Psychological Bulletin
1996, Vol. 119, No. 1, 3-22

0033-2909/96/$3.00

The Empirical Case for Two Systems of Reasoning

Steven A. Sloman
Brown University

Distinctions have been proposed between systems of reasoning for centuries. This article distills properties shared by many of these distinctions and characterizes the resulting systems in light of recent findings and theoretical developments. One system is associative because its computations reflect similarity structure and relations of temporal contiguity. The other is "rule based" because it operates on symbolic structures that have logical content and variables and because its computations have the properties that are normally assigned to rules. The systems serve complementary functions and can simultaneously generate different solutions to a reasoning problem. The rule-based system can suppress the associative system but not completely inhibit it. The article reviews evidence in favor of the distinction and its characterization.

One of the oldest conundrums in psychology is whether people are best conceived as parallel processors of information who operate along diffuse associative links or as analysts who operate by deliberate and sequential manipulation of internal representations. Are inferences drawn through a network of learned associative pathways or through application of a kind of "psycho-logic" that manipulates symbolic tokens in a rule-governed way? The debate has raged (again) in cognitive psychology for almost a decade now. It has pitted those who prefer models of mental phenomena to be built out of networks of associative devices that pass activation around in parallel and distributed form (the way brains probably function) against those who prefer models built out of formal languages in which symbols are composed into sentences that are processed sequentially (the way computers function).

An obvious solution to the conundrum is to conceive of the mind both ways—to argue that the mind has dual aspects, one of which conforms to the associationistic view and one of which conforms to the analytic, sequential view. Such a dichotomy has its appeal: Associative thought *feels* like it arises from a different cognitive mechanism than does deliberate, analytical reasoning. Sometimes conclusions simply appear at some level of awareness, as if the mind goes off, does some work, and then comes back with a result, and sometimes coming to a conclusion requires doing the work oneself, making an effort to construct a chain of reasoning. Given an arithmetic problem, such as figuring out change at the cash register, sometimes the answer springs to mind associatively, and sometimes a person has to do mental arithmetic by analyzing the amounts involved and operating on the resultant components as taught to do in school. This distinction has not been missed by philosophers or psychologists; it can be traced back to Aristotle and has been discussed, for example, by James (1890/1950), Piaget (1926), Vygotsky (1934/1987), Neisser (1963), and Johnson-Laird (1983), amongst others mentioned later.

However, the distinction is not a panacea; it is problematic for a couple of reasons. First, characterizing the two systems involved in a precise, empirically consequential way raises a host of problems. Distinctions that have been offered are not, in general, consistent with each other. Second, characterizing the systems themselves is not enough; their mode of interaction must also be described. A psychologically plausible device that can integrate computations from associative networks and symbol-manipulating rules has proven elusive.

In this article, I review arguments and data relevant to the distinction and, in light of this evidence, provide an updated characterization of the properties of these two systems and their interaction. A fair amount of data exists concerning how people reason about various problem domains, but little research has directly attempted to dissociate the systems that underlie such reasoning. Some researchers have made this effort, and I review their results below. However, a larger part of the data that provide evidence for such a dissociation was not originally collected for that purpose. Much of my empirical discussion therefore involves a reinterpretation of data, some of which are fairly old and well known. To preview, several experiments can be interpreted as demonstrations that people can simultaneously believe two contradictory answers to the same reasoning problem—answers that have their source in the two different reasoning systems. Before proceeding, I try to clarify the kind of distinction that I am arguing for.

Two Forms of Computation

The most lucid expression of the distinction and its psychological reality is, not unexpectedly, that of William James (1890/1950). He described *associative thought* or "empirical thinking" as "trains of images suggested one by another" (p. 325). A person reasons this way when, for example, creating a design. Images, new and old, come to mind, providing ideas and standards of comparison. James believed that associative thought is "only reproductive," in that the objects of associative thought are all elements of or abstractions from past experience,

This work was funded by a grant from Brown University. I am indebted to Boaz Keysar for his insightful and helpful comments on a prior version of this article.

Correspondence concerning this article should be addressed to Steven A. Sloman, Department of Cognitive and Linguistic Sciences, Brown University, Box 1978, Providence, Rhode Island 02912. Electronic mail may be sent via Internet to sloman@cog.brown.edu.

but the data I review suggest otherwise. True reasoning is "productive" according to James, for it can deal with novel data: "Reasoning helps us out of unprecedented situations" (p. 330). In a strange city, for example, a person can generally find where he or she is going because he or she has the ability to reason about maps and systems of transportation.

Associative System

Today, one might describe James (1890/1950) as distinguishing between two systems that implement different computational principles. Roughly, one system is associative, and its computations reflect similarity and temporal structure; the other system is symbolic, and its computations reflect a rule structure.

The associative system encodes and processes statistical regularities of its environment, frequencies and correlations amongst the various features of the world. For example, a symmetric association between two features can be interpreted as a kind of correlation between those features. In some formal associative systems, an association from Feature A to Feature B can be interpreted as the conditional probability of B given A (e.g., Hinton & Sejnowski, 1986), and some can be shown to generate optimal statistical estimates (e.g., Jordan & Jacobs, 1994). Generally speaking, associative systems are able to divide perceptions into reasonable clusters on the basis of statistical (or at least quasi-statistical) regularities. They treat objects in similar ways to the extent that the objects are perceived as similar (e.g., J. A. Anderson, Gately, Penz, & Collins, 1990; Carpenter & Grossberg, 1987; Rumelhart & Zipser, 1985) because the degree to which an association is operative is proportional to the similarity between the current stimulus and previously associated stimuli. On this view (ignoring such considerations as aesthetics), associative thought uses temporal and similarity relations to draw inferences and make predictions that approximate those of a sophisticated statistician. Rather than trying to reason on the basis of an underlying causal or mechanical structure, it constructs estimates based on underlying statistical structure. Lacking highly predictive causal models, this is the preferred mode of analysis for many forecasters, such as weather and economic ones.

In summary, I claim that associative reasoning inherits a property of associative systems: It computes on the basis of similarity and temporal structure.[1] Evidence for this claim appears below. Also, because the study of similarity—the respects in which objects and events are common and distinctive—has a correspondence to the study of statistical structure—the study of variability and covariability—the associative system computes on the basis of information of the same kind as that used to draw statistical inferences.

My central claim is about the principles that govern computation in the two systems, not the details of the systems' processing. I have claimed that the associative system computes similarity and statistical structure, but rules are also able to compute such structure. Indeed, statistics textbooks are about computation using statistical rules. However, such rules are at a more specific level of description than my analysis. They describe detailed structural models. Whereas such detail is ultimately desirable, it requires models richer than current data on reasoning can support, other than for a few well-studied laboratory tasks. The concept of *association* permits analysis at a more abstract level of description, a level whose structure is closer to that of the empirical domain represented. Associations can exist between representations of elements in a domain. Rules in statistics textbooks are not about the domain itself but about statistical concepts and procedures. They do not encode, for example, that wings are associated with flight, rather they encode how to, for example, conceptualize or calculate a correlation coefficient. Associations capture structure not by indicating how to calculate it but by representing it directly. Similarity structure need not even be represented explicitly; it can be implicit in a set of associations.

Of course, rules of a different kind could be used to describe an empirical domain. A person could have a rule that states, for example, "If X has wings, then X can probably fly." Such a rule could be construed as an association. A system embodying rules of this sort may be computationally equivalent to a system of associative reasoning. Whether it is depends on other aspects of the system—the details of its computations—and not merely on its means of representation. The strength and the weakness of rules is their generality. They can represent any proposition. However, this very representational power makes an unconstrained notion of rule empirically vacuous. In the next section, I hone down my intended meaning for *rule-based reasoning.*

Taken alone, the notion of association is equally empty empirically. Any relation can be described as a complex association; thus, associative systems are also very general representational devices.[2] I therefore limit my use of the term *associative system in reasoning* to mean a cognitive system that draws inferences on the basis of similarity and contiguity. To illustrate, later I review evidence that sometimes people reason in a fashion inconsistent with a rule of probability but in agreement with their judgments of similarity. Reasoning performance is accurately predicted by judgments of similarity taken out of the problem context in the absence of any further assumptions about the knowledge that people bring to bear on the task. I attribute such behavior to associative reasoning. I motivate my attribution below by describing associative devices, types of connectionist systems, that some have argued can compute the kind of similarity structure to which people are sensitive. Again, the data I review do not allow a choice between detailed computational models. My descriptions serve merely as evidence that the kind of associative computation that I characterize arises naturally in some simple, well-studied systems.

Rule-Based System

The computational principles underlying rule-based reasoning are more opaque and more controversial than those of associative reasoning. One such principle, mentioned by James (1890/1950) and reasserted by Fodor and Pylyshyn (1988), is

---

[1] By associative, I do not mean to imply tabula rasa. Associative networks can come with complex constraints and predispositions. Relatedly, I am not claiming that similarity is not dependent on prior knowledge, biases, and current context. Goldstone (1994) showed how similarity maintains explanatory force despite such dependencies.

[2] Recurrent connectionist networks can be shown to be computationally powerful enough to exceed the processing power of a Turing machine (Siegelmann & Sontag, 1995).

productivity. Rule-based systems are productive in that they can encode an unbounded number of propositions (Chomsky, 1968); that is, rules can be composed with each other to generate an ever larger set of propositions. To see this, consider arithmetic in which a person can always generate a new number by adding 1 to the largest number in a set. A second principle is that rules are systematic, in the sense that their ability to encode certain facts implies an ability to encode others. For example, if one can reason about John loving Mary, one also has the capacity to reason about Mary loving John. Fodor and Pylyshyn argued that the productivity, systematicity, and hence compositionality of mental representations necessitate that human reasoning is generated by a language of thought that has a combinatorial syntax and semantics. My claim is that their argument is only relevant to one form of reasoning.

I call this form of reasoning *rule based* because rules are the form of representation that exhibit the properties of productivity and systematicity most transparently. Rules are abstractions that apply to any and all statements that have a certain well-specified, symbolic structure. Most important, they have both a logical structure and a set of variables. For instance, the conjunction rule of probability states that $Pr(A) \geq Pr(A \& B)$ where $Pr(A)$ means the probability of Event A, so the rule states that no two events can be more probable than either one alone. Elements of the logical structure include *Pr*, *&*, and $\geq$, all of which have a fixed role. The set of variables here is (A, B), which are arbitrary as long as they signify some sort of event; they can be bound to any event, so the rule can be applied to any pair of events. The rule is productive in that, given a new Event C, a person can infer that $Pr(A \& B) \geq Pr(A \& B \& C)$, and so on, for any number of other events. The rule is systematic in that a person could rewrite the rule as $Pr(B) \geq Pr(B \& A)$. A is not special, relative to B, in any sense relevant to Pr. The relation is purely formal or syntactic, in the sense that correct application of the rule is determined by relations amongst symbols and not by any sort of meaning that is attributed to the symbols.

Variables *vary;* that is, they can be instantiated in more than one way. Because they assume a class of possible values, they are necessarily abstract. My discussion concerns rules that contain variables, and therefore, rules that are abstract, they can be instantiated in more than one way. This does not imply that rules have to be content independent. For instance, Cheng and Holyoak (1985) discussed sets of rules for reasoning, which they called "pragmatic reasoning schema," that are associated with particular content domains. They suggested that certain rules are associated with reasoning about situations involving domains such as permission. An example of such a rule is, "If the action is to be taken, then the precondition must be satisfied." Such rules involve both variables (like *precondition* and *action,* which must be specified on each occasion of use) and logical structure (the form "If, then"), so therefore I count them as rules.

So far I have said two forms of reasoning exist and can be differentiated by the computational principles that they implement, one based on principles of similarity and contiguity and the other on rules, and I have tried to specify what rules are. I noted that if rules are the kind of general representational medium that I just described, then any representation could be expressed using rules, so a system of reasoning that does not use rules could not be empirically distinguished from one that does. Smith, Langston, and Nisbett (1992) pointed out that a distinction must be made between a system that *follows* rules from one that simply *conforms* to rules: "For rule following to occur, there must be a correspondence between the rule and a mental event" (p. 3). A computer follows rules that are written in a program, whereas a ball falling to the ground only conforms to the laws of physics. Following Smith et al., I limit the term *rule based* to systems that explicitly follow rules. Also following Smith et al., I limit myself to rules in *reasoning.* I exclude consideration of rules hypothesized to describe perception, motor control, language use, or the kind of linguistic competence studied by formal linguistics (cf. Smolensky, 1988) because these may all be special skills.

Rules come in different kinds. Some rules are instructions, such as statements in a computer program or a recipe; others are laws of nature or society or rules of logic. People are capable of following all of these sorts of rules (and of disobeying some). Rules can be normative, by telling how a person should behave to reach some prespecified goal (such as the conjunction rule if a person wants to maintain a coherent set of probabilities), or descriptive, by telling how a person does behave in certain contexts (such as the superstitious rule about which shoe to tie first). In contexts in which a normative rule obviously applies, it usually becomes descriptive as well. So, some rules are handed down through culture, others are made up, and some are discovered in nature or logic. Humans can understand and apply all of these rules without external support as long as they have become internalized, as long as their analytical machinery has access to and mastery of them.

Johnson-Laird and his colleagues (e.g., Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991) have argued that deductive reasoning is not a matter of applying either formal or content-specific rules of inference. Rather, deduction consists of applying "procedures that search for interpretations (or mental models) of the premises that are counterexamples to conclusions" (1991, p. 23). According to their account, deduction consists of three stages: (a) Language and general knowledge are used to comprehend a situation; this results in one or more mental models of the situation. (b) A parsimonious description of the models is generated; this description is a putative conclusion. (c) Alternative models are searched for in which the putative conclusion is false; this constitutes a search for counterexamples. These researchers have shown that this theory is compatible with a fair quantity of data from a variety of reasoning tasks. The mental model theory has the additional virtue that it captures strong intuitions about the representations that we use when we think about relations amongst categories and events.

Johnson-Laird has argued persuasively that much of everyday deduction is unlike theorem proving. He has posited a procedure to arrive at a determination of the logical relation between the premises and conclusion of an argument that has a fundamentally different rationale and design than the sequential application of inference rules advocated by theorists such as Braine (1990) and Rips (1994). Nevertheless, the mental models theory shares key assumptions with rule-based theories. Both approaches depend heavily on symbols. Like rules, mental models consist entirely of symbols. Some symbols are tokens, which refer to entities in the statements of an argument. Other symbols represent negation, and still others represent superset–subset relations.

One version of the mental model theory concerns reasoning with propositional connectives such as *and, or,* and *if, then* (Johnson-Laird, Byrne, & Schaeken, 1992). The authors themselves pointed out that the mental model notation they used is isomorphic to a standard form of logical notation called *disjunctive normal form.* To illustrate, the sentence "p or q"—which means that either p or q or both are true—would appear in the mental model theory as three models, each of which describes one possibility:

p not q

not p q

p q

This representation is equivalent to the standard logical expression "(p *and* [*not* q]) or ([*not* p] *and* q) or (p *and* q)." Writing it down in standard logical form makes it apparent that mental models of propositional reasoning enjoy the criterial properties that I have assigned to rules: They have both logical structure (exemplified by the key words *and, or,* and *not*), and they have variables (exemplified by p and q). The criterial properties of the rule-based system are thus sufficiently general to encompass central aspects of Johnson-Laird's mental model theory. Rips (1994), in a much fuller analysis of the relation between rules and mental models, concluded that "both types of theories are, at heart, methods for transforming configurations of (syntactically structured) symbols in ways that preserve truth" (p. 376).

I do not intend to try to resolve the rather contentious debate between mental model and formal rule theorists (see, e.g., Rips, 1986; Johnson-Laird & Byrne, 1993, and the accompanying commentaries) in favor of one or the other. Indeed, the two views are not incompatible. The heterogeneity of human thought suggests that people have multiple procedures to establish the validity of inferences. So, like all good dichotomies, the current one might eventually decompose into a trichotomy of associative, rule, and mental model systems. The data that I focus on, however, support only a single distinction. I stick to the associative versus rule-based terminology because it seems specific enough to be informative and general enough to capture the various forms of reasoning.

Discussion

Human reasoning seems to be performed by two systems, two algorithms that are designed to achieve different computational goals. One is associative, and it seems to operate reflexively. It draws inferences from a kind of statistical description of its environment by making use of the similarity between problem elements, interpreted (as seen below) using such aspects of general knowledge as images and stereotypes. The other is rule based and tries to describe the world by capturing different kinds of structure, structure that is logical, hierarchical, and causal-mechanical. Traditionally, only relations of this latter kind have proven able to support coherent *explanations*—in contrast to predictions—in science (Salmon, 1989). To the extent that such relations are also required for explanatory coherence in daily life (see Pennington & Hastie, 1992, for suggestive evidence), a rule-based structure provides a more compelling justification for a response than does an associative one (Brooks, Norman, & Allen, 1991; Rips, 1990).

Table 1 summarizes my characterization of the two systems. The point of this article is not that both systems are applied to every problem that a person confronts, or that each system has an exclusive problem domain. Rather, the forms have overlapping domains, domains that differ depending on the individual reasoner's knowledge, skill, and experience. Table 1 lists some functions that show off each system's capacities. The common mode of operation of the two systems however is clearly interactive. Together, they lend their different computational resources to the task at hand; they function as two experts who are working cooperatively to compute sensible answers. One system may be able to mimic the computation performed by the other, but only with effort and inefficiency and even then not necessarily reliably. The systems have different goals and are specialists at different kinds of problems. When a person is given a problem, however, both systems may try to solve it: Each may compute a response, and those responses may not agree. Below, I argue that cases can be found in every domain of reasoning that has been studied in detail in which they do not. Because the systems cannot be distinguished by the problem domains to which they apply, deciding which system is responsible for a given response is not always easy. It may not even be possible because both systems may contribute to a particular response.

One tentative rule of thumb to help identify the source of an inference has to do with the contents of awareness. When a response is produced solely by the associative system, a person is conscious only of the result of the computation, not the process. Consider an anagram such as *involnutray* for which the correct answer likely pops to mind associatively (involuntary). The result is accessible, but the process is not. In contrast, a person is aware of both the result and the process in a rule-based computation. Consider a harder anagram such as *uersoippv.* If you figured out the answer (purposive), and even if you did not, you likely applied various rules (like put different letters in the first position) which were consciously accessible. If you did get the answer, you were aware not only of the process but also of the result.

Awareness provides only a fallible heuristic for identifying systems, not a necessary or sufficient condition. The heuristic encounters two problems. First, both systems may contribute to a response. For example, Ross (1989) showed that the successful use of a reminding to aid learning of probability theory required the iterative application of (a) similarity-based retrieval and generalization and (b) rule-based inferencing for reconstruction and analogy. Introspection alone cannot be expected to tease apart these subtle mutual influences. Second, some reasoning is not obviously associative and yet apparently occurs without conscious awareness (Nisbett & Wilson, 1977). For example, mathematicians have reported having the solutions to difficult problems leap to mind, even though their thoughts were elsewhere (e.g., Hadamard, 1945). The nature of the reasoning that underlies this kind of creative insight is unknown. I can only speculate that it is not identical to that which underlies rule-based reasoning of the sort I describe. I can report more definitely that all the rules encountered below can be reported by those individuals who use them.

A Connectionist Proposal

To make the distinction more concrete, I describe an approach to reasoning that embodies the dichotomy just outlined.

Table 1
Characterization of Two Forms of Reasoning

Characteristic	Associative system	Rule-based system
Principles of operation	Similarity and contiguity	Symbol manipulation
Source of knowledge	Personal experience	Language, culture, and formal systems
Nature of representation		
Basic units	Concrete and generic concepts, images, stereotypes, and feature sets	Concrete, generic, and abstract concepts; abstracted features; compositional symbols
Relations	(a) Associations (b) Soft constraints	(a) Causal, logical, and hierarchical (b) Hard constraints
Nature of processing	(a) Reproductive but capable of similarity-based generalization (b) Overall feature computation and constraint satisfaction (c) Automatic	(a) Productive and systematic (b) Abstraction of relevant features (c) Strategic
Illustrative cognitive functions	Intuition Fantasy Creativity Imagination Visual recognition Associative memory	Deliberation Explanation Formal analysis Verification Ascription of purpose Strategic memory

Neisser argued in 1963 that distinctions like those between intuitive and rational thought, primary and secondary process, autistic and realistic thinking, and the like are best thought of as alternative modes of organizing computer programs for pattern recognition. Computers can be programmed to do multiple processing, to examine many different properties of a pattern simultaneously (parallel processing), or to examine properties and make decisions sequentially (conventional computer programming). Smolensky (1988) made a proposal that, effectively, brought Neisser's distinction up-to-date. Smolensky argued that humans make inferences using one of two mechanisms, a conscious rule interpreter or an intuitive processor. The conscious rule interpreter algorithmically processes knowledge expressed in the form of rules that could be completely described using a traditional symbolic language. Specifically, it processes knowledge in the form of production rules. These are instructions that have the general form "*If* a condition is satisfied, *then* perform some action" and have been used as a representational medium by a variety of cognitive theorists (e.g., J. R. Anderson, 1993; Holland, Holyoak, Nisbett, & Thagard, 1986; Newell, 1990). According to Smolensky, only a subset of an individual's knowledge base makes contact with this interpreter: knowledge that is publically accessible, that can be reliably verified, and that is formal and symbolic. In short, rules represent and operate sequentially on knowledge that exists as part of a cultural community.

To see this, consider a prototypical case of rule application: writing down a proof of a mathematical theorem. Such a proof is written down using knowledge of an abstract system, a mathematical theory. Statements licensed by the system enjoy a certainty that goes beyond any individual's own authority; they are justified by the mathematical community's acceptance of the system. However, because the system is shared knowledge, its rules must be expressed in terms that are communicable, terms for which there is common reference. When an individual is reasoning, the common reference is to the task being performed. The rules of any mathematical theory refer to elements of the statements that are written down, just as the rules of chess refer to the elements that constitute the game of chess. In other words, rules refer to objects and events that are at the same level of abstraction as the concepts of the task itself. For this reason, Smolensky (1988) claimed that the kind of knowledge on which rules operate is fully describable at the conceptual level of analysis.

The intuitive processor is implemented in the same hardware as the rule interpreter, but the types of knowledge on which they operate are different.[3] The intuitive processor can only be fully analyzed at the subconceptual level, a level of knowledge representation more molecular than concepts that are symbolically represented. Representations at the subconceptual level are distributed in the sense that concepts are represented by one or more patterns, each of which includes many features, and each feature participates in many patterns. The advantages of subconceptual representations are threefold. First, they allow more information to be coded in a single representation. They not only symbolize a concept but represent some of its internal structure. They constitute an analysis of a concept. The advantage of including such analyses in a representation is to permit simpler and faster processes of reasoning. Smolensky's (1988) simple associative systems mainly just associate and generalize.

[3] Smolensky (1988) argued that symbolic rules can be implemented on connectionist hardware, a claim that has elicited substantial opposition (see the commentaries that follow Smolensky's article).

Instead of doing analysis by applying sophisticated reasoning processes, much of the analysis is part of the representation; the required information is not in the processor, but in what is processed. The disadvantage of this approach is that reasoning can only consider analyses that are already represented. Rules are not so limited; they can perform arbitrary and complex operations and, therefore, novel analyses of concepts. They put the information in the processor and not the representation. Consequently, applying rules is relatively complex and slow. The second advantage of subconceptual representations is that they generalize automatically on the basis of feature overlap (cf. Sloman, 1993). By associating the features of a concept with other features, one is automatically also associating the features of related concepts—namely, those that share the first concept's features. Finally, subconceptual representations are context dependent. Concepts are represented by a set of features, so any features that the context brings along are automatically included in the concept representation. The subconceptual idea is summarized by Smolensky as "the subsymbolic hypothesis: The intuitive processor is a subconceptual connectionist dynamical system that does not admit a complete, formal, and precise conceptual-level description" (p. 7).

Thus, intuition is seen as an associative mechanism in which the associations are not between concepts but between components or attributes of concepts. According to Smolensky (1988), these associations comprise a key distinguishing characteristic of subconceptual knowledge: It embodies a large set of soft constraints. Soft constraints need not be satisfied; unlike the hard constraints that traditionally characterize symbolic computation, they do not have necessary consequences. Associations are traversed in parallel, so they cooperatively contribute to a state that is maximally consistent with all the units and associations. This resultant state can be thought of as an associative network's inference. Inference becomes a dynamic process involving a large set of parallel constraints, which are satisfied simultaneously as well as they can be. Notably, this is precisely what recurrent connectionist systems do.

The preferred metaphor amongst connectionists for describing this type of recurrent parallel computation and for describing the process of reasoning is that of settling into a stable state. The idea is that a reasoning problem can be modeled by representing the attributes of each goal, subgoal, fact, belief, hypothesis, and other relevant piece of information about the problem with units or nodes. A network can then be constructed to solve the problem by putting connections (associations) between these nodes to represent the relations between pairs of problem features. If two attributes are mutually supportive, such as a hypothesis and a supporting piece of evidence, then positive or excitatory connections are put between the units representing them. If two attributes are contradictory, such as two mutually exclusive hypotheses, then negative or inhibitory connections link them. Each unit is a variable that takes a numerical value, and the set of all units is therefore a vector of numbers. The network is dynamic, which means that the values of the units change over time; the vector evolves according to a set of activation equations. The network is put into an initial state by activating all those units that represent problem-relevant knowledge, such as facts about the world, goals, and the current context. An inference is the result of a constraint satisfaction process (defined by the activation equations) in which the network dynamically sends activation back and forth until it stops because it is in a state that is alternately referred to as the point of *minimum energy* (Hopfield, 1984) or of *maximum harmony* (Smolensky, 1986) or *coherence* (Thagard, 1989). If all goes well, this final state of the network includes a representation of the desired inference. Examples of such systems can be found in Holyoak and Thagard (1989), Schultz and Lepper (1992), Sloman (1990), and Thagard (1989).

Working from this metaphor, Hinton (1990) proposed a distinction between intuitive and rational inference. An intuitive (what I call *associative*) inference corresponds to a single settling of a network, the process just described. Rational (or *rule-based*) inferences are more complex and "require a more serial approach in which parts of the network are used for performing several different intuitive inferences in sequence" (p. 50). Hinton argued that people typically perform computations that they are good at in a few sequential steps, each involving a computationally intensive intuitive inference.

The hallmark of a rule-based inference is that it involves traversal of a conceptual hierarchy. This follows from Hinton's (1990) definition of "*symbol* [italics added]: It is a small representation of an object that provides a 'remote access' path to a fuller representation of the same object" (p. 49). Often, these fuller representations include other symbols. In conventional digital computers, symbols are essentially pointers to locations in which more information about the object can be found. That information frequently consists of more pointers to other objects. In the connectionist systems described by Hinton, symbols are not literally pointers, rather they are "reduced representations" that contain some of the internal structure of the represented object but also, like pointers, serve as remote access paths. Tracing through these "symbolic links" is equivalent to traversing a hierarchy of objects mapped out by the pointers.

Such a part–whole hierarchy might represent the outline of a paper or the structure of an argument. Traversing such a hierarchy presents special problems because it requires that each entity in the domain being modeled have multiple representations, one at each level of the hierarchy. For instance, a paragraph in a paper outline must be represented both as a component of a section of the paper and as an aggregate of sentences. The same paragraph may sometimes serve as a role and sometimes as a filler, sometimes as a part and sometimes as a whole, sometimes as a pointer and sometimes as a rich schematic structure. Hinton (1990) argued that the mechanism responsible for leaping amongst these various levels of representation, deciding which to pack and unpack, and drawing conclusions in the process (rule-based inference) is distinct from the mechanism that, at each step, actually does the packing and unpacking, fills in default information, and satisfies the multiple constraints involved (associative inference).

By providing concrete proposals about how two reasoning mechanisms could operate simultaneously and cooperatively, Smolensky (1988) and Hinton (1990) have provided evidence that a dual mechanism idea is computationally feasible. I now turn to the psychology of reasoning to see if the idea is psychologically plausible.

Two Forms of Categorization

The contemporary arena for debate in experimental psychology on the dual nature of thought is the study of conceptual structure—how an individual mentally represents categories. The debate turns on the very question at issue: Is processing associative or rule based? The property of associative thought that is prominent in the study of categorization is similarity. Theories ascribing categorization to similarity-based processes have propagated in large part because of Rosch's work on the structure of natural categories (reviewed in Rosch, 1978). She showed that within-category structure is graded in the sense that people treat some category members as more central than others (Rosch & Mervis, 1975). This work fueled the development of a class of prototype models that assumed that categories were mentally represented by the instance that was most similar to members of the same category and least similar to members of different categories (see Smith, 1989, for a review). Exemplar models of categorization (see Medin & Ross, 1989, for a review) are also defined in terms of similarity. In these models, people are imagined to retain all instances that they observe, categorizing similar ones together and different ones separately only when confronted by a task requiring categorization. Recurrent networks models (e.g., Knapp & Anderson, 1984; Schyns, 1991), although sharing some of the properties of both prototype and exemplar models, are also similarity based. Unlike most other models, similarity relations are transformed nonlinearly in these models. Each of these models is consistent with a variety of data which together convincingly show that similarity does play a role in categorization.

Theory-Based Categorization

Opposing this "original sim" (Keil, 1989) view of concepts are those who ascribe categorization to lay theories (Carey, 1985; Murphy, 1993; Murphy & Medin, 1985). These theorists have, following Goodman (1955), noted that similarity is inadequate to explain concept use. One problem is that similarity provides no account of which predicates are "projectible." An individual projects a predicate whenever ascribing a feature to an object or category by virtue of the feature's relation to other objects or categories, such as when ascribing a personality trait to someone by virtue of the person's group membership. The problem is illustrated by Murphy. Imagine viewing a new animal at the zoo, a zork. Having never seen a zork before, which zork predicates can be projected to other zorks? Most people are willing to make generalizations about the zork's size, shape, mode of locomotion, and so on but are not willing to generalize the zork's location, age, sex, and so on. Similarity cannot explain this differential willingness. If similarity were the only relevant factor, then all known zork properties would be projected to other (similar) zorks. Presumably, willingness is determined by biological knowledge that, for example, members of a species tend to have a common shape. Some people call this knowledge a *theory.* Another problem with similarity is that it is inconsistent. Judgments of similarity depend on how the context of use determines the relative weighting of category features (A. B. Markman & Gentner, 1993; Medin, Goldstone, & Gentner, 1990; Tversky, 1977). For example, a mechanical monkey is similar to a real monkey because they have common perceptual features, but a mechanical monkey is similar to a mechanical horse because they have common internal features. Pure similarity-based theories offer no account of this.

Simple similarity structures cannot explain concept use, and neither can the classical view (Smith & Medin, 1981) that concepts refer to stored definitions composed of necessary and sufficient conditions (Quine, 1977; cf. Wittgenstein's, 1953, famous demonstration of the irreducibility of *game* to either a set of necessary or sufficient truth conditions). Instead, concepts are posited to be central elements of an interconnected web of beliefs (Quine, 1977), or a lay theory. The idea is that psychological concepts have a status analogous to that of scientific concepts (Carey, 1985). They exist by virtue of the explanations they provide, in homeostatic combination with other concepts, for observed causal relations. "Technically, a theory is a set of causal relations that collectively generate or explain the phenomena in a domain" (Murphy, 1993, p. 177).

Much of the support for the theory view of categorization comes from studies of children's categorizations by Keil (1989). In one set of experiments, Keil used the discovery and transformation paradigms. In both paradigms, children and adults were asked to categorize an ambiguous stimulus, something that looks like one object (e.g., a horse) but has the insides of another (e.g., a cow). In the discovery paradigm, one object is described to the child who is then told that scientists have studied this object and found that its insides are actually those of the second object. In the transformation paradigm, the first object is converted to the second by means of surgery on its external, perceptible features. Keil observed a strong developmental trend. Kindergarteners categorized both artifact and natural kind categories on the basis of their appearance (the first object in the discovery paradigm and the second object in the transformation paradigm). However, whereas older children and adults also categorized artifacts on the basis of their appearance, they categorized natural kinds according to their internal constitution. Furthermore, Keil showed that cross-ontological transformations, in which natural kinds were converted into artifacts or artifacts into natural kinds, were not deemed acceptable by people of any age; the internal structure of the object determined how it was categorized. Rips (1989b) replicated the natural kind results in a similar experiment with adults.

Keil's (1989) interpretation of these experiments is that people develop theories of biological entities from which they derive the critical features for categorizing them, features that outweigh an apparent similarity to members of other categories. Keil went so far as to suggest that even in preschoolers inductions over natural kinds have nothing to do with the "original sim." Apparently, people are not satisfied to throw objects and states of affairs into categorical bins on the basis of similarity; rather, they want to understand their origins and effects in causal terms (and, when animate, their purpose). Keil recognized that similarity does play a role and concluded, in line with the thesis of this article, that concepts have a dichotomous structure in which theoretical relations sit amongst associative ones. He clearly believed however that theoretical relations dominate during development and are therefore primary for categorization. For a supporting view, see the work of S. Gelman (e.g., Gelman & Markman, 1986; Gelman & Medin, 1993); for a contrary one, see Jones and Smith (1993).

The theory-based view of categorization is a type of rule-based reasoning in three senses. First, the rule-based system is uniquely qualified to construct explanations by virtue of the trace of rules that rule-based reasoning leaves behind. The rules used during an episode of reasoning constitute an explanation. Such explanations are intrinsic to theory-based categorization, which posits that categories are chosen to the extent that the categorizer can explain their appropriateness. Relatedly, the application of a lay theory seems to require a rule. A person may decide that an object with the internal features of a cow is a cow, but doing so requires applying a rule derived from an understanding of the causal relations that constitute cows (Smith & Sloman, 1994). Finally, the theory-based view is inconsistent with the associative view, which assumes that processing is similarity based.

Dissociating Judgments: Categorization Versus Similarity

Evidence for the independence of similarity and categorization is provided by Rips (1989b) and follow-up research by Smith and Sloman (1994). Consider one of Rips's stimuli. He gave participants a sparse description of an object, such as "a circular object with a 3-in. diameter," and asked them to imagine it. He then asked one group of University of Chicago students whether it was more similar to a pizza or a quarter. The majority of students reported that it was more *similar* to a quarter (after all, its diameter was much closer to the diameter of the average quarter than to that of the average pizza). He asked another group of students whether it was more *likely to be* a pizza or a quarter. Then, the majority of students chose pizza (after all, pizzas come in a variety of diameters, whereas quarters are all just about the same size and less than 3 in.). In conclusion, similarity is not always the basis of categorizations; that is, categorizations are not always determined associatively. Similarity can be less important than rules, such as "if its diameter is not close to $^3/_4$ in., then it's not a quarter." For nominally defined categories with nonvariable dimensions such as quarters, people surely can and sometimes do apply such commonsense rules.

Nevertheless, since Aristotle, scholars have been aware that concepts reflect similarity structure (for more recent evidence, see Allen & Brooks, 1991; Brooks, Norman, & Allen, 1991; Ross, 1987). Whether this means relative magnitude on a single dimension or one of an infinite variety of ways of aggregating values across dimensions, the relative similarity ascribed to pairs of objects is useful for making predictions about concept use. Smith and Sloman (1994) have shown that even in Rips's (1989b) paradigm, which was designed to demonstrate the existence of rule structure, most responses are determined by similarity. Smith and Sloman used sparse descriptions as well as rich ones that contained features characteristic of the nonvariable category, such as "a circular object with a 3-in. diameter that is silver colored." With University of Michigan students serving as participants, Smith and Sloman found that the proportions of students choosing a given category were almost identical in the categorization and similarity conditions with both kinds of descriptions. Only by requiring participants to think aloud while making their decisions—as Rips did—did they find Rips's dissociation between similarity and categorization, and even then only with sparse items; with rich descriptions, categorization again tracked similarity. Clearly, similarity is important to the process of categorization. Nevertheless, Rips was able to demonstrate that, under the right conditions, people do categorize using rules. Moreover, the conclusions that they come to on the basis of such rules can override their associatively based conclusions.

This dissociation between similarity and categorization is a form of *functional independence* in which measures of two systems are affected differently by some independent variable. This implies the existence of at least one cognitive process associated with one system and not the other, that is, that the two systems are distinct with respect to this process. Functional independence has been used in the study of memory for many years as a criterion to distinguish memory systems (Tulving, 1983).

Two Forms of Reasoning

The Case for Rules

Many authors have posed similar distinctions, only to argue for one or the other side (cf. Rips, 1990, for an insightful exposition of the differences between what he calls the "strict" and "loose" views of reasoning). Recent proponents of the rule-based view, who argue that reasoning consists of sequential operations on symbolic structures, include Braine (1990), who argued for a "natural logic" approach to reasoning, and Newell (1990) who offered a "unified theory of cognition" that includes a rule-based theory of syllogistic reasoning. Evidence for systems of rules is provided by Osherson (1975), Rips (1983, 1989a, 1994), and Braine, Reiser, and Rumain (1984), all of whom had participants judge the validity of arguments. In each case, a set of rules from propositional logic was proposed, and experiments were reported that showed that participants' error rate, speed of response, or both were proportional to the number of rules needed to determine an argument's validity. These results are consistent with the claim that participants made validity judgments by using rules of the sort proposed. Rips (1990) extracted the valid arguments used by Braine et al. (1984), which people found extremely easy to evaluate (they were correct at least 97% of the time). For example, given the premise A & B, people consistently and validly conclude that A is true. Rips pointed out that very short proofs of these deductions can be constructed using Braine's rule set, which is consistent with the claim that people are using such a set to evaluate arguments. However, these data can be interpreted differently. These deductions may be so easy because the act of representing the premises requires the reasoner to also represent the conclusion (cf. Johnson-Laird et al., 1992).

Smith et al. (1992) consolidated the case for rules in reasoning. Using a set of eight criteria, they reviewed evidence that people explicitly apply abstract rules when reasoning. The criteria consisted of predictions such as that performance on rule-governed items is as accurate with unfamiliar as with familiar material and a rule, or components of it, may be mentioned in a verbal protocol. Other predictions were that the abstractness of material should not influence performance, rules will sometimes be overextended, performance will decrease as the number of rules required increases, application of a rule will prime its subsequent use, training on the rule will improve performance, and the effectiveness of such training should not depend on the similarity of training and target domains. Applying these criteria to studies of reasoning, they found evidence for four rules: *modus ponens* (if A then B, together with A implies B), the contractual rules of permission and obligation, and the statistical law of large numbers. In summary, a body of evidence has accumulated that makes a strong case for the psychological reality of a small set of rules some of the time.

The case for associative processes in reasoning makes for a thicker portfolio (e.g., Margolis, 1987), but it is also less compelling. Below, a small part of the abundant evidence demonstrating the important role played by memory and similarity in reasoning is reviewed. However, the claim that these data demonstrate associative processing is limited in principle. As I established earlier, any apparently associative process can be described as rule based because of the representational power of rules. Memory access and the application of a similarity-based heuristic are often modeled with rules (see, e.g., Kolodner, 1983, for memory; Collins & Michalski, 1989, for similarity). The case for two forms of reasoning therefore largely rests on other types of evidence.

Simultaneous Contradictory Belief

A body of data rich enough to provide substantial support for the hypothesis of two reasoning systems does exist. The data are drawn from a diverse set of reasoning tasks that share a single crucial characteristic. They all satisfy what I call *Criterion S.* A reasoning problem satisfies *Criterion S* if it causes people to simultaneously believe two contradictory responses. By "believe," I mean a propensity, a feeling or conviction that a response is appropriate even if it is not strong enough to be acted on. A taste of this form of evidence, though one that may not entail rule application, can be found in statements such as "Technically, a whale is a mammal" (Lakoff, 1972). The statement makes sense, more sense than "Technically, a horse is a mammal" because a common mode of conceiving of whales has them more similar to fish. A whale is simultaneously both a mammal (technically) and a fish (informally, of course). Situations abound in which people first solve a problem in a manner consistent with one form of reasoning and then, either with or without external prompting, realize and admit that a different form of reasoning provides an alternative and more justifiable answer. Judges are often forced to ignore their sense of justice to mete out punishment according to the law. These instances provide evidence for two forms of reasoning if, and only if, the tendency to provide the first response continues to be compelling irrespective of belief in the second answer, irrespective even of certainty in the second answer.

The logic of this form of evidence is easily illustrated by considering how perceptual illusions provide evidence for a dichotomy in a domain other than reasoning. The Müller–Lyer illusion (Figure 1) suggests that perception and knowledge derive from distinct systems. Perception provides one answer; the horizontal lines are of unequal size, although knowledge (or a ruler) provides quite a different one—they are equal. The knowledge that the two lines are of equal size does little to affect the perception that they are not. The conclusion that two independent systems are at work depends critically on the fact that the perception and the knowledge are maintained simultaneously. Even when I tell myself that the lines are of equal length, I see lines of different lengths. If the knowledge changed the perception, then the only valid conclusion would be that I had, at different times, contradictory responses. The responses however could have been generated by the same system; the system may have simply changed its output.[4] For this reason, the criterion is not one of *bistability.* The Müller–Lyer illusion is not bistable in the sense that people first resolve to one interpretation and then to another (as people do with the Necker, 1832, cube). Rather, at each point in time two contradictory opinions are held; one provided by the perceptual system and another by a system of abstract comprehension. Of course, usually perception and knowledge do not contradict one another, but that does not mean that they constitute a single system. Similarly, the failure of a reasoning problem to satisfy Criterion S is not evidence against two reasoning systems. The associative and rule-based systems may converge to the same answer, in which case no contradictory beliefs would arise.

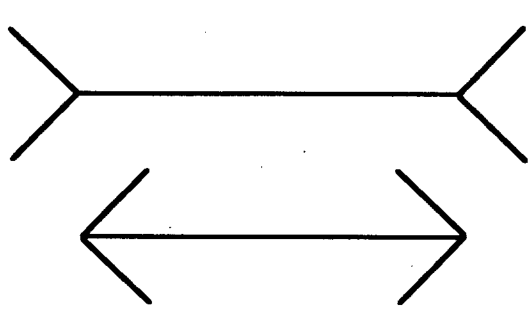

Figure 1. Müller–Lyer illusion.

The criterion is also not that people will affirm both responses. I would refuse to affirm that one line in Figure 1 is longer than the other, even though that is the conclusion of my perceptual system. So the criterion offered involves a leap of faith. Psychologists have to trust participants' and their own intuitions that some part of their minds are compelling them to believe something that some other part of their minds knows to be false.

Judgment. A variety of phenomena in the field of judgment satisfy Criterion S, many of which are reviewed by Kahneman, Slovic, and Tversky (1982). Perhaps the best known and most compelling example of simultaneous contradictory belief is an example of the conjunction fallacy of Tversky and Kahneman (1983), the Linda-the-bank-teller problem. They gave their participants the following paragraph that describes the hypothetical person Linda:

> Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. (p. 297)

Then, they asked the participants to rank order eight statements about Linda according to the statement's probability. The statements included the following two:

> Linda is a bank teller. (T)
> Linda is a bank teller and is active in the feminist movement. (T & F) (p. 297)

[4] A single process interpretation of cases in which an intuitive judgment conflicts with a critical judgment can be found in Margolis (1987).

Three groups of participants, including a group of graduate and medical students with statistical training and a group of doctoral students in the decision science program of the Stanford Business School, more than 80% of the time ranked Statement T & F as more probable than Statement T. A general principle participants used to make this judgment is similarity, as shown by Tversky and Kahneman. A more complete demonstration of the role of similarity in this context can be found in Smith and Osherson (1989; Shafir, Smith, & Osherson, 1990) using typicality judgments. Similarity is embodied in a heuristic that Tversky and Kahneman called *representativeness.* Evidence that participants use the representativeness heuristic for this kind of judgment is strong, although they are also influenced by other factors (Gavanski & Roskos-Ewoldsen, 1991; Shafir et al., 1990). The paragraph describing Linda is more similar to that of a feminist bank teller than it is to a stereotypical bank teller (participants' ratings confirm). One can more easily imagine Linda as a feminist bank teller, which leads one to conclude that she is more likely to be one. Of course, Statement T & F could not possibly be more probable than Statement T because it presupposes T; the truth of T & F entails that T be true. A conjunction can never be more probable than one of its constituents.

Apparently, two mechanisms exist that lead to divergent conclusions. On one hand, an intuitive heuristic leads to the conclusion that T & F is more probable. On the other hand, a probabilistic argument leads to the conclusion that T is more probable. Both mechanisms have psychological force. Most researchers (though not all, see Cohen, 1981; Gigerenzer, 1991) are willing to assent to the logic of the conjunction rule of probability in this case and, therefore, believe that T is more likely. Indeed, Tversky and Kahneman (1983) reported that few participants attempted to defend their responses. Nevertheless, a compulsion remains to respond that T & F describes a possible world that seems more likely.[5] I can trace through the probability argument and concede its validity, while sensing that a state of affairs that I can imagine much more easily has a greater chance of obtaining. As one participant who acknowledged the validity of the conjunction rule said, "I thought you only asked for my opinion" (Tversky & Kahneman, 1983, p. 300). Fortunately, opinions and reasoned conclusions do not usually diverge.

The conjunction fallacy is a robust effect that has been demonstrated with a variety of materials and in a variety of situations that make the relation between T and T & F transparent. Tversky and Kahneman (1983) showed that the same result is obtained even if no filler items are used, participants are simply asked which of T and T & F is more probable. Although the effect was reduced by having participants bet on their responses, a majority still chose the conjunction over its constituent. The effect is not attributable to a misunderstanding of the meaning of the statements (Crandall & Greenfield, 1986).

The incidence of the fallacy can be reduced by asking participants to make an assessment of frequency rather than a probability judgment (Fiedler, 1988; Tversky & Kahneman, 1983). Fielder found that 91% of his participants committed the conjunction fallacy when he asked them to rank order statements about Linda's profession with respect to their probability. However, when he asked them how many out of 100 people who are like Linda the statements applied to, he found that only 22% of the participants' estimates violated the conjunction rule. The implication of the conjunction rule seems to be more transparent and similarity relations less influential when participants evaluate the frequency of conjunctions within concrete sets rather than the probability of combinations of properties. One way to understand the conjunction rule is in *set-theoretic* terms (the set of things with properties T and F is a subset of the set of things with property T). Describing the options in terms of sets may successfully cue the extentional relation described by the conjunction rule.

In any case, the conclusion holds that, when not so cued, people tend to make judgments on the basis of representativeness that violate a rule, a rule which most are happy to grant. Even after granting the rule, we feel a compulsion to report an answer that violates it. We may not report such an answer, but the fact that we are able to inhibit the response suggested by similarity is evidence for two systems.

Argument strength. Other demonstrations that satisfy Criterion S can be found by observing how people project unfamiliar properties amongst categories. Sloman (1993) found that people tend to project properties from a superordinate category to a subordinate only to the extent that the categories were similar (the inclusion-similarity phenomenon). For example, when asked to rate the convincingness of the following argument on a 10-point scale, on which 10 indicated maximal convincingness, participants gave it a mean rating of 9.6; they found it highly convincing.

> All birds have an ulnar artery.
> Therefore, all robins have an ulnar artery.

A second argument however received a rating of only 6.4, significantly lower statistically.

> All birds have an ulnar artery.
> Therefore, all penguins have an ulnar artery.

This pattern of response does not conform to set theoretic logic in that penguins are birds, so if all birds have a property then penguins must have it. A survey and model of argument strength phenomena (Sloman, 1993) provides evidence that a measure of feature overlap plays a dominant role in determining participants' judgments. Penguins apparently have little enough in common with other birds that they are not thought to necessarily exhibit a property held by all birds, even though they themselves are birds. When participants however were told in debriefing interviews that a good reason was available to assign both arguments the maximal convincingness rating, namely the obvious category inclusion rule, they consistently agreed. They were also adamant (some more than others) that their responses were also sensible, though they inevitably failed to express why. I conclude that, after debriefing, participants had two answers in mind to the given problem, one associative and one symbolic. The associative or similarity-based one was generated automatically on presentation of the question, but the symbolic- or inclusion-based one, arrived at later, was able to inhibit the associative response.

A related demonstration called the *inclusion fallacy* is re-

[5] Gould (1991) shared this intuition: "I know that the [conjunction] is least probable, yet a little homunculus in my head continues to jump up and down, shouting at me—'but she can't just be a bank teller; read the description'" (p. 469).

ported by Osherson, Smith, Wilkie, Lopez, and Shafir (1990). They asked people to choose which of the following two arguments seemed stronger:

> Robins have an ulnar artery.
> Therefore, birds have an ulnar artery.
>
> Robins have an ulnar artery.
> Therefore, ostriches have an ulnar artery.

The majority of participants chose the first argument because robins and birds are more similar than robins and ostriches. However, most people also concede that the second argument is at least as strong because ostriches are birds, so any evidence that increases belief that all birds have some property should necessarily increase belief to at least the same extent that all ostriches have the property.[6] This is a striking example in which a compelling logical argument fails to erase an even more compelling intuition: How much evidence can a fact about robins provide for an animal as dissimilar as an ostrich?

Syllogistic reasoning. A *syllogism* is a kind of deductive argument with two premises and a conclusion consisting of quantified categories assigned to predicates, such as the following famous one:

> All men are mortal.
> Socrates is a man.
> Therefore, Socrates is mortal.

Demonstrations abound that willingness to affirm the conclusion of a syllogism, even an invalid syllogism, varies with prior beliefs. A person is more likely to consider a syllogism valid if he or she agrees with its conclusion (e.g., Janis & Frick, 1943; Markovits & Nantel, 1989; Rips, 1990) or desires the conclusion (McGuire, 1960). Reasoning is not based on formal considerations alone; it is affected by content or *belief bias* effects. Revlin, Leirer, Yopp, and Yopp (1980) asked participants to "decide which of five possible conclusions have to follow unambiguously from the given premises" of the following:

> No members of the ad-hoc committee are women.
> Some U.S. senators are members of the ad-hoc committee.
> Therefore:
> a. All U.S. senators are women.
> b. No U.S. senators are women.
> c. Some U.S. senators are women.
> d. Some U.S. senators are not women.
> e. None of the above is proven.
>
> No U.S. governors are members of the Harem Club.
> Some Arabian sheiks are members of the Harem Club.
> Therefore:
> a. All Arabian sheiks are U.S. governors.
> b. No Arabian sheiks are U.S. governors.
> c. Some Arabian sheiks are U.S. governors.
> d. Some Arabian sheiks are not U.S. governors.
> e. None of the above is proven. (p. 589)

In the first case, syllogistic logic agrees with the belief that "Some U.S. senators are not women." As a consequence, 83% of responses were correct. In the second case, logic conflicts with belief. Logic dictates, as it did in the first case, that again Answer d is correct. A more appealing conclusion however is the one known to be empirically true, that is, in the second case Answer b: "No Arabian sheiks are U.S. governors." Only 67% of participants chose Answer d. Participants do not ignore logical entailments; they accept more valid syllogisms than invalid ones (e.g., Evans, Barston, & Pollard, 1983).[7] Nevertheless, belief bias effects do occur as in the case at hand. These effects may not all be due to only differential availability in memory (see Rips, 1994, p. 343). Determining belief often requires more than simple memory access. Still, the current example shows that empirical belief obtained fairly directly through associative memory can inhibit the response generated by psycho-logic.

Belief-bias effects motivate a distinction between belief and deduction even if syllogistic reasoning is ascribed to mental models, as Johnson-Laird and Byrne (1991) did. They supposed that "reasoners will search for refuting models more assiduously if their initial conclusion is unbelievable than if it is believable" (p. 125). This idea presupposes a distinction between the determinants of belief and the search for refuting models, the latter constituting much of the process of deduction according to the mental model view.

Conditional reasoning. In the famous four-card selection task (Wason, 1966), participants are shown a display of four cards and told that each card has a letter on one side and a number on the other. They see, for example, a card with an *A*, a card with a *C*, a card with a *4*, and a card with a *3*. They are asked to choose only those cards necessary to decisively determine whether the following rule holds: If a card has a vowel on one side, then it has an even number on the other side. Under these conditions, the majority of participants choose the card marked *A* and the card marked *4*. The next largest group of participants choose only the *A* card, which is a good choice because the rule is falsified if an odd number appears on the other side. Similarly, assuming the rule is an instance of the standard formal logical relation of conditional implication, the *3* card should be turned over. The rule is falsified if a vowel appears on its opposite side. However, turning over the *4* card or the *C* card is harder to justify; whatever appears on the opposite side of these cards would be consistent with the rule. Only a small minority of participants choose the cards dictated by formal logic: the *A* and *3* cards.

Performance has been greatly facilitated by embedding the task in certain meaningful contexts (e.g., Johnson-Laird, Legrenzi, & Legrenzi, 1972) and by framing the task in different ways (e.g., Cheng & Holyoak, 1985). Rips (1994) provided one account of the conditions under which context facilitates performance. However, psychologists still do not understand why performance diverges with the dictates of standard logic on the original version of the task. At least part of the answer has been suggested by Evans (1982) who argued that participants *match* terms in the rule with states of affairs in the cards. They choose the *A* and *4* cards because the rule being tested mentions vowels

---

[6] This effect does not depend on the absence of explicit universal quantification (such as the word *all* or *every single* preceding each category name). Shafir, Smith, and Osherson (1990) found inclusion fallacies using the *every single* wording even after elaborate efforts to explain to participants the set-theoretic meaning of *every single*.

[7] Evans et al. (1983) observed the influence of syllogism validity under the same conditions in which they found belief-bias effects. This implies that participants agree with the experimenter on the task they are asked to perform; they are not simply reporting whether conclusions are true or false.

and even numbers but not consonants or odd numbers. To account for performance with negated terms, Evans (1989) also posited that the word *if* directs attention to the card that satisfies the antecedent, so if told "if no vowel, then no even number," participants pick the *C* card (no vowel) and, because of the matching bias, the *4* card (even number).

Matching is an associative process; it involves a computation of similarity. Indeed, Evans (1982) speculated that two competing psychological processes determine performance in the selection task, a perceptually based matching process and a linguistic–logical process. He implemented this idea in a stochastic model of participants' reasoning. The model assumes that participants respond on the basis of either *interpretation* (logic) or a *response bias,* which in this case amounts to matching. Choice probability for a given card was taken to be a linear function of these two tendencies. He showed that the model fits choice data closely. Evans's characterization of the dual mechanisms is admirably more specific than simply contrasting associative with rule-based processes. He argued that the logical mechanism is verbal, whereas the matching mechanism is nonverbal and imagistic. Unfortunately, the evidence he presented relevant to these particular characterizations is scarce and sometimes contrary to prediction.

Many different variables affect selection task performance (Evans, Newstead, & Byrne, 1993). One variable suggests an associative influence different from matching. Consider the rule "If a person is drinking beer, then the person must be over 21 years of age" and the four cards "drinking beer," "drinking ginger-ale," "22 years of age," and "19 years of age." (A useful shorthand for the rule is "If P, then Q," which allows one to refer to the cards as P, not-P, Q, and not-Q, respectively.) Griggs and Cox (1982) had participants imagine themselves as security officers enforcing the rule and found that 73% of them selected only the two cards dictated by standard logic, P and not-Q. To explain these selections, Cheng and Holyoak (1985) claimed that the rule and its associated context elicit a *permission* schema, a set of content-dependent rules that elicit responses consistent with standard logic. However, Kirby (1994) showed that the frequency with which the not-Q card was picked depended on how not-Q was instantiated. Kirby gave participants not only "19 years of age" but also the not-Q options of "12 years of age" and "4 years of age." Kirby predicted that participants would not expect 12-year-olds, and certainly not 4-year-olds, to be drinking beer and therefore participants would be less likely to choose these cards to test the rule. His prediction was confirmed; participants chose "19 years" most often and "4 years" least often. Participants' willingness to pick not-Q was proportional to the plausibility of P appearing on the other side. Pollard and Evans (1983) demonstrated a parallel effect using abstract materials of the sort used in the original selection task. They increased the proportion of not-Q selections by training participants to expect P on presentation of not-Q.

These effects of expectedness are associative inasmuch as they are related to the strength of association of various instantiations of not-Q with P. They cannot be attributed to the application of standard rules of implication or the "pragmatic" rules of Cheng and Holyoak (1985) because such rules distinguish only between true and false values of the consequent and not among different ways of instantiating not-Q (such as 19 vs. 12 vs. 4 years of age). Nor are they attributable to the explicit use of a probabilistic rule. Conceivably, participants could have used a rule equivalent to "If not-Q suggests P *with high probability,* then choose not-Q." However, in related Wason-task experiments, Kirby (1994) found that participants' estimates of the relevant probabilities were uncorrelated with card selection. Moreover, such a rule begs the question, How would the relevant probability judgments be generated? One way would be by evaluating the strength of association between not-Q and P, in which case performance is controlled by the association.

In conclusion, the selection task offers another case that satisfies Criterion S. On one hand, responses on the abstract version of the task seem to be governed in part by an associative matching process and in part by an associative process that generates expectations. On the other hand, participants were willing to assent to the logic that suggests a different set of responses. Indeed, the majority of participants believed they made an error when provided with the standard logical formulation. Wason (1977) found that 21 of 34 participants arrived at the logical answer themselves when they were engaged in a kind of Platonic dialogue about the task. I suspect that one reason the Wason task has been so thoroughly studied is that the quick and dirty associative answer is so compelling, even to those who have discussed the task countless times in classes and seminars. I invariably have to slowly and deliberately work through the logic of the task to convince myself of the logical answer because I am always tempted to give a different one.[8]

Logan's Instance Theory

Logan (1988) described and tested a model of automatization consistent with my conclusion. His model applies to tasks, such as arithmetic, for which an answer can be obtained in two ways: either by using an algorithm (by rule) or automatically by retrieving an answer from memory (by similarity between the current problem and a previous one). Logan assumed that performance results from a race between these two processes. As participants gained experience with the task, their base of instances stored in memory increased, which increased the probability that automatic memory retrieval provides an answer before the completion of the algorithm. His statistical model of these competing processes successfully fit both lexical decision and alphabet arithmetic reaction time data. He also confirmed some qualitative predictions of the model. Logan's model made the strong assumption that the effect of practice is to increase the associative knowledge base without affecting the processing of rules at all.

The evidence that alphabet arithmetic has an associative component suggests that arithmetic also does. This insight helps to make sense of data showing that arithmetic has much the same character as other kinds of associative retrieval. For example, people give correct answers more quickly to arithmetic problems that they have recently practiced (Campbell, 1987; Stazyk, Ashcraft, & Hamman, 1982). A connectionist model of these effects is provided by J. A. Anderson, Spoehr, and Bennett (1994).

[8] Oaksford and Chater (1994) and Over and Evans (1994) argued that this temptation is quite rational: What I am calling the associative response can have the effect of maximizing the information gained about the hypothesis.

Empirical Conclusions

I have provided direct evidence for dual systems in categorization and in several domains of reasoning. In categorization, I noted the large amount of evidence showing the crucial role played by similarity, but I also presented evidence for principles distinct from similarity. One such principle was explanatory coherence, which accounts for findings that features are weighted in categorization decisions concerning natural kinds in proportion to their theoretical centrality (Keil, 1989; Rips, 1989b). I also discussed a dissociation between similarity and categorization produced by considering objects with dimension values close to those of members of a category but nevertheless excluded from the category because they violated a rule concerning lack of variability of that dimension within the category (Rips, 1989b).

In reasoning, I reviewed direct evidence for a small number of rules (Smith et al., 1992) and indirect evidence in the form of verified predictions from models based on rules. I also presented evidence for associative processing. This included demonstrations of response interference in reasoning caused by the intrusion of similarity-based processing. The conjunction, inclusion-similarity, and inclusion fallacies all resulted from similarity-based processing, whereas some performance on the original Wason (1966) task can be attributed to a matching process. The belief-bias effect in syllogistic reasoning and the expectedness effect in the Wason task are attributable to availability, that is, to relatively greater retrievability of relevant instances from memory. Most current models of memory attribute greater availability to spatial or temporal contiguity or to similarity between a retrieval cue and a target. Similarity and contiguity are the hallmarks of associative relations.

The bulk of the evidence for two forms of reasoning comes from the abundant and varied evidence of reasoning tasks that satisfy Criterion S. I reviewed evidence from four different domains of reasoning in which people were simultaneously compelled to believe two contradictory answers to a reasoning problem, in some cases with more than one demonstration from the same domain. Notice that the reader need only accept my conclusion in a single demonstration for the thesis of this article to hold. These data are complemented by the evidence for Logan's (1988) instance theory, which assumes that certain cognitive tasks can be performed either algorithmically or through associations to an instance memory.

Associative Intrusion and Rule-Based Suppression

These data help to characterize the interaction between the two systems. In all the demonstrations of simultaneous contradictory belief, associative responses were shown to be automatic, in that they persisted in the face of participants' attempts to ignore them. Despite recognition of the decisiveness of the rule-based argument, associative responses remained compelling (see Allen & Brooks, 1991, for analogous effects in categorization). Both systems seem to try, at least some of the time, to generate a response. The rule-based system can suppress the response of the associative system in the sense that it can overrule it. The associative system however always has its opinion heard and, because of its speed and efficiency, often precedes and thus neutralizes the rule-based response. In Freud's (1913) terms, repression sometimes fails.

Epstein, Lipson, Holstein, and Huh (1992) came to a closely related conclusion. In research directed at a distinction alike in many respects to the current one, they asked participants to consider vignettes describing people's reactions to negative outcomes. The vignettes described different actors suffering identical consequences for which they were equally responsible. Participants assumed both a self-orientation (how foolish would you feel if you had reacted that way?) and a "rational" orientation (how foolishly did the person in the vignette actually behave?). The rational orientation asked participants to make a more objective response than the self-orientation that asked them only to guess at a subjective feeling. By demanding objectivity, the rational orientation demanded responses that participants could justify; the self-orientation asked only that they report their impressions. Rules provide a firmer basis for justification than do impressions, and therefore participants were more likely to respond on the basis of rules in the rational than in the self-orientation condition. On the assumption that subjective impressions emerge from associative computation, participants were more likely to respond associatively in the self-orientation condition. Epstein et al. found that self-orientation judgments differed for different vignettes, depending on such causally irrelevant factors as whether actors had behaved as usual or unusually. Most pertinent here, judgments made with a rational orientation reduced but did not eliminate this effect. In conformity with Epstein et al., I conclude that even when a person is attempting to be rule governed, associative responses encroach on judgment. The force of the evidence is to support not only the conclusion that people have and use two computationally distinct systems of reasoning but also that the associative system intrudes on the rule-based one.

Representation in the Associative System

All the associative responses discussed above were based on fairly global correspondences between concepts represented as (more or less structured) sets of features. Concepts were not first distilled into one or two relevant features. For example, participants had to use features to compute the similarity between Linda and feminists because they did not have information about Linda other than a feature list. Little task-specific selection or differential weighting of features took place because performance was predictable from similarity judgments taken out of the problem context. In the inclusion-similarity phenomenon of argument strength, people preferred to project a property from a superordinate category (e.g., birds) to a similar subordinate (e.g., robins) than to a dissimilar subordinate (e.g., penguins). These similarity judgments have nothing to do with the argument strength rating task. In most contexts, robins are more similar to birds than penguins are.

The conspicuous feature of the data that I have reviewed is the extent to which people's modal inferences involved computations that considered only similarity structure and associative relations. This claim might appear contrary to work showing that associative judgments of similarity and probability can depend on hierarchical and causal structure. I mentioned earlier that A. B. Markman and Gentner (1993) and Medin, Goldstone, and Gentner (1993) have shown that similarity judgments can be strongly influenced by structured relations. The point is buttressed by Tversky and Kahneman (1983) who showed that the presence of a causal relation can increase a statement's representativeness. Their participants judged that John P. was less likely to kill one of his employees than he was to kill one of his employees to prevent him from talking to the police (a conjunction fallacy). The added motivation produced by the causal relation made the proposition seem more likely. The causal statement is more representative than the noncausal one of the standard model of murderers; a person tends to think of murderers as motivated. In short, people seem to be sensitive to both hierarchical and causal structure when performing associative operations.

On one hand, I have argued that certain judgments are associative, and yet they are sensitive to hierarchical and causal structure. On the other hand, I have argued that only rules, and not associations, can represent such structure. These arguments are not contradictory because mere sensitivity does not imply representational capacity. Similarity and probability judgments could be sensitive to hierarchical and causal relations because they depend on representations constructed by rules, and those rules could construct different representations depending on hierarchical and causal knowledge. To illustrate, I conclude from the example of the suspected murderer John P. that the similarity of an action to the actions expected of a murderer is increased by providing a cause for the action, in particular a motivation for a murder. Such a conclusion has two conditions. First, it requires comprehension that killing an employee to prevent him or her from talking to the police is a causal relation from the motive of preventing to the action of killing. Because it involves a causal relation, I claim that such comprehension involves at least one rule. Second, it requires a decision that a description that includes a motivation is more similar to the standard model of a murderer than a description that does not. I claim that this operation is associative. In this fashion, the causal analysis can be taken out of the associative part of the computation. Hinton's (1990) notion of *reduced representation* (alluded to above) assumes a separation of this form. Hinton suggested that intuitive inference occurs using associations between "compact" representations that contain some, but not all, of the internal structure of an object—such as the result of a causal analysis.

Analogous analyses apply to the other demonstrations that apparently associative judgments are sensitive to nonassociative structure (e.g., A. B. Markman & Gentner, 1993; Medin et al., 1993). I cannot prove that such analyses are correct. Their possibility however defeats the claim that the sensitivity of an associative judgment to rule-based relations must contradict the distinction between associative and rule-based processing. Demonstrations that structural relations influence associative judgments nevertheless retain value. For one, they are a reminder of just how intricately coordinated associative and rule-based processing can be.

Two further properties of associative thought are noteworthy. The first is attributable to James (1890/1950), who pointed out that, although associative thought often deals in concrete images, it can also deal in abstract concepts. For instance, a person can easily think about water or sheep as general categories, not only as particular instances. When thinking about wool, a person might make use of an association to sheep, but not to any sheep in particular, rather to sheep as a category. Second, and contrary to James, the associative system is not simply reproductive but can deal with novel stimuli. The similarity judgments underlying the conjunction fallacy, the inductive argument strength phenomena, and the matching effect in the Wason (1966) task were not retrieved from memory. The comparison process took place on-line.

Representation in the Rule-Based System

The data support the supposition that rule-based inference involves the administration of hard constraints between symbols (Smolensky, 1988). In each demonstration that I reviewed, participants were shown to have the capacity to appreciate relations of necessity or sufficiency between variables, a capacity that in each case manifested (I argued) as rule application. Explanation of the logic of the Wason (1966) task resulted in participants understanding and acknowledging a rule that amounted to " 'If not-Q, then not-P' necessarily follows from 'If P, then Q'." In both argument strength demonstrations, participants came to appreciate, after explanation, class-inclusion rules, in which subclasses inherit a property of a class. I also saw that participants were able to comprehend the probabilistic conjunction rule, which may imply a capacity to comprehend certain set theoretic relations (although the rule could have been understood in other ways as well). Rips's (1989b) pizza-quarter example showed that people are able to appreciate that an object cannot be a member of a category that has a nonvariable dimension, if the object's value on that dimension is outside the range for that category. These relations are all instances of hard constraints.

The data also support the claim that rule-based inference tends to involve a small number of features. The rules alluded to above considered at most two properties of any object or event. This observation is consistent with James's (1890/1950) claim that true reasoning consists of two stages. The first involves *sagacity*, the ability to discover through analysis and abstraction the particular attribute of an object or fact that is essential to the problem at hand. (The second is *learning*, the act of recalling those properties or consequences.) Sagacity allows a person, for instance, to open a box of crackers by picking out the aspect of the box that is openable.

Quine (1977) presupposed that the analysis and abstraction of features is a key aspect of reasoning; he called a process that depends on feature extraction the fundamental problem of induction, namely, determining the "projectibility" of a predicate. A predicate has to be identified and selected before its projectibility can be determined. Quine showed that this general problem subsumes the two paradoxes most widely discussed in the philosophical literature on induction. The first paradox, from Hempel (1965), is that the statement "All ravens are black" implies that all nonblack things are nonravens (by *modus tollens*). Therefore, evidence that a nonblack thing is a nonraven provides inductive support for the hypothesis that all ravens are black. So the fact that my shoes are white supports the conclusion that ravens are black. This however is absurd. Quine solved the paradox by arguing that the predicates *nonblack* and *nonraven* are not projectible; they do not generalize to other categories. Of course his solution requires an account of the conditions under which a predicate is projectible, an account that has yet to be offered.

The second riddle of induction is the problem of grue

(Goodman, 1955). Clearly, the fact that all emeralds that I have observed are green increases my willingness to affirm that all emeralds are green. Now, I construct a new property called *grue* which means "green before tomorrow but blue afterward." Because I cannot predict the future, as far as I know all the emeralds that I have observed are grue. Therefore, I should be willing to affirm that all emeralds are grue for the same reason that I am willing to affirm that all emeralds are green. But I am not. Again, Quine (1977) pointed out that the issue is projectibility. Green is projectible, but grue is not.

Nobody can say with certainty how projectibility is determined. Quine (1977) argued that the problem of projectibility constitutes one description of the general problem that scientific theories confront. In mature scientific disciplines, domain theories exist that identify causal mechanisms. These causal mechanisms identify relevant features that, according to Quine, eliminate any problem of projectibility. In everyday reasoning, determining projectibility is a function of the rule-based system if the following is accepted (a) that, like scientific theories, lay theories have the function of identifying relevant features, and (b) my earlier conclusion that theory-based reasoning is a type of rule-based reasoning.

Automatic–Controlled Processing and Development

I have characterized associative inference as reflexive and rule-based inference as a deliberate form of symbol manipulation. The deliberate quality of rule-based reasoning suggests that it is accomplished through goal-oriented, "optional" strategies (Posner & Snyder, 1975). These characterizations suggest a parallel between, on one hand, associative and rule-based reasoning and, on the other hand, automatic and controlled processing (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Historically, the automatic–controlled distinction has been applied to perceptual–motor tasks, such as visual search, and not to reasoning, but it may turn out to subsume the associative–rule distinction.

Associative processes may be shown to satisfy the two criteria laid out for automatic processes (Shiffrin, Dumais, & Schneider, 1981): "1: Any process that does not use general, nonspecific processing resources and does not decrease the general, nonspecific processing capacity available for other processes. . . ." (p. 227) "2: Any process that demands resources in response to external stimulus inputs, regardless of subjects' attempts to ignore the distraction." (p. 228). The previous section of this article argued that associative processes satisfy the second criterion. No evidence I know of speaks to the first criterion. The untested prediction is that cognitive load should place a greater burden on rule-based than associative processes. Tasks which have a large rule-based component, such as theorem proving, should be more adversely affected by a secondary conceptual task than tasks which are mostly associative, such as similarity judgment.

Theoretical discussion of the automatic–controlled distinction has focused on learning, in particular the nature of the transformation of controlled processes into automatic ones (see the discussion of Logan, 1988, above), which is analogous to the transformation of rule-based processes into associative ones. The existence of such transformations follows from a modification of an argument of Vygotsky's (1934/1987): The rule-based system must developmentally precede the associative system because an organism with only an associative system would not have the resources to develop analytic thinking skills. Unstructured associative devices are unlikely to find descriptions of their environment that obey rule-based principles such as productivity and systematicity. An organism however that can analyze its environment by generating useful and descriptive rules can internalize those rules by using them to nominate features to be associated.

Most associationists take this position. Hinton (1990) stated that rational inferences become intuitive over time: "People seem to be capable of taking frequently repeated sequences and eliminating the sequential steps so that an inference that was once rational becomes intuitive" (p. 51; see also Smolensky, 1988). Rumelhart (1989) claimed that a person develops formal skills such as mathematics by internalizing the symbolic manipulations that he or she learns to do externally. A person starts doing algebra by manipulating marks that are put on blackboards and paper but eventually can do simple manipulations mentally. The claim is that people first figure the world out deliberately and sequentially, and only with time and practice does the knowledge become integrated into the associative network. The idea is not that people are born with a fully functioning system of abstract comprehension, only that they try to analyze the world from the beginning (Carey, 1985).

However, the developmental story is not that simple; effects between reasoning systems are not unidirectional. Evidence also suggests that people rely on associative processes when they do not have knowledge of or access to rule-based ones (Quine, 1977, said that a person falls back on an "animal sense of similarity" when a lay theory is not available).[9] This is one interpretation of Keil's (1989) discovery and transformation results, reviewed above. The youngest children may have categorized animals by appealing to perceptual similarity because they had no preferable basis for their decision, such as a theory of biology. In summary, associative and rule-based reasoning are interwoven in development, just as they are in task performance. People need some rule-based reasoning to know what features to begin with in domains that they are neither phylogenetically nor ontogenetically adapted to, but they reason associatively when they do not have access to rules that might prove more definitive or certain.

General Discussion

What the Distinction Is Not

The distinction between associative and rule-based reasoning is not the same as the one between induction and deduction, although that distinction is often assumed to be the key psychological one. Induction and deduction are not well-defined psychological processes; they are only well defined as argument types (Skyrms, 1986). Very roughly, *inductive* arguments are those in which the premises make the conclusion more probable; *deductive* ones are those in which the conclusion is necessarily true if the premises are. (Rips, 1990, pointed out that

[9] Freud (1913) also claimed that associative thought—what he called "primary process thought"—developmentally preceded purposive thought, or secondary process thought. In fact, he named the thought processes on this basis.

even the set of arguments cannot be independently partitioned into deductive and inductive ones. The definition given only distinguishes methods of *assessing* the strength of an undifferentiated set of arguments.) The distinction is actually orthogonal to the current one because both reasoning systems influence people's judgments of the validity of both kinds of arguments. I have described examples of both inductive arguments (e.g., the inclusion fallacy) and deductive arguments (e.g., belief-bias effects) that are assessed, and in contradictory ways, by the two reasoning systems. Both kinds of arguments are influenced by at least one common process, namely, a matching process that reflects similarity structure.

The distinction is also not the same as the one between analytic and nonanalytic cognition (e.g., Allen & Brooks, 1991). That distinction focuses on the dual influences in perception, categorization, and reasoning of instance- or exemplar-based processing and processing based on abstract information. According to this distinction, processing is analytic if responses are made on the basis of a stored abstraction, whether that abstraction is in the form of a prototype or a rule. I am distinguishing prototypes from rules. Prototypes are indeed abstract, but reasoning from them is essentially similarity based in that, according to prototype models, decisions are based on similarity to a prototype. Exemplar processes are also similarity based. Therefore, I group exemplar and prototype-based processes together and contrast them to rule-based processes. My distinction happens to fit more comfortably with the connectionist paradigm, in which exemplars, prototypes, and combinations of the two are all stored together (McClelland & Rumelhart, 1985).

Systems' Functions

Why should human beings need two systems of thought? One answer is that the systems serve complementary functions. The associative system is able to draw on statistical structure, whereas a system that specializes in analysis and abstraction is able to focus on relevant features. A different sort of complementarity is that associative paths that are followed without prejudice can be a source of creativity, whereas more careful and deliberative analyses can provide a logical filter guiding thought to productive ends. Mathematics, law, and probably all disciplines demand this combination of creativity and rigorous rule application.

Freud (1913) supplied an answer of a completely different sort. He suggested that the two forms of thought, or psychic processes, have their source in two aspects of human experience. On one hand, a person desires gratification and avoidance of pain. According to Freud, a person is driven by the pain principle. He described a primary process in which energy spreads around the psyche, collecting at ideas that are important to the individual and making them more intensive. He held this process responsible for channeling wish fulfillment and pain avoidance. On the other hand, a person must try to satisfy these urges in a world full of obstacles and boundaries. Gratification must sometimes be delayed. Inhibiting this primary process, and thus making both gratification more likely in the long run and behavior more socially acceptable, is secondary process thought, governed by the reality principle. Freud called such inhibition *repression,* which helps the individual behave in accordance with logical, physical, and cultural constraints. Primary process thought sets the stage for fantasy and imagination; secondary process, for purposive activity.

Freud (1913), indeed every theorist who has discussed the issue, believed the source of most rule-based knowledge is cultural. Consistent with this claim, all the rule-based reasoning detailed above reflects cultural knowledge (probability theory, class-inclusion logic, etc.) imparted by the experimenter to the participant. This notion of internalizing rules was axiomatic to Vygotsky (1934/1987), who emphasized the role of language in the cultural diffusion of rules. He believed that learning to think analytically is mostly a process of internalizing speech. He argued for a detailed description of a process according to which the child's thinking begins with social speech, passes through a stage of egocentric speech, and then crystallizes in the form of inner speech and logical thought. A recent example of an empirical analysis that has this flavor is due to E. M. Markman (1989) who showed how linguistic cues help children learn rules concerning class-inclusion hierarchies.

Implications

Conceptual Structure

Associationists and rule-based theorists tend to have different views concerning the determinants and extent of conceptual coherence. Associationists tend to believe that beliefs are usually consistent with each other because they reflect the world and the world is necessarily coherent, for it must obey the laws of nature. People may have contradictory beliefs because different aspects of their experience may provide evidence for opposing views. Experience in the home may suggest that people tend to be generous, but experience on the highway may suggest that people tend to be selfish. On this view, coherence is a property of concepts by virtue and to the extent that experience in the world is coherent.

Rule-based theorists tend to believe that people possess a more potent urge for coherence. Rules can reflect structure in the world as well as conform to their own syntax and semantics, which may impose further structure. Any formal calculus of belief embodies assumptions about which beliefs are consistent with each other. For example, the probability calculus assumes that the probability of an event is equal to 1 minus the probability of the event not occurring, an assumption which may not always be optimal (Shafer, 1976). Thus, rules enforce their own principles of coherence and, accordingly, rule-based theorists tend to believe that people try to conform. Some of them (e.g., Keil, 1989; Murphy, 1993) imply that people try to construct a global rule-based theory, which causes them to try to be globally coherent in their everyday lives (and not just when doing philosophy or science).

Allowing humans to be both associationists and rule governed suggests a way to reconcile these views. People may have an urge for coherence, but that urge is for *local* coherence. People apply rules in such a way that current explanations, the temporary contents of working memory, are internally consistent and consistent with the long-term knowledge deemed relevant. The demand for coherence does not go beyond that; a person does not expect his or her beliefs to constitute a grand, unified theory that pertains to every aspect of existence. This point is most vivid when one considers whether there are theories that determine which features of a concept are essential. Concepts surely have some attributes that serve as essential more often than others. For example, the attributes of my computer screen that allow it to emit light are more essential more of the time than the attributes that cause it to reflect ambient light. I am more likely to conceive of my computer screen as a light emitter than as a source of glare. I can do either, however, and indeed I can conceive of my computer screen in many other ways too—as an expense, as indispensable, and as well constructed. My computer does not have essential properties, I do not ascribe essential properties to it, and I do not have a theory of computer screens that I store away for use when I discuss or use them. Rather, my current goal makes certain properties relevant, and I am able (usually) to focus attention on them. To emphasize the goal dependency of the way human beings determine what is essential, I say we aim for *explanatory* coherence (Sloman, 1994), not conceptual coherence. For the most part, a person can rely on the world to maintain coherence across situations (unless perceptions are terribly distorted). Because they reflect objects and events in the world fairly directly, the associative system can do some of that work.

Education

The distinction is relevant to educational practices in two ways. First, it suggests that teachers should be aware that students have two tasks: They must both master the rules of the domain because rules provide productivity, systematicity, and a means to verify conclusions, and they must develop useful associations between elements of the domain to allow reasoning to become less effortful and more flexible. The necessity of learning both of these skills does not increase the burden placed on the learner; usually it decreases it. Useful associations guide the rule learner in the right direction; rule training provides a means to check and correct performance. Rule training also provides skills for the associative system to master inasmuch as rule application becomes associative with practice. Both rules and associations play a role in reasoning, therefore in learning, and can be mutually supportive (cf. Ross, 1989).

Second, the distinction may help teachers predict which concepts learners find easy and which they find difficult. Concepts should be easy to learn when the rules that govern them are compatible with students' natural associations. Concepts should be harder to learn when the two conflict (ask anyone who has tried to teach the logic of *modus tollens* or the meaning of statistical significance). The distinction between rules and associations may prove most valuable in such situations because it highlights the need to focus on these cases and to engage in what can be a difficult process of explanation.

Everyday Reasoning

These cases of inconsistency between rules and associations are one of the primary sources of conflict both within and between individuals. Decisions that a person makes everyday are made more difficult by opposing recommendations from the two systems. Those who feel safe and secure when driving a car do not always feel that wearing a seatbelt is worth the trouble and discomfort, particularly when travelling short distances. After all, they may never have had their lives saved by a seatbelt. Often they wear a seatbelt anyway. When the car beeps to remind the occupant to put it on, coming up with a plausible justification for not wearing it can prove difficult. In this case, the conflict involved is minimal and easily ignored. However, analogous situations arise in which the conflict is much greater and the result less predictable, such as whether to wear a condom. The thorough analysis and decision to wear one the day before may become insignificant in the face of a compelling reason not to wear one the moment before, a reason that may stem from a sense of invulnerability arising from previous occasions.

This sort of conflict dominates much of choice behavior. Choices consumers make are often between products that conjure up strong associations because of effective advertising or market longevity and products whose value can be analytically justified. Choosing between brand names, with which a person has had a long experience, and generic products, which sometimes have identical ingredients and a lower price, has this characteristic. This type of conflict is even more palpable when considering political options. A politician may seem attractive when expressing particular values or promising to solve particular problems, but analysis may suggest that enacting the candidate's policies is either impractical, immoral, or both. More generally, a person can be torn between descriptions that he or she resonates to and descriptions that he or she finds to be analytically more accurate.

Conclusions

People are renowned for their willingness to behave in ways that they cannot justify, let alone explain. Instead of performing a complete analysis of their interests, people vote for a politician because they have always voted for that person; they buy an item because it is associated with an image that they would like to project. Most people however only go so far. They would not do something that would be considered irrational if it entailed a real penalty or cost. They would not buy the item if it had been linked to cancer. So, on one hand, people are compelled to "follow their noses" by allowing associations to guide them; but, on the other hand, they are compelled to behave in a manner that they believe to be more justifiable. The fact that people are pulled in two directions at once suggests two forces pulling.

Evidence from the literature on animal learning suggests that organisms are likely to have a variety of special-purpose mechanisms (Gallistel, Brown, Carey, Gelman, & Keil, 1991). The application of an associative system to reasoning may represent the development of just such a special-purpose mechanism. Associative systems can capitalize on the ability of memory and similarity-based generalization to usually draw reasonable inferences, while maintaining the flexibility to do so in uncountable varieties of situations. Such a system would complement one that reasons by rules.

A lot of effort has been spent on arguing whether the human mind is best conceived as an associative system, especially in its modern connectionist guise, or as a classical symbol-manipulating device. The answer seems to be that the mind is both. This answer is consistent with a wave of interest that has recently developed in hybrid systems: computational systems that combine the precision and productive power of symbolic rules with the learning, automatic generalization, and constraint satisfaction power of connectionist associations (e.g., McMillan, Mozer, & Smolensky, 1992; Mozer & Das, 1993; see the collections in Bookman & Sun, 1993; Hinton, 1991). These efforts however have yet to be guided by psychological facts. In this article, I have begun a review of such facts by attempting to characterize two reasoning systems and say something about their interaction.

Neisser began his 1963 article by pointing out that "the psychology of thinking seems to breed dichotomies" (p. 1). A dichotomy is only as valuable as the explanatory power of the hypothetical systems that it distinguishes. Of course, as most theoretical entities do, mine generate more questions than explanations. Can theorists specify the systems' computational capacities with both mathematical precision and empirical reference? Can such specifications help researchers to understand cognitive pathology and more about learning and systematic human error? Answers to these and other questions await further reasoning and discovery.

References

Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General, 120,* 3–19.

Anderson, J. A., Gately, M. T., Penz, P. A., & Collins, D. R. (1990). Radar signal categorization using a neural network. *Proceedings of the IEEE, 78,* 1646–1657.

Anderson, J. A., Spoehr, K. T., & Bennett, D. J. (1994). A study in numerical perversity: Teaching arithmetic to a neural network. In D. S. Levine & M. Paricio, IV (Eds.), *Neural networks for knowledge representation and inference* (pp. 311–335). Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1993). *Rules of the mind.* Hillsdale, NJ: Erlbaum.

Bookman, L. A., & Sun, R. (Eds.). (1993). Architectures for integrating neural and symbolic processes [Special issue]. *Connection Science, 5.*

Braine, M. D. S. (1990). The "natural logic" approach to reasoning. In W. F. Overton (Ed.), *Reasoning, necessity, and logic: Developmental perspectives* (pp. 133–157). Hillsdale, NJ: Erlbaum.

Braine, M. D. S., Reiser, B. J., & Rumain, B. (1984). Some empirical justification for a theory of natural propositional logic. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and thinking* (Vol. 18, pp. 313–371). New York: Academic Press.

Brooks, L. R., Norman, G. R., & Allen, S. W. (1991). Role of specific similarity in a medical diagnostic task. *Journal of Experimental Psychology: General, 120,* 278–287.

Campbell, J. I. D. (1987). The role of associative interference in learning and retrieving arithmetic facts. In J. A. Sloboda & D. Rogers (Eds.), *Cognitive processes in mathematics* (pp. 107–122). Oxford, England: Oxford University Press.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: MIT Press.

Carpenter, G. A., & Grossberg, S. (1987). ART 2: Self-organization of stable category recognition codes for analog input patterns. *Applied Optics, 26,* 4919–4930.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17,* 391–416.

Chomsky, N. (1968). *Language and mind.* New York: Harcourt, Brace & World.

Cohen, J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences, 4,* 317–331.

Collins, A., & Michalski, R. (1989). The logic of plausible reasoning: A core theory. *Cognitive Science, 13,* 1–50.

Crandall, C. S., & Greenfield, B. (1986). Understanding the conjunction fallacy: A conjunction of effects? *Social Cognition, 4,* 408–419.

Epstein, S., Lipson, A., Holstein, C., & Huh, E. (1992). Irrational reactions to negative outcomes: Evidence for two conceptual systems. *Journal of Personality and Social Psychology, 62,* 328–339.

Evans, J. St. B. T. (1982). *The psychology of deductive reasoning.* London: Routledge & Kegan Paul.

Evans, J. St. B. T. (1989). *Bias in human reasoning.* Hillsdale, NJ: Erlbaum.

Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition, 11,* 295–306.

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction.* Hillsdale, NJ: Erlbaum.

Fiedler, K. (1988). The dependence of the conjunction fallacy on subtle linguistic factors. *Psychological Research, 50,* 123–129.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28,* 3–71.

Freud, S. (1913). *The interpretation of dreams.* London: George Allen & Unwin.

Gallistel, C. R., Brown, A. L., Carey, S., Gelman, R., & Keil, F. C. (1991). Lessons from animal learning for the study of cognitive development. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind* (pp. 3–36). Hillsdale, NJ: Erlbaum.

Gavanski, I., & Roskos-Ewoldsen, D. R. (1991). Representativeness and conjoint probability. *Journal of Personality and Social Psychology, 61,* 191–194.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23,* 183–209.

Gelman, S. A., & Medin, D. L. (1993). What's so essential about essentialism? A different perspective on the interaction of perception, language, and conceptual knowledge. *Cognitive Development, 8,* 157–167.

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review, 93,* 254–267.

Goldstone, R. L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition, 52,* 125–157.

Goodman, N. (1955). *Fact, fiction, and forecast.* Cambridge, MA: Harvard University Press.

Gould, S. J. (1991). *Bully for brontosaurus: Reflections in natural history.* New York: Norton.

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology, 73,* 407–420.

Hadamard, J. (1945). *The psychology of invention in the mathematical field.* Princeton, NJ: Princeton University Press.

Hempel, C. G. (1965). *Aspects of scientific explanation and other essays.* New York: Free Press.

Hinton, G. E. (1990). Mapping part–whole hierarchies into connectionist networks. *Artificial Intelligence, 46,* 47–76.

Hinton, G. E. (1991). *Connectionist symbol processing.* Cambridge, MA: MIT Press.

Hinton, G. E., & Sejnowski, T. J. (1986). Learning and relearning in Boltzmann machines. In J. L. McClelland, D. E. Rumelhart, & PDP Research Group (Eds.), *Parallel distributed processing* (pp. 282–317). Cambridge, MA: MIT Press.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction.* Cambridge, MA: MIT Press.

Holyoak, K. J., & Thagard, P. R. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13,* 295–355.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Science, USA, 81,* 3088–3092.

James, W. (1950). *The principles of psychology.* New York: Dover. (Original published 1890)

Janis, I., & Frick, P. (1943). The relationship between attitudes toward conclusions and errors in judging logical validity of syllogisms. *Journal of Experimental Psychology, 33,* 73–77.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction.* Hillsdale, NJ: Erlbaum.

Johnson-Laird, P. N., & Byrne, R. M. J. (1993). Précis of *Deduction. Behavioral and Brain Sciences, 16,* 323–380.

Johnson-Laird, P. N., Byrne, R. M. J., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review, 99,* 418–439.

Johnson-Laird, P. N., Legrenzi, P., & Legrenzi, M. (1972). Reasoning and a sense of reality. *British Journal of Psychology, 63,* 395–400.

Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development, 8,* 113–139.

Jordan, M. I., & Jacobs, R. A. (1994). Hierarchical mixtures of experts and the EM algorithm. *Neural Computation, 6,* 181–214.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge, England: Cambridge University Press.

Keil, F. C. (1989). *Concepts, kinds and cognitive development.* Cambridge, MA: MIT Press.

Kirby, K. N. (1994). Probabilities and utilities of fictional outcomes in Wason's four-card selection task. *Cognition, 51,* 1–28.

Knapp, A. G., & Anderson, J. A. (1984). Theory of categorization based on distributed memory storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 616–637.

Kolodner, J. L. (1983). Reconstructive memory: A computer model. *Cognitive Science, 7,* 281–328.

Lakoff, G. (1972). Hedges: A study in meaning criteria and the logic of fuzzy concepts. In *Papers from the Eighth Regional Meeting, Chicago Linguistic Society* (pp. 183–228). Chicago: Chicago Linguistic Society.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Margolis, H. (1987). *Patterns, thinking, and cognition.* Chicago: University of Chicago Press.

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology, 25,* 431–467.

Markman, E. M. (1989). *Categorization and naming in children.* Cambridge, MA: MIT Press.

Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition, 17,* 11–17.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114,* 159–188.

McGuire, W. J. (1960). A syllogistic analysis of cognitive relationships. In C. I. Hovland & M. J. Rosenberg (Eds.), *Attitude organization and change* (pp. 65–111). New Haven, CT: Yale University Press.

McMillan, C., Mozer, M. C., & Smolensky, P. (1992). Rule induction through integrated symbolic and subsymbolic processing. In J. E. Moody, S. J. Hanson, & R. P. Lippmann (Eds.), *Advances in neural information processing systems* (pp. 969–976). San Mateo, CA: Morgan Kaufmann.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1990). Similarity involving attributes and relations: Judgments of similarity and difference are not inverses. *Psychological Science, 1,* 64–69.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100,* 254–278.

Medin, D. L., & Ross, B. H. (1989). The specific character of abstract thought: Categorization, problem solving, and induction. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 5, pp. 189–223). Hillsdale, NJ: Erlbaum.

Mozer, M. C., & Das, S. (1993). A connectionist symbol manipulator that discovers the structure of context-free languages. In C. L. Giles, S. J. Hanson, & J. D. Cowan (Eds.), *Advances in neural information processing systems* (pp. 863–870). San Mateo, CA: Morgan Kaufmann.

Murphy, G. L. (1993). Theories and concept formation. In I. Van Mechelen, J. Hampton, R. Michalski, & P. Theuns (Eds.), *Categories and concepts: Theoretical views and inductive data analysis* (pp. 173–200). London: Academic Press.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92,* 289–316.

Necker, L. A. (1832). Observations on some remarkable optical phenomena seen in Switzerland; and an optical phenomenon which occurs on viewing a figure of a crystal or geometrical solid. *The London and Edinburgh Philosophical Magazine and Journal of Science* (3rd series), *1*(5), 329–343.

Neisser, U. (1963). The multiplicity of thought. *British Journal of Psychology, 54,* 1–14.

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

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101,* 608–631.

Osherson, D. N. (1975). *Logical abilities in children* (Vol. 3). Hillsdale, NJ: Erlbaum.

Osherson, D., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97,* 185–200.

Over, D. E., & Evans, J. St. B. T. (1994). Hits and misses: Kirby on the selection task. *Cognition, 52,* 235–243.

Pennington, N., & Hastie, R. (1992). Reasoning in explanation-based decision making. *Cognition, 49,* 123–163.

Piaget, J. (1926). *The language and thought of the child.* London: Routledge & Kegan Paul.

Pollard, P., & Evans, J. St. B. T. (1983). The effect of experimentally contrived experience on reasoning performance. *Psychological Research, 45,* 287–301.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Quine, W. V. (1977). Natural kinds. In S. P. Schwartz (Ed.), *Naming, necessity, and natural kinds* (pp. 155–175). Ithaca, NY: Cornell University Press.

Revlin, R., Leirer, V., Yopp, H., & Yopp, R. (1980). The belief-bias effect in formal reasoning: The influence of knowledge on logic. *Memory & Cognition, 8,* 584–592.

Rips, L. J. (1983). Cognitive processes in propositional reasoning. *Psychological Review, 90,* 38–71.

Rips, L. J. (1986). Mental muddles. In M. Brand & R. M. Harnish (Eds.), *Problems in the representation of knowledge and belief* (pp. 258–286). Tucson: University of Arizona Press.

Rips, L. J. (1989a). The psychology of knights and knaves. *Cognition, 31,* 85–116.

Rips, L. J. (1989b). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21–59). Cambridge, England: Cambridge University Press.

Rips, L. J. (1990). Reasoning. *Annual Review of Psychology, 41,* 321–353.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking.* Cambridge, MA: MIT Press.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 629–639.

Ross, B. H. (1989). Remindings in learning and instruction. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 438–469). Cambridge, England: Cambridge University Press.

Rumelhart, D. E. (1989). Towards a microstructural account of human reasoning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 298–312). Cambridge, England: Cambridge University Press.

Rumelhart, D. E., & Zipser, D. (1985). Feature discovery by competitive learning. *Cognitive Science, 9,* 75–112.

Salmon, W. C. (1989). Four decades of scientific explanation. In P. Kitcher & W. C. Salmon (Eds.), *Minnesota studies in the philosophy of science. Vol. 13: Scientific explanation* (pp. 3–195). Minneapolis: University of Minnesota Press.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1–66.

Schultz, T., & Lepper, M. (1992). A constraint satisfaction model of cognitive dissonance phenomena. *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society,* 462–467.

Schyns, P. G. (1991). A modular neural network model of concept acquisition. *Cognitive Science, 15,* 461–508.

Shafer, G. (1976). *A mathematical theory of evidence.* Princeton, NJ: Princeton University Press.

Shafir, E., Smith, E. E., & Osherson, D. (1990). Typicality and reasoning fallacies. *Memory & Cognition, 18,* 229–239.

Shiffrin, R. M., Dumais, S. T., & Schneider, W. (1981). Characteristics of automatism. In J. B. Long & A. Baddeley (Eds.), *Attention and performance* (Vol. 9, pp. 223–238). Hillsdale, NJ: Erlbaum.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190.

Siegelmann, H. T., & Sontag, E. D. (1995). On the computational power of neural networks. *Journal of Computer and System Sciences, 50,* 132–150.

Skyrms, B. (1986). *Choice and chance: An introduction to inductive logic* (3rd ed.). Belmont, CA: Wadsworth.

Sloman, S. A. (1990). *Persistence in memory and judgment: Part-set inhibition and primacy.* Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Sloman, S. A. (1993). Feature-based induction. *Cognitive Psychology, 25,* 231–280.

Sloman, S. A. (1994). When explanations compete: The role of explanatory coherence on judgments of likelihood. *Cognition, 52,* 1–21.

Smith, E. E. (1989). Concepts and induction. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 501–526). Cambridge, MA: MIT Press.

Smith, E. E., Langston, C., & Nisbett, R. (1992). The case for rules in reasoning. *Cognitive Science, 16,* 1–40.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Smith, E. E., & Osherson, D. N. (1989). Similarity and decision-making. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 60–75). Cambridge, England: Cambridge University Press.

Smith, E. E., & Sloman, S. A. (1994). Similarity- versus rule-based categorization. *Memory & Cognition, 22,* 377–386.

Smolensky, P. (1986). Information processing in dynamical systems: Foundations of harmony theory. In J. L. McClelland, D. E. Rumelhart, & PDP Research Group (Eds.), *Parallel distributed processing* (pp. 194–281). Cambridge, MA: MIT Press.

Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences, 11,* 1–23.

Stazyk, E. H., Ashcraft, M. H., & Hamman, M. S. (1982). A network approach to mental multiplication. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 320–335.

Thagard, P. R. (1989). Explanatory coherence. *Behavioral and Brain Sciences, 12,* 435–502.

Tulving, E. (1983). *Elements of episodic memory.* Oxford, England: Oxford University Press.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352.

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

Vygotsky, L. S. (1987). Thinking and speech. In R. W. Rieber & A. S. Carton (Eds.), *The collected works of L. S. Vygotsky. Vol. 1: Problems of general psychology.* New York: Plenum Press. (Original published 1934)

Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), *New horizons in psychology, I* (pp. 135–151). Harmondsworth, England: Penguin.

Wason, P. C. (1977). Self-contradictions. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking* (pp. 114–128). Cambridge, England: Cambridge University Press.

Wittgenstein, L. (1953). *Philosophical investigations.* New York: Macmillan.

Received October 18, 1993
Revision received January 3, 1995
Accepted January 10, 1995 ■

New Editor Appointed

The Publications and Communications Board of the American Psychological Association announces the appointment of Kevin R. Murphy, PhD, as editor of the *Journal of Applied Psychology* for a six-year term beginning in 1997.

As of March 1, 1996, submit manuscripts to Kevin R. Murphy, PhD, Department of Psychology, Colorado State University, Fort Collins, CO 80523-1876.