At their best, comparative methods can also reveal commonalities between human and animal perception that had escaped earlier analyses. We think of the human's ability at one-trial learning as a uniquely human achievement, based on human evolutionary innovations such as episodic memory, verbal mediation, and new forms of motivation. But Burton Slotnick's chapter, "Olfactory Perception," reveals that the humble laboratory rat can also perform one-trial learning—if the task is an olfactory discrimination and the rat is accustomed to using this modality in discrimination tasks. Mistakes are made only if the positive and the negative stimuli have an odor component in common, making the discrimination more difficult, or if the rat has had experience only in discriminations based on other sensory modalities. Discrimination had seemed so laborious for the rat because we were usually asking it to use vision or audition, senses that are not as important to a rat as they are to a human. In short, the tasks lacked ecological validity from the rat's standpoint.

The book's other chapters apply the principles outlined in these examples to such varied and fascinating topics as ecolocation in dolphins (Whitlow Au), auditory and visual control of saccadic eye movements (Martha Jay and David Sparks), monkey models of human amblyopia (Ronald Boothe), and perception of musical patterns by human infants (Sandra Trehub). The reader will come away with a heightened appreciation for the variety and power of new approaches in comparative perception.

Bruce Bridgeman
Program in Experimental Psychology, Clark Kerr Hall
University of California, Santa Cruz
Santa Cruz, CA 95064

References

Artim, J., & Bridgeman, B. (1989). The physiology of attention: Participation of cat striate cortex in behavioral choice. *Psychological Research, 50,* 223–228.

Bridgeman, B. (1972). Visual receptive fields in single cells of monkey visual cortex during visual tracking. *Science, 178,* 1106–1108.

Bridgeman, B. (1980). Temporal response characteristics of cells in monkey striate cortex measured with metacontrast masking and brightness discrimination. *Brain Research, 196,* 347–364.

Tyrrell, R., & Owens, D. A. (1988). A rapid technique to assess the resting states of the eyes and other threshold phenomena: The modified binary search (MOBS). *Behavior Research Methods, Instruments, & Computers, 20,* 137–141.

The Adaptive Character of Thought

By John R. Anderson. Hillsdale, NJ: Erlbaum, 1990. 276 pp. Text ed., $45.00.

The rational choice approach, long dominant in economics, has invaded political science, law, sociology, and other social sciences in recent years (Hirshleifer, 1985). The approach has won some influential new converts but, not surprisingly, many established scholars have fought against it. Cognitive psychologists, most notably Amos Tversky and Daniel Kahneman (1983), have provided the most effective ammunition for the resistance. Recently, the 1990 meetings of the American Political Science Association in San Francisco featured two well-attended panel discussions on the extent to which the cognitive data undermine the assumption of rationality.

With his new book, Anderson opens a new front in this struggle for the hearts and minds of social scientists. A respected and productive cognitive psychologist, Anderson states that his book "develops what I hope is both a clear and strong case for a rational analysis of human cognition" (p. xi). Anderson recognizes that advocating the rational approach in the home base of its opponents creates the "danger . . . that I might be expelled from the intellectual school that treated me so well. I have to confess that there were times when I ran away from ideas that seemed too threatening. . . . But I was always driven back to the rational analysis, because it was just too good of an idea to be abandoned" (p. xi). Anderson's almost Biblical conversion (reminiscent of the stories of Balaam or of St. Paul) is followed by a strong defense of his monotheistic new faith despite accusations of "apostasy" from polytheistic colleagues.

In this review, we (a mainstream but agnostic economist and a cognitive psychologist who has for some time recognized the descriptive power of some normative models) will try to sort out Anderson's potential impact on the rational choice debate in social sciences and the value of Anderson's specific rational choice models for cognitive psychology.

Thesis

In his first and last chapters, Anderson distinguishes the rational approach from the "architectural" (or process) approach. The architectural approach regards the mind as an "arbitrary configuration of . . . mechanisms" (or internal processes) to be inferred from observed behavior. Although this approach has dominated cognitive science for decades,

> it is just not possible to use behavioral data to develop a theory . . . in the concrete and specific terms to which we have aspired. . . . [T]here is an infinite number of mechanisms that compute the same input-output functions . . . and, consequently, identifying the behavioral function will not identify the mechanism. So, behavioral data will never tell us what is in the mind at the implementation level. It is time we stopped fooling ourselves. (p. 24)

The rational approach assumes that cognition is "optimized to its environment by evolution" (p. 26) and that therefore behavior is predictable on the basis of primarily environmental considerations. The identifiability problem disappears because internal processes are relevant in the rational approach only to the extent that they suggest constraints such as general sorts of computational limitations. (In particular, the internal processes need not literally follow the complex mathematics of Bayes's theorem, etc., p. 251.) Implementing the rational approach requires only the specification of evolutionarily relevant goals and appropriate constraints; the assumption of rationality then predicts that observed behavior will be optimal. Thus the rational approach promises both parsimony and precision.

Metatheoretical issues

Learning about what goes on inside the head is not easy, because there are too many possible process models of behavior. Cognitive psychologists have usually focused on the processes inside the head rather than outside. Anderson has been no exception in this regard. The turnaround in this book for Anderson is that now he seeks to understand human cognition by "considering in detail what is outside the human head and try to determine what would be optimal behavior given the structure of the environment and the goals of the human" (p. 3). This endeavor appears to be worthwhile enough, although we hope that Anderson did not miss the fact that the goals of the human, even though they are biologically viable, are probably as much inside the head as outside the head.

Given this new metatheoretical approach, it is only reasonable that Anderson would try to articulate his framework. He begins with discussing the levels of a cognitive theory and, as is persistent in the literature, he begins with David Marr's three levels. He attributes to Marr an unstated adaptationist principle. After discussing Chomsky's well-known characterization of competence and performance, Anderson considers Pylyshyn's distinction between algorithm and functional architecture. Anderson also acknowledges the connectionist movement well underway in psychological inquiry. He criticizes the connectionists primarily in terms of their claim to neurological modeling. Anderson's point, already in the literature, is that neuroconstraints are sterile at this stage of the game. He describes Newell's knowledge level, a level that corresponds perhaps to Marr's computational level, to illustrate the fact that much of human behavior can be predicted without knowing how the behavior occurred.

Given this background, Anderson claims that there are only two levels of analysis that are psychologically real: the algorithm level and the biological level. It is not clear what Anderson means by algorithm level. For Pylyshyn, it is a flexible level penetrable by the perceiver. It seems that Anderson should really be saying that the two levels are the biological and the computational in Marr's terminology or the knowledge level in Newell's terminology. He equates the implementation level with what seems to be the algorithm level (the functional architecture in Pylyshyn's framework). In any event, he claims that his algorithm level overcomes problems of identifiability that persist at the implementation level. He adds that the physical symbol hypothesis is the appropriate language to use at the algorithm level, arguing that symbols are the only way to achieve intelligence. Before evaluating Anderson's thesis, we describe his application of the rational framework to psychological inquiry.

Memory

Anderson applies an adaptive (rational) analysis to four areas of inquiry: memory, categorization, causal inference, and problem solving. We limit our discussion to just the first two areas because these areas include the most demanding mathematical analyses and clearcut attempts at predictions based on a rational analysis. Anderson bases both his memory analysis and his analysis of categorization on well-formulated developments in machine learning. This is not the first time that psychologists have sought viable models in the domain of artificial intelligence. However, investigators seldom, if ever, justified their use of these models in terms of rationality. The models were attractive because they were computationally adequate and did not appear to violate any known psychological principles such as neural plausibility.

For Anderson, the goal of memory is to get access to needed information from the past. Thus Anderson's currency is need probability, which replaces more traditional currencies such as memory strength, activation, or association. Our need for information is constrained primarily by the structure of the environment and, therefore, the properties of the environment are the dominant influence on what information is needed. Internal processes are considered only in the assumption that there is some cost in retrieving information from memory. (Anderson does not appear to address the cost associated with storage and retention.) At any moment, there is a need probability $P(A)$ associated with each memory A. There is a cost C for an attempted retrieval from memory, and there is a gain G for successful retrieval. The unobservable parameters C and G are assumed constant across different situations. Anderson postulates that a rationally designed system would order the memories by their probabilities $P(A)$, initiate a serial search, and stop retrieving when $P(A) < C/G$. Memories for which the expected gain is less than the cost will not be retrieved.

Given this framework, it is necessary to estimate the need probability. To do so, Anderson assumes two sources of information: history independent of context, and current contextual cues. He assumes the sources are correctly combined according to the rules of probability theory (i.e., using Bayes's theorem). Although rationality dictates a unique way of combining the sources of information, it allows greater latitude in specifying each source. For the history source, he adapts a model by Burrell previously applied to library and computer-file retrieval systems. The Burrell model assumes that the objects to be retrieved (memory traces for Anderson) have a probability distribution whose parameters are estimated from historical data. Specifically, the distribution of retrieval requests for an object is assumed to be exponential, and the conjugate prior on the exponential parameter λ is assumed to have the gamma distribution with conjugate parameters $v + n$ and $b + t$. History (the number of previous retrieval requests n during the time interval t) affects the expected need probability because $(v + n)/(b + t)$ is the conditional expected value of the need probability.

Later Anderson uses an augmented version of the Burrell model which weights recent time more heavily. Anderson estimates the other source of information, cue strength Q, by breaking down the context into component cues and combining these using a Bayesian analysis. The multiplicity of cues are assumed to be independent of one another.

Anderson should be applauded for attempting a rational analysis of memory. We offer two specific criticisms of the present attempt, which we hope will spur the evolution of this approach. First, Anderson wants power function predictions for delay and practice effects (Figure 2-1 to 2-4). He obtains approximate power functions only by carefully tuning the augmented Burrell model. Thus the desired predictions arise primarily from auxiliary assumptions, and not from the rationality assumption or even from the basic Burrell model (which predicts exponential functions rather than power functions). Second, Anderson derives an optimal deterministic model and then adds noise to it without considering what this means for the optimality analysis. In Anderson's model, memory retrieval fails if the need probability is below threshold. Thus, behavior should be deterministic. Given that we seldom, if ever, see deterministic behavior, Anderson tacks on variance, appealing to the presence of noise. The problem is that he does not question what this assumption means for the optimality of the model, which was predicated on a deterministic environment and noiseless processing capabilities. There is no presumption that the model remains optimal (or even approximately optimal) once noise of this sort is introduced. And most of the striking predictions of the model rely heavily on the noise parameter s (or B) and other auxiliary assumptions such as Zipf's law, which are not derived from optimizing behavior.

Categorization

For Anderson, the goal of categorization is to predict the features of objects. In this regard, it is surprising that Anderson did not articulate the similar functions of memory and categorization. Is that creature out there dangerous? Is this a memory problem or a categorization problem? It is a pattern recognition or inference problem—which also ties in the analysis of causal inference and problem solving. Again, exploiting the work in machine learning, Anderson derives a unique treatment of the categorization literature. Most models of categorization assume that the number of categories being considered by the subject are given by the experimenter and are fixed throughout the experiment. Anderson, on the other hand, assumes that the system partitions the class of objects into an open-ended number of categories. The optimal partition is not computationally feasible, unfortunately, and Anderson utilizes an algorithm developed in machine learning by Fisher and Leibowitz.

The machine learning algorithm proceeds roughly as follows. The m previous objects have been categorized into $k < m$ categories. The algorithm looks at the $(m + 1)$st object and calculates probabilities that the object belongs to each of the k current categories and also a probability that it belongs to a new category. The object is assigned to the most probable category, old or new. The key ingredients in the probability calculation are a free parameter called the coupling probability, and conditional probabilities given by his Equations 3-6 and 3-7. It can be shown that, aside from a small-sample adjustment, these equations are equivalent to the special case of Medin's exemplar model with the similarity parameter equal to zero.

Anderson's treatment of categorization suffers from some of the same shortcomings as his treatment of memory. A case was made for optimal categorization, but in the end he settled for a nonoptimal algorithm, in this case because of computational difficulties. The algorithm locks into the sequence of objects being presented, and cannot change its partitioning once it is established. Such rigidity hardly seems adaptive—the system should stay as flexible as possible to survive in natural environments. Moreover, Anderson's model relies heavily on the notion of *coupling probability*, which is defined as the prior probability that two objects come from the same category. This probability is a free parameter with no rational underpinnings, yet much of the model's predictive power seems to come from choosing this parameter.

Evaluation

Puzzling through the nuts and bolts of Anderson's implementations reveals to us that rational analysis does not provide an easy solution to the identifiability problem confronting experimental psychologists. Making assumptions about computational limitations reintroduces issues of identifiability and psychological reality even in an optimal system (Massaro & Friedman, in press). Anderson promises us that a rational level analysis will illuminate psychological mechanisms. Additionally, these psychological mechanisms are going to be informed by what is outside the head rather than what is inside the head. Anderson does not live up to this claim, however, because in his analysis, what is inside the head plays a larger role than what is outside in developing predictions of optimal behavior in different situations. He notes throughout that he has been consistently plagued by the profound lack of identifiability in psychological theorizing. By identifiability, there are several ways to skin a cat or several different models that can predict the same performance. Anderson believes that he can determine whether behavior is optimal without being concerned with psychological mechanisms. However, Anderson depends on what is inside the head in terms of defining the constraints on human performance such as short-term memory limitations or serial processing limitations. These are clearly statements about psychological mechanisms, and we do not see how they avoid the problem of identifiability. One cannot really argue with Anderson's disillusionment with the identifiability issue. On the other hand, one can challenge whether his rational analysis overcomes problems of identifiability.

It appears that Anderson has not been completely converted because he wants to play in both leagues (p. 31). He hopes to follow up the analysis in this book with a new theory within the adaptive control of thought (ACT*) framework. Thus, although his goal is still a psychological theory, he wants to emerge from the scientific endeavor with an answer of what is happening, in addition to why it is happening because of evolutionary history and biological constraints.

Anderson's work is reminiscent of Herbert Simon's because the implementation of his constraints in determining whether or not behavior is optimal seems highly analogous to Simon's idea of bounded rationality. In both cases the investigators are interpreting apparently irrational or nonadaptive behavior as locally rational. Anderson ends by proposing three new practices for cognitive psychology: to develop rational theories of various cognitive phenomena, to study the actual structure of the environment, and to construct psychological models guided by the rational analysis. Taking a strong inference point of view, however, one would also want to develop nonrational models and ask if these can be in fact falsified. Within any rational model, auxiliary assumptions seem necessary to achieve any sort of precision, but they come at the expense of parsimony. A second generation of rational models for memory and categorization could improve the trade-off Anderson has made, but in the end we believe that Anderson has promised more for rational analysis than it can deliver.

Some comparative and historical perspective may help in evaluating the role rational analysis can play. At least from Aristotle's time (and probably from the beginning of our species) people have sought theoretical explanations based on tangible processes, and at least initially have strongly resisted more austere approaches. The point could be made in many fields including religious cosmology, but perhaps it is safe to take examples from physics. Newton's contemporaries found his theory of gravitation very mysterious: *What process* could produce force at a distance as called for in his inverse square law? Several now-discredited concepts (e.g., ether and phlogiston) were invented to provide sorely missed process explanations (of electromagnetic wave transmission and of combustion, respectively). Thermodynamics provides a cogent example. Clausius, Kelvin, and other mid-nineteenth century scientists were able to organize a large body of data on the basis of two (later three) general optimization and equilibrium principles that to this day suffice for most bread-and-butter applications. A later generation (Gibbs, Boltzmann, and others) explained the general principles of classical thermodynamics in terms of the underlying molecular-level processes. The process theory, statistical mechanics, does not replace classical thermodynamics but rather provides rigorous but intuitive foundations and extends its applicability. Similarly, more than 250 years after Newton's discoveries, Einstein's general relativity theory finally provided a process model (if one can use that term for non-Euclidean space-time) underlying Newtonian mechanics.

Turning to social sciences, we view the present state of economic theory as similar to that of classical thermodynamics or Newtonian mechanics. The optimization principle of rational choice together with mutual consistency principles (such as market-clearing or Nash equilibrium) provide austere but empirically very useful explanations of large bodies of economic data. However, anomalies remain and the standard theoretical models are indeterminate in some cases (e.g., multiple Nash equilibria). As yet no process model has come along to undergird standard theory and extend its applicability.

The present state of cognitive psychology is essentially the reverse: There are many process models but no unifying principle. In our view, Anderson's main contribution to his discipline is to remind us of the success that constrained optimization has had in organizing data in other fields and to show that its techniques can be applied to understanding how the mind works.

For that reason we recommend that cognitive psychologists read Anderson's new book. He claims too much for the power and parsimony of rational models, and his specific models probably will not stand the test of time. Nevertheless, Anderson has made an important methodological contribution by forcefully bringing the rational approach to the attention of cognitive psychologists. We do not expect to see any wholesale conversions (Anderson in the end is no St. Paul), but we do expect psychologists to use rational models more frequently in the future. In our view, social science will be stronger when process models are consistent with adaptive rationality, and when rational choice models have consistent process model foundation. We salute Anderson for moving us a step closer to that goal.

Dominic W. Massaro
Program in Experimental Psychology, Clark Kerr Hall

Daniel Friedman
Department of Economics
University of California, Santa Cruz
Santa Cruz, CA 95064

internet: massaro@fuzzy.ucsc.edu
bitnet: massaro@ucscc

References

Hirshleifer, J. (1985). The expanding domain of economics. *American Economic Review, 75*(6), 53–68.

Massaro, D. W., & Friedman, D. (in press). Adaptive rationality and identifiability of psychological processes. *Behavioral and Brain Sciences.*

Tversky, A., & Kahneman, D. (1983). Extension versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90,* 293–315.