The winner gets retrieved in the time it takes to reach the boundary. There is normally distributed trial-to-trial variability, or noise, in the activation of the chunks, and therefore in the rate at which each chunk

¹An additional parameter f may be used to scale the activation: $RT = F * e^{-f * A_i} + t_{er}$. This parameter is typically held constant at 1, and we make the same simplification here as it has no bearing on the outcomes.

accumulates evidence: $A_i \sim N(\mu_i, \sigma_i)$. As such, drift rates follow a lognormal distribution. The resulting RT distributions can be shown to be lognormal too:

$$RT_i \sim LN(\mu_i + \ln(F), \sigma_i) + t_{er} \quad (8)$$

Figure 1B demonstrates that ACT-R and the LBA generate identical response time distributions for a given set of parameters when using the mapping in equations (5)–(7). Interactive versions of these figures, in which the model parameters can be freely adjusted, are available at <https://cogmod.shinyapps.io/actr-lba/>.

Simulation: Recovering ACT-R Parameters

Given this mapping, it should be possible to identify ACT-R memory parameters from response data (RT and choice) using existing methods for fitting the LBA. Therefore, we performed a simulation study with two goals: to investigate whether the LBA can recover ACT-R memory parameters from a typical participant sample completing a reasonable number of trials, and to ensure that parameter recovery works regardless of specific parameter values. The code required to reproduce this simulation study is available at <https://osf.io/wpvj7/>.

Data

ACT-R was used to simulate 25 distinct model participants, each performing a sequence of retrieval trials. Retrieval was modelled as a competition between two chunks, c and f , representing a correct and incorrect response to a retrieval cue, respectively. For each model participant, ACT-R parameters were sampled randomly from plausible distributions, listed in Table 1. To ensure that parameters recovered by the LBA would all be on the same scale, we fixed the standard deviation of the activation of the correct response (σ_c) to 1, both in ACT-R and in the LBA².

We repeated the process with differently sized data sets, ranging from 25 to 50,000 trials per participant, to gauge the effect of data set size on recovery accuracy.

Model fitting

The LBA was fitted separately to each model participant’s responses using the *nlm* optimiser in R (version 3.6.3; R Core Team, 2020). We ran this optimiser 250 times with randomly generated starting values, and kept only the parameter values that yielded the highest summed log-likelihood across all runs. The *dLBA* density function from the *rt* package (version 0.11-2; Singmann, Brown, Gretton & Heathcote, 2020) served as the objective function. For each model participant, we derived individual ACT-R parameters from the best-fitting LBA using the mapping in equations (5)–(7).

Results

The results of the parameter recovery process are shown in Figure 2. As Figure 2A indicates, original parameter values

²See Brown and Heathcote (2008) for alternative solutions to the scaling problem in accumulator models.

Table 1: ACT-R parameters used in the simulation study.

	Description	Distribution
μ_c	Mean activation of correct answer	$\mu_c \sim N_-(-.5, .5)$
μ_f	Mean activation of incorrect answer	$\mu_f \sim N_-(-1.5, .5)$
σ_c	SD of activation of correct answer	$\sigma_c = 1$
σ_f	SD of activation of incorrect answer	$\sigma_f \sim N_+(1.5, .5)$
F	Latency factor	$F \sim N_+(1, .5)$
t_{er}	Non-retrieval time	$t_{er} \sim N_+(.75, .5)$

Note: N_+ and N_- are truncated normal distributions, limited to positive and negative values, respectively.

could already be recovered with reasonable accuracy from a data set with 100 trials per participant. Some parameters (e.g., σ_f and t_{er}) appear easier to recover than others, but even the larger errors do not appear to show systematic over- or underestimation.

Figure 2B shows how recovery accuracy changed as a function of data set size. Recovery accuracy, measured as the absolute error of recovered parameter values relative to the original values, is shown separately per parameter (coloured points) as well as across parameters (black points). Unsurprisingly, recovery accuracy improved when there were more trials constraining the fit, though the current fitting method reached a plateau once there were at least 250 trials per participant.

Example Application: Modelling Changing Retrieval Performance in Empirical Data

To demonstrate how the method may be used to explain dynamic memory performance in terms of cognitively meaningful constructs, we fitted the LBA to empirical data from a multi-session retrieval practice task.

Data

We use data from a retrieval practice task completed by recruits of the Commando Corps, Royal Netherlands Army (*Korps Commandotroepen*), in which participants learned the names of fictitious safehouses on a map. On first presentation, a safehouse was shown with its name, while subsequent repetitions required participants to select the correct name themselves from a set of four answer options. Participants completed three 8-minute sessions over the course of a week. They studied a different map in every session, and maps were counterbalanced between participants. The task was presented within an adaptive learning system that schedules each item to be repeated whenever its activation is expected to hit a threshold (van Rijn, van Maanen & van Woudenberg, 2009; van der Velde, Sense, Borst & van Rijn, 2021). As such, we could expect the activation of the chunks being retrieved to be fairly stable across trials, despite the novelty of the materials. Response accuracy and response time were recorded in every trial.

Session 1 was scheduled on the first day of the week, while the second and third sessions took place several days later and were scheduled immediately before and after a high-intensity loaded speed march of about 40 minutes. We

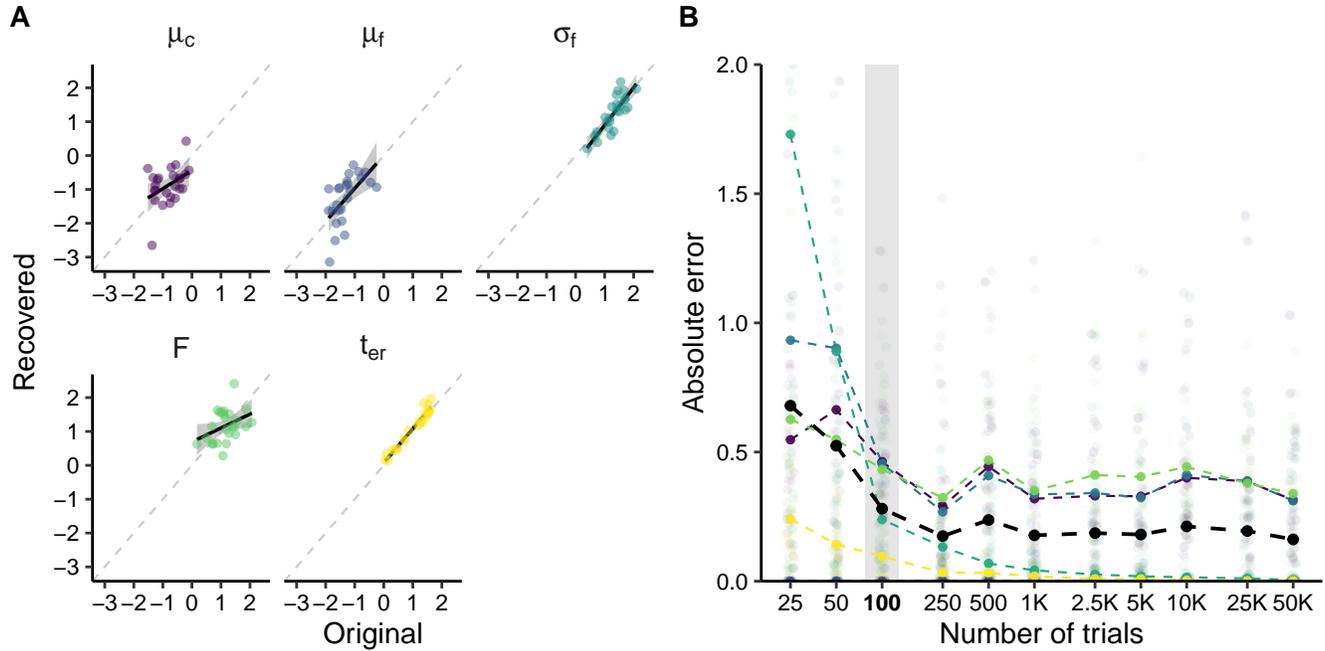


Figure 2: Recovery of ACT-R parameters using the LBA. **A**: Original versus recovered parameter values for a data set with 100 trials per participant. Parameter descriptions are given in Table 1. **B**: Recovery accuracy (absolute error) for data sets with different numbers of trials per participant. Light coloured points show individual errors, dark coloured points show the mean error per parameter, and black points show the mean error across parameters.

expected performance to change for two reasons: increased familiarity with the task might lead to better performance after the first session, and the physical exertion of the speed march might affect performance in the third session.

For the analysis, we removed the first trial for each item (in which the answer was shown on screen), trials in which participants did not respond within 30 s, and trials in which the recorded response time was lower than 300 ms. Since the simulation study showed that recovery was worse in small data sets, participants had to have completed at least 50 practice trials per session to be included. In addition, we required that participants made at least 5 error responses per session, to give the model a chance of fitting the error response distribution. These criteria struck a balance between ensuring a sufficient number of observations per participant and including as many participants as possible. They yielded a data set with 12,568 usable observations (out of 29,441) from 50 (out of 127) participants.

Model fitting

We fitted the LBA separately to each of the three retrieval practice sessions for each participant. The fitting procedure was the same as in the simulation study. The analysis code is available at <https://osf.io/wpvj7/>.

Results

Figure 3 shows participants' performance on the task over the three sessions. Despite the task difficulty being the same in all three sessions, performance improved in two ways. Firstly, response accuracy increased and then plateaued: a logistic

mixed-effects model with a main effect of session and random intercepts for participants showed that accuracy increased from the first to the second session ($z = -4.680, p < .001$), but found no evidence for a change from the second to the third session ($z = -0.253, p = .8$). Secondly, responses became faster: a generalised mixed-effects model with a Gamma link function and with a main effect of session and random intercepts for participants found a decrease in response times on correct trials from session 1 to session 2 ($t = 2.250, p = .0244$), and from session 2 to session 3 ($t = -7.182, p < .001$).

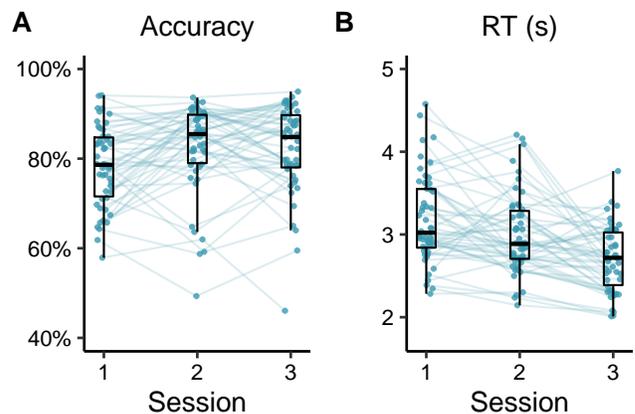


Figure 3: Performance on the retrieval practice task by participant. **A**: percentage correct responses per session. **B**: median response time on correct responses per session.

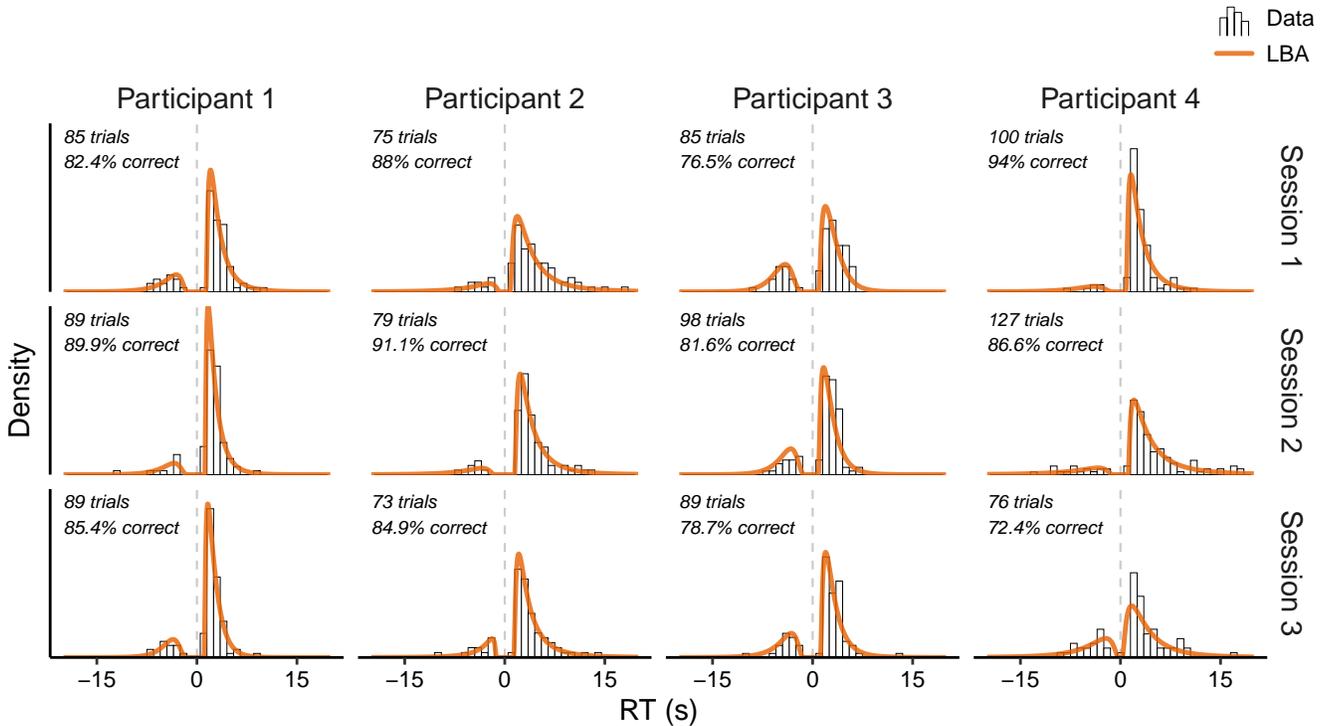


Figure 4: Best fits of the LBA to the response data of four participants over three retrieval practice sessions. Error responses are shown as negative RTs. The number of available trials and the response accuracy are shown in the top left corner of each plot.

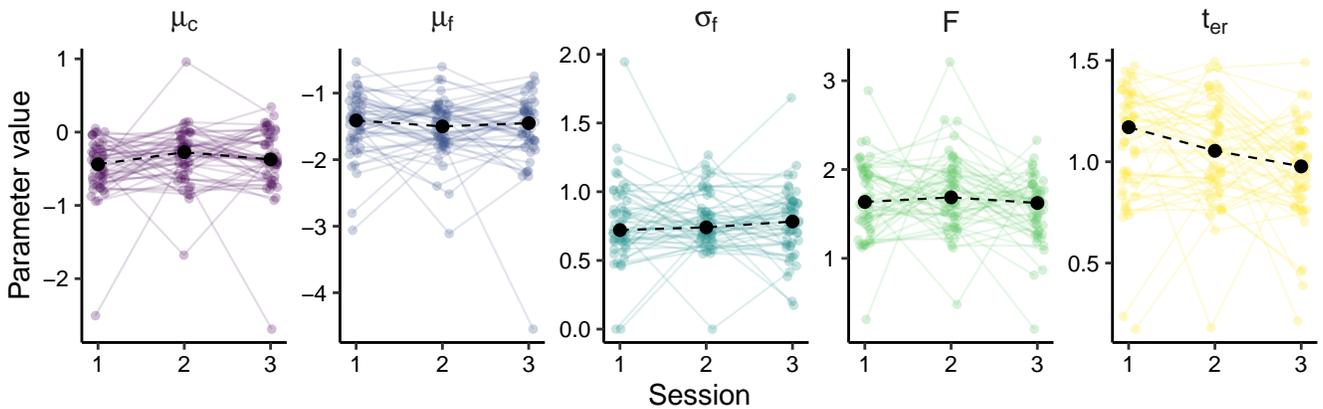


Figure 5: ACT-R memory parameters inferred from the data. Coloured points show individual estimates; large black points indicate the median value across participants. Parameter descriptions are given in Table 1. Note: Y-axes differ between plots.

Figure 4 shows the best fit of the LBA to the response time distributions of four randomly selected participants. These examples suggest that the model captured the shape of the data quite well, although the low number of trials and high response accuracy did make it challenging to fit the error responses.

The inferred ACT-R parameters are shown in Figure 5. There is substantial variation in the parameter values for individual participants, but they are nonetheless clustered quite neatly around the sample averages. As one would expect, the activation of the correct answer (μ_c) tended to exceed the activation of the incorrect answer (μ_f), reflecting participants'

better-than-chance performance. To explore possible changes in parameter values over time, we fitted separate linear mixed-effects models to each parameter, testing whether there was a session effect on the parameter value, with random intercepts for participants. These models suggested that the parameters generally remained fairly constant between sessions³.

³Aside from the reported effects, there was some evidence for a decrease in the F parameter from session 2 to session 3, though the corresponding model failed to converge. More generally, these results should be interpreted with a degree of caution, as repeated LBA fits yield slightly different parameter estimates due to random variation.

However, the activation of the correct answer (μ_c) did appear to increase from session 1 to session 2 ($t(98) = -2.050$, $p = .043$). Furthermore, since the outcome of the retrieval depends on which of the two candidate chunks has the highest activation, rather than on the individual activation of either chunk, we also fitted a linear mixed-effects model with the difference in activation $\mu_c - \mu_f$ as the dependent variable. This model similarly suggested that the activation difference was higher in session 2 than in session 1 ($t(98) = -3.133$, $p = .00228$), indicating that, on average, participants' chances of retrieving the correct answer improved. Finally, the non-retrieval time t_{er} showed a significant decrease from session 2 to session 3 in particular ($t(98) = -2.351$, $p = .0207$), reflecting a speed-up in perceptual and/or motor functions.

In conclusion, exploratory analysis of the inferred ACT-R parameter estimates suggests that the observed increase in accuracy and response speed from session 1 to session 2 could be the result of a higher mean activation of the correct answer and a greater difference in activation between the correct and incorrect answer, while the drop in response times from session 2 to session 3 may be attributable to a decrease in non-retrieval time t_{er} .

Discussion

We have demonstrated a mapping of the parameters of the linear ballistic accumulator onto parameters governing declarative memory retrieval in ACT-R. By fitting the LBA to retrieval data and mapping the inferred LBA parameters onto ACT-R memory parameters, we can arrive at a mechanistic explanation for observed performance changes, without needing to build and fit an ACT-R model directly. The resulting ACT-R parameters—activation of chunks, duration of non-retrieval processes, and latency factor—have cognitively meaningful interpretations within the wider context of the architecture, enhancing the interpretation that could be given by the LBA alone. The mapping extends upon an earlier mapping between the lognormal race model and ACT-R (Nicenboim & Vasissth, 2018), by adding the ability to fit the latency factor. From a theoretical standpoint, ACT-R benefits from this connection to the LBA too: the latency factor is given a more concrete meaning, namely as a measure of response caution.

The method described here allows one to disentangle several factors contributing to memory retrieval performance. In many settings, inside and outside the laboratory, the parameters governing our behaviour are inevitably in flux: we learn and forget, we become tired or impatient, our goals and desires change, we let our minds wander. There is clear explanatory power in being able to capture such changes within a mathematical model. Linking the terms of that mathematical model to constructs defined in a cognitive architecture can further aid the interpretation of observed behaviour.

An important limitation of this method is that it assumes that the distribution of drift rates—and therefore the activation of memory chunks—remains constant across a block of trials. This assumption is most likely to be met when information

is so ingrained that there is no appreciable decay in its activation (e.g., sentence processing; Nicenboim & Vasissth, 2018), or when retrieval attempts are timed such that they occur whenever a particular activation is reached (e.g., adaptive scheduling, as used in our empirical example).

We used a relatively simple procedure for fitting the LBA. Extending this approach to a hierarchical Bayesian LBA may be beneficial (e.g., Nicenboim, 2018). It would enable modelling multiple participants and sessions simultaneously, improving the ability to estimate and compare participant-level and group-level effects, while also capturing the uncertainty in those estimates (Fisher et al., 2020). This could be particularly valuable in smaller data sets, where our current approach still struggles.

In summary, we have demonstrated how a mapping between the linear ballistic accumulator and the ACT-R cognitive architecture can aid in capturing dynamic performance in a cognitive model, thereby contributing to growing efforts to integrate formal modelling approaches.

Acknowledgements

We thank Leendert van Maanen for his helpful guidance on accumulator modelling, and Maurits Baatenburg de Jong and Ruud den Hartigh for their help in collecting the empirical data.

References

- Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* Oxford, UK: Oxford University Press.
- Brown, S. D. & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, 57(3), 153–178. doi: 10.1016/j.cogpsych.2007.12.002
- Fisher, C. R., Houpt, J. W. & Gunzelmann, G. (2020). Developing memory-based models of ACT-R within a statistical framework. *Journal of Mathematical Psychology*, 98, 102416. doi: 10.1016/j.jmp.2020.102416
- Nicenboim, B. (2018). The implementation of a model of choice: The (truncated) linear ballistic accumulator. In *StanCon*. Helsinki, Finland. doi: 10.5281/zenodo.1465990
- Nicenboim, B. & Vasissth, S. (2018). Models of retrieval in sentence comprehension: A computational evaluation using Bayesian hierarchical modeling. *Journal of Memory and Language*, 34. doi: 10.1016/j.jml.2017.08.004
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59–108. doi: 10.1037/0033-295X.85.2.59
- Ratcliff, R., Smith, P. L., Brown, S. D. & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, 20(4), 260–281. doi: 10.1016/j.tics.2016.01.007
- Rouder, J. N., Province, J. M., Morey, R. D., Gomez, P. & Heathcote, A. (2015). The lognormal race: A cognitive-

- process model of choice and latency with desirable psychometric properties. *Psychometrika*, 80(2), 491–513. doi: 10.1007/s11336-013-9396-3
- Singmann, H., Brown, S., Gretton, M. & Heathcote, A. (2020). *Rtdists: Response time distributions*.
- Smith, P. L. & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, 27(3), 161–168. doi: 10.1016/j.tins.2004.01.006
- Usher, M. & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108(3), 550–592. doi: 10.1037/0033-295X.108.3.550
- van der Velde, M., Sense, F., Borst, J. & van Rijn, H. (2021). Alleviating the cold start problem in adaptive learning using data-driven difficulty estimates. *Computational Brain & Behavior*. doi: 10.1007/s42113-021-00101-6
- van Maanen, L. & van Rijn, H. (2007). An accumulator model of semantic interference. *Cognitive Systems Research*, 8(3), 174–181. doi: 10.1016/j.cogsys.2007.05.002
- van Maanen, L., van Rijn, H. & Taatgen, N. (2012). RACE/A: An architectural account of the interactions between learning, task control, and retrieval dynamics. *Cognitive Science*, 36(1), 62–101. doi: 10.1111/j.1551-6709.2011.01213.x
- van Rijn, H., van Maanen, L. & van Woudenberg, M. (2009). Passing the test: Improving learning gains by balancing spacing and testing effects. In *Proceedings of the 9th International Conference on Cognitive Modeling* (pp. 110–115). Manchester, UK.

A Hidden Semi-Markov Model Classifier for Strategy Detection in Multiplication Problem Solving

Ernö Groeneweg (ernogroeneweg@gmail.com)
Utrecht University

Kim Archambeau (k.c.archambeau@uva.nl)
University of Amsterdam & Université Libre de Bruxelles

Leendert Van Maanen (l.vanmaanen@uu.nl)
Utrecht University

Abstract

Self-report as a tool to understand different cognitive processing strategies has been criticised for decades, but to date there have not been many alternatives. To remedy this hiatus, we propose to apply a recently developed method for processing stage analysis (Hidden semi-Markov Model Multivariate Pattern Analysis, HsMM-MVPA) to a cognitive strategy prediction task. HsMM-MVPA uses specific patterns in EEG data to determine the most likely number of sequential processing stages. Under the assumption that cognitive processing strategies differ in the number of stages, we constructed a classifier using fitted HsMM-MVPA to try and differentiate between two cognitive strategies in unseen data. The method is applied to data from a multiplication verification task, in which participants are asked to verify the truth of a solution to a multiplication problem (3×9). We asked participants to indicate via self-report whether they knew the answer by heart (Strategy 1, Retrieval) or needed to compute the answer (Strategy 2, Procedural). The classifier could predict the self report labels above chance, suggesting that the number of processing stages identified using EEG can be used to track the cognitive processing strategy that are in use throughout a task.

Keywords: cognitive strategies; cognitive processing stages; classification; HsMM-MVPA; EEG

Introduction

When different people apply different cognitive strategies to solve the same problem, a question arises: if people use different strategies, how do these strategies differ? It seems that we could differentiate between cognitive strategies by way of the underlying processing stages. A robust method for decomposing a trial into processing stages should then be able to do a differentiation by strategy. To test this hypothesis, we constructed an EEG classifier based on a novel modelling method with which we will try to predict what strategy is used in unseen EEG data.

Processing Stage Decomposition

A recently proposed framework for detecting cognitive processing stages by Anderson et al. (2016), called Hidden semi-Markov Model multivariate pattern analysis (HsMM-MVPA), has been suggested to give nuanced insight into processing stages that might be present in EEG, MEG, or fMRI data. Within this framework, processing stages are modelled as a hidden Markov chain using distributed peaks in activity. By additionally modelling the cognitive processing stages as a semi-Markov chain, we can get insight into the temporal on- and offset as well as the durations of processing stages.

Such nuanced insight into the characteristics of processing stages allows for differentiation between strategies, under the assumption that these characteristics differ between strategies. The HsMM-MVPA framework has been shown to be a versatile method for detecting processing stages in a variety of conditions and tasks (Anderson et al., 2018; Berberyan et al., 2021; Borst & Anderson, 2015; Portoles et al., 2018; van Maanen et al., 2021; Walsh et al., 2017; Zhang, van Vugt, et al., 2018; Zhang, Walsh, & Anderson, 2017, 2018; Zhang, Borst, et al., 2017).

In many implementations, HsMM-MVPA has fitted models that explain EEG data from multiple participants very well. This suggests that there could be some commonality between participants in how the HsMM-MVPA method represents these processing stages. Therefore, it stands to reason that there is overlap in cognitive processing stages between participants employing the same cognitive strategy for the same task. In the current paper, we aim to understand whether an HsMM-MVPA model can be used to distinguish between cognitive strategies in EEG data from unseen participants. Concretely, we collected EEG data from people performing a multiplication task, while we also collected self-reports of the strategies that people use during the task. Then, we estimated the optimal HsMM-MVPA model for the self-reported strategies. We hypothesise that a classifier based on these HsMM-MVPA models predicts which strategy unseen participants used on a particular multiplication problem. If this prediction is above chance, this will support the hypothesis that processing stages can be used to differentiate between strategies.

Hidden semi-Markov Model Multivariate Pattern Analysis Standard Hidden Markov Models consist of two stochastic finite-time chains. One is a hidden Markov chain X and the other is an observable chain Y whose behaviour depends on X . For every pair (x, y) where $x \in X$, $y \in Y$ there is a probability that x happens when y is observed (Visser et al., 2009). In a HMM, the duration of a state corresponds to the duration of a single observation. In contrast, in a hidden *semi*-Markov model, it is possible to have multiple observations per hidden state, which allows for variable state durations (Yu, 2010). Since processing stages are not assumed to have the same duration, HsMMs are best suited for this analysis (Anderson et al., 2016). In this study, extracted components from

EEG data are the observations Y , with the underlying processing stages being modelled by the most likely sequence of hidden states X .

To discover different processing stages, HsMM-MVPA relies on the assumption that a processing stage onset is signified by a cognitive event that can be discovered in the EEG signal by looking for positive or negative peaks, distributed across different brain regions. This assumption is shared by two main theories explaining the generation of event-related potential (ERP); the classical theory (Shah et al., 2004) and the synchronised oscillations theory (Makeig et al., 2002). These theories agree that a cognitive event is signified by a positive or negative peak in the EEG signal, although they disagree on the exact origin of this peak (see Anderson et al. (2016) for a more extensive discussion).

HsMM-MVPA searches for positive or negative peaks in the EEG signal (called *bumps*), with the subsequent *flats* denoting a processing stage. The HsMM-MVPA model consists of a number of bumps, as well as a set of gamma distributions of stage durations across trials. The algorithm first attempts to find bumps that represent the onset of a cognitive stage and the flats that separate these bumps. The goal of HsMM-MVPA is to identify the topography and temporal location of each bump on each trial. The method allows for variability within the duration of cognitive processes for each trial, so bumps can occur at different time points per trial. The trials are analysed individually, but all trials of all participants are taken into account simultaneously. A model is then fitted to various participants simultaneously.

Classification

A challenge with classifying EEG data is that there is a high degree of variability from participant to participant (Saha & Baumert, 2020). Implementations of HsMM-MVPA as described above are able to discover good-fitting models across multiple participants, as long as the participants' EEG signal is collected under the same conditions. This suggests that the number of processing stages is equal across participants using the same strategy, yielding the hypothesis that the EEG signal of different participants performing a task under the same condition can be modelled with the same HsMM-MVPA model. These models will be used to distinguish between the same strategies in unseen EEG data, making classification possible.

We will fit models to EEG data collected from participants verifying single-digit multiplication problems, differentiating between two self-reported strategies: *retrieval* for memory retrieval and *procedural* for procedural strategies. Using these models, we will attempt to classify these same strategies in unseen participants. We determine the sensitivity and specificity of the models to self-reported trials using an receiver operating characteristic (ROC) curve based on the likelihood of the Retrieval strategy. If the area under that curve is larger than 0.5, the model is correctly identifying self-reported Retrieval-trials higher than chance.

We chose to focus our analysis on the retrieval strategy, since previous work suggested that what people interpret as a simple memory retrieval is much more homogenous (Archambeau et al., unpublished). In fact, there are many different solution strategies that all can be described as a procedural strategy (LeFevre et al., 1996; Ashcraft, 1992). For example, to calculate that 6×4 is equal to 24, one can retrieve from memory the related fact that $6 \times 5 = 30$, and then subtract 6. An alternative procedural strategy involves consecutively adding 6 while simultaneously counting the number of additions. When this number reaches 4, you have arrived at the answer.

Method

We collected EEG, accuracy, and response times from individuals verifying single-digit multiplication problems (Archambeau et al., 2019). Participants were shown a single-digit multiplication problem with an answer and asked to verify whether the given answer was correct. Next, they were asked to self-report whether they knew the answer from memory (i.e., the retrieval strategy) or computed the answer another way (Procedural).

Design & Procedure

Forty-two undergraduate students from the Université Libre de Bruxelles (ULB) between the ages of 17 and 52 ($m = 22.24$) took part in the multiplication experiment. The study was approved by the local Ethical Review Board of the ULB, Faculty of Psychological and Educational Sciences. All participants provided informed consent and received course credit for their participation. Each trial started with the presentation of a fixation point of 500 ms. A multiplication problem containing two operands in Arabic format and the multiplication operator "x" (e.g., 6×4) was displayed in the centre of the screen for 200 ms, followed by a blank screen of 120 ms. Then, an answer was shown (24) until a response was provided. The participants were asked to indicate via button press whether the proposed solution of the problem was correct or not. Participants were asked to be as fast and as accurate as possible. When a response is given, a 300 ms interval occurred, after which participants were prompted to report what strategy they used to verify the multiplication problem; "memory" or "calculation strategy". Then, the next trial was initiated with an inter-trial interval of 1000ms. The task consisted of 4 blocks of 248 trials, for a total of 992 trials. There were three trial types. Besides trials where the given solution was correct (positive or P), there were trials two trial types where the given solution was incorrect: interfering solutions (I) and non-interfering solutions (NI). With an interfering solution, the answer given is table-related to one of the operands. The given interfering solution of a problem $a \times b$ could be the correct solution of $(a \pm 1) \times b$ or $a \times (b \pm 1)$. Multiplication problems with a single digit as the correct solution were removed, as well as 9×9 . Half of both I and NI were smaller and half were larger than the correct solution. Although the I and NI split is not relevant for the cur-

rent study, the impact interference has on ERP is accounted for in this analysis. There were 496 positive trials and 496 negative trials with 248 interfering solution trials and 248 unrelated solution trials. Each multiplication problem was repeated 32 times. On 16 of these repetitions the provided solution was correct, on 8 it was an I solution, and on 8 it was an NI solution. The multiplication problems were presented in a pseudo-randomised order, ensuring that successive problems never had the same operands. The multiplication task was run on a 17-inch laptop computer, using the Psychophysics Toolbox extension (Brainard, 1997) in MATLAB (version R2013a, The Mathworks Inc., Natick, Massachusetts, USA). EEG data was collected using a BioSemi interface with 72 channels, at a sampling rate of 2048 Hz. EMG activity was also recorded for other purposes, beyond the scope of the current study.

Behavioural Analysis

Four data subsets were created. Three of these based on the three experimental conditions positive (P), related (I), and unrelated (NI). For the fourth subset, we collapsed all data under the assumption that the two self-report strategies would be shared across the three experimental conditions, increasing the sample size. Participants 32-42 were split off as a test set, making participants 1-31 the training set. For this analysis, incorrect responses were removed (6.5% of trials). Additionally, we removed response time outliers from all subsets of the data to remove any trials where the participant may have been confused or distracted. When matching the behavioural data to the EEG data for epoching, four participants from the training set were removed due to incorrect event numberings in the EEG data. In total, 15% of the data was removed in this step.

After cleaning our data sets, class imbalance was computed. As seen in Table 1, in all subsets the majority of trials is labelled as 'retrieval'. These numbers will be considered as a baseline for classification accuracy. Although the standard deviation in RT in the test data subsets is much lower (potentially due to some participants having generally slower RTs in the training set), HsMM-MVPA can account for variations in the temporal offset and stage durations (Anderson et al., 2016).

Data Preprocessing

The data was processed in MATLAB (version 2020a, The Mathworks Inc., Natick, Massachusetts, USA) using the open source EEGLab plugin (Delorme & Makeig, 2004). First, the data was re-referenced to all-electrode average and high-pass filtered at 1 Hz and low-pass filtered at 40 Hz, as oscillations outside of this range are not commonly associated with brain activity (Henry, 2006). The data was resampled to 512 Hz and flatlines and overly noisy sections of data were removed automatically using built-in EEGLab functions, before applying Independent Component Analysis (ICA) using the FastICA algorithm (Hyvarinen, 1999). Next, the ICLabel plugin was used to automatically flag non-brain related components

Condition	% Retrieval	Mean RTs (ms)	SD RTs (ms)
Training data			
All	87.1%	1398	2211
P	88.5%	1312	2051
I	83.4%	1628	2101
NI	87.9%	1339	2581
Test data			
All	88.4%	1110	874
P	91.9%	1040	819
I	81.3%	1280	998
NI	84.0%	1099	837

Table 1: Overview of trials labelled 'retrieval' after error and outlier removal. Standard deviation is computed within every subset (Positive (P), Interfering (I), and non-interfering (NI), across participants.

from the data (Pion-Tonachini et al., 2019). These flagged components were then removed from the data. In total about 10% of the data was removed in this step.

Fitting HsMM-MVPA

To fit the HsMM-MVPA models, the data were resampled to 100 Hz and then epoched between stimulus onset and the participants' response. Spatial principal component analysis (PCA) was applied to all datasets to extract the 10 principal components from the data channels. In all subsets, the 10 principal components account for more than 97% of variance in the data.

We consider our bumps to have a duration of 50 ms, as this duration produces robust results even if the actual durations are slightly longer or shorter (Anderson et al., 2016). The duration of the subsequent flats was modeled with a gamma distribution with a shape parameter of 2. The results are not sensitive to the exact choice of shape parameter, except that it simplifies the estimation of flat distributions (Anderson et al., 2016). In a model, n bumps results in $n + 1$ flats (or $n + 1$ processing stages), since the first stage starts with a flat when the stimulus is applied.

We constructed different HsMM-MVPA models for every subset of the data and for both self-reported strategies, resulting in eight different models. Model estimation begins with a 1-bump model and creates models for an increasing number of bumps until a number of bumps n_{max} is reached, with n_{max} being the maximum number of 50 ms bumps that fit in the duration of the shortest epoch. During estimation, two parameters of each hidden state are obtained: (1) the amplitudes of the bumps that mark the onsets of the processing stages and (2) the scale parameter of a gamma distribution describing the stage durations. Data from all trials and all participants in a training set were taken into account simultaneously. The match between the EEG data and the model was maximised using a standard expectation-maximisation (EM)

algorithm (Moon, 1996).

The fitting process begins by defining initial amplitudes for both the bumps and the gamma distributions for stage durations. Since the convergence of the EM algorithm can be sensitive to the choice of starting point, ending up in a local maximum (Wu, 1983), we used a process based on work by Zhang, Walsh, & Anderson (2018). Per subset, we first fit separate HsMM-MVPA models for each condition on n_{max} bumps, obtaining bump amplitudes and gamma distributions. Next, we used those parameters for models with $n_{max} - 1$ bumps, iteratively leaving out each of the bumps in n_{max} , selecting the model with best fit. These bumps become the new n_{max} before the above process is repeated until only a one-bump model n_1 is left. This way, we can find all potential bump topologies while avoiding local maxima.

We used a leave-one-out cross-validation (LOOCV) procedure to prevent overfitting. For every training subset, we estimate an HsMM-MVPA model on all participants minus one and then test the fit of this model on the omitted participant. This process is repeated for all participants. Finally, we used a sign test to test for how many participants the log-likelihoods of the models with $n + 1$ bumps increased significantly compared to an n -bump model. If a model with one additional bump outperforms the previous model for a sufficiently large number of participants, we can say that the additional complexity of that model is warranted. This step is crucial for fitting models that generalise well across participants (Anderson & Fincham, 2014). To select the best models, sign tests were used on every n -bump model to see for how many participants its fit improved compared to the $n - 1$ bump model. The best model is one that improves significantly for more than half of the training participants. For a more detailed mathematical description and code for HsMM-MVPA we refer to Anderson et al. (2016) and Berbery et al. (2021).

Classification

After estimating the most likely parameters for all models, we used our preprocessed test data to estimate the likelihood of every trial per subset under the models. We also estimated the likelihoods of all test trials per subset under models of different subsets to further test how well the models generalise. As this is a binary classification task, we compute true positive rates (TPR) and false positive rates (FPR) to plot ROC curves. Then, we determine the area under the curve (AUC), where 0.5 denotes the model classifying according to chance. We also report classification accuracy (the proportion of trials classified as their corresponding self-reports) and F1-values (computed as $F_1 = \frac{TPR}{TPR+(FPR+FNR)/2}$) for all test trials classified under all models.

Results

Model Selection

As can be seen in Figure 1, Retrieval models are consistent. In all four subsets, the models fit to the Retrieval strategy show

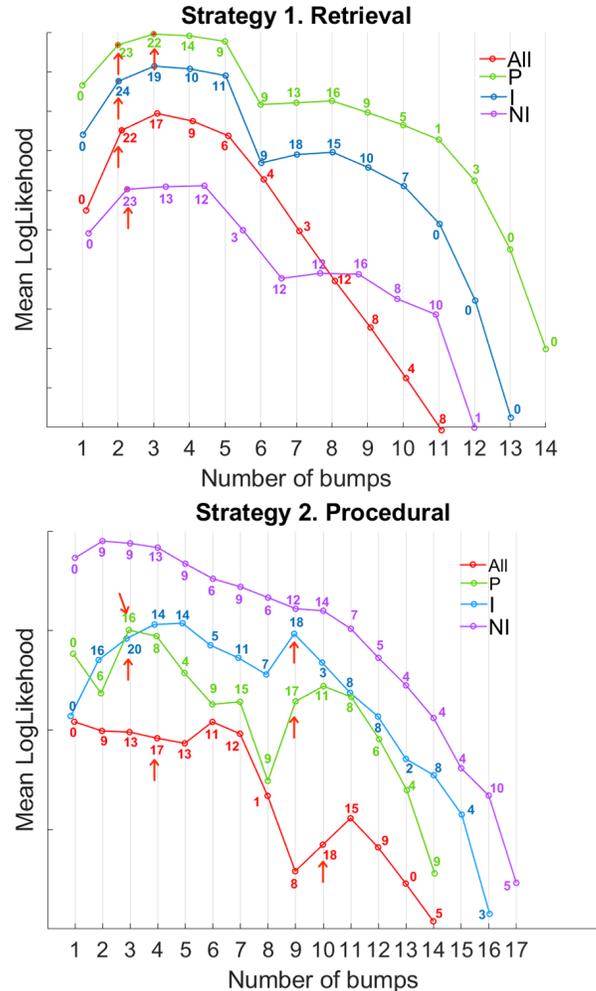


Figure 1: Model selection curves for both strategies. The numbers beside the points denote for how many out of 27 participants the log-likelihood of an n -bump model significantly increases over $n - 1$ bumps. The arrows highlight the optimal models according to a sign-test for each subset of the data (see the section Fitting HsMM-MVPA for details).

that the 2-bumps model improves significantly over the 1-bump model in more than half of participants. As this result is consistent across all four subsets, the assumption seems justified that the Retrieval-strategy is generally well described by a 2-bump, 3-stage Hidden semi-Markov model. As there is no significant improvement going to a 3-bump, 4-stage model, 2-bump 3-stage models were selected. In the Procedural strategy, results are far less consistent. As a different number of bumps seems to perform best in all four subsets, we classify using the Retrieval-model only.

Classification Results

After estimating log-likelihood of every trial in the test set under all four Retrieval-models, we constructed Receiver Operating Characteristic-curves with an increasing discrimination

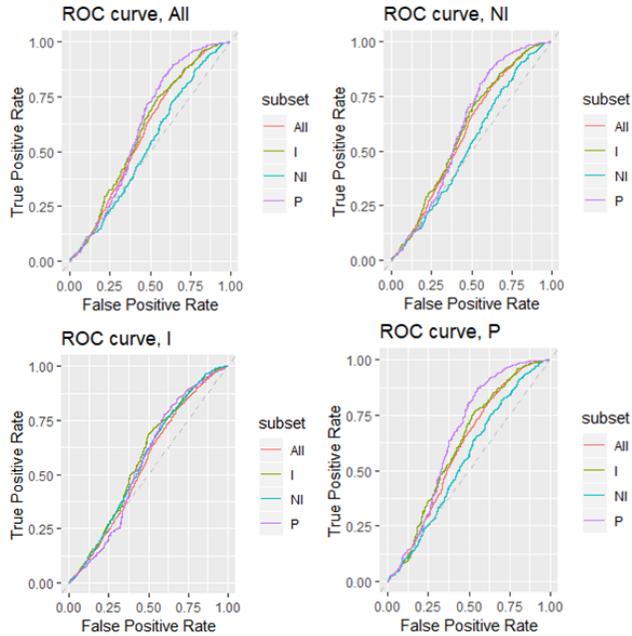


Figure 2: ROC curves for all test sets predictions under all models

threshold (Figure 2). Accuracies and F1 values can be seen in Table 2. The highest accuracy and F1 score per subset were selected. We see that all classifiers perform better than random (curve above the grey dashed line, $AUC > 0.5$), with P-trials in the test set performing best overall, averaging at 0.61. AUC's were computed per test participant for significance testing, showing that under all models All and P-trials are significantly higher than chance. We also see that F1-scores tend to be higher than accuracies, denoting that the classifiers are better at correctly identifying Retrieval-trials as Retrieval than they are at identifying non-Retrieval trials as non-Retrieval.

Discussion

The goal of the current study was to discover whether cognitive strategies can be differentiated between based on the number of processing stages. In general we can say with confidence that HsMM-MVPA when applied to EEG data can do so in a way that finds processing stages across participants. In other words, an HsMM-MVPA model fit to a specific cognitive strategy can recognise most unseen trials of that same strategy, even when that unseen trial is from a participant whose EEG data the model has not seen at all. This implies that, when different people report using the same cognitive strategy under the same experimental condition, their EEG patterns and consequently the processing stages are similar as well. There is also consistency between the models fit to the four experimental conditions with respect to their classification on the test data subsets, but almost everywhere this consistency is proportional to the variation in the class im-

Subset	AUC	Max F1	Max acc
Model trained on All			
All	0.584‡	0.943	89.3%
P	0.612‡	0.974	94.9%
I	0.597†	0.902	82.4%
NI	0.527	0.916	84.5%
Model trained on P			
All	0.603‡	0.944	89.4%
P	0.653‡	0.974	95.0%
I	0.618	0.902	82.6%
NI	0.550	0.917	84.7%
Model trained on I			
All	0.548‡	0.942	89.1%
P	0.557‡	0.974	94.9%
I	0.577†	0.902	82.2%
NI	0.568	0.943	89.3%
Model trained on NI			
All	0.586†	0.943	89.3%
P	0.610‡	0.974	95.0%
I	0.598†	0.902	82.4%
NI	0.529	0.916	84.7%

Table 2: Area Under Curve (AUC) of the ROC curves, as well as the highest F1 scores and accuracies for every test set.

†: significant below $p=0.05$

‡: significant below $p=0.01$

balance of our data. F1 scores everywhere tend to be higher than classification accuracy, which is closer to random performance. This means that a retrieval-based classifier is accurate at identifying trials that were self-reported as retrieval, but less accurate with respect to trials that were self-reported as procedural.

In our data, the retrieval strategy is the more consistent one. All 2-bumps, 3-stage models were well fitting on all four data subsets. This means that it is highly likely that participants used 3 processing stages when using the retrieval strategy. In contrast, the procedural-strategy seems to be much less cohesive. A first possible explanation is that participants can calculate the answer to a problem in different ways, which could lead to a different number of processing stages. This means that "procedural" encompasses a number of strategies (LeFevre et al., 1996; Ashcraft, 1992). Some of these strategies might be closer in number of stages to our 3 memory retrieval stages, which could partially account for the number of false positives our retrieval-classifier finds. For instance, a procedural strategy to verify 6×7 could involve memory retrieval of 6×6 as part of the strategy, leading this hypothetical trial to fit well with our retrieval models.

To complicate matters further, there is the potential of noise or biases to be introduced when using self-reports as a tool for

setting our ground-truth (Kirk & Ashcraft, 2001). This means that wrongly labeled trials might be present. Alternative ways of ascertaining cognitive strategies, like a mixture modelling approach could be considered for this purpose (Archambeau et al., unpublished; Thevenot et al., 2007; van Maanen et al., 2014, 2016). There is also a discrepancy in the means and standard deviations of our response times between our training and test sets. HsMM-MVPA is able to account for variance in temporal onset and duration of processing stages (Anderson et al., 2016), so in theory this is no problem. However, this is a variable that could be corrected for in future.

Future work might encompass cross-validating this analysis with different train-test data splits. In addition, a comparison between scalp topographies of the best models and both correctly-labeled and mislabeled test set trials could give insight into trials that might have been incorrectly self-reported. Another avenue would be to apply this analysis to a different experimental task, such as a division task instead of multiplication. Finally, further insight into using model-based classifiers instead of more traditional machine learning approaches could improve the explainability of cognitive data classifiers. HsMM-MVPA provides a nuanced understanding of processing stages that other machine learning methods often do not.

In conclusion, this first investigation into using HsMM-MVPA as a tool for classification of cognitive strategies shows promise. The next step would be to investigate how far this promise can lead.

References

- Anderson, J. R., Borst, J. P., Fincham, J. M., Ghuman, A. S., Tenison, C., & Zhang, Q. (2018). The common time course of memory processes revealed. *Psychological science*, 29(9), 1463–1474.
- Anderson, J. R., & Fincham, J. M. (2014). Extending problem-solving procedures through reflection. *Cognitive psychology*, 74, 1–34.
- Anderson, J. R., Zhang, Q., Borst, J. P., & Walsh, M. M. (2016). The discovery of processing stages: Extension of sternberg's method. *Psychological review*, 123(5), 481.
- Archambeau, K., De Visscher, A., Noël, M.-P., & Gevers, W. (2019). Impact of ageing on problem size and proactive interference in arithmetic facts solving. *Quarterly Journal of Experimental Psychology*, 72(3), 446–456.
- Archambeau, K., Molenaar, D., Forstmann, B., Noël, M.-P., Gevers, W., & Van Maanen, L. (unpublished). Age-related differences in the resolution of interference in simple arithmetic depends on the strategy used.
- Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition*, 44(1-2), 75–106.
- Berbery, H. S., van Maanen, L., van Rijn, H., & Borst, J. (2021). Eeg-based identification of evidence accumulation stages in decision-making. *Journal of Cognitive Neuroscience*, 33(3), 510–527.
- Borst, J. P., & Anderson, J. R. (2015). The discovery of processing stages: Analyzing eeg data with hidden semi-markov models. *NeuroImage*, 108, 60–73.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, 10(4), 433–436.
- Delorme, A., & Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1), 9–21.
- Henry, J. C. (2006). Electroencephalography: basic principles, clinical applications, and related fields. *Neurology*, 67(11), 2092–2092.
- Hyvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE transactions on Neural Networks*, 10(3), 626–634.
- Kirk, E. P., & Ashcraft, M. H. (2001). Telling stories: The perils and promise of using verbal reports to study math strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(1), 157.
- LeFevre, J.-A., Bisanz, J., Daley, K. E., Buffone, L., Greenham, S. L., & Sadesky, G. S. (1996). Multiple routes to solution of single-digit multiplication problems. *Journal of Experimental Psychology: General*, 125(3), 284.
- Makeig, S., Westerfield, M., Jung, T.-P., Enghoff, S., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2002). Dynamic brain sources of visual evoked responses. *Science*, 295(5555), 690–694.
- Moon, T. K. (1996). The expectation-maximization algorithm. *IEEE Signal processing magazine*, 13(6), 47–60.
- Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). Iclabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, 198, 181–197.
- Portoles, O., Borst, J. P., & van Vugt, M. K. (2018). Characterizing synchrony patterns across cognitive task stages of associative recognition memory. *European Journal of Neuroscience*, 48(8), 2759–2769.
- Saha, S., & Baumert, M. (2020). Intra- and inter-subject variability in eeg-based sensorimotor brain computer interface: A review. *Frontiers in Computational Neuroscience*, 13, 87. Retrieved from <https://www.frontiersin.org/article/10.3389/fncom.2019.00087> doi: 10.3389/fncom.2019.00087
- Shah, A. S., Bressler, S. L., Knuth, K. H., Ding, M., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2004). Neural dynamics and the fundamental mechanisms of event-related brain potentials. *Cerebral cortex*, 14(5), 476–483.
- Thevenot, C., Fanget, M., & Fayol, M. (2007). Retrieval or non-retrieval strategies in mental arithmetic? an operand recognition paradigm. *Memory & Cognition*, 35(6), 1344–1352.
- van Maanen, L., Couto, J., & Lebreton, M. (2016). Three boundary conditions for computing the fixed-point property in binary mixture data. *Plos one*, 11(11), e0167377.
- van Maanen, L., de Jong, R., & van Rijn, H. (2014). How to assess the existence of competing strategies in cognitive tasks: a primer on the fixed-point property. *PloS one*, 9(8), e106113.
- van Maanen, L., Portoles, O., & Borst, J. P. (2021). The discovery and interpretation of evidence accumulation stages. *Computational brain & behavior*.
- Visser, I., Raijmakers, M. E., & van der Maas, H. L. (2009). Hidden markov models for individual time series. In *Dynamic process methodology in the social and developmental sciences* (pp. 269–289). Springer.
- Walsh, M. M., Gunzelmann, G., & Anderson, J. R. (2017). Relationship of p3b single-trial latencies and response times in one, two, and three-stimulus oddball tasks. *Biological psychology*, 123, 47–61.
- Wu, C. J. (1983). On the convergence properties of the em algorithm. *The Annals of statistics*, 95–103.
- Yu, S.-Z. (2010). Hidden semi-markov models. *Artificial intelligence*, 174(2), 215–243.
- Zhang, Q., Borst, J. P., Kass, R. E., & Anderson, J. R. (2017). *Inter-subject alignment of meg datasets in a common representational space* (Tech. Rep.). Wiley Online Library.
- Zhang, Q., van Vugt, M., Borst, J. P., & Anderson, J. R. (2018). Mapping working memory retrieval in space and in time: A combined electroencephalography and electrocorticography approach. *NeuroImage*, 174, 472–484.
- Zhang, Q., Walsh, M. M., & Anderson, J. R. (2017). The effects of probe similarity on retrieval and comparison processes in associative recognition. *Journal of Cognitive Neuroscience*, 29(2), 352–367.

Zhang, Q., Walsh, M. M., & Anderson, J. R. (2018). The impact of inserting an additional mental process. *Computational Brain & Behavior, 1*(1), 22–35.

Individualizing a Biomathematical Fatigue Model with Attention Data

Bella Z. Veksler (b.veksler@tier1performance.com)

Tier1 Performance Solutions
Covington, KY 41011 USA

Megan B. Morris (megan.morris.3@us.af.mil) and Glenn Gunzelmann (glenn.gunzelmann@us.af.mil)

Air Force Research Laboratory
WPAFB, OH 45433 USA

Abstract

Fatigue is a problematic factor in many workplace environments, resulting in safety and health risks that require monitoring and management. One means to monitor and manage fatigue is through the use of tools implementing biomathematical fatigue models to create assessment and predictions of operator fatigue based on sleep habits. Unfortunately, these models tend to provide assessments and predictions for an “average” operator given work schedules, lacking individualization. One way in which these models can be individualized is through the use of at-the-moment performance data that can modulate the model estimates. In the current effort, we describe an initial attempt at developing an algorithm to individualize fatigue assessments and predictions from a widely-used biomathematical fatigue model with performance data from a common attention task. We discuss the sleep datasets used for the effort, scaling procedure, and model fitting using a genetic algorithm. We then discuss future directions we will take to further increase the effectiveness and efficiency of the individualization capability and its implications.

Keywords: Psychomotor Vigilance Test; Genetic Algorithm

Fatigue is a problematic factor in several workplace environments such as aviation (Caldwell & Caldwell, 2016), commercial motor vehicle (National Academies of Sciences, Engineering, and Medicine, 2016), railroad (Gertler, DiFiore, & Raslear, 2013), and medical (Kancherla et al., 2020) operations. Given the resulting safety and health risks associated with fatigue, it is crucial that organizations implement fatigue risk management (FRM) programs, policies, and other mitigation efforts to combat fatigue. Traditionally, organizations have commonly implemented policy limits regarding work/duty hour limits and rest breaks to allay fatigue. Increasingly, organizations have implemented various types of FRM programs that provide resources and tools to help mitigate fatigue, document fatigue, and examine incidents involving fatigue (Gander et al., 2011). One tool found within some high-risk operational setting programs is the use of biomathematical fatigue

models to create assessments and predictions of operator fatigue. Biomathematical fatigue models include homeostatic regulation and circadian rhythm processes, among other factors, to create predictions of fatigue for operators (Mallis et al., 2004).

One particular model that is used by organizations such as the United States Air Force (USAF) Air Mobility Command (AMC), the U.S. Federal Railroad Administration, among others, is the Sleep, Activity, Fatigue, and Task Effectiveness biomathematical fatigue model (SAFTE; Hursh, Redmond, et al., 2004). This model is typically used as the basis of the Fatigue Avoidance Scheduling Tool (FAST; Hursh, Balkin, et al., 2004), a tool that provides fatigue predictions based on prescriptive sleep schedules given work and rest times. One issue with the SAFTE model and other similar biomathematical models is that the model provides predictions for an “average” operator, lacking individualization. Some researchers have had success individualizing predictions of biomathematical fatigue models. Recently, Liu et al. (2017) had success in individualizing the Unified Model of Performance (UPM; Rajdev et al., 2013) with Psychomotor Vigilance Test (PVT; Dinges & Powell, 1985) reaction times. Since the SAFTE model is the basis of several FRM programs and research has provided support of its effectiveness (Hursh, Redmond, et al., 2004; Van Dongen, 2004), we believe it is advantageous to implement a similar technique as Liu et al. (2017). In the current effort, we develop an algorithm to modulate the SAFTE model fatigue estimates with PVT data. This will provide more valid fatigue assessments from the biomathematical model through individualization gained from use of PVT data.

SAFTE Model

The SAFTE model is a three process model that includes homeostatic regulation, circadian rhythm, and sleep inertia processes to calculate general performance effectiveness

(fatigue) predictions. The model also includes a process to account for chronic sleep deprivation. The SAFTE model embedded in FAST is proprietary and includes additional features to take time zone changes and light into account. In the current effort we utilize the non-proprietary version of the SAFTE model (Hursh, Redmond, et al., 2004). The SAFTE model includes 16 parameters. These are listed in Table 1, along with their general mechanism within the model and effects on the output of the model when modified.

Psychomotor Vigilance Test

The PVT is one of the most widely used tasks to assess fatigue due to its sensitivity to sleep decrements and robustness to learning effects (Arsintescu et al., 2017; Balkin et al., 2000; Basner & Dinges, 2011). In the PVT, participants wait for a rolling reaction time indicator in milliseconds to appear on a computer screen in a known location. When this indicator appears, the participant must respond as fast as possible. The PVT is traditionally 10 minutes in length and has a random inter-stimulus interval (ISI) of 2 to 10 seconds. Mean and median reaction time and number of lapses (reaction times greater than 500 ms) are the most common metrics examined to assess alertness or fatigue, but there are several other metrics that are also sensitive to fatigue (e.g., mean 1/RT, slowest 10% 1/RT, etc.) (Basner & Dinges, 2011).

Current Effort

In the remainder of the paper, we will describe our process for individualizing the SAFTE biomathematical fatigue model using PVT data. First, we will describe the archival sleep deprivation dataset used to develop and test the algorithm. We will then describe the process to scale model outputs to the PVT outcomes and how specific SAFTE parameters were chosen. We will then demonstrate the predictive capability of fitting the chosen parameters to individuals. Lastly, we will discuss implications of this work and future plans.

Table 1: SAFTE Model Parameters

Par	Rep	Effects	DV	RE
p	24h acrophase	Shifts effectiveness curve left and right	18	[1,24]
pp	12h acrophase	Changes shape of effectiveness curve	3	[1,12]
beta	Relative amplitude of 12h rhythm	If both circadian peaks are at the same height	0.5	[0,1]
m	Sleep propensity mesor	Positive values increase sleep inertia	0	[-5,10]
as	Sleep propensity amplitude	Higher values increase effectiveness	.55	[-5,5]
a1	Performance rhythm amplitude (fixed %)	Height of peak of circadian component	7	[0,20]
a2	Performance rhythm amplitude (variable %)	Height of peak of circadian component	5	[0,20]
f	Feedback amplitude	How gradually sleep increases reservoir	.00262 43	[0,1]
k	Performance use rate	Depletion rate while awake	.5	[0,1]
k1	Down-regulation time constant	Only during sleep	.22	[0,5]
k2	Reference level for SI regulation	Only during sleep	0.5	[.01,5]
k3	Recovery time constant	Only during sleep	.0015	[0,5]
SI max	Max sleep accumulation per minute	Only effect when sleep <=3 hours	3.4	
I	Sleep inertia time constant	Only effect 2 hours following awakening	.04	
I max	Max inertia following awakening	Only effect 2 hours following awakening	5	
RC	Reservoir capacity	Kept constant across participants	2880	

Note. Par = Parameter; Rep = Represents; DV = Default Value; RE = Range Explored

Method

Dataset

To test fits from the model we utilized PVT data from two 62-hour sleep deprivation studies run at Washington State University (Tucker, Whitney, Belenky, Hinson, & Van Dongen, 2009; Whitney, Hinson, Jackson, & Van Dongen, 2015). The first dataset (Whitney, Hinson, Jackson, & Van Dongen, 2015) included 26 participants ($M_{age} = 25.92$, $SD_{age} = 4.05$, $Range_{age} = 22-37$, 16 males and 10 females) from the surrounding Washington State University community. Participants were randomly assigned to a sleep deprivation ($n = 13$) or control group ($n = 13$). The second dataset (Tucker, Whitney, Belenky, Hinson, & Van Dongen, 2009) included 23 participants ($Range_{age} = 22-38$, 12 males and 11 females) also from the surrounding Washington State University community. Participants were randomly assigned to a sleep deprivation ($n = 12$) or control group ($n = 11$).

The following description of the protocol was common to both studies, except where noted. Participants spent 6 consecutive days (7 in the first study) and 6 nights in the lab. The first two days were a baseline period where participants had 10 hours time in bed from 22:00 to 08:00 each night. The control group continued this sleep schedule in the following days, but the sleep deprivation group was deprived of sleep for 62 continuous hours. The last two days were a recovery period where both groups had 10 hours time in bed each night. Participants completed several different tasks during the studies, but we only focus on the PVT task in the current effort. The PVT task was 10 minutes in length with a random ISI of 2 to 10 seconds. PVT bouts were collected about every 2 hours during scheduled time awake. This resulted in 8 baseline bouts for both groups, 24 bouts for the sleep deprivation group and 14 bouts for the control group during the sleep deprivation period, and 10 recovery bouts for both groups. We specifically focused on the sleep deprivation groups from both studies for this modeling effort. For fitting the model to the human data, we aggregated each participant's median RT by bout.

Scaling Model Outputs to PVT

Sleep schedule input into SAFTE followed the protocol described above. Output from the SAFTE model produces an effectiveness value on a scale of 0 to 1. As is the case in many biomathematical models, the output requires scaling and inversion to reflect the dependent measure of interest (in this case, the median RT per bout) (Van Dongen, 2004). We linearly transformed this value using the following formula: $Model = scale + scale * (1-EV)$, where $scale$ is determined for each participant and is the minimum median RT from all

the bouts that went into fitting the model, EV is the effectiveness value output from SAFTE.

Finalizing Parameters to Modulate

After reviewing the effects of each SAFTE parameter on model output, we found that 12 of the parameters were good potential candidates for the individualization of model output. The culling of the original 16 parameters to 12 was done by visually inspecting the effects of each parameter independently with respect to a 62h sleep deprivation sleep schedule. We found that $Slmax$, i , and $Imax$ all had minimal effects on the Effectiveness values output by the model during periods of wakefulness. We chose to also keep the Reservoir Capacity (RC) constant across all participants as its magnitude is directly related to the k parameter which controls the rate at which the reservoir is depleted during wakefulness. Rather than vary both parameters, we chose the k parameter to vary. Table 1 also lists the ranges we used in exploring parameter effects.

The parameter space we wanted to explore in this work was fairly large (as seen in Table 1) and rather than try to run a brute force exhaustive search for each participant, we turned to genetic algorithms. Genetic algorithms have been used in many domains in order to find sets of parameters that minimize some fitness functions fairly efficiently (Fogel, 2006). We used the GA package in R to run a genetic algorithm with a population size of 50 and convergence determined by 50 generations with no change in fitness (Scrucca, 2013, 2017). The fitness function used in the GA was the root mean squared error (RMSE) between the human data and model output after scaling of the median RT. This initial parameter exploration was done using all of the bouts of data for each participant.

Although there are likely significant interactions between the various parameters, as a first pass at determining how much each parameter contributes to individually fitting the human data, we ran the genetic algorithm while varying 11 parameters and keeping the 12th constant. As a control, we used a model with default SAFTE parameters (green dotted line in all figures). Figure 1 shows the resultant average error across all participants when each parameter was held constant at its default value while the rest were explored. The red line in the figures indicates the error when all 12 parameters were varied. As the figure suggests, maintaining the k parameter at default had a considerable effect on the model's error. Using the results derived here, we compared the average error to the 12 parameter model and conservatively culled any parameters whose exclusion (keeping them at default values) either resulted in better performance than the 12 parameter model or were within .5

error units. From this point forward we kept the p , pp , m , f , and $k3$ parameters at default values.

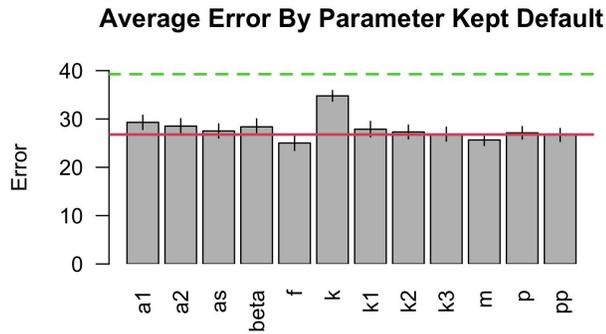


Figure 1: Average RMSE across all participants when fitting by keeping each parameter constant while the other 11 are varied.



Figure 2: Average RMSE across all participants when fitting by keeping each parameter constant while the other 6 are varied.

We then repeated the above procedure with the 7 remaining parameters and ran the genetic algorithm while varying 6 parameters and keeping the 7th at default. The results are shown in Figure 2. Given these results, we found that we can further keep as default the parameters $a1$ and $beta$. In an attempt to further reduce the number of parameters, we ran the GA only varying the $a2$, as , and k parameters as those seemed to provide the largest improvement to fit, as well as only varying the k parameter. Figure 3 shows the average error based on the number of parameters compared to the fully default model (green line) and the 12 parameter model (red line). Based on these results, the 5 parameter model which varies $a2$, as , k , $k1$, and $k2$ produces the best individual fits to our dataset. These parameters correspond to how high the peak of the circadian component is ($a2$), the amplitude of sleep propensity with higher values resulting in higher effectiveness values (as),

how quickly the reservoir is depleted (k), and how quickly the reservoir is refilled ($k1$ and $k2$).

There was a statistically significant difference in error in the number of parameters in the model as determined by a linear mixed effects model, ($F(5, 836) = 224.68, p < .001$). Post-hoc tests indicated that both the 5-parameter model and the 7-parameter models have significantly less error than the 12-parameter or default models ($p < .05$). In the interest of simplicity, we used the 5-parameter model for predicting performance for each PVT bout based on all preceding bouts.



Figure 3: Average RMSE across all participants of the best fitting models by number of free parameters.

Results

Having settled on the 5 parameters we found to be the most appropriate for capturing individual differences in our dataset, we subsequently ran the genetic algorithm to find the best fitting parameters for each participant up to each bout time in order to predict the next bout's performance. Since the goal of the current work is to be able to adjust parameters in real-time to predict future performance, this approach should establish the validity of using the 5 parameters to fit individuals. Figure 4 depicts the average error across all participants during each bout based on the parameter set which minimizes the error of all previous bouts. For comparison, we also used the SAFTE model with default parameters and scaled each individual's performance as before. There was a statistically significant interaction between model type and hour on the error between the model's predicted median RT and the human data as determined by a linear mixed effects model, ($F(33, 1611) = 6.03, p < .001$) as well as both the main effect of type ($F(1, 1611) = 167.54, p < .001$) and hour ($F(33, 1611) = 20.28, p < .001$). Post-hoc tests indicated that there were significant differences between the default model and the 5-parameter individualized model in all hours between 105 and 141 into

the study ($p < .01$). These hours corresponded to being awake for 25 to 61 hours.

It should be noted that there was considerable variability between participants in terms of how accurately the individualized model was capable of predicting performance. In particular, the majority of participants ($n = 19$) had an average error of 30 ms or less across all bouts. However, in some participants the later bouts which occurred during the sleep deprivation period were not as well fit by the model, as shown in Figure 4 in which bouts occurring during hours 120-140 have high variability and higher error. After inspecting the poorer fitting participants, we found that there was a difference in goodness of fit between the participants from the first study and those of the second. It was unclear why this would be the case as both studies used the same protocol. However, after filtering out the participants from the second study, we found a reduction in prediction error which mimicked that of the error we see when fitting the entire data set, see Figure 5. We again found a significant interaction between the model type and hour, ($F(33, 795) = 12.27, p < .001$) as well as both the main effect of type ($F(1, 795) = 318.88, p < .001$) and hour ($F(33, 795) = 20.68, p < .001$). As in the above analysis, post-hoc tests revealed significant differences between the two models for hours 105 to 141.

Taken together, our initial individualized modeling effort resulted in much better predictions of next bout performance than the default parameter model despite the large variability inherent during sleep deprivation bouts.

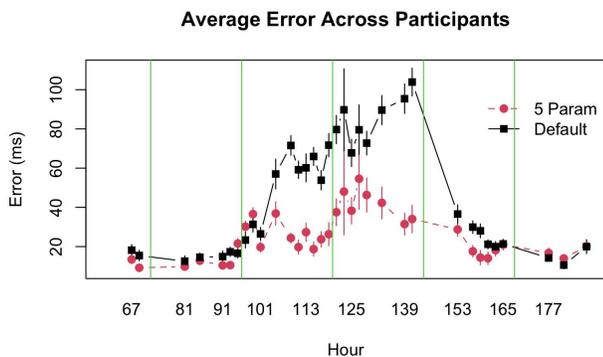


Figure 4: Predicted bout’s median RT error based on all previous bout data, includes 25 participants in the sleep deprivation condition. Green lines are day boundaries.

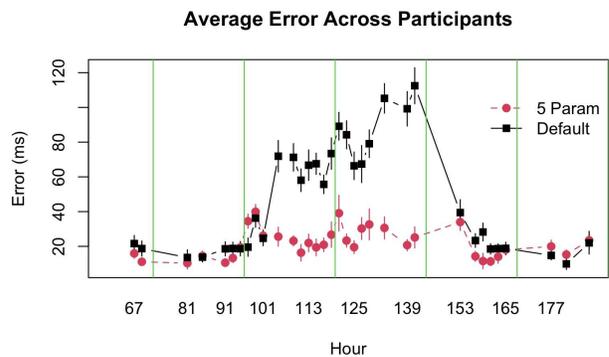


Figure 5: Predicted bout’s median RT error based on all previous bout data, includes only the 13 participants from the first 62-h study who were in the sleep deprivation condition.

Discussion

In the current effort, we have demonstrated an initial attempt to develop an algorithm to individualize SAFTE biomathematical fatigue model estimates with PVT performance data. Although SAFTE includes 16 parameters which could theoretically all be manipulated in order to fit individual performance data, we attempted to cull the number of parameters down, both to avoid overfitting and to more efficiently find best fits, while still maintaining the ability to both fit the data and predict future performance. Out of the 16 parameters, 3 were negligible in their contribution to effectiveness values while awake and a 4th highly correlated with another parameter. Further exploration of the remaining 12 parameters found that 5 more could be culled without appreciably affecting the individual fits. Furthermore, a 5-parameter model was capable of fitting individual data as well as a 7-parameter model. Further reducing the number of parameters, however, produced worse fits. This suggests that these five parameters are associated with important individual differences regarding fatigue. The $a2$ parameter is likely associated with differences in circadian typology, the as parameter is associated with how quickly individuals fall asleep and their ability to stay asleep, and $k, k1,$ and $k2$ parameters are likely associated with differences in sleep need.

We then used the 5-parameter model to fit individual performance up to a particular bout and predict performance on the subsequent bout. The error between the model’s predicted median RT and that of each individual participant’s was within a range commensurate with using the entire data set to fit the parameters, suggesting that this approach may allow us to update parameter estimates with

limited data and provide a more individualized model of performance than the default SAFTE model. Although the individualized predicted fits get somewhat worse during the sleep deprivation period, they are still much better than using the default parameters.

The traditional PVT implementation is not practical in operational contexts due to the length of the task (10 minutes) and the hardware used to collect the reaction times (e.g., desktop computer or laptop) (Lamond et al., 2005). As a result, researchers have examined the validity of shorter PVT implementations (e.g., 5 or 3 minutes) on handheld devices (e.g., Basner et al., 2011, Grant et al., 2017; Lamond et al., 2005). Overall, these studies have found these implementations to be valid assessments of fatigue when the traditional PVT implementation is not possible. Three minute smart-phone based PVTs are an especially attractive alternative for operational environments as they are short in duration and operators commonly carry these devices on their person. As a result, real-time performance from a smartphone PVT can be used to individualize biomathematical fatigue models within FRM programs. The current effort is a first step in allowing us to use the output from these shorter duration PVT implementations to individualize predictions.

We will continue to improve upon the algorithm by testing with additional sleep deprivation, restricted sleep, and shift-work datasets to demonstrate performance in various sleep impairment-related contexts. Future work will also explore other scaling mechanisms as well as different dependent measures such as number of lapses and false alarms as those have also typically been used to evaluate PVT performance. We will also work toward being able to predict performance further than one bout in the future. Finally, the ultimate goal of this work is to provide efficient real-time parameter estimation on an individual basis allowing us to predict future performance.

Acknowledgments

The opinions expressed herein are solely those of the authors and do not necessarily represent the opinions of the United States Government, the U.S. Department of Defense, the U.S. Air Force, or any of their subsidiaries, or employees. This research was supported by funding from the 711th Human Performance Wing Studies and Analysis Intramural Proposal (Defense Health Program) program through the Aerospace Medicine Studies and Analysis Council. Distribution A. Approved for public release. Case number AFRL-2021-1848.

References

- Arsintescu, L., Mulligan, J. B., & Flynn-Evans, E. E. (2017). Evaluation of a psychomotor vigilance task for touch screen devices. *Human Factors*, 59(4), 661-670.
- Balkin, T., Thorne, D., Sing, H., Thomas, M., Redmond, D., Wesensten, N., Williams, J., Hall, S., & Belenky, G. (2000). Effects of sleep schedules on commercial motor vehicle driver performance (Report No. MC-00-133). Springfield, VA: National Technical Information Service, U.S. Department of Transportation.
- Basner, M., & Dinges, D. F. (2011). Maximizing sensitivity of the Psychomotor Vigilance Test (PVT) to sleep loss. *Sleep*, 34(5), 581-591.
- Basner, M., Mollicone, D., & Dinges, D. F. (2011). Validity and sensitivity of a brief psychomotor vigilance test (PVT-B) to total and partial sleep deprivation. *Acta Astronautica*, 69, 949-959.
- Caldwell, J. A., & Caldwell, J. L. (2016). Fatigue in aviation: A guide to staying awake at the stick. New York, NY: Routledge.
- Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods Instruments and Computers*, 17, 652-655.
- Gander, P., Hartley, L., Powell, D., Carbon, P., Hitchcock, E., Mills, A., & Popkin, S. (2011). Fatigue risk management: Organizational factors at the regulatory and industry/company level. *Accident Analysis and Prevention*, 43, 573-590.
- Gertler, J., DiFiore, A., & Raslear, T. (2013). Fatigue status of the U.S. railroad industry. U.S. Department of Transportation Federal Railroad Administration (report num: DOT/FRA/ORD-13/06). Office of Research and Development Washington, DC.
- Grant, D. A., Honn, K. A., Layton, M. E., Riedy, S. M., & Van Dongen, H. P. A. (2017). 3-minute smartphone-based and tablet-based psychomotor vigilance tests for the assessment of reduced alertness due to sleep deprivation. *Behavior Research Methods*, 49(3), 1020-1029.
- Fogel, D. (2006). *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence* (3rd ed.). Piscataway, NJ: IEEE Press.
- Lamond, N., Dawson, D., & Roach, G. D. (2005). Fatigue assessment in the field: Validation of a hand-held electronic psychomotor vigilance task. *Aviation, Space, and Environmental Medicine*, 76(5), 486-489.
- Hursh, S. R., Balkin, T. J., Miller, J. C., & Eddy, D. R. (2004). The fatigue avoidance scheduling tool: Modeling to minimize the effects of fatigue on cognitive performance. *SAE Transactions*, 113(1), 111-119.

- Hursh, S. R., Redmond, D. P., Johnson, M. L., Thorne, D. R., Belenky, G., Balkin, T. J., Eddy, D. R. (2004). Fatigue models for applied research in warfighting. *Aviation, Space, and Environmental Medicine*, 75(3), A44-A53.
- Kancherla, B., Upender, R., Collen, J. F., Rishi, M. A., Sullivan, S. S., Ahmed, O., Peters, B. R. (2020). Sleep fatigue and burnout among physicians: An American Academy of Sleep Medicine position statement. *Journal of Clinical Sleep Medicine*, 16(5), 803-805.
- Liu, J., Ramakrishnan, S., Laxminarayan, S., Balkin, T. J., & Reifman, J. (2017). Real-time individualization of the unified model of performance. *Journal of Sleep Research*, 26, 820-831.
- Mallis, M. M., Mejdal, S., Nguyen, T. T., & Dinges, D. F. (2004). Summary of the key features of seven biomathematical models of human fatigue and performance. *Aviation, Space, and Environmental Medicine*, 75(3), A4-A14.
- National Academies of Sciences, Engineering, and Medicine. (2016). *Commercial motor vehicle driver fatigue, long-term health, and highway safety: Research needs*. Washington, DC: National Academies Press.
- Rajdev, P., Thorsley, D., Rajaraman, S., Rupp, T. L., Wesensten, N. J., Balkin, T. J., & Reifman, J. (2013). A unified mathematical model to quantify performance impairment for both chronic sleep restriction and total sleep deprivation. *Journal of Theoretical Biology*, 331, 66-77.
- Ratcliff, R., & Van Dongen, H. P. A. (2018). The effects of sleep deprivation on item and associative recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(2), 193-208.
- Scrucca, L. (2013). GA: A package for genetic algorithms in R. *Journal of Statistical Software*, 53(4), 1-37.
- Scrucca, L. (2017). On some extensions to GA package: Hybrid optimisation, parallelisation and islands evolution. *The R Journal*, 9/1, 187-206.
- Tucker A.M., Whitney P., Belenky G., Hinson J.M., Van Dongen H.P. (2010). Effects of sleep deprivation on dissociated components of executive functioning. *Sleep*, 33(1), 47-57.
- Van Dongen, H. P. A. (2004). Comparison of mathematical model predictions to experimental data of fatigue and performance. *Aviation, Space, and Environmental Medicine*, 75(3), A15-A36.
- Whitney, P., Hinson, J. M., Jackson, M. L., & Van Dongen, H. P. A. (2015). Feedback blunting: Total sleep deprivation impairs decision making that requires updating based on feedback. *Sleep*, 38, 745-754.

Covering Strategy Changes: From System 1 to System 2 in Syllogistic Reasoning

Evelyn Wiens (evelynwiens1@gmail.com)

Cognitive Computation Lab, University of Freiburg, Germany

Alice Ping Ping Tse (alice@hum.ku.dk)

Department of Nordic Studies and Linguistics, University of Copenhagen, Denmark

Marco Ragni (ragni@sdu.dk)

Danish Institute for Advanced Studies, South Denmark University, Odense, Denmark

Cognitive Computation Lab, Technical Faculty, University of Freiburg, Germany

Abstract

Most cognitive models for human syllogistic reasoning aim to explain an *average* reasoner, i.e., the responses given by aggregating the response of the majority of reasoners. Studies show that individuals can deviate a lot from this *average* reasoner. So far, there have been very few models to explain and predict the responses of individual reasoner. In empirical studies, it can be observed that participants often rely on heuristic strategies (System 1 processes) to solve syllogistic problems but participants switch to analytical strategies (System 2 processes) occasionally. The study by Tse et al. (2014) demonstrated that inhibition of the matching heuristic is necessary to switch to the analytical processes in conflict problems that the output from the heuristic does not agree with that from analytical processes. This paper presents four mechanisms to incorporate individual differences in reasoning strategies and effect induced by problem type of the syllogism in predictive computational models built according to the mental model theory, mReasoner, and verbal models theory. Among these models, the composite model, which takes the highest accuracy model for individual reasoner, can reach a median accuracy of 86% in predicting the conclusions given by individual reasoner in the study.

Keywords: predictive cognitive modeling; syllogistic reasoning; strategy changes; dual process theory; System 1 and 2

Introduction

Consider the following syllogistic reasoning example:

All snakes are reptiles. [Premise 1]

No rabbit is a snake. [Premise 2]

Therefore, no rabbit is a reptile. [Conclusion]

For the example above, 93% of the 107 participants responded that the conclusion follows logically, which is termed “the conclusion is valid” (Tse, Ríos, García-Madruga, & Bajo-Molina, 2014). This example denotes a traditional syllogistic deduction consisting of two premises, featuring one of the four categorical quantifiers each (*All*, *Some*, *No*, or *Some ... not*, which are usually abbreviated as A, I, E, and O, respectively). Together, the two premises provide information about three terms (*reptile*, *snake*, *rabbit*), two of which only occur in one of the premises – the so-called end-terms, or the major terms (*reptiles* and *rabbit*). The goal of syllogistic reasoning is to connect the information conveyed by the premises, i.e. the categorical relationship between the two end-terms and the middle term (i.e. the term which occur in both premises), to deduce the relationship between the two end-terms. The middle term (also known as the minor

term) does not appear in the conclusion. In the example, in addition to the premises, a conclusion candidate is presented below the horizontal line for the reasoner to verify. The term “mood” is used to describe the combination of the quantifiers in the premises and conclusion. The example above is of the mood AE-E for using the abbreviations above. A syllogism can be not only by using the four quantifiers for each premise and conclusion, the terms itself in the premises can be organised in four different ways, called figures:

	Figure 1	Figure 2	Figure 3	Figure 4
Premise 1	A-B	B-A	A-B	B-A
Premise 2	B-C	C-B	C-B	B-C

By replacing *reptiles* with A (or a), *snakes* with B (or b) and *rabbits* with C (or c), the example above is a syllogism of figure 2, and can be denoted by AE-E2 and the premises and conclusion can be denoted by *Aba*, *Ecb* and *Eca* respectively. There are 256 possible syllogisms (64 different mood times 4 different figures) but only 27 of them have at least one valid conclusion.

Like many other common daily reasoning processes, humans tend to employ some heuristics when they want to solve a syllogistic problem, unless under certain circumstances. As proposed by Evans and his colleagues in the dual processing theory (Wason & Evans, 1974; Evans, 2006, 2008, 2011; Evans & Over, 2013), humans use the unconscious, intuitive, cognitive-resources-undemanding and rapid System 1 processes by default. The use of heuristics is among these processes. The output from these processes can be prone to biases arise from common sense, beliefs and previous experience. A classic example in human syllogistic reasoning is the belief bias effect that humans tend to accept more (invalid) conclusions which agree with their own beliefs and prior knowledge than otherwise (Morley, Evans, & Handley, 2004). However, humans can switch to the System 2 processes which are conscious, analytical, cognitive-demanding and rule-based under specific conditions, such as when they are told to solve the problems carefully (Evans, 2007).

Another example to illustrate dual processing processes in reasoning is the use of the matching heuristic to solve syllogistic problems. Humans choose the conclusion quantifier which matches the quantifier of the premise with a *lower*

number of entities, i.e. the more "conservative" premise, favouring the order $E > O = I \gg A$. Therefore, the conclusion quantifier is the same as (matches) at least one of the premise' quantifiers. Therefore, for the AE example above, humans tend to accept or produce the E-conclusion (No A-C or No C-A conclusion). As mentioned, 93% of the participants accepted that the conclusion followed from the two premises but the conclusion is indeed invalid. That means, only 7% of the participants made the correct response – to reject the conclusion. The matching heuristic is syntactic in nature as it involves merely "matching" the conclusion quantifier with quantifiers of the two premises. Unlike the belief bias effect that the heuristic depends on the semantic of the conclusion, reasoners do not have to process the semantic information (i.e., reptiles, snakes, rabbits) in the premises and the conclusion when using the matching heuristic.

The matching heuristic is one out of 12 cognitive theories that aim to explain the aggregated response of participants in syllogistic reasoning and many of the 12 theories can cover a large number of responses given (Khemlani & Johnson-Laird, 2012). In this paper, we will extend a previous analysis conducted by Riesterer, Brand, and Ragni (2020) and Bischofberger and Ragni (2020) that focused on predicting individual reasoner, to test if incorporating individual reasoner's strategy change from system 1 to system 2 processes during reasoning can yield an adapted model of these theories with better predictive power.

The paper is structured as follows: In the next section we will introduce the data from the study by Tse et al. (2014) cognitive models on syllogistic reasoning. We will introduce how we adapted the models in Section 3 and report the results in Section 4, followed by a discussion (Section 5) which concludes the paper.

The Data

The data are from the study by Tse et al. (2014). 107 students from the University of Granada (mean age = 22.34 years, SD = 4.43; 89 females and 18 males) participated in the experiment. They were rewarded with course credits. The experiment was conducted in Spanish. The participants were all native speakers of Spanish and did not have any training in logic before.

Each participant had to judge the validity of 16 syllogisms, with each followed by a lexical decision task. The syllogisms were chosen to test the interplay between the use of matching heuristics (system 1 processes) and analytical strategy (system 2 processes). Therefore, conflict problems which are either match-invalid (the conclusion quantifier matches with the quantifier of the more conservative premise but it is logically invalid) or mismatch-valid (the conclusion does not agree with the matching heuristic but it is logically valid) and no-conflict problems which are either match-valid and mismatch-invalid were constructed. That means, participants can reach the same conclusion (accepting the conclusion) for no-conflict problems using both the matching heuristic or the

analytical strategy; but for the conflict problems, participants have to inhibit the use of the matching heuristic (System 1 non-analytical default approach) and switch to the analytical strategy (System 2). Due to the aforementioned constraint, the only possible options are AE, EA and AA problems which allow the construction of conflict and no-conflict problems. AA problems were not chosen as the two premises have the same quantifier and the matching heuristic is based on matching the conclusion qualifier with the premise quantifiers. The syllogisms with the E conclusion were used as the matched syllogisms (AE-E and EA-E) while syllogisms with the O conclusion were used as the mismatched syllogisms (AE-O and EA-O). In order to prevent participants from guessing the experimental manipulation and as only one of the AE-O2 and EA-O1 syllogisms are invalid (it is not possible to have two mismatch-invalid AE-O and EA-O syllogisms), the AE-A1 and EA-A2 syllogisms were included as fillers to replace a AE-O and EA-O syllogism respectively. There were eight conflict problems, six no-conflict problems and two fillers, see Table 1. Half of the syllogisms (i.e. eight) had a valid conclusion while the other half were invalid.

Table 1: Types of problem used in the experiment Tse et al. (2014).

<i>Problem Type</i>	<i>Conclusion Type</i>	<i>Syllogism</i>
8 Conflict Problems (multiple-model)	4 Match-invalid	2 AE-E2
		2 EA-E1
	4 Mismatch-valid	2 AE-O1
		2 EA-O2
6 No-conflict Problems (single-model)	4 Match-valid	2 AE-E1
		2 EA-E2
	2 Mismatch-invalid	AE-O2
		EA-O1
2 fillers	invalid	AE-A1
		EA-A2

In the lexical decision task (LDT) after each syllogistic problem, participants were asked to judge whether 24 letter strings were real words in Spanish or not one by one. Half of them (i.e. twelve) were non-words while the other half were words in Spanish, with six of them related to the two terms in the conclusion while the other six were unrelated to the terms in the syllogisms.

The Predictive Model Task & Individualization

We use the CCOBRA-framework¹ to ensure a modeling evaluation standard as proposed by Riesterer et al. (2020). The model has then to predict the conclusion which should be drawn by the individual participant, before the he/she responds. In a predictive analysis, cognitive models need to be able to adapt to the individual they need to predict. This is in most cases done by a parameter optimization process.

¹<https://github.com/CognitiveComputationLab/ccobra>

CCOBRA uses a leave-one-out cross validation method and each test run generates automatically both the test- and training data. The output (the predicted response) from the model is then compared with participant’s response.

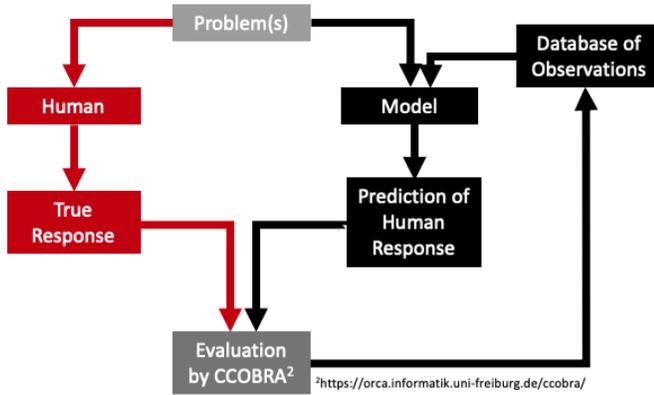


Figure 1: CCOBRA a system to evaluate the predictive accuracy of models.

To capture individual differences, the parameters of the implemented models have to be fitted to an individual. The best fit is determined by iterating over the prediction-response pairs and performing a grid search over the parameter space of the individual models. The parameter optimization for each individual is done before the actual prediction of the answers and therefore evaluate the overall ability of the model to account for individual data.

Cognitive Models

For an in-depth presentation of the existing cognitive models see (Khemlani & Johnson-Laird, 2012). We will present here briefly 3 core theories for System 2 (mental model theory, mReasoner, verbal models theory) that our models are based on.

Mental Model Theory

The mental model theory (MMT) (Johnson-Laird & Philip, 1983) postulates that people draw inferences with the help of mental models. A mental model consists of abstract tokens that reflect the situation asserted by the premises. For example, a (or an initial) mental model of the syllogism *All A are B. No B are C* can be:

- a b
- a b
- ¬ b c
- ¬ b c

A conclusion is either drawn based on the initial mental model or refuted with a search of alternative models. This implementation is based on a formalization of the classical theory of mental models (Sugimoto, Sato, & Nakayama, 2013) with some parameter adjustments (Bischofberger & Ragni,

2020). A parameter determines how likely certain individuals search for alternative models (i.e. search for counterexamples). Thus, the implementation can distinguish between people who consider alternative models and those who do not.

mReasoner

mReasoner (Khemlani & Johnson-Laird, 2013) is a more powerful model that follows the assumptions of the mental model theory. The construction of the initial mental model is individualized by two parameters. The size of the initial model is controlled through the parameter λ and the completeness of the encoded information is determined with the parameter ϵ . Conclusions are generated with heuristics and validated by the mental model. The parameter σ specifies the probability that individuals seek counterexamples after formulating a tentative conclusion. Furthermore, a parameter ω determines if falsified conclusions are weakened. Weakened conclusions are also validated with a search for counterexamples.

Verbal Models Theory

The verbal models theory (Polk & Newell, 1995) assumes that syllogistic reasoning is fundamentally verbal. The model implements multiple parameters that allow high adaptability to an individual. The relations of the premises are encoded into a mental model. Identifying tokens mark more easily accessible information that are derived from the subject of the premises. A conclusion is formulated from the marked tokens. If the program cannot formulate a conclusion, the mental model is reencoded. For this purpose, additional information is extracted from non-identifying tokens. Depending on the type of premise, internal parametrization, and reference token, a new premise is formulated to extend the mental model. The process is repeated until a conclusion can be formulated or reencoding fails. In the latter case, the program returns *no valid conclusion* (NVC).

Making Cognitive Models Adaptive

The cognitive models we just presented need to be made adaptive to predict individual reasoner in the aforementioned CCOBRA-framework. In the following, we describe four mechanisms on how the three models were made adaptive to the response of individual participant.

Probability for searching for counterexamples is adjusted individually

The first set of models are adjusted to model individual reasoning behaviour. To achieve this, parameter settings of the three theories (MMT, mReasoner, Verbal Models) were fitted to individual responses using CCOBRA. The flow structure is shown in Figure 2. In this adaption process, the possible effect of *matching bias* on reasoning, effect of problem type (conflict vs. no-conflict) and individual reasoning strategy (System 1 vs. System 2 processes) were not implemented. Therefore, these models adjust their internal parameterization

according to the response(s) of individual participant. The parameters are updated according to whether a participant just accept the initial conclusion or try to search for counterexamples. The probability for searching for counterexamples (i.e. the System 2 processing) of a participant would be higher if that particular participant gave more System 2 responses, i.e. a higher likelihood of System 2 responses.

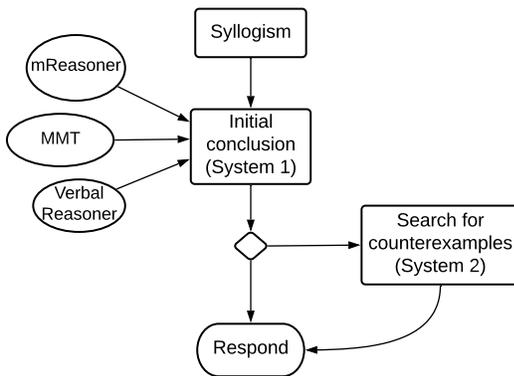


Figure 2: System 2 through individually adjusted probabilities

Probability for searching for counterexamples according to the selective processing model

The critical difference between conflict and no-conflict problems in the (Tse et al., 2014) study is that for no-conflict problems, participants can get the correct responses by either System 1 or System 2 processes (as the conclusions are congruent with the output from the matching heuristic); while for the conflict problems, participants must switch to System 2 processes in order to get the correct responses (as the conclusions are incongruent with the output from the matching heuristic). To resolve conflicts between the outputs from System 1 and System 2 processes in conflict-problems, the selective processing model (Evans, 2000) is generalized so that a conclusion is accepted as soon as there is at least one mental model (initial or alternative) that supports it. Originally selective processing is used in belief bias study to for the conflict between the belief bias and validity of the conclusion; while we have adapted it for the conflict resolution between the matching bias and validity of the conclusion. For match-invalid and match-valid syllogisms alternative models are not searched because the initial model (System 1) already support the given conclusion. In mismatch-problems alternative models (considered as the analytical System 2 processes) are searched. These set of models do not take into account individual differences, it focused on the property of the syllogisms.

Probability for searching for counterexamples is adjusted according to individual strategy and the selective processing model

These models take into account the response time of the participants to the syllogistic tasks and the results of the subsequent lexical decision task when predicting whether the participant would search for alternative models for MMT and mReasoner or reencode the mental models for verbal models. After an initial mental model is constructed by the respective cognitive model, the tentative conclusion is returned or refuted depending on the individual. If the response time of a particular trial is above a specified threshold (9000 ms in the implementation) and the semantic priming effect is diminished (the difference between the response times of the unrelated words and related words is smaller than 15 ms in the lexical decision task) the model tends to refute the initial conclusion (System 2 processes). For switching to System 2 processes, a longer response time is expected as System 2 processes are cognitive resources demanding and it also takes time to inhibit the output from System 1 processes. The missing of semantic priming effect indicates the inhibition of the heuristic (bias) processes (Tse et al., 2014; De Neys & Frassens, 2009). The threshold values were chosen based on evaluation results after a few tests.

The parameters of the respective theories are adjusted according to the behavioral information (response time and priming effect) to implement switches between System 1 and System 2. Additionally, the parameter settings are individualized (see "Probability for searching for counterexamples is adjusted individually" section) to ensure high predictive accuracy outside the threshold conditions.

The search for alternative models in the MMT and mReasoner models incorporated also the aforementioned selective processing model method. An overview over the reasoning process is shown in Figure 3.

Additive probability model

Another model that can describe the conflict between heuristic and analytic processes is the additive probability model (Evans, 2007). In this model, the underlying cognitive process is separated by an heuristic process and an analytical one. This implementation uses individual strategy (as mentioned in the previous paragraph) to determine from which process the predictions are computed. If the response time of the syllogistic task was low or the semantic priming effect occurred (in the LDT), a heuristic process computes the answer. In this case, a parameter determines if an answer consistent with the matching hypothesis is returned or the conclusion is blindly accepted. The analytical process is modeled using the mReasoner. Depending on the internal parameters of an individual, a conclusion can be either accepted or rejected. The selective processing model is not implemented here. For instance, selective processing always predicts acceptance of the conclusion in match-invalid syllogisms. That is because System 1 processes support the given conclusion. Without

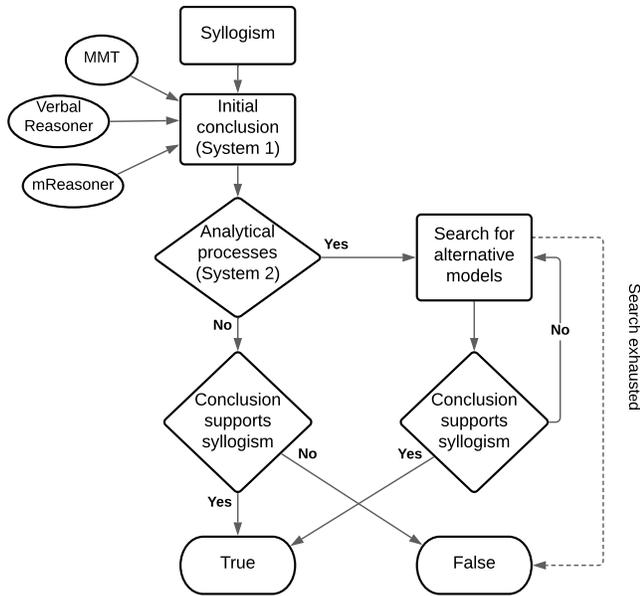


Figure 3: Alternative model search with individual strategy and selective processing.

selective processing there is an individual chance for System 2 processes for those syllogisms in this mechanism.

Composite Model

The previous models each describe individual strategies and theories of how people resolve conflicts between analytic and heuristic responses. While the previous models and extensions already implement individual differences, some models represent an individual’s reasoning process better than others. To take advantage of the strengths of each model, a composite model was implemented. The composite model predicts the participants’ responses based on the model that could achieve the highest accuracy for that person.

Evaluation and Discussion

Adaptive processes described in the last section were implemented in the three cognitive models built according to the mental model theory, mReasoner, and verbal models theory and were then evaluated in CCOBRA. To compare the overall performance, state-of-the-art models, including PSYCOP, Matching, Conversion, and PHM; and two benchmark models, Uniform and Most Frequent Answer (MFA), were added. The Uniform Model assumes a uniform distribution over the set of responses. Any cognitive model should recognize basic patterns in the data set and outperform this model. Due to the noise in the data, it is unrealistic to assume that the models can perfectly predict participants’ responses. Therefore, the MFA model serves as an empirical upper bound for models that do not implement inter-individual differences. The MFA returns the most frequent response given by the participants for each syllogism. A higher predictive power than MFA indicates that the specific models are able to capture different

strategies of individuals (Riesterer et al., 2020).

Figure 4 shows the predictive power of the models. All models, except PSYCOP, make more accurate predictions than the uniform model and are therefore able to capture some properties of the data set. The low performance of PSYCOP is due to the high number of possible predictions for the syllogisms and the lack of adaptation to individuals in the implementation. The heuristic models PHM and Conversion achieve prediction accuracies of about 55%. The mental model theory and verbal models theory can predict 62% and 66% of the responses, respectively. Without individual strategies, the search for alternative models (counterexamples) is randomly determined and this leads to comparatively poor results. The matching heuristic, as well as atmosphere, achieve a prediction accuracy of 79%. Since many participants rarely considered alternative models, the accuracy of these heuristics is comparatively high. Moreover, these heuristics can perfectly replicate the answers for some participants as the experimental materials were designed to test the matching heuristic. Matching hypothesis is a modified version of the atmosphere hypothesis and thus the results of these two heuristics were expected to be similar.

The adaptive models all achieve higher accuracies than their static implementations. While the adaptive implementation of the verbal models is able to predict 68% of the data correctly, the adaptive version of mental model theory can achieve 77% accuracy. However, both of them cannot outperform the matching heuristic model. The adaptive implementation of mReasoner (80%) is able to outperform that of atmosphere and matching heuristics, and almost achieves MFA benchmark accuracy (81%).

The selective processing model achieves the same accuracy as the MFA benchmark. Although individual differences was not implemented in this model, it is able to predict most responses in the dataset, as well as the most frequent answers for a syllogism. The verbal models theory with individual strategy implemented (58%) performs worse than the static implementation. Although the re-encoding process of this theory relies on semantic processes, the participants’ response times, as well as the results of the lexical decision task, cannot improve the predictive accuracy. The adaptive implementation of mReasoner and MMT with selective processing can outperform the results of the MFA. mReasoner with adaptive parameters and selective processing achieves 84% accuracy and MMT 81%. Thus, parameters of these models be adapted with individual strategies of the participants to obtain a better predictive power. They can capture individual responses and do not merely predict the majority response (average reasoner).

The additive probability model (80%) has about the same predictive power as the MFA model. Although the overall performance is worse compared to the other models, the performance of fewer participants can be improved. The composite model achieves a median accuracy of about 86%. The verbal models performs worse than MMT and mReasoner, in-

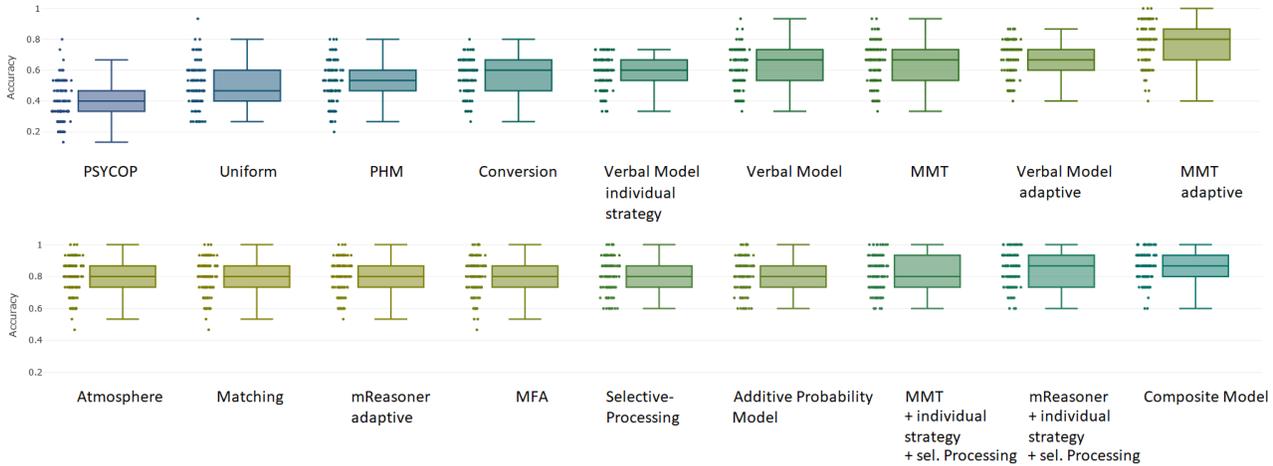


Figure 4: The predictive accuracies of the adjusted models and state-of-the-art models as given in Khemlani and Johnson-Laird (2012)

cluding the adaptive version. The mReasoner obtains the best performance in general. Comparing the four mechanisms, the cognitive models with individual strategy and selective processing implemented obtain the best results. Also, the models with individual strategy (based on the LDT and response times of the syllogisms) perform almost the same as those with the selective processing model.

In summary, the results demonstrate that incorporating individual strategies and effect of problem types can improve the performance of cognitive models to predict the responses of individual reasoners. This shows that the integration can improve state-of-the-art models to predict the responses of individual reasoner substantially and surpasses the MFA benchmark. Finding more optimized ways to integrate individual properties of a participant such as using results from the LDT and the response times to estimate individual strategy in reasoning, and to identify more individual differences can push the models to predict individual reasoners even further.

Acknowledgements

Support to MR by the Danish Institute of Advanced Studies and the DFG (RA 1934/4-1, RA1934/9-1) is gratefully acknowledged.

References

Bischofberger, J., & Ragni, M. (2020). Improving cognitive models for syllogistic reasoning. In *Proceedings of the 42th annual conference of the cognitive science society*.

De Neys, W., & Franssens, S. (2009). Belief inhibition during thinking: Not always winning but at least taking part. *Cognition*, *113*(1), 45–61.

Evans, J. S. B. T. (2000). Thinking and believing. *Mental models in reasoning*.

Evans, J. S. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review*, *13*(3), 378–395.

Evans, J. S. B. T. (2007). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning*, *13*(4), 321–339. doi: 10.1080/13546780601008825

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol.*, *59*, 255–278.

Evans, J. S. B. T. (2011). Dual-process theories of reasoning: Contemporary issues and developmental applications. *Developmental Review*, *31*(2-3), 86–102.

Evans, J. S. B. T., & Over, D. E. (2013). *Rationality and reasoning*. Psychology Press.

Johnson-Laird, P. N., & Philip, N. (1983). Mental models: Towards a cognitive science of language, inference, and consciousness. *Harvard University Press*(6).

Khemlani, S., & Johnson-Laird, P. N. (2012). Theories of the syllogism: A meta-analysis. *Psychological bulletin*, *138*(3), 427.

Khemlani, S., & Johnson-Laird, P. N. (2013). The processes of inference. *Argument & Computation*, *4*(1), 4–20. doi: 10.1080/19462166.2012.674060

Morley, N. J., Evans, J. S. B. T., & Handley, S. J. (2004). Belief bias and figural bias in syllogistic reasoning. *The Quarterly Journal of Experimental Psychology Section A*, *57*(4), 666–692.

Polk, T. A., & Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review*, *102*(3), 533–566. doi: 10.1037/0033-295x.102.3.533

Riesterer, N., Brand, D., & Ragni, M. (2020). Predictive modeling of individual human cognition: Upper bounds and a new perspective on performance. *Topics in Cognitive Science*, *12*(3), 960–974. doi: 10.1111/tops.12501

Sugimoto, Y., Sato, Y., & Nakayama, S. (2013). Towards

a formalization of mental model reasoning for syllogistic fragments. *Proceedings of the 1st International Workshop on Artificial Intelligence and Cognition (AIC 2013)*, 1100, 140-145.

- Tse, P. P., Ríos, S. M., García-Madruga, J. A., & Bajo-Molina, M. T. (2014). Inhibitory mechanism of the matching heuristic in syllogistic reasoning. *Acta Psychologica*, 153, 95–106. doi: 10.1016/j.actpsy.2014.08.001
- Wason, P. C., & Evans, J. S. B. T. (1974). Dual processes in reasoning? *Cognition*, 3(2), 141–154.

Is Similarity-based Interference Caused by Lossy Compression or Cue-based Retrieval? A Computational Evaluation

Himanshu Yadav (hyadav@uni-potsdam.de)

Department of Linguistics, University of Potsdam
14476 Potsdam, Germany

Garrett Smith (gasmith@uni-potsdam.de)

Department of Linguistics, University of Potsdam
14476 Potsdam, Germany

Shravan Vasishth (vasishth@uni-potsdam.de)

Department of Linguistics, University of Potsdam
14476 Potsdam, Germany

Abstract

The similarity-based interference paradigm has been widely used to investigate the factors subserving subject-verb agreement processing. A consistent finding is facilitatory interference effects in ungrammatical sentences but inconclusive results in grammatical sentences. Existing models propose that interference is caused either by misrepresentation of the input (representation distortion-based models) or by mis-retrieval of the interfering noun phrase based on cues at the verb (retrieval-based models). These models fail to fully capture the observed interference patterns in the experimental data. We implement two new models under the assumption that a comprehender utilizes a lossy memory representation of the intended message when processing subject-verb agreement dependencies. Our models outperform the existing cue-based retrieval model in capturing the observed patterns in the data for both grammatical and ungrammatical sentences. Lossy compression models under different constraints can be useful in understanding the role of representation distortion in sentence comprehension.

Keywords: Similarity-based interference; lossy memory representation; cue-based retrieval

Introduction

Similarity-based interference in subject-verb agreement dependencies has played an important role in understanding the mechanisms underlying sentence comprehension (Wagers, Lau, & Phillips, 2009; Lago, Shalom, Sigman, Lau, & Phillips, 2015). In this paradigm, a noun phrase matching in agreement features with the verb—called a distractor—is presented along with the subject noun. For example, in the following sentences (a) and (c), the distractor noun phrase *the cabinet(s)* matches the number feature of the verb in contrast to conditions (b) and (d), where it does not.

(a) Grammatical, interference condition

The key to the cabinet unsurprisingly was rusty.

(b) Grammatical, no-interference condition

The key to the cabinets unsurprisingly was rusty.

(c) Ungrammatical, interference condition

* The key to the cabinets unsurprisingly were rusty.

(d) Ungrammatical, no-interference condition

* The key to the cabinet unsurprisingly were rusty.

A consistent finding is that of facilitation in ungrammatical conditions: reading times at the verb ‘were’ in condition (c) are, on average, faster than in condition (d) (Jäger, Engelmann, & Vasishth, 2017; Wagers et al., 2009; Lago et al., 2015; Dillon, Mishler, Sloggett, & Phillips, 2013; Jäger, Mertzen, Van Dyke, & Vasishth, 2020). By contrast, the results are inconclusive in grammatical conditions: reading times at the verb in condition (a) can be faster, slower, or comparable to condition (b). Figure 1 shows the observed interference effects in the grammatical and ungrammatical conditions from 11 published datasets.

Several models have been proposed to explain the facilitatory interference effect in the ungrammatical conditions, but these models cannot explain the range of effects in the grammatical conditions. Most of these models can be placed into one of two categories, cue-based retrieval accounts, and representation distortion-based accounts.

The cue-based retrieval account (Lewis & Vasishth, 2005) assumes that dependency completion between the subject and the verb is driven by a cue-based retrieval process: encountering a verb triggers a content-addressable search in memory using feature specifications such as [+subject] or [+plural], called retrieval cues. The cue-based retrieval model correctly predicts the facilitatory effect in ungrammatical conditions. But the model predicts an inhibitory effect in grammatical conditions: a slowdown in condition (a) compared to (b). This prediction is not supported by the interference effect data in the grammatical conditions shown in Fig. 1.

Representation distortion-based accounts assume that the representation of the pre-verbal sentence material—subject noun and/or distractor noun—undergoes distortion with time. One of the representation distortion-based accounts—the encoding-based model (Bock & Eberhard, 1993; Eberhard, 1997)—maintains that the plural feature of the distractor noun percolates up to the subject noun phrase causing a *misrepresentation* of the subject in a proportion of trials. The encoding-based model predicts facilitatory effect in both grammatical and ungrammatical conditions, which is not supported by the observed pattern of effects (see Fig. 1).

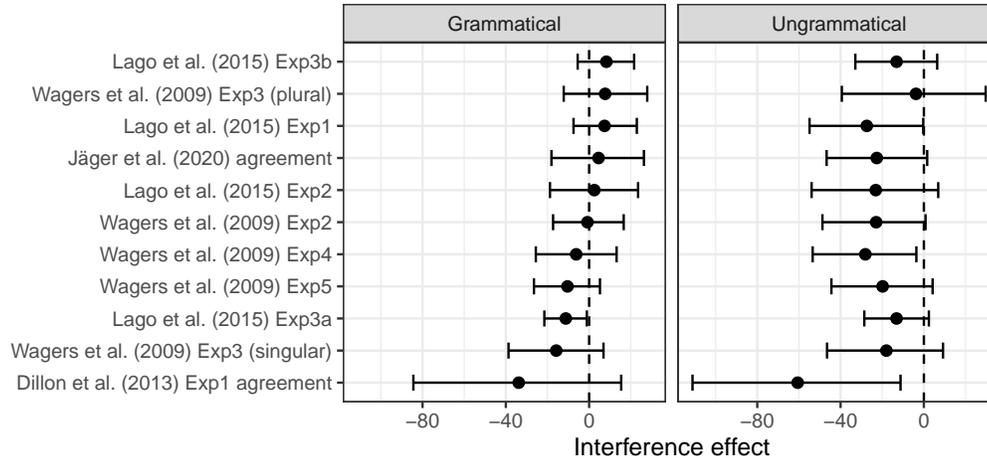


Figure 1: The pattern of interference effects in subject-verb agreement dependencies. Here, “interference effect” means the difference in reading times at the verb between the interference and no-interference conditions.

Another class of representation distortion-based models are based on lossy compression of the linguistic input (Futrell, Gibson, & Levy, 2020). These models assume that a comprehender obtains a distorted representation of the true intended message due to lossy memory encoding, and they reconstruct a set of possible true representations using their prior linguistic knowledge. A well-tested model of this type is the lossy-context surprisal model of Futrell et al. (2020). The model captures working memory effects within an expectation-based framework. It assumes that after reading or hearing a series of words, the words can be corrupted by deleting content words (e.g., nouns or verbs) at a constant rate. Processing difficulty at a new word is the expected surprisal of the word given this lossy memory representation of its preceding context. Futrell et al. (2020) show that the model explains structural forgetting effects (Vasishth, Suckow, Lewis, & Kern, 2010), and Futrell (2019) explain information locality across languages using the lossy compression model.

An important limitation of the literature on similarity-based interference effects is that researchers either invoke cue-based retrieval or some kind of lossy compression model to explain the data. Moreover, the two classes of model have never been pitted against each other in any systematic quantitative evaluation, even though a considerable amount of benchmark data are available on interference effects. A further intriguing possibility, which needs to be quantitatively evaluated, is that both lossy compression and cue-based retrieval could play a role in a hybrid model.

We address these open issues by implementing two lossy compression models of similarity-based interference to try to capture the observed effects in both grammatical and ungrammatical conditions in subject-verb agreement dependencies. We compare the performance of our models against the cue-based retrieval model of Lewis and Vasishth (2005); Vasishth, Nicenboim, Engelmann, and Burchert (2019) on interference

effect data from the 11 publicly available datasets shown in figure 1.

A lossy compression model of interference effects

We implement a lossy-context surprisal model as described in Futrell et al. (2020) with some additional assumptions to model interference effects in subject-verb agreement dependencies.

Assumptions

Consider the sentence “The key to the cabinets unsurprisingly was rusty”. The observed pre-verbal noun phrase in this sentence is *the key to the cabinets*. We call this input I . The lossy compression model assumes the following:

1. The linguistic input received by the comprehender has undergone lossy compression: there was some true representation r ; due to lossy memory encoding, the true representation r distorts to the observed input I such that the plural marker on the nouns can either be deleted or inserted or left unchanged at constant rates
2. The comprehender reconstructs a set of possible true representations from input I conditioned on their prior linguistic knowledge and the rates of deletion/insertion in the system
3. The processing difficulty at the verb is the expected (average) surprisal of encountering the verb given all possible true representations of the input I

Next, we derive the processing difficulty and reading times at the verb in subject-verb agreement dependencies.

Calculating processing difficulty and reading times at the verb

In the sentence “The key to the cabinets unsurprisingly was rusty”, the input is

$$I = N \ P \ N.pl$$

where N represents a noun, P represents a preposition, and $.pl$ represents a plural marker on a noun.

The possible true representations, r_i , that can lead to input I due to lossy compression are,

$$\begin{aligned} r_1 &= N.pl \ P \ N.pl & r_2 &= N.pl \ P \ N \\ r_3 &= N \ P \ N.pl & r_4 &= N \ P \ N \end{aligned}$$

The processing difficulty for the upcoming verb will be proportional to the *expected surprisal* of the verb given all possible true representations r_1, r_2, \dots, r_N :

$$D(V|I) \propto \sum_{i=1}^N -\log P(V|r_i) \cdot P(r_i|I) \quad (1)$$

where $-\log P(V|r_i)$ is the surprisal — negative log conditional probability — of seeing a plural/singular verb after the context r_i ; we compute conditional probabilities from the COW corpora (Schäfer, 2015; Schäfer & Bildhauer, 2012). And, $P(r_i|I)$ is the probability density of reconstructing a representation r_i from the given input, I . We can derive $P(r_i|I)$ using Bayes’ rule,

$$P(r_i|I) \propto P(I|r_i)P(r_i) \quad (2)$$

where $P(r_i)$ is the prior probability density of a possible true representation r_i and can be estimated from corpus data. $P(I|r_i)$ represents the lossy memory encoding function: the likelihood that a representation r_i distorts to I given a constant deletion rate d and constant insertion rate a :

$$I|r_i \sim Memory(r_i, d, a) \quad (3)$$

where d is the rate of deleting a plural marker and a is the rate of inserting a plural marker. Table 1 shows the likelihood of obtaining I from each possible representation r_i . Finally, we transform processing difficulty into reading times using a linear linking function. Reading times in j^{th} trial, RT_j , will be:

$$RT_j = S \cdot D(V|I) + \epsilon_j \quad (4)$$

where S is a scaling parameter and ϵ_j is the random noise in the j^{th} trial such that $\epsilon_j \sim Normal(0, 20)$. The model has thus 3 free parameters: deletion rate d , insertion rate a and scaling parameter S .

Possible true representation	Likelihood of generating I from r_i
r_i	$P(I r_i)$
$N.pl \ P \ N.pl$	$d(1-d)$
$N.pl \ P \ N$	da
$N \ P \ N.pl$	$(1-a)(1-d)$
$N \ P \ N$	$(1-a)a$

Table 1: The lossy memory encoding function: the likelihood of obtaining the observed input I ($N \ P \ N.pl$) from lossy compression of a possible true representation r_i

Prior predictions

We use the model equations stated in the previous section and generate prior predictions from the model. This allows us to determine the range of effects the model can generate and compare them against the observed interference effect data. The joint distribution of interference effects in grammatical and ungrammatical conditions — $\{E_{gram}, E_{ungram}\}$ — is assumed to come from the lossy compression model conditional on its free parameters, the deletion rate d , the insertion rate a , and the scaling parameter S

$$\{E_{gram}, E_{ungram}\} \sim Model(d, a, S) \quad (5)$$

We specified the priors as follows. For deletion rate d and insertion rate a , we choose a weakly informative prior because we do not want to make any strong assumptions about these parameters:

$$d \sim Normal_{lb=0, ub=1}(0, 0.25) \quad (6)$$

$$a \sim Normal_{lb=0, ub=1}(0, 0.25) \quad (7)$$

where $lb = 0$ and $ub = 1$ indicate a lower bound of 0 and upper bound of 1 respectively. For the scaling parameter S , we choose a Gaussian prior centered at 25 and with standard deviation of 5; this range was chosen so that the model does not generate unreasonably large or small reading times (see Jäger et al., 2017, for meta-analysis estimates of reading times):

$$S \sim Normal_{lb=0}(25, 5)$$

Figure 2 shows the prediction space of the lossy compression model against the observed interference effect data. The model is able to predict a facilitatory effect in ungrammatical conditions and positive, zero or negative effects in grammatical conditions. Thus, the prior predictions of the model are consistent with qualitative pattern of the observed interference effects, but the magnitudes of predicted effects do not often align with the human data.

The lossy compression model, presented here, assumes that the link between the lossy memory representations and reading time effects is the average surprisal of the upcoming word. However, one is free to choose a different linking function. In the next section, we introduce a hybrid model that integrates the lossy compression and cue-based retrieval mechanisms in order to predict reading times at the verb.

Lossy-compression-plus-retrieval model

Recent work has shown that a model unifying representation distortion- and retrieval-based mechanisms shows a better fit to interference effect data from subject-verb agreement dependencies (Yadav, Smith, & Vasishth, 2021). Given this modeling evidence, it is interesting to explore a model combining lossy compression and cue-based retrieval in a single set of processes. We implement the lossy compression-plus-retrieval model with the idea that the cue-based retrieval at the verb is preceded by lossy memory representation of the intended message. Here, reading times are determined by the cue-based retrieval mechanisms, but the retrieval process operates on a noisy version of the intended input.

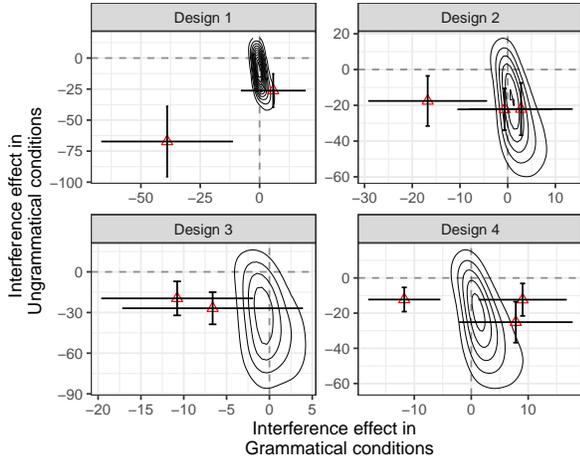


Figure 2: The prior predictive interference effect (in milliseconds) generated by the lossy compression model is shown as a contour of the joint distribution of effects in the grammatical and ungrammatical conditions. The red triangular points and errors bars around them represent the observed interference effects and their 95% credible intervals obtained from published datasets. The predictions differ across experimental designs because prior density of possible true representations $p(r_i)$ is estimated to be different for each design. Design 1: English subject relative clause constructions; Dillon et al. (2013), Exp 1 and Jäger et al. (2020). Design 2: English object relative clause constructions; Wagers et al. (2009) Exp 2, Exp 3, and Lago et al. (2015) Exp 2. Design 3: English prepositional phrase constructions; Wagers et al. (2009) Exp 4, Exp 5. Design 4: Spanish relative clause constructions; Lago et al. (2015) Exp1, Exp 3a, and Exp 3b.

Assumptions

The lossy compression-plus-retrieval model assumes that

1. Dependency completion between the subject and the verb is driven by a cue-based retrieval process
2. Cue-based retrieval is affected by changes in representation of the subject and the distractor nouns due to lossy compression of the intended message (as described in the previous section)

Next, we derive the updated retrieval time equation to account for representation change due to lossy compression.

Calculating retrieval times

The retrieval time at the verb in the j^{th} trial, RT_j , is an exponential function of the activation of the retrieved chunk,

$$RT_j = Fe^{-A_{j,\text{retrieved}}} \quad (8)$$

where F is a scaling parameter called the latency factor which reflects overall processing speed.

Under cue-based retrieval, the chunk with the highest activation gets retrieved in each trial. The activation of the chunk

retrieved in j^{th} trial would be the maximum of the activation of the subject and the distractor noun:

$$A_{j,\text{retrieved}} = \max\{A_{j,\text{subject}}, A_{j,\text{distractor}}\} \quad (9)$$

The activation of the subject and the distractor in a trial is determined by the amount of activation they receive via cue-feature match. The noun phrase that matches more cues receives a higher activation. Thus, activation of the subject and the distractor in j^{th} trial is a function of their representation,

$$\{A_{j,\text{subject}}, A_{j,\text{distractor}}\} \sim \text{Activation}(r_j) \quad (10)$$

where r_j is the representation of subject and distractor noun in the j^{th} trial. The lossy compression-plus-retrieval model assumes that the representation in the j^{th} trial is sampled from probability density of reconstructing r from input I ,

$$r_j \sim P(r|I, a, d) \quad (11)$$

where a and d are the insertion and deletion rates, respectively. The probability density function $P(r|I, a, d)$ can be derived in the same way as in equation 2. Using these equations, the lossy compression-plus-retrieval model allows us to make reading time predictions at the verb, which we now compare to reading time data from 11 experiments.

Prior predictions

We generate prior predictions from the lossy compression-plus-retrieval model conditional on its three free parameters, the deletion rate d , the insertion rate a , and the latency factor F . For the deletion rate and the insertion rate, we specify the same priors as in equation 6 and 7. For the latency factor, we used a truncated normal distribution:

$$F \sim \text{Normal}_{lb=0.1}(0.15, 0.03)$$

where $lb = 0.1$ indicates a lower bound of 0.1 on latency factor values. We choose this lower bound because a latency factor of less than 0.1 generates unreasonably fast reading times.

Figure 3 shows the prediction space of the lossy compression-plus-retrieval model against observed interference effect data. The model predictions are consistent with the facilitatory effect in ungrammatical conditions, but inconsistent with the range of effects in grammatical conditions.

Model comparison

We compare the performance of the lossy compression models (which assume that representation undergoes distortion due to information loss) against the cue-based retrieval model (which assumes that representation is intact and a retrieval process drives processing) on 11 published datasets. We use stratified k-fold cross-validation for model comparison: (1) We split each dataset into 6 folds (subsets) such that each fold contained observations from all participants for all conditions, (2) we prepared 6 sets of training and test data by leaving out one fold as test data and taking other 5 as training

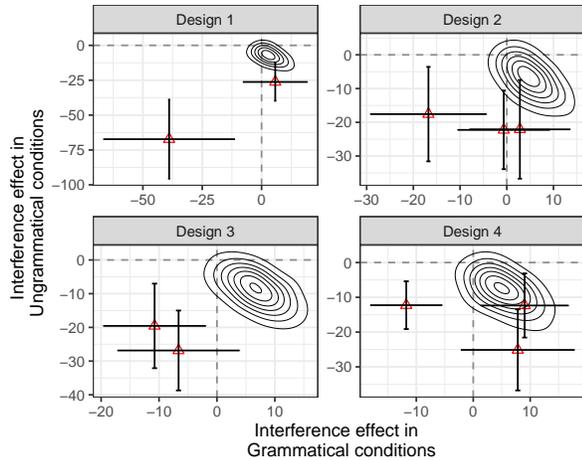


Figure 3: The prior predictive interference effect (in milliseconds) generated by the lossy compression-plus-retrieval model is shown as a contour of joint distribution of effects in the grammatical and ungrammatical conditions. The red triangular points and errors bars around them represent observed interference effects.

data, (3) in each iteration, we fit the models on training data using Approximate Bayesian Computation¹ (Sisson, Fan, & Beaumont, 2018) and computed the predictive accuracy of the fitted model on the test data in terms of log pointwise predictive density. Figure 4 shows the comparison of estimated log pointwise predictive density (\widehat{elpd}) of the models on 11 datasets. We find that:

1. The \widehat{elpd} values for the lossy compression model are larger than the cue-based retrieval model for 6 out of 11 datasets suggesting stronger evidence in the favor of lossy-compression model. The models are indistinguishable for the remaining five datasets.
2. The lossy-compression-plus-retrieval model shows higher predictive accuracy than the cue-based retrieval model for six out of 11 datasets.
3. The lossy-compression-plus-retrieval model and the lossy compression model show comparable performance.

Overall, the results suggest that a lossy compression model or a lossy compression-plus-retrieval model can explain the data better than the standard cue-based retrieval model.

Discussion

We have implemented two models—a lossy compression model and a lossy compression-plus-retrieval model—and investigated whether they can outperform the cue-based retrieval model. More specifically, we investigated whether,

¹Approximate Bayesian Computation (ABC) allows us to fit complex models when the likelihood of a model cannot be expressed mathematically. We use a particle filtering-based ABC algorithm to estimate posterior distributions of free parameters in the models.

compared to the cue-based retrieval model, these two models can furnish a better account for the pattern of interference effects in grammatical and ungrammatical subject-verb agreement dependencies. Both models are based on the idea of lossy memory representations of the intended message. The lossy compression model assumes that the linguistic input received by a comprehender is subject to information loss, and that the comprehender infers a set of possible true representations from the given input using their prior linguistic knowledge. Reading times are then predicted to be proportional to the expected surprisal of the next word given the set of possible true representations. By contrast, the hybrid lossy compression-plus-retrieval model assumes that dependency completion is driven by a cue-based retrieval process which is affected by a change in the representation of memory chunks due to lossy compression. Reading time predictions here are derived from the assumptions of cue-based retrieval (Lewis & Vasishth, 2005).

The evaluation of the three models’ predictive performance shows that both lossy compression and lossy compression-plus-retrieval models are better at explaining the interference effect data than the cue-based retrieval model of Lewis and Vasishth (2005). An important implication of the modeling results is that the cognitive processes underlying dependency completion in sentence comprehension might involve representation distortion due to lossy compression of the intended message.

An interesting open question is whether the deletion and insertion rates assumed in the lossy memory encoding function are sensitive to factors like the syntactic position of the nouns and the distance between the nouns and the verb. For example, a noun that appears earlier in the sentence may enjoy a primacy advantage (Häussler & Bader, 2015), and therefore be less likely to be distorted by deletion/insertion noise. Similarly, there could be a subject advantage in memory such that the representation of subject nouns is distorted at slower rates than other noun phrases (Futrell et al., 2020). Another reasonable assumption can be that the memory representation of nouns is susceptible to only deletion noise and not insertion noise. Our model is currently agnostic to these factors. But they can be explored by developing constraints on deletion and insertion rate for different noun phrases in our model. We plan to take this up in future work.

In sum, the modeling presented here demonstrates, for the first time, that a prominent and well-accepted explanation for interference effects—cue-based retrieval—is outperformed by models that assume lossy compression. The fact that the two lossy compression models (the one with and without cue-based retrieval) show comparable fits raises an interesting question: is the cue-based retrieval assumption needed at all to explain interference effects? This is an important open question that should be addressed in future work.

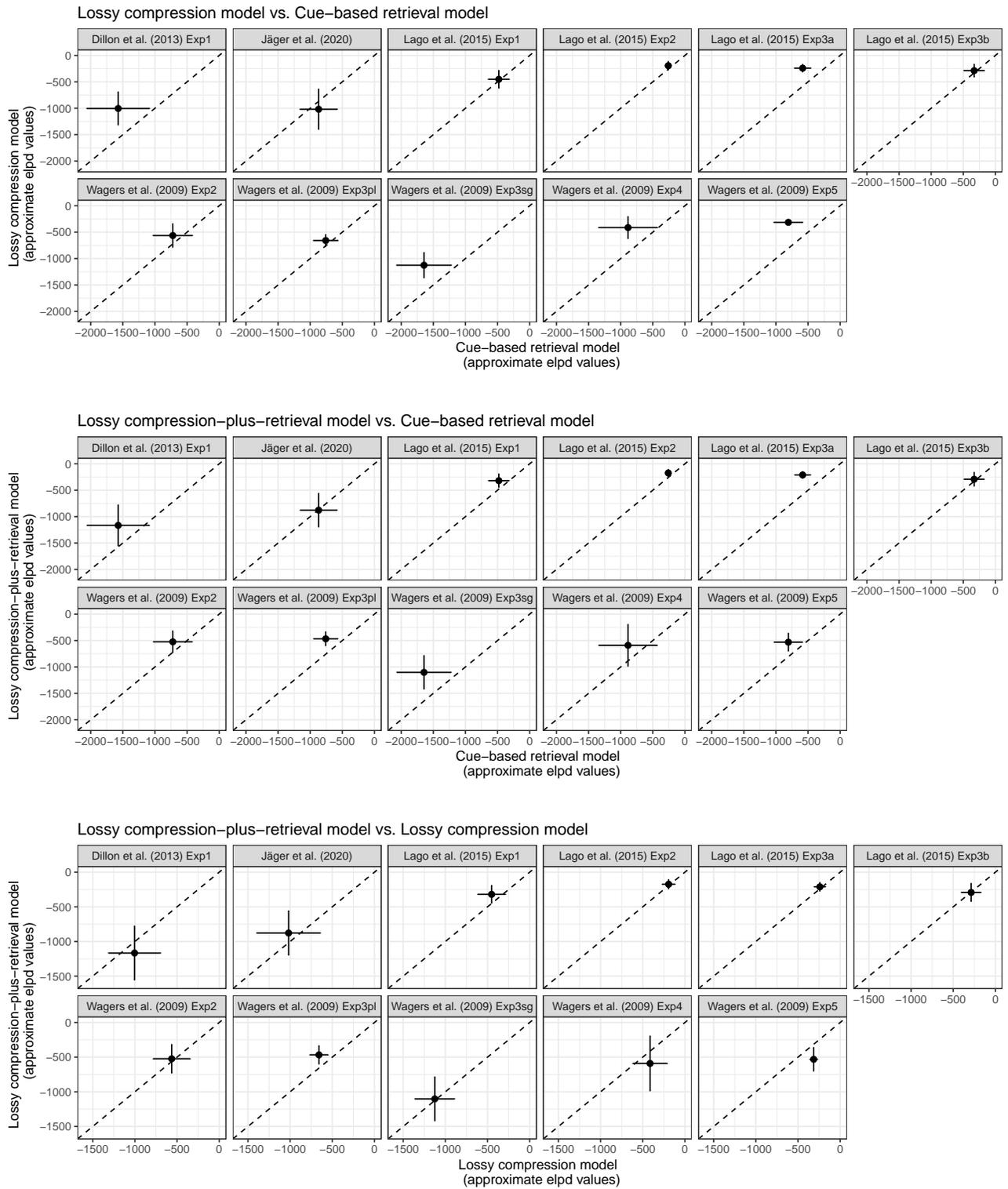


Figure 4: Estimated log pointwise predictive density for each model for each dataset based on stratified k-fold cross validation.

Acknowledgments

We thank the anonymous reviewers for helpful suggestions.

References

- Bock, K., & Eberhard, K. M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, 8(1), 57–99.
- Dillon, B. W., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, 69, 85–103.
- Eberhard, K. M. (1997). The marked effect of number on subject–verb agreement. *Journal of Memory and Language*, 36, 147–164.
- Futrell, R. (2019). Information-theoretic locality properties of natural language. In *Proceedings of the first workshop on quantitative syntax (quasy, syntaxfest 2019)* (pp. 2–15).
- Futrell, R., Gibson, E., & Levy, R. P. (2020). Lossy-context surprisal: An information-theoretic model of memory effects in sentence processing. *Cognitive science*, 44(3), e12814.
- Häussler, J., & Bader, M. (2015). An interference account of the missing-*vp* effect. *Frontiers in Psychology*, 6, 766.
- Jäger, L. A., Engelmann, F., & Vasishth, S. (2017). Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language*, 94, 316–339.
- Jäger, L. A., Merten, D., Van Dyke, J. A., & Vasishth, S. (2020). Interference patterns in subject-verb agreement and reflexives revisited: A large-sample study. *Journal of Memory and Language*, 111.
- Lago, S., Shalom, D. E., Sigman, M., Lau, E. F., & Phillips, C. (2015). Agreement processes in Spanish comprehension. *Journal of Memory and Language*, 82, 133–149.
- Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29(3), 375–419.
- Schäfer, R. (2015). Processing and querying large web corpora with the cow14 architecture.
- Schäfer, R., & Bildhauer, F. (2012). Building large corpora from the web using a new efficient tool chain. In *Lrec* (pp. 486–493).
- Sisson, S. A., Fan, Y., & Beaumont, M. (2018). *Handbook of approximate bayesian computation*. CRC Press.
- Vasishth, S., Nicenboim, B., Engelmann, F., & Burchert, F. (2019). Computational models of retrieval processes in sentence processing. *Trends in Cognitive Sciences*, 23, 968–982. doi: <https://doi.org/10.1016/j.tics.2019.09.003>
- Vasishth, S., Suckow, K., Lewis, R. L., & Kern, S. (2010). Short-term forgetting in sentence comprehension: Crosslinguistic evidence from verb-final structures. *Language and Cognitive Processes*, 25(4), 533–567.
- Wagers, M., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61, 206–237.
- Yadav, H., Smith, G., & Vasishth, S. (2021). *Feature encoding modulates cue-based retrieval: Modeling interference effects in both grammatical and ungrammatical sentences*. PsyArXiv. Retrieved from <https://psyarxiv.com/76aex>

Individual Differences in Decision Making Strategies Can be Predicted by Resting-State Functional Connectivity

Yuxue Cher Yang (chery@uw.edu)

Department of Psychology, University of Washington, Seattle, WA 98195 USA

Catherine Sibert (sibert@uw.edu)

Department of Psychology, University of Washington, Seattle, WA 98195 USA

Andrea Stocco (stocco@uw.edu)

Department of Psychology, University of Washington, Seattle, WA 98195 USA

Abstract

As the study of individual differences becomes more widespread, questions arise about the reasons that a particular individual might adopt a particular strategy. Using both the behavioral and functional neuroimaging data of healthy adults from Human Connectome Project (HCP) we examined decision making in an incentive processing task (Delgado et al. 2000). A pair of distinct ACT-R models, representing a Declarative strategy and a Procedural strategy, were used to classify subjects as either Declarative or Procedural decision makers based on their behavioral data. A machine learning Lasso analysis was performed on each subject's resting state functional connectivity, and was able to match the ACT-R model classifications to a high degree of accuracy. This suggests that the strength of connections between brain regions may play an important role in shaping the decision making process of a given individual.

Keywords: Decision Making; Strategy; Computational modeling; Functional connectivity; Procedural Memory; Declarative Memory; ACT-R

Introduction

It has been argued that, to be effective, computational cognitive models need to switch from nomothetic, group-level descriptions to idiographic, individual-level ones (Zhou et al., 2021). A promising framework in this sense was proposed by Ritter and Gobet (2000), who argued that an architecture can be used to successfully capture the invariant part of the mind, while different parameter values can be used to model variations across individuals. This approach was tested successfully by Daily and Lovett (2001), who succeeded in capturing individual differences in working memory through a single parameter in the ACT-R architecture (spreading activation W), and more recently, by Xu and Stocco (2021) using behavioral data. Recent work has also shown that individual parameter values are associated with different signatures of neural activity in EEG data (Zhou et al., 2021) and fMRI (Rice & Stocco, in press). These neural signatures were identified from "resting-state" recordings, that is, task-free sessions in which participants are not asked to do anything in particular, and which offer the opportunity to observe spontaneous but highly organized brain activity (Fox et al., 2005). The fact that parameter values that capture individual differences are

reflected in resting state imaging data suggests a biological underpinning for these parameters.

Despite its successes, the approach of identifying individual differences with parameter values still runs into conceptual roadblocks. While a cognitive architecture can be assumed to reflect an invariant, innate blueprint (Taatgen, 2020), participants are typically measured when performing a *specific task*, and, even with the same architecture, participants might perform the same tasks in the same way. For example, simple association learning tasks can be modeled using two strategies, a procedural-based reinforcement learning strategy and a memory-based, decision-by-sampling or instance-based learning strategy. Haile et al. (2020) showed that different participants are best fit by different strategies. This implies that attempts to measure single parameters across participants is ultimately doomed to fail: it does not make sense to estimate learning rate (a reinforcement learning parameter) from participants who rely on memory, and it does not make sense to measure rate of forgetting (a successfully decodable parameter) from individuals who follow a memory-less, procedural learning strategy.

Through computational models, it is possible to make inferences about which strategy a participant is using (Haile et al., 2020). But what makes participants *prefer* a strategy over another? In principle, strategy selection could be a function of personal preference, habit, or cost-benefit analysis (Payne, Bettman, & Johnson, 1993). One enticing possibility is that strategy selection might reflect bounded rationality (Lewis et al., 2014): individuals choose the strategy that plays to their strengths, yielding the best results given the computational costs involved. If this is the case, then it follows that preference for a strategy over another would also ultimately depend on identifiable stable characteristics of their brain activity.

To test this hypothesis, we analyzed a dataset including almost 200 participants for whom performance on a simple decision-making task and resting-state fMRI data were available. Computational models implementing alternative strategies were fit to individual behavioral data to determine the most likely strategy used by each participant. Machine learning techniques were then employed to identify the

facets of spontaneous neural activity that best predict which strategy will be used by each individual. We expected to find that decision-making strategies associated with the use of memory resources (such as retrieving the previous success history of an option) would be associated with increased functional connectivity in fronto-parietal regions responsible for cognitive control. Conversely, we expected that decision-making strategies associated with habitual and reward-based learning would be associated with increased functional connectivity in sensorimotor cortices responsible for automatic stimulus-response behaviors and with the basal ganglia circuit responsible for feedback-driven learning (Yin & Knowlton, 2006).

Methods

This study analyzed both behavioral and neuroimaging data obtained from the Human Connectome Project (HCP) dataset (Van Essen et al., 2013). A total of 199 participants (111 females, 85 males, and 3 did not disclose) who completed both sessions of the task-based fMRI gambling game were included in this study. All participants were healthy adults with no neurodevelopmental or neuropsychiatric disorders. The experimental protocol, subject recruitment procedures, and consent to share de-identified information were approved by the Institutional Review Board at Washington University.

The Incentive Processing Task in the HCP

This incentive decision making task was adapted from the gambling paradigm developed by Delgado and Fiez (2000). Participants were asked to guess if the number on a mystery card (represented by a “?”, and ranging from 1-9) was more or less than 5. After making a guess, participants were given feedback, which could take one of three forms, *Reward* (a green up arrow and \$1), *Loss* (a red down arrow and -\$0.50), or *Neutral* (a gray double headed arrow and the number 5). The feedback did not depend on the subject’s response, but was determined in advance; the sequence of pre-defined feedback was identical for all participants. The task was presented in two runs, each of which contains 64 trials divided into eight blocks. Blocks could be *Mostly Loss* (6 loss trials pseudo-randomly interleaved with either 1 neutral and 1 reward trial, 2 neutral trials, or 2 reward trials) or *Mostly Reward* (6 reward trials pseudo randomly interleaved with either 1 neutral and 1 loss trial, 2 neutral trials, or 2 loss trials). In each of the two runs, there were two *Mostly Reward* and two *Mostly Loss* blocks, interleaved with 4 fixation blocks (15 seconds each).

Resting-State fMRI Analysis

This study employed the “minimally preprocessed” version of resting-state fMRI data, which has already undergone a minimal number of standard preprocessing steps including artifact removal, motion correction, normalization, and registration to the standard MNI ICBM152 template. Additional preprocessing steps were performed using the AFNI software (Cox RW, 1996),

including despiking, spatial smoothing with an isotropic Gaussian 3D filter FWHM of 8 mm, and removal of linear components related to the six motion parameters and their first-order derivatives.

Functional connectivity measures were constructed from the HCP resting-state data using Power et al. (2011)’s whole brain parcellation. This parcellation was used to construct a 264 Region of Interest (ROI) functional atlas, with each ROI containing 81 voxels. This parcellation atlas is defined in the MNI space and was applied to all participants in HCP dataset. The extraction of the time series and calculation of the connectivity matrices was performed using R (RStudio Team, 2016) and Python. Pearson correlation coefficients and partial correlation coefficients between the time series of each brain region were calculated for each participant, resulting in a 264×264 symmetric connectivity matrix for each session for each subject. The averaged correlation coefficients across subjects were calculated by first transforming each r value into a z -value, and then retransforming the average z value back into an equivalent r value using the hyperbolic tangent transformation (Silver & Dunlap, 1987).

Response Switch Analysis

Because in the Incentive Processing task the feedback is scheduled in advance and does not depend on actions taken by participants, it is impossible to define participant’s performance in terms of either accuracy or learning. This poses a challenge when trying to determine if participants are responding to feedback. The most meaningful way to check whether participants change their behavior in response to feedback is through analyzing their Win-Stay, Lose-Shift (WSLS) probabilities. Thus, our main dependent variable was the tendency to switch responses after a Loss feedback and after a Reward feedback. This response switch is coded as 0 if the current response is the same as the next response, and coded as 1 if the current response is not the same as the next response. Because the response switch is a binary variable, the analysis was conducted with logistic mixed-effects models using orthogonal contrast coding as implemented in the “lme4” package in R. Given that Neutral trials make up only a small proportion of total trials, they were excluded in statistical tests. In the mixed-effect model, Block Type (Reward or Loss) and Trial Type (Reward or Loss) were treated as fixed effects, and individual participants were treated as random effects. The parameters were estimated based on the maximum likelihood.

On the group-level, there is no significant effect of feedback nor Block Type on the probability of switching responses. However, and critically for this study, on an individual-level, individuals do exhibit different behavioral response profiles. Figure 1 demonstrates the mean probability of response switching as a function of Trial Type (feedback received) and Block Type.

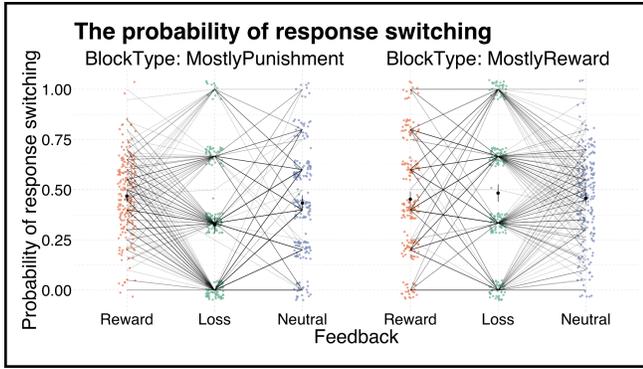


Figure 1: The mean probability of response switching as a function of feedback and block type. Each color dot and grey line represents the mean probability of response switching of a single participant, and the black dot represents the mean and 95% confidence interval across participants.

We also examined whether the response times change as a function of previous feedback (the Trial Type of the previous trial) and Block Type. Excluding neutral trials in the statistical analysis, on average, participants tend to take longer when making decisions in Mostly Reward blocks than in Mostly Loss blocks ($\beta = 15.21$, $SE = 6.39$, $p = 0.017$), regardless of previous feedback. Figure 2 shows the mean response time (RT) as a function of Previous Feedback and Block Type. Compared to the probability of response switching, however, the pattern of RTs was found to be noisier and less consistent across individuals, and was therefore not included in the following modeling analysis.

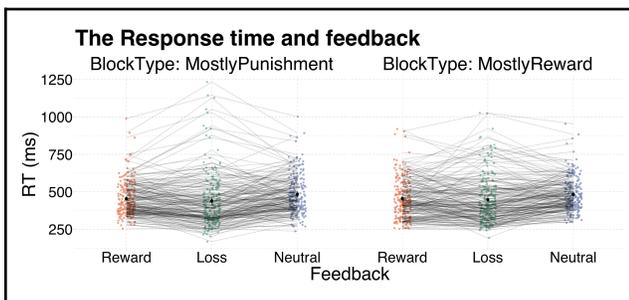


Figure 2: Mean response time as a function of previous trial feedback and block type. Each colored dot and grey line represents the RT of an individual, and the black dot represents the mean and confidence interval (95%) of RT across participants.

ACT-R Model Design

While the behavioral data does not reveal major effects across subjects, it offers an exciting opportunity from a modeling perspective. There exist two competing explanations of how decision making occurs in a repeated choice paradigm, one based on episodic memory of previous choices (Gonzalez et al., 2003) and one based on reinforcement learning (Daw et al., 2011). Each explanation

is dependent on different mechanisms, and, ultimately, reliant on different strategies. Both explanations were implemented as two computational models in the ACT-R cognitive architecture (Anderson, 2007): as a *Declarative Model*, reliant on memory retrieval, and a *Procedural Model*, which makes use of reinforcement learning.

Declarative Model The Declarative Model relies on the declarative module to retrieve a memory of prior actions and their corresponding feedback. When presented with a mystery card, the model selects an action, LESS or MORE, for evaluation, and makes a retrieval of the prior history of feedback associated with that action. If the retrieved history contains a WIN result of the chosen action, it will execute that action, but if the history contains a LOSE or NEUTRAL result, the model will execute the alternate action. If no history is retrieved, an action will be executed at random. After making a guess, the model is presented with feedback, which is encoded as a new memory chunk associated with the selected action. In ACT-R, memory chunks are retrieved based on their activation, calculated with a base-level learning function that reflects the degree to which a chunk matches the context of the retrieval request, and the recency of prior retrievals (Eq 1). If the activation surpasses a specified threshold, the chunk is possible to retrieve and if multiple chunks meet this threshold, the chunk with the greatest activation will be selected. The model functions by remembering the results of previous actions to guide future actions.

$$A_i = \ln\left(\sum_{j=1}^n t_j^{-d}\right) + \epsilon \quad (1)$$

Procedural Model By contrast, the Procedural Model represents the possible actions of the decision-making processes as competing rules, and reinforcement learning is used to increase the use of the rule that leads to the best outcomes. Instead of encoding each trial as a memory of an action and associated feedback, the model has two competing production rules that execute the MORE and LESS actions. When presented with the mystery card, the model will choose one of the rules to execute based on their utility. Initially, both rules have equal utility, and one will be chosen at random. After making a guess, the model is presented with a WIN, LOSE, or NEUTRAL response, and this feedback is encoded as an adjustment to the utility of the selected production rule (+1 for a WIN result, -1 for a LOSE result, and no change for a NEUTRAL result). At any time point t , the utility U of production p is calculated using Eq 2, where α indicates the learning rate, R_t is the reward the production received for at time t . Previous rewards will encourage the model to repeat the associated action, while a pattern of losses will decrease the utility of the action and encourage the selection of the alternate action.

$$U_t = U_{t-1} + \alpha(R_t - U_{t-1}) + s \quad (2)$$

Individual Fit and Model Evaluation

To examine the predictions of our model, we used a grid-search approach to find the best possible parameters within the parameter space shown in Table 1. Each model simulates 64 trials, the same as the experimental paradigm for participants, repeated over 50 runs. The simulated stimuli were presented in the same order as the real experimental stimuli to avoid any potential noise from sequence effects in the simulation. Following the six conditions (Reward, Loss, Neutral trials in Mostly Reward Block and Reward, Loss, Neutral trials in Mostly Loss Block), the mean probability of response switching, P(Switch), and its standard deviation are computed.

Table 1: Model parameter space in the simulations.

Models	Parameter	Value	Meaning
Declarative	ϵ	0 - 0.5	activation noise
	d	0.2 - 0.85	memory decay
Procedural	s	0 - 0.5	utility noise
	α	0.05 - 0.5	learning rate

In order to evaluate the goodness-of-fit for individual fitting, we estimated maximum Log-Likelihood across the parameter space. The likelihood function of a particular model with parameters θ , $L(m, \theta | x)$, is the probability that, given the parameterized model and set of observed data to fit, the model would produce that data: $L(m, \theta | x) = P(x|m, \theta)$. Here, m and θ refers to the model and its parameters, and x refers to the observations. Common comparison metrics, such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), are both based on likelihood, but rely on closed-form likelihood functions. While it is possible to derive such functions for simple models (such as logistic models or linear models), they can be incredibly difficult to derive for more complex models and impossible for arbitrarily complex models based on ACT-R and other high level architectures. Some attempts have been made to evaluate complex models with basic likelihood metrics: Stocco and Haile (2018), Prat and Stocco (2020), and Yang and Stocco (2019) have all used BIC to compare competing ACT-R models. However, the equation used to estimate BIC is a closed-form approximation that is based on Residual Sum of Squares and was originally derived for linear models; as such, it does not necessarily hold for ACT-R.

In this paper, we followed the computationally expensive but more accurate solution of empirically calculating the likelihood function by simulating each model and set of parameters multiple times, and calculating the empirical probability distribution of each set of results (Yang, Karmol, Stocco, in press). Knowing the mean and standard deviation of this distribution, the value of $P(x|m, \theta)$ can then be calculated directly. If a model is designed to predict n data points (corresponding, for instance, to different experimental conditions), its likelihood can be expressed as

the joint probability that any of those data points can be produced. For simplicity, and assuming independence, this can be expressed as the product of the probability of observing each individual data point in the empirical data, i.e., $L(m, \theta | x_1, x_2, \dots, x_n) = \prod_i L(m, \theta | x_i)$. Finally, to avoid computational problems with vanishing small probabilities, it is common to express this value in terms of *log* likelihood:

$$\log L = \log P(x | m, \theta) = \sum_i \log z[(x_i - x_{i,m}) / \sigma_{i,m}] \quad (3)$$

Results

Decision-Making Strategy Identification

By excluding participants who did not complete the gambling task and two sessions of resting state fMRI scanning, a total of 199 participants were fit by ACT-R models. Of these, 127 (63.82%) were best fit by the Declarative Model, and thus were identified as Declarative decision makers. The remaining 72 (36.18%) individuals were best fit by the Procedural Model, and identified as Procedural decision makers. The logistic mixed-effects model was conducted using orthogonal contrast coding as implemented in the lme4 package in R. ACT-R Model Type (Declarative vs. Procedural), Block Type (Mostly Reward vs. Mostly Loss), and Feedback (Reward vs. Loss) were treated as fixed effects, and individual subjects were treated as random effects. Full statistical results are shown in Table 2. In contrast to the lack of significant effects across the behavioral data, the probability of response switching was found to be statistically different between the two groups identified as either Declarative decision makers or Procedural decision makers ($z = -6.11, p < 0.001$), supporting the validity of the ACT-R model identification.

Table 2: Results of the Logistic Mixed Effects Model of the Probability of Response Switch

Statistical Test	odds ratio	se	z	p
(Intercept)	0.88*	0.05	-2.30	0.022
Model Group	0.71***	0.04	-6.11	<0.001
Block Type	0.98	0.03	-0.80	0.423
Trial Type	1.08**	0.03	2.84	0.005
Model Group by Block Type	1.07*	0.03	2.46	0.014
Model Group by Trial Type	0.8***	0.02	-8.10	<0.001
Block Type by Trial Type	1.02	0.03	0.78	0.434
Model Group by Block Type by Trial Type	1.10***	0.03	3.47	0.001
Random Effect				
σ^2	3.29			
ICC	0.12			

<i>N HCPID</i>	199
<i>observation</i>	9746
<i>Marginal R²</i>	0.030/0.144
<i>/Conditional R²</i>	
<i>Log-Likelihood</i>	-4010.323

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

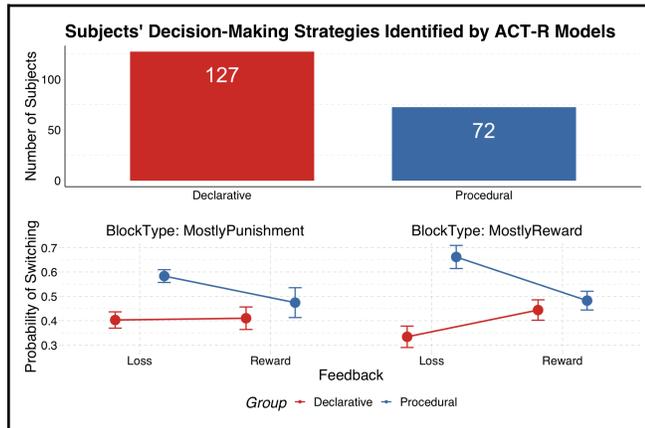


Figure 3. (Top) Counts of individuals identified by ACT-R models. The red bar represents the number of participants best fit by the Declarative model; the blue bar represents the number of participants best fit by the Procedural Model. (Bottom) The probability of response switching by two groups of individuals identified as either Declarative or Procedural decision makers.

Supervised Classification with Logistic Model

To explore if the behavioral differences between Declarative and Procedural decision makers are indicated by an individual's underlying brain structure, we trained a Logistic Regression model using resting state functional connectivity as its variable, and predicted the probability of a participant being labeled as either Declarative-based or Procedural-based decision maker by the ACT-R model classification. In order to handle an imbalanced dataset with unequal target labels, upsampling was applied by randomly adding data from the minority class. Having 69,696 (264 ROI \times 264 ROI) connections, we want to select only the most important connections contributing to the prediction, therefore, Lasso regularization was applied to the Logistic Model. Lasso is a machine learning regression analysis technique that performs both variable selection and regularization in order to improve the prediction accuracy and interpretability of the computational model. It can reduce model complexity by penalizing large numbers of coefficients and also prevents overfitting which may result from simple linear regression. Lasso minimization is calculated using Eq 4, where the tuning parameter λ controls the degree of penalty: for greater values of λ , more coefficients are forced to become 0.

$$\sum_{i=1}^n (y_i - \sum_j x_{ij} \beta_j)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (4)$$

To account for the large disparity between the number of participants and the number of predictors, we performed a Grid search in sklearn (Pedregosa et al., 2011) with 20-fold cross-validation to determine the best fit hyper-parameter λ (6.73). To maximize the penalty to coefficients, the highest value of λ with validation accuracy with one standard error of the maximum accuracy was chosen (as recommended by Krstajic et al (2014)). Instead of splitting the entire dataset into training and testing sets, we refit the model using Leave-one-out (LOO) cross validation. The model is trained on all samples except one and the prediction is made on that one sample, then the process is repeated across the full dataset. The mean score (accuracy), true positive rate (TPR), true negative rate (TNR), false positive rate (FPR), false negative rate (FNR) are calculated across all folds to evaluate the performance of the model. By definition, the receiver operating characteristic curve (ROC) demonstrates the performance of a classification model by plotting the relationship between TPR vs. FPR at different classification thresholds. We calculated the AUC (Area under the curve), which is one of the most important metrics for evaluating a classification model's performance; as the AUC of a model approaches 1, the model approximates an ideal, perfect classifier. It provides information about how well a classification model is capable of distinguishing between classes. The overall classification accuracy is 0.88 and the ROC-AUC is 0.94, indicating that predicting from an individual's resting state functional connectivity, the Lasso Logistic model is capable of matching ACT-R's prediction about whether an individual is a Declarative-based or Procedural-based decision maker.

Connectivity Map

With Lasso regularization, approximately 1.4% of β estimates in the Logistic model are not zero, suggesting a relatively sparse neuro functional connectivity of the resting brain. In a 264×264 β coefficients matrix, the β_{ij} value indicates the weight of connectivity between the i -th and the j -th region in classifying whether the human subject is a Declarative decision maker or a Procedural decision maker from the resting state functional connectivity. The ultimate effect of β on the predicted group assignment depends on the polarity of the underlying functional connectivity. A positive β value has different implications if applied to a positive or negative partial correlation between two regions. To make the interpretation of the values unambiguous, we multiplied the β matrix with the averaged partial correlation coefficient matrix A , obtaining a group-level weighted averaged correlation matrix W . Figure 4 demonstrates the brain connectivity map of W , thresholded so that only the most predictive 68 connections (corresponding to 0.01% of the initial pool of regressors) are shown. In this figure, red lines represent functional connections that are predictive of

a Declarative decision maker, and blue lines represent functional connections that are predictive of a Procedural decision maker. Color shades suggest the strength of predictability.

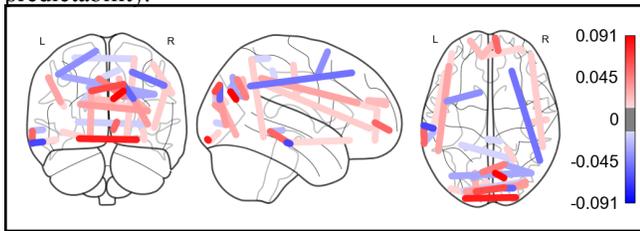


Figure 4. The group-level weighted averaged brain connectivity plot.

As we anticipated, the results show a dissociation between the types of connectivity associated with Declarative or Procedural strategies. Using the Power et al (2011) functional classification of these regions as a guideline, the results show that the use of a Declarative strategy was mostly associated with increased functional connectivity in the networks of regions associated with task control (fronto-parietal networks and attention networks) and episodic memory (default mode network and memory retrieval network), while the use of Procedural strategy was mostly linked to increased functional connectivity in sensorimotor and subcortical networks.

Discussion

This paper shows that individual preferences for using a declarative or a procedural strategy can be decoded from patterns of resting state functional connectivity data. The specific connectivity values suggest that an individual's preference for a particular strategy might be adaptive and rational. Specifically, individuals exhibiting a stronger fronto-parietal connectivity play to their strengths, and tend to use declarative strategies that are more reliant on controlled memory retrieval, while individuals with stronger sensorimotor connectivity tend to use procedural strategies. In general, the patterns of functional connectivity are compatible with ACT-R's regions.

Although our results are encouraging, a number of limitations must be acknowledged. First, the Declarative vs. Procedural classification of individuals' probability of switching is based on a log-likelihood model fitting procedure, and thus, no ground-truth labels were available. Moreover, the optimal parameter was searched from a finite grid, and determined by the highest log-likelihood value compared to empirical data. Second, the task is highly unusual, in that it provides no real opportunity for learning from feedback. Further study could model the learning effect and investigate whether different learning mechanisms could also be predicted by the neuro-functional connectivity.

These limitations notwithstanding, we believe that our results have some important implications. First, they provide a new and deeper way to connect individual differences in task performance with individual

neurobiology, showing how the latter might provide constraints on the specific strategies that are selected.

Second, they have implications for ACT-R. Procedural knowledge has been traditionally associated, in ACT-R, with the function of the basal ganglia. While the role of the basal ganglia in learning procedural knowledge is well supported (Knowlton et al., 2006 etc.), it is not clear that the basal ganglia are also the ultimate seat of procedural knowledge. In fact, both modeling work (Stocco, Lebiere, & Anderson, 2010) and experimental work using neurostimulation (Rice & Stocco, 2019) point to procedural knowledge being ultimately encoded in a set of cortico-cortical connections that directly link stimulus-response associations. This interpretation is compatible with our findings that find greater likelihood of using procedural strategies in individuals with stronger perceptuo-motor connectivity.

Acknowledgements

This research was supported by grant FA9550-19-1-0299 from the Air Force Office of Scientific Research to AS.

References

- Anderson, D. R. (2007). Model based inference in the life sciences: a primer on evidence. *Springer Science & Business Media*.
- Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research*, 29(3), 162-173.
- Daily, L. Z., Lovett, M. C., & Reder, L. M. (2001). Modeling individual differences in working memory performance: A source activation account. *Cognitive Science*, 25(3), 315-353.
- Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience*, 8(4), 429-453.
- Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., & Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of neurophysiology*, 84(6), 3072-3077.
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, 102(27), 9673-9678.
- Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making'. *Cognitive Science*, 27(4), 591-635.
- Haile, T., Prat, C. S., & Stocco, A. (2020) One size doesn't fit all: Idiographic computational models reveal individual differences in learning and meta-learning strategies. *Proceedings of the 18th International Conference on Cognitive Modeling*, pp. 75-81.
- Lewis, R. L., Howes, A., & Singh, S. (2014). Computational rationality: Linking mechanism and behavior through

- bounded utility maximization. *Topics in cognitive science*, 6(2), 279-311.
- Krstajic, D., Buturovic, L. J., Leahy, D. E., & Thomas, S. (2014). Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of Cheminformatics*, 6(1), 1-15.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The Adaptive Decision Maker*. Cambridge university press.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *The Journal of machine Learning research*, 12, 2825-2830.
- Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., ... & Petersen, S. E. (2011). Functional network organization of the human brain. *Neuron*, 72(4), 665-678.
- Rice, P., & Stocco, A. (2019). The role of dorsal premotor cortex in resolving abstract motor rules: Converging evidence from Transcranial Magnetic Stimulation and cognitive modeling. *Topics in cognitive science*, 11(1), 240-260.
- Rice, P., & Stocco, A. (in press). Estimating individual differences in working memory through ACT-R modeling and resting state connectivity. *Proceedings of the 19th International Conference on Cognitive Modeling*.
- Silver, N. C., & Dunlap, W. P. (1987). Averaging correlation coefficients: should Fisher's z transformation be used?. *Journal of applied psychology*, 72(1), 146.
- Stocco, A., Lebiere, C., & Anderson, J. R. (2010). Conditional routing of information to the cortex: A model of the basal ganglia's role in cognitive coordination. *Psychological review*, 117(2), 541.
- Taatgen, N. A. (2020). Cognitive Architectures: Innate or Learned?. *Common Model of Cognition Bulletin*, 1(2), 476-480.
- The R Team (2016). *RStudio: integrated development environment for R*. RStudio, Inc., Boston, Massachusetts, USA.
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., & Wu-Minn HCP Consortium. (2013). The WU-Minn human connectome project: an overview. *Neuroimage*, 80, 62-79.
- Xu, Y., & Stocco, A. (2021). Recovering reliable idiographic biological parameters from noisy behavioral data: The case of basal ganglia indices in the Probabilistic Selection Task. *Computational Brain & Behavior*, 1-17.
- Yang, Y.C. and Stocco, A. (2019). Syntactic priming depends on procedural, reward-based computations: Evidence from experimental data and a computational model. *Proceedings of the 17th International Conference on Cognitive Modeling*, pp. 307-313.
- Yang, Y. C., Karmol, A. M., & Stocco, A. (in press). Core cognitive mechanisms underlying syntactic priming: A comparison of three alternative models. *Frontiers in Psychology*.
- Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, 7(6), 464-476.
- Zhou, P., Sense, F., Van Rijn, H., & Stocco, A. (2021). Reflections of idiographic long-term memory characteristics in resting-state neuroimaging data. *Cognition*, 212, 104660.