Gobet, F., & Ritter, F. E. (2000). Individual Data Analysis and Unified Theories of Cognition: A methodological proposal. In N. Taatgen & J. Aasman (Eds.), *Proceedings of the 3rd International Conference on Cognitive Modelling*. 150-157. Veenendaal (NL): Universal Press.

Individual Data Analysis and Unified Theories of Cognition: A Methodological Proposal

Fernand Gobet (frg@psyc.nott.ac.uk)
ESRC CREDIT, School of Psychology
University of Nottingham
Nottingham NG7 2RD UK

Frank E. Ritter (ritter@ist.psu.edu)
School of Information Sciences and Technology
Penn State University
State College, PA 16801 USA

Abstract

Unified theories regularly appear in psychology. They also regularly fail to fulfil all of their goals. Newell (1990) called for their revival, using computer modelling as a way to avoid the pitfalls of previous attempts. His call, embodied in the Soar project has so far, however, failed to produce the breakthrough it promised. One of the reasons for the lack of success of Newell's approach is that the methodology commonly used in psychology, based on controlling potentially confounding variables by using group data, is not the best way forward for developing unified theories of cognition. Instead, we propose an approach where (a) the problems related to group averages are alleviated by analysing subjects individually; (b) there is a close interaction between theory building and experimentation; and (c) computer technology is used to routinely test versions of the theory on a wide range of data. The advantages of this approach heavily outweigh the disadvantages.

1. Introduction

What is the best way to make theoretical progress in the study of behaviour in general and of cognition in particular? To develop micro-theories explaining a small domain or to aim at a higher goal, and develop an overarching theory covering a large number of domains—a unified theory? Modern psychology, as a field, has tended to prefer micro-theories. It is true that unified theories have regularly appeared in psychology—think of Hull's, Piaget's or Skinner's theories—but it is generally admitted that such unified theories have failed to offer a rigorous and testable picture of the human mind. Given this relatively unsuccessful history, it was with interest that cognitive science has observed Newell's (1990) call for a revival of unified theories in psychology.

Newell, who focused on cognition, was quite aware of the problems that plagued previous attempts: vagueness, lack of specific predictions, and untestability, to cite the most damaging. He was also aware that psychology now has a tool that was not available to Hull, Piaget, or Skinner: computers. Newell embodied his call for unified theories of cognition (UTCs) in the Soar project, where computer modelling plays a dominant role. In the ten years or so of the Soar project, the verdicts by observers have ranged from mild support (e.g. Norman, 1991) to strong disagreement (Cooper & Shallice, 1995).

One of the reasons for the limited success of Newell's own brand of UTC is that the methodology commonly used in psychology, based on controlling potentially confounding variables by using group data, is not the best way forward for developing UTCs. Instead, we propose an approach, which we call individual data modelling, where (a) the problems related to

group averages are alleviated by analysing subjects individually on a large set of tasks; (b) there is a close interaction between theory building and experimentation; and (c) computer technology is used to routinely test versions of the theory on a wide range of data. The discussion of the advantages and disadvantages of this approach will show that there are significant advantages and that this approach will also help traditional approaches progress. The main potential disadvantage—lack of generality—may be taken care of by adequate testing procedures.

1.1. Unified Theories

A common criticism of unified theories of the past is that they were formulated in rather vague terms, and that, as a consequence, both their internal consistency and their testability were open to serious doubt. This criticism also applies to some extent to theories such as Piaget's and Hull's, which, although formulated formally (i.e. logic and mathematics respectively), were too unspecified and awkward to make direct, testable empirical predictions. The strength of Newell's (1990) argument for UTCs is that it avoids the danger of lack of specificity by showing that we now have the necessary technology (the computer) to build theories complex enough to match human intelligence. The most advanced computational theories of cognition, including Soar (Newell, 1990) and ACT-R (Anderson & Lebière, 1998), are an existence proof that psychological theories can be formulated in a way that satisfies the exigencies of scientific rigor and specificity while predicting intelligence by exhibiting it themselves.[1]

Newell (1990) proposed a special type of UTC as a new methodological way of studying cognition. The key idea is that a single architecture should be used to account for as many regularities in empirical data as possible, and that this architecture should be implemented in a computer program. This avoids the vagueness of verbal theories and makes it possible to simulate complex behaviour. Newell's insight is that multiple constraints are brought to bear with UTCs, allowing one to limit the number of degrees of freedom in the theory and to converge on a theory that accounts satisfactorily for most of the regularities of human behaviour. Newell is clear (1990, p. 15-17) that this approach does not imply that a single mechanism must unify, but that the set of mechanisms must work and exist together as a unified whole.

To make Newell's ideas more concrete, let us consider a simple, idealised example. Researcher A develops a theory of memory, and manages to estimate two parameters: capacity of working memory (WM), and time to create a new node in long-term memory. On her own, Researcher B develops a theory of problem solving in arithmetic, and uses a parameter for the capacity of WM. This is essentially a free parameter that can be adjusted to fit the data. Now consider Researcher UTC, who is simultaneously interested in both domains, memory and problem solving in arithmetic. For her, the capacity of WM is not a free parameter anymore, as it was 'set' with simulations on memory. The parameter estimated for WM has constrained the space of possible theories for researcher UTC. Were researcher UTC to change values of the WM parameter, perhaps because its current value does not allow problem solving in arithmetic to be carried out at all, she would know that the two theories were contradictory, and have to retest, and perhaps revise her theory of memory with the new value.

Without a UTC, it is much harder to propagate constraints across subtheories, both because of the amount of information to be processed by the theorists (i.e. translating the restrictions between formalisms), and because of the biases that they may hold for or against their theory's features. UTCs are, therefore, the necessary vehicle to propagate constraints across subtheories. In fact, attempts to propagate constraints across subtheories start to create unified theories.

1.2. Soar as a Candidate UTC

[1] We focus on symbolic cognitive architectures here, but our proposal also applies to non-symbolic architectures.

Newell illustrates his UTC methodology with Soar, a 'candidate Unified Theory of Cognition.' Soar, which is both a cognitive theory and an AI system, represents intelligence as a function of problem solving and learning, and essentially describes cognition as search in problem spaces. In Soar, all knowledge is encoded as productions and all learning is done by chunking.

How does Soar fare with empirical data? Reasonably well, as it has been tested in detail against various domains and types of data. A partial list includes reaction tasks, episodic memory, typing, categorical learning, sentence comprehension, skill acquisition, reasoning with syllogisms, problem solving (e.g. cryptarithmetic), human-computer interaction, development (balance-beam task), and driving (Aasman & Michon, 1992; Altmann & John, 1999; Howes & Young, 1996; Lewis, 1996; Polk & Newell, 1995). In particular, the chunking mechanism used by Soar offers a parsimonious explanation of the ubiquitous power law of learning (Newell, 1990). In spite of this long list of achievements, which even its opponents acknowledge (e.g. Cooper & Shallice, 1995), Newell's UTC approach has been subjected to a tide of criticisms. These may be classified into two categories: criticisms against Soar and criticisms against UTCs as a general research methodology.

Norman (1991), while positive overall, is a good example of a set of criticisms aimed at Soar as a UTC. Norman finds implausible the level of theoretical unification proposed by Soar: a single learning mechanism, a single knowledge representation, and a uniform problem state. Soar's lack of an explicit working memory capacity limit is also seen as a problem. Norman also regrets that Soar does not take into account more neuropsychological evidence, and notes that there may be non-symbolic intelligence not captured by Soar's symbolic mechanisms.

A different line of criticism against Newell's project is to identify the methodological difficulties faced by UTCs in general. This line is adopted, among others, by Cooper and Shallice (1995). They assert that any theory can be implemented in Soar, and that, as a consequence, Soar can be seen as 'just' a powerful computer language. They also deplore the gap between theoretical descriptions and computational implementations of the theory, and consider that UTCs do not adequately address the problem of irrelevant specification (what aspects of a program make psychological claims, and what aspects are present just to have the program run?). They also note that start-up assumptions can be tailor-made for each task and that Soar modellers may use different assumptions. This can weaken the potential benefit of UTCs, namely bringing to bear multiple constraints.

Even though Cooper and Shallice (1995) identified some genuine difficulties faced by UTCs and illustrated the uneasiness that traditional approaches in psychology have with Newell's approach to a UTC, many of their criticisms are unwarranted, such as how they misrepresent the nature of the empirical illustrations in Newell (1990). (See Young, Ritter & Gobet, in preparation, for a more detailed discussion.) However, the focus of this paper is to address a set of other difficulties not addressed by Cooper and Shallice that seriously hamper the usefulness of UTCs, and, in particular, the supposed strength of bringing together multiple constraints. These difficulties have to do with the way constraints are (or are not) efficiently used to prune the search for possible theories. We can identify four difficulties that we address with our new approach:

1. How to deal with differences between subjects in their architecture and knowledge (between-subject variability)? Is averaging the data a good solution?
2. How to deal with differences in difficulty between task instances and task types (between-task variability)?
3. How to control for subjects' (multiple) strategies (between-strategy variability)?
4. How to estimate quantitative and qualitative (e.g. strategies) parameters using typical data?

A significant part of these difficulties comes from using group data. The pitfalls associated with such groupings have been known for a long time (e.g. Newell & Simon, 1972; Siegler,

1987), but are rarely dealt with satisfactorily in current research. It has been noted several times that data averaged over people may not accurately reflect the behaviour of any one person. The same point has been made for data averaged over a task, where subjects may use different strategies. Finally, what is the meaning of cognitive parameters estimated from the 'average' subject, and what constraining power do they have?

A solution may be to use ranges of parameter values. For example, Card, Moran, and Newell (1983) estimated the range of the possible values associated with the capacity and the decay rate of various information stores, such as visual short-term memory. Using ranges has the disadvantage that UTCs may lose one of their strongest aspects (using multiple constraints), because ranges do not constrain the theory as well as point measurements. Another solution is to dispose with group data completely, and to turn to Individual Data Modelling. The second solution is explored here.

1.3. Individual Data Analysis (IDA)

The obsession of modern psychology with statistical testing has led its practitioners to hold strong prejudices against IDA. However, research using IDA has a long history in psychology (for a review see Dukes, 1968), including Freud's efforts to develop psychoanalysis and Piaget's 'clinical method' to understand children's development. Even now, this methodology is not uncommon in neuropsychology, where the rarity of patients with specific brain damage almost compels it, in clinical psychology, which has developed methodological tools based on individual data to study the effect of therapies, and in psychophysics, where only two subjects are typically needed, the experimenter and a 'naive' subject to test for experimenter biases.

IDA has had a lasting impact on cognitive psychology (broadly defined) as well. Ebbinghaus (1885) started the field of verbal learning by experimenting on himself. Bryan and Harter (1899), documenting how Morse code is learnt, and Seibel (1963), analysing a choice reaction-time task, have studied single subjects tested over long periods (for up to 70,000 trials) on a given task. De Groot (1978/1946), launched the modern study of expertise using analyses that focused on the detailed description of individual subjects. In developmental psychology, Siegler (e.g. 1987) has several times warned about the dangers of averaging data, and has developed techniques to study the development of each child separately using a microgenetic methodology. The information-processing approach to problem solving (Newell & Simon, 1972) has also tended to focus on subjects individually. Finally, Lovett, Reder, and Lebière (1997) and Miwa and Simon (1993) have investigated ways of using computer modelling to estimate individual parameters. This influence remains in current experiments on expertise, where a single subject's development is studied for a long period of time, using intensive data collection, or where the same (few) subjects are observed undertaking a variety of tasks (e.g. Chase & Ericsson, 1982; Gobet & Simon, 1996).

2. IDM A Proposed Solution to Some UTC Difficulties

A way to keep the multiple-constraint advantage offered by UTCs while making their development tractable is to do Individual Data Modelling (IDM). That is,

> *to gather a large number of empirical/experimental observations on a single subject (or a few subjects analysed individually) using a variety of tasks that exercise multiple abilities (e.g. perception, memory, problem solving), and then to use these data to develop a detailed computational model of the subject that is able to learn while performing the tasks. The model will be checked and refined as new data is added.*

The combination of both UTCs and using individual subject's data is the key, as we shall argue below. This approach solves the problem of variability between subjects, strategies, and tasks,

so detrimental if one wants to use the full constraints offered by the data. After elaborating on the main advantage offered by this approach, we discuss several potential difficulties of our approach, and show that none of them is critical.

2.1. Rapid Interaction between Theory Development and Data Collection

The main advantage offered by IDM is to allow hypotheses to be generated and tested in a rapid cycle, and therefore to rapidly improve the theory. This methodology also allows one to collect converging evidence, and to rapidly test and retest theoretical mechanisms and parameters. The use of previous values immediately informs and constrains the theory that is being developed. Because data constraints become usable, they become real.

Interacting with the model in such a way will suggest new experiments or new variations of old experiments. The experiments may then been carried out rapidly (there is only a single subject to arrange and run!), and the data gathered may be compared swiftly to the theory. This rapid collection of data, which is similar to other fields of research such as biochemistry, allows both rapid feedback and a close interaction between theory building and data collection. This pace is in stark contrast with the relatively slow collection of data typical in psychology.

2.2. Practicality of IDM

The first obvious objection to the proposed approach is that it is not implementable in practice. Gathering a large variety of data is difficult, and having a subject participate in multiple experiments is costly.

Can we hope to gather the amount and variety of data necessary? We believe it to be possible. Many experiments on perception, memory, and problem solving are now computer-based. Computer-based display of experiments speeds up acquisition of data and allows collection of detailed data (e.g. reaction times). It also makes it possible to interface the UTC model with the software used to run experiments (Ritter, Baxter, Jones, & Young, in press). As the theory may guide the selection of further experiments, the subject(s) should be kept available so that they can participate in new experiments. We contend that the financial cost of 'reusing' a subject is less than running dozens of subjects on various experiments. A good example that this approach is practicable is the digit-span research (Chase & Ericsson, 1982; Staszewski, 1990).

2.3. Controlling for Strategies

In most current theories, strategies are essentially free parameters, and as such impede the effective use of multiple constraints. It is one of the strongest features of our approach that it offers a solution to this problem. In addition, it makes it possible to capture strategy changes within a task by a single subject (Delaney et al., 1998). This is possible for two reasons. The first is that our approach encourages the collection of detailed and varied data (e.g. eye movements, verbal protocols, reaction times, and so on)—allowing cross-validation between data types. The second is that a simulation model provides fine-grained predictions. Experiments can also be carried out where strategies are systematically and specifically varied (Gobet et al., 1996; Medin & Smith, 1981). In the long term, this set of converging evidence will constrain strategies into fixed parameters for each subject, therefore reducing the theoretical degrees of freedom.

2.4. Controlling for Learning and Other 'Confounds'

A legitimate worry is that subject learning during the sessions will corrupt the data of later sessions. But here, because current UTCs include mechanisms for learning, the effect of learning is no longer a confound. Learning can be the object of theoretical investigation and simulation, for example, by analysing how the estimated parameters are affected by practice and by simulating the results of learning.

Another advantage of our approach, in particular its emphasis on coupling the simulation model to the experimental apparatus, is that experiments can be modelled in detail—the model can have access to the same situations the subjects saw. Interestingly, this includes experiments that would not be seen as 'perfect' from a traditional methodological point of view. Consider the example of an experiment where the stimuli order presented to a subject has not been properly randomised. This infelicity in the design does not matter within our approach, because the model can be presented with the same stimuli order as the human subjects, and we may actually predict, or postdict, what is the effect of the order of presentation confound (e.g. Ritter & Bibby, 1997).

2.5. Requirements for Software Development

An important aspect of our approach is that the (computational) theory must be regularly tested against the previously modelled empirical data to make sure that any change in theory (both in mechanisms and parameters) does not vitiate previous matches. This was one of the ideas behind Newell's program, but it has not been carried out systematically within the Soar community. This regular testing of previous simulations, that the ACT-R and EPAM communities have now started to implement, is not a trivial task. Various technical problems (changes in the programming language, in the hardware, and so on) conspire against it. Two actions seem imperative: (a) to develop programs to test simulations in batch; and (b) to interface the task simulations and the cognitive UTC model in a way that is robust against modifications of the model. Ideally, the task simulations should be reusable for another theory. Finally, the search through the space of possible theories should be supported through the use of optimisation techniques (e.g. genetic algorithms) to search parameter sets to better fit the data (e.g. Ritter, 1991).

2.6. Averaging over Theoretical Parameters

The methodology proposed here is not antithetical to group summaries or aggregates. However, instead of computing aggregate values using observed data (such as reaction times, errors, etc.), we propose first to estimate UTC-parameters for each subject, and then to compute aggregate values over these parameters. We believe that this 'between-subject analysis of parameters' offers a method for estimating aggregate values that is more robust and theoretically more meaningful than the traditional way of aggregating data (see Table 1).

As noted above, UTC-parameters need not be necessarily numeric, and can represent various types of knowledge, such as strategies. In this case, summaries may take the form of probability distributions over possible strategies. If one wishes to extract one single value from the distribution, one may, for example, take the parameters that occur most often for a given goal (modal strategy). Clearly, for some parameters, taking a summary value could be meaningless. For example, subjects may use, for the same goal, strategies so idiosyncratic as to make overlap between subjects non-existent, and averaging strategies does not make sense.

Table 1. Two ways of summarising data. On the left, the traditional approach, where observables are summarised across subjects. On the right, the IDM approach, where theoretical parameters are first estimated for each subject using observables, and only then summarised across subjects.

Subjects	Subjects' Observables (over tasks)			Estimated UTC Parameters
S_1	$o_{11}, o_{12}, o_{13} \ldots o_{1t}$	$\rightarrow$ IDM $\rightarrow$		$p_{11}, p_{12}, p_{13} \cdots p_{1n}$
S_2	$o_{21}, o_{22}, o_{23} \ldots o_{2t}$	$\rightarrow$ IDM $\rightarrow$		$p_{21}, p_{22}, p_{23} \cdots p_{2n}$
S_m	$o_{m1}, o_{m2}, o_{m3} \ldots o_{mt}$	$\rightarrow$ IDM $\rightarrow$		$p_{m1}, p_{m2}, p_{m3} \cdots p_{mn}$
Summary values	$O_1, O_2, O_3 \ldots O_t$			$P_1, P_2, P_3, \ldots P_n$

2.7. Difficulties

In spite of these advantages, or perhaps because of them, there are a few difficulties that this approach may face. First, it can be argued that the experiments that can be used both by subjects and by the model currently represent only a subset of human activities, typically activities that can be hosted by a computer. This is not an important problem, we think, because this subset contains quite a large number of behaviours; in addition, the advances in virtual reality and robotics continue to extend the range of activities that can be simulated.

Second, the theory may be overfit to a single subject (or to a few subjects), and thus not be generalisable (Spada & Plötzner, 1994). This is probably true to some extent (in particular the knowledge and the strategies used by the subject), but we believe that the fundamental parameters constraining cognition do not vary to the point that they make generalisation (within certain bounds) impossible. On the other hand, if the results of fitting a single subject are not generalisable across subjects, creating models of average data is futile for these models do not correspond to natural phenomena. The within-subject danger of overfitting may be alleviated by the many experiments done with each subject. In addition, and perhaps most tellingly, the resulting theory can simply be tested with other subjects taken individually. Our approach should make it clear which experiments will yield the most information with respect to the key theoretical parameters, and therefore allow the development of an 'optimal' subset of tests. This reduction in the number of experiments will make it easier to test further subjects, and therefore show whether and how parameters vary across a given population. This, of course, is a more powerful way of studying cognition than to limit oneself to average values of observables (see Table 1).

Third, another concern is that the data may be difficult to analyse because of their density. This is certainly a cost, but it should be born in mind that this density carries more information than a low density of data, and represents a savings in time running subjects. Current technology (e.g. SPA: Ritter & Larkin, 1994) facilitates this analysis to a certain extent.

Finally, is also necessary to consider the difficulty of how to estimate the fit of the model to the data (a classical problem in computer simulation). A practical approach is to use a convergence of measures of fit, such as the amount of variance accounted for or the mean squared error explained.

3. Conclusion

The approach we have outlined in this paper contrasts with the traditional hypothesis-testing, Popperian bent of psychology, by emphasising theory building and refining, almost taking an engineering stance. It recognises the great insights historically gained by IDA—even though these analyses are criticised by the dominant, statistics-driven experimental tradition in psychology. It also recognises the power of using computers to help build psychological theories, as exemplified in Anderson's ACT-R and Newell's Soar. It aims at taking the best of these two methodologies—IDA and UTC's computer modelling—and combining them to create IDM.

References

Aasman, J., & Michon, J. A. (1992). Multitasking in driving. In J. A. Michon & A. Akyürek (Eds.), *Soar: A cognitive architecture in perspective (p. 169-198).* Dordrecht (NL): Kluwer.

Altmann, E. M., & John, B. E. (1999). Episodic indexing: A model of memory for attention events. *Cognitive Science, 23*(2), 117-156.

Anderson, J. R., & Lebière, C. (1998). *The atomic components of thought*. Mahwah, NJ: LEA.

Bryan, W. L., & Harter, N. (1899). Studies on the telegraphic language. The acquisition of a hierarchy of habits. *Psychological Review, 6*, 345-375.

Card, S., Moran, T., and Newell, A., (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: LEA.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. *The Psychology of Learning and Motivation,* 16, p. 1-58.
Cooper, R., & Shallice, T. (1995). Soar and the case for unified theories of cognition. *Cognition, 55,* 115-149.
de Groot, A. D. (1978). *Thought and choice in chess.* The Hague: Mouton Publishers. (Original in Dutch, 1946).
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. *Psych. Sci., 9,* 1-8.
Dukes, N. F. (1968). N=1. *Psychological Bulletin, 64,* 74-79.
Ebbinghaus, H. (1964/1885). *Memory, a contribution to experimental psychology*. New York: Dover. (Original in German, 1885).
Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology, 31,* 1-40.
Gobet, F., Richman, H., Staszewski, J., & Simon, H. A. (1997). Goals, representations, and strategies in a concept attainment task: The EPAM model. *The Psychology of Learning and Motivation, 37,* 265-290.
Howes, A. & Young, R. M. (1996). Learning consistent, interactive, and meaningful task-action mappings: A computational model. *Cognitive Science, 20,* 301-356.
Lewis, R. L. (1996). Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of Psycholinguistic Research, 25,* 93-115.
Medin, D. L. & Smith, E. E. (1981). Strategies and classification learning. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 241-253.
Miwa, K., & Simon, H. A. (1993). Production system modeling to represent individual differences: Tradeoff between simplicity and accuracy in simulation of behavior. In *Prospects for artificial intelligence: Proceedings of AISB'93* (p. 158-167). Amsterdam: ISO Press.
Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
Norman, D. A. (1991). Approaches to the study of intelligence. *Artificial Intelligence, 47,* 327-346.
Polk, T. A. & Newell, A. (1995). Deduction as verbal reasoning. *Psych. Rev., 102,* 533-566.
Ritter, F. E. (1991). Towards fair comparisons of connectionist algorithms through automatically generated parameter sets. In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society* (p. 877-881). Hillsdale, NJ: LEA.
Ritter, F. E., Baxter, G. D., Jones, G., & Young, R. M. (in press). Supporting cognitive models as users. *ACM Transactions on Computer-Human Interaction.*
Ritter, F. E., & Bibby, P. A. (1997). Modelling learning as it happens in a diagramatic reasoning task (Tech. Report No. 45). ESRC CREDIT, Psychology, U. of Nottingham.
Ritter, F. E., & Larkin, J. H. (1994). Using process models to summarize sequences of human actions. *Human-Computer Interaction, 9,* 345-383.
Seibel, R. (1963). Discrimination reaction time for a 1,023-alternative task. *Journal of Experimental Psychology, 66,* 215-226.
Spada, H., & Plötzner, R. (1994). Multiple mental representations of information. In R. Lewis & P. Mendelsohn (Eds.), *Lessons from learning* (p. 1-11). North-Holland: Elsevier.
Young, R. M., Ritter, F. E., & Gobet, F. (in preparation). *Soar-ly pressed: Revisiting Cooper & Shallice on Soar and Unified Theories of Cognition.*