22. Two Systems of Reasoning

Steven A. Sloman

THE EMPIRICAL CASE FOR TWO SYSTEMS OF REASONING

The stimulation from a classic paper in the heuristics and biases tradition does not come only from the insights provided into processes of judgment and decision making; it also comes from anxiety, from the tension introduced between immediate intuition and more measured rational belief. Clearly, there is a limit to how much one’s interest is piqued by reading about other people’s mistakes. It is our own mistakes, and the insights they bring, that are so arresting and compelling. The tension is revealing because it reflects a gap within our own heads between, on one hand, our intuitions and, on the other hand, those of our beliefs that we consider rational. The classic demonstrations often suggest two minds at work: one following the “natural assessment methods” like representativeness and availability; and the other working to form coherent, justifiable sets of beliefs and plans of action. As Tversky and Kahneman have repeatedly shown, the two minds do not always agree.

The distinction between these two minds can be construed in terms of one of the central puzzles in experimental psychology – 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. Do we draw inferences through a network of learned associative pathways or by applying some kind of psychologic that manipulates symbolic tokens in a rule-governed manner? The debate has raged (again) in cognitive psychology for well over a decade now. It pits 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 as though it arises from a different cognitive mechanism than does deliberate, analytical reasoning. Sometimes conclusions simply appear at some level of awareness, as if our minds go
off, do some work, and then come back with a result, and sometimes coming
to a conclusion requires doing the work yourself, making an effort to construct
a chain of reasoning. 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), Piaget (1926), Vygotsky (1934/1987), Neisser (1963),
and Johnson-Laird (1983) among others, as shown above.

This distinction is implicit in the notion of “transparency” (Tversky &
Kahneman, 1983). A manipulation that reduces bias by making an extensional
probabilistic or logical-relation transparent is in essence providing a representa-
tion that affords rule-based inference, allowing people to go beyond associative
reasoning, but the presence of two systems for judgment and choice can make
life difficult. Cases of inconsistency between rules and associations are one of
the primary sources of conflict both within and between individuals. Decisions
that we make every day are made more difficult by opposing recommendations
from the two systems. This sort of conflict dominates much of choice behavior.
Consumer choices are often between products that conjure up strong associa-
tions due to effective advertising or market longevity and products whose value
can be justified analytically. Choosing between brand names, with which we
have had long experience, and generic products, which sometimes have identi-
cal ingredients and a lower price, tests this character. This type of conflict is
even more palpable when we consider political options. A politician may seem
attractive when expressing our values or promising to solve our problems, but
analysis may suggest that enacting the candidate’s policies is impractical, im-
moral, or both. More generally, people can be torn between descriptions that
they “restate” to and descriptions that they find analytically more accurate.

TWO FORMS OF COMPUTATION

The most lucid expression of the distinction and its psychological reality is that
of William James (1890/1950), who describes associative thought or empirical
thinking as “trains of images suggested one by another.” James believed that
associative thought is “only reproductive” in that the objects of associative
thought are all elements of or abstractions from our past experience, although
the data reviewed here suggest otherwise. True reasoning is “productive” ac-

According to James because it can deal with novel data: “Reasoning helps us out
of unprecedented situations” (p. 330).

Associative System

Today, we might describe James as distinguishing between two systems that
implement different computational principles. Roughly, one system is associ-
ative 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. Generally speaking, associative systems are able to divide perceptions
into reasonable clusters on the basis of statistical (or at least quasi-statistical) reg-
ularities. They treat objects in similar ways to the extent the objects are perceived
as similar (e.g., Rumelhart & Zipser, 1985). The primary reason for this is that
the degree to which an association is operative is proportional to the similarity
between the current stimuli and previously associated stimuli. On this view,
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 mechan-

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 productivity. Rule-based systems are productive in that they
can encode an unbounded number of propositions (i.e., rules can be combined
with each other to generate an ever-larger set of propositions). To understand
this, consider arithmetic, in which we can always generate a new number by
adding 1 to the largest number in our 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
(1988) argue that the productivity, systematicity, and, therefore, composition-
ality 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 relevant only to one form of reasoning. I
call this form of reasoning rule-based because rules are the form of representa-
tion 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 importantly, they have both a logical
structure and a set of variables. 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 we attribute to the symbols.

Variables vary (i.e., they can be instantiated in more than one way). Because
they assume a class of possible values, they are necessarily abstract. My discus-
sion concerns rules that contain variables, and therefore rules are abstract; they

In the next section, I discuss two kinds of rules for reasoning that they call
"mathematical reasoning schemas" associated with particular content domains. They suggest 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 (e.g., precondition and action, which must be specified on each occasion of use) and, they involve logical structure (the form if-then); therefore, I count them as rules.

Rules come in different kinds. Some rules are instructions, like statements in a computer program or a recipe; other rules are laws of nature or society or rules of logic. People are capable of following all of these rules (and of disobeying some). Rules can be normative, telling us how we should behave to reach some prespecified goal (such as the conjecture rule in order to maintain a coherent set of probabilities); or descriptive, telling us how we do behave in certain contexts. In contexts in which a normative rule obviously applies, it usually becomes descriptive as well. So, some rules are handed down to us by our culture, others we make up ourselves, 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 analytic machinery has access to and mastery of them.

Johnson-Laird and his colleagues 1973; Johnson-Laird & Byrne, 1991 have argued persuasively that much of everyday deduction is unlike theorem-proving. They have 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 applications of inference rules advocated by theorists such as Braine (1990) and Rips (1994). Nevertheless, their "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 that refer to entities in the statements of an argument. Other symbols represent negation and still others represent super-set or subset relations. Mental models enjoy the critical properties that I have assigned to rules: They have both logical structure and variables. The critical properties of the rule-based system are thus sufficiently general to encompass central aspects of the mental models theory.

Evans and Over (1996) come to a similar conclusion to mine about the dual process nature of thought. They draw their conclusion from a thorough survey of the deductive reasoning literature and related decision-making phenomena, and from the evidence supporting the distinction between implicit and explicit learning. They argue that the two systems of thought are driven by different types of rationality. The associative system is generally useful for achieving one's goals; the rule-based system is more adept at ensuring that one's conclusions are sanctioned by a normative theory.

Table 22.1 summarizes my characterization of the two systems. The point of this chapter is not that both systems are applied to every problem a person confronts, nor that each system has an exclusive problem domain; rather, the forms have overlapping domains that differ depending on the individual reasoner's knowledge, skill, and experience. Table 22.1 lists some functions that show off each system's capacities. However, the common mode of operation of the two systems 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. However, when a person is given a problem, both systems may try to solve it; each may compute a response, and the responses may not agree. 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 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, we are conscious only of the result of the computation, not the process. In contrast, we are aware of both the result and the process in a rule-based computation.
Various sources of evidence support the associative/rule-based distinction. One source is the study of conceptual structure. On one hand, scholars have been aware since the time of Aristotle that our concepts reflect similarity structure (for recent evidence, see Brooks, Norman, & Aller, 1993). Similarity is one of the hallmarks of associative processing. On the other hand, concepts serve analytic ends: they are used by rules to construct explanations. Some theorists argue that psychological concepts have a status analogous to that of scientific concepts (Carey, 1985; Quine, 1977). In support of this claim, experimental evidence makes clear that explanatory principles can guide the way even young children make classifications (Keil, 1989). The strength of the evidence on both sides of this issue suggests that both types of principles are operative in categorization. Indeed, Ashby, Alfonso-Reese, Turken, and Waldron (1998) propose a dual process neuropsychological theory of category learning that parallels the associative/rule-based dichotomy.

The case for the two systems of thought can be made on several grounds (see Sloman, 1996). For the purposes of this chapter, I focus on one—the existence of simultaneous, contradictory beliefs.

**TWO FORMS OF REASONING**

**Simultaneous Contradictory Belief**

Among the most compelling evidence for the hypothesis of two reasoning systems are data 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 believe two contradictory responses simultaneously. By “believe,” I mean a propensity, 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, although one that may not entail rule application, can be found in statements like, “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 (intuitionally). 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 in order 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–Lytes illusion 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 while I tell myself that the lines are of equal length, I see lines of different length. At each point in time, we hold two contradictory opinions: one provided by our perceptual system, 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.

**Judgment.** A variety of phenomena in the field of judgment satisfy Criterion S (many of which are reviewed in this book). Perhaps the best-known and most compelling example of simultaneous contradictory belief is an example of the conjunction fallacy reported by Tversky and Kahneman (1983), the Linda—bank-teller problem. They gave their participants the following paragraph describing 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.

Then they asked 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)

More than 80% of groups of graduate and medical students with statistical training and a group of doctoral students in the decision science program of the Stanford Business School ranked statement T & F as more probable than statement T. A general principle participants used to make this judgment is similarity or representativeness. In the paragraph description, Linda is more similar to a feminist bank teller than she is to a stereotypical bank teller (as participants’ ratings confirm). We can more easily imagine Linda as a feminist bank teller, which leads us 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.

Apparently, we have two mechanisms that lead us to divergent conclusions. On one hand, an intuitive heuristic leads us to conclude that T & F is more probable; on the other hand, a logical argument leads us to conclude that T is more probable. Both mechanisms have psychological force. Most researchers are willing to assert 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)
report 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. (Gould, 1991, shares 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). 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.

To buttress my claim, I ran a small, informal experiment. The three-prisoners problem (formally equivalent to the “three doors” or “Monty Hall” problems), like the Linda problem, is a problem of uncertain reasoning that involves two different responses: one intuitively compelling and one probabilistically correct. However, unlike the Linda problem, it does not satisfy Criterion S; people do not simultaneously believe the two responses in this problem. Indeed, this is an extremely difficult problem for both laypersons and experts in uncertain reasoning, who often do not resolve on even one interpretation (Fak. 1992; Shimno & Ichikawa, 1999). When they are able to resolve more than one interpretation, they do it sequentially (in analogy to the sequential resolution of the Necker cube), not simultaneously (in analogy to the Müller-Lyer illusion). They convince themselves of one response, or resolve to a different interpretation that elicits a different response that undermines the first.

I described both the Müller-Lyer and Necker cube illusions to five colleagues (faculty and graduate students in the Department of Cognitive and Linguistic Sciences at Brown University) and pointed out the difference in the temporal relation between the conflicting responses in the two cases (simultaneous vs. serial). I then described the Linda-the-bank-teller problem and discussed the two responses (she is more likely to be a bank teller versus she is more likely to be a feminist bank teller). Next, I asked participants whether the temporal relation between the points in time in which the responses seemed compelling was more analogous to the Müller-Lyer illusion or to the Necker cube. Every participant said the Müller-Lyer illusion. I also gave participants the three-prisoners problem, obtained their responses, and discussed the Bayesian reasoning that leads to a different response. Finally, I asked them whether the temporal relation between the interpretations that led to the two responses was more analogous to the Müller-Lyer illusion or to the Necker cube. Every participant said the Necker cube. In summary, everyone I asked (p = 1/32) agreed that the Linda-the-bank-teller problem satisfied Criterion S and that the three-prisoners problem violated it. The credibility of the intuition that beliefs are simultaneous in some problems is increased by the observation that the intuition is not ubiquitous.

People tend to make judgments on the basis of representativeness that violate a rule, a rule that most of us 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 (1998) found that people tended to project properties from a superordinate category to a subordinate only to the extent that the categories were similar. For example, consider the following argument:

**Fact:** Every individual piece of electronic equipment exhibits magnetic picofluctuation.

**Conclusion:** Every individual piece of audio equipment exhibits magnetic picofluctuation.

Participants were invited to assume that the fact (premise) was true and to rate the conditional probability of the conclusion given the fact. Set-theoretic logic dictates that if all audio equipment is electronic equipment, then audio equipment should inherit whatever properties electronic equipment has. In other words, someone who agrees that audio equipment is electronic should assign a conditional probability of 1 to the conclusion. I considered only the probability judgments of those participants who affirmed the category inclusion relation between audio and electronic equipment. Their mean conditionalized probability judgment was .89. Not only was there a substantial proportion of judgments less than 1, but when the conclusion category was atypical of the premise category, judgments were even lower. For example, the argument

**Fact:** Every individual piece of electronic equipment exhibits magnetic picofluctuation.

**Conclusion:** Every individual kitchen appliance exhibits magnetic picofluctuation.

received a mean probability judgment (again, conditional on affirmation of the category inclusion relation) of only .76. Sloman (1993) provides evidence that a measure of feature overlap plays a dominant role in determining judgments. When told in debriefing interviews that a good reason was available to assign both arguments the maximal probability judgment — namely, the obvious category inclusion rule - participants consistently agreed. They were also adamant (some more so than others) that their responses were also sensible, although 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 rule-based. The associative or similarity-based answer was generated automatically upon presentation of the question, but the rule- or inclusion-based answer arrived at later, was able to inhibit the associative response.

A related demonstration, the inclusion fallacy, was reported by Osherson, Smith, Wilkie, Lopez, and Shafir (1990), who asked people to choose which of
In the first case, syllogistic logic agrees with our 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 tells us, as it did in the first case, that again, it is correct. However, a more appealing conclusion is the one we know to be empirically true, conclusion b: No Arabian sheiks are U.S. governors. Only 67% of participants chose d in the second case. Participants do not ignore logical entailments; they accept more fallacious syllogisms than invalid ones (e.g., Evans, Barston, & Pollard, 1983). Nevertheless, belief bias effects do occur, as in the case at hand. The current example shows that empirical belief obtained fairly directly through associative memory can inhibit the response generated by psycho-logic.

RELATED EVIDENCE

Windschitl and Weber (1999) point out that expressions of certainty are differentially sensitive to context depending on whether they are expressed on precise numeric scales or imprecise nonnumeric ones. If given a precise numeric forecast (e.g., “Experts believe that the chance of rain is 70%”) and asked to make a numeric probability judgment (e.g., “What’s the probability of rain?”), people normally provide a judgment that matches the expert’s forecast. Windschitl and Weber show that given the same forecast but asked to make their judgment using a nonnumeric scale (e.g., a verbal scale or one represented by line length), people’s judgments are affected by the representativeness of the event. Representative events (e.g., rain in London) are judged more likely than unrepresentative events (e.g., rain in Madrid), despite identical expert forecasts. Their interpretation is that numeric judgments tend to reflect rule-based responding; nonnumeric judgments are more often products of associative reasoning.

Logan (1988) describes and tests a model of automatization consistent with my conclusion. His model applies to tasks like 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 assumes that performance results from a race between these two processes. As participants gain experience with the task, their memory base of instances increases, which increases the probability that automatic memory retrieval will provide an answer prior to the completion of the algorithm. His statistical model of these competing processes successfully fit both lexical decision and alphabet arithmetic (e.g., A = 1, B = 2, . . ., Z = 26; C + F = ?) reaction time data. He also confirmed some qualitative predictions of the model. Logan’s model makes 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 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 practiced recently (Campbell, 1987; Stazyk, Ashcraft, & Hamman, 1982).

A host of evidence exists in the literature on social cognition that suggests that dual processes control impression formation and attitude change (cf. Chaiken & Trope, 1999). For example, a variety of “correction” phenomena have been observed wherein participants are presented with a stimulus that they believe biases their impressions. People tend to correct for the bias unless they are distracted, suggesting that correction involves an optional, controlled process that is independent of some more automatic impression formation process. For example, Martin, Seta, and Crella (1990) primed participants with positive or negative concepts and then gave information about an individual that could be construed in positive or negative terms. When tested without distraction, participants’ impressions contrasted with the priming stimuli: they were corrected for the primes. But when tested with distraction, impressions showed an assimilation effect; participants failed to correct sufficiently for the primes. Wegener and Petty (1995) showed that the direction of a correction is flexible, and depends on how participants understand and explain the source of bias, suggesting that correction is governed by deliberate rule application. These differential effects of distraction amount to a dissociation between a controlled, analytic correction process and more automatic ones. However, the research does not imply a specific characterization of these automatic processes.

Two dual process models have emerged in the social cognition literature that contrast with the associative/rule-based distinction: Petty and Cacioppo’s (1986) elaboration likelihood model and Chaiken, Liberman, and Eagly’s (1989) heuristic/systematic model (both reviewed in Eagly and Chaiken, 1993, chapter 7). The elaboration likelihood model assumes dual routes for constructing attitudes: a central one based on analysis and elaboration of arguments and a peripheral one defined by all other variables that affect attitudes such as attractiveness, group identification, and conditioning. The heuristic/systematic model differs from elaboration likelihood mainly in the degree to which it constrains the nonelaborative route. It proposes that this mode of processing makes use of a set of heuristics—learned procedural knowledge—that must be accessed in order to affect judgment. A heuristic cue refers “to any variable whose judgmental impact is hypothesized to be mediated by a simple decision rule” (Eagly & Chaiken, 1993, p. 327). Thus, the model departs from the associative/rule-based distinction in assuming that nonanalytic processing makes use of rules as opposed to associations. Unlike the elaboration likelihood model, it shares with the current distinction the assumption that the dual processes occur in parallel, not sequentially.

EMPIRICAL CONCLUSIONS

I reviewed evidence from multiple domains 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 paper to hold. These data are complemented by Windschitl and Weber’s (1999) demonstration of the effect of associative inference on nonnumeric forecasting, the evidence for Logan’s (1988) instance theory, which establishes that certain cognitive tasks can be performed either algorithmically or through associations to an instance memory, and the evidence that correction processes in social judgment can be dissociated from other, less theory-bound, processes.

Associative Intrusion and Rule-Based Suppression

These data help us to characterize the interaction between the two systems. In all 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. However, the associative system always has its opinion heard, and, because of its speed and efficiency, often precedes and thus neutralizes the rule-based responses. Epstein, Lipson, Holstein, and Hub (1992) come 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 (e.g., “How foolish would you feel if you had reacted that way?”) and a “rational” orientation (e.g., “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 to guess only at a subjective feeling. By demanding objectivity, the rational orientation demanded responses that participants could justify; the self-orientation asked only that participants 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. 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 failed to eliminate this effect. In conformity with Epstein et al., I conclude that even when one is attempting to be rule-governed, associative responses encroach on judgment. The force of the evidence is to support not only the conclusion that human beings have and use two computationally distinct systems of reasoning, but also that the associative system intrudes on the rule-based one.
Two Systems of Reasoning

Representation in the Associative System

All the associative responses discussed previously 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 the concept of a typical female because they had no 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 outside of the problem context.

The conspicuous feature of the data 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. Markman and Gentner (1993) and Medin, Goldstone, and Gentner (1993) have shown that similarity judgments can be strongly influenced by structural 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 is more likely to kill one of his employees than he was to kill one of his employees who is suspicious about a missing elephant. 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 murderer; we tend to think of murderers as motivated. In short, people seem to be sensitive both hierarchically and causally when performing associative operations.

On one hand, I argue that certain judgments are associative and yet they are sensitive to hierarchical and causal structure. On the other hand, I argue that only rules, and not associations, can represent such structure. These arguments are not contradictory because more 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, we can conclude from the example of our suspected murderer John P. that the similarity of an action to the actions we expect 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, we must be able to comprehend that killing an employee to prevent him from talking to the police is a causal relation from the motive of preventing the firing of a murderer. Because it involves a causal relation, I claim that such comprehension involves at least one rule. Second, we must decide that a description that includes a motivation is more similar to our model of a murderer than a description that does not. I claim that this operation is associative. In this fashion, we can take the causal analysis out of the associative part of the computation.

Two further properties of associative thought are noteworthy. The first is attributable to Newell (1980), who pointed out that although associative thought often proceeds in concrete images, it can also deal in abstract concepts. For instance, we can easily think about water or sheep as general categories, not as particular instances. When thinking about wool, we might make use of an image of a 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 and the subsumption argument strength phenomena were not retrieved from memory. The comparison process took place on-line.

Automatic/Controlled Processing and Development

I have characterized associative inferences as a 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, on one hand, between associative and rule-based reasoning and, on the other hand, between automatic and controlled processing (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Historically, the automatic/controlled distinction has been applied to perceptual-motor tasks, like 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, Dumas, & Schneider, 1981). One: Any process that does not use general, nonspecific processing resources and does not decrease the general, nonspecific processing capacity available for other processes. . . . Two: Any process that demands resources in response to external stimulus inputs, regardless of participants' attempts to ignore the task. (p. 227-228). I argued earlier that associative processes satisfy the second criterion. No evidence is available directly to the first, although the correction phenomena in social judgment described earlier are suggestive. The prediction is that cognitive load should place a greater burden on rule-based than associative processes.

Tasks that have a large rule-based component, like theorem-proving, should be more adversely affected by a secondary conceptual task than tasks that are mostly associative, like similarity judgment. Theoretical discussions of the automatic/controlled distinction have focused on learning, in particular the interaction of the transformation of controlled processes in a automatic ones, analogous to the transformation of rule-based processes in associative ones. The existence of such transformations follows from a modification of an argument made by 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 like productivity and systematicity. But an organism 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) states 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). Rumelhart (1989) claims that we develop formal skills like mathematics by internalizing the symbolic manipulations that we learn to do externally. We start doing algebra by manipulating marks that we put on blackboards and paper, but eventually we do simple manipulations in our heads. 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 our associative network. The idea is not that we are born with a fully functioning system of abstract comprehension, only that we 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 we fall back on our “animal sense of similarity” when a lay theory is not available). In summary, associative and rule-based reasoning are intertwined in development just as they are in task performance. Some rule-based reasoning is needed to know what features to begin with in domains that humans are not phylogenetically nor ontogenetically adapted to, but associative reasoning predominates when rules that might prove more definitive or certain are inaccessible.

GENERAL DISCUSSION

What the Distinction Is Not

The distinction between associative and rule-based reasoning is not the same as one that is often assumed to be key psychologically, that between induction and deduction. 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) points out that even the set of arguments cannot be partitioned independently into inductive and deductive ones. Our definition only distinguishes methods of assessing the strength of an undifferentiated set of arguments. The distinction is 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 assessed 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 non-analytic cognition (e.g., Allen & Brooks, 1991). That distinction focuses on the dual influences of abstractions and more concrete exemplars in perception, categorization, and reasoning. According to it, processing is analytic if responses are made on the basis of a stored abstraction, whereas: 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 am grouping exemplar and prototype-based processes together and contrasting 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).

The 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 our two forms of thought, or psychic processes, have their source in two aspects of human experience. On one hand, we desire gratification and avoidance of pain. On the other hand, we 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. Primary process thought sets the stage for fantasy and imagination; secondary process, for purposive activity.

Freud — indeed, every theorist who discusses the issue — believes the source of most rule-based knowledge is cultural. Consistent with this claim, all the rule-based reasoning detailed previously reflects cultural knowledge (e.g., probability theory, class-inclusion logic) imposed by the experimenter to the participant. This notion of internalizing rules was axiomatic to Vygotsky, 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 that a child’s thinking begins with social speech, passes through a stage of egocentric speech, and crystallizes in the form of inner speech and logical thought.
Implications for 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 our 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 our experience may provide evidence for opposing views. Our experience in the home may suggest that people tend to be generous but our 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 we 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. 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 that 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. Rules are applied 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. For the most part, people can rely on the world to maintain coherence across situations (unless our perceptions are terribly distorted). Because they reflect objects and events in the world fairly directly, the associative system can do some of that work.

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 or buy an item because it is associated with an image that they would like to project. However, most people only go so far. They would not do something that they consider irrational if it entailed a real penalty or cost. Fewer people buy an item after it has been linked to cancer. So, on one hand, people "follow their noses" by allowing associations to guide them; on the other hand, they are compelled to behave in a manner more justifiable. The fact that people are pulled in two directions at once suggests two forces pulling.

23. The Affect Heuristic

Paul Slovic, Melissa Finucane, Ellen Peters, and Donald G. MacGregor

This chapter introduces a theoretical framework that describes the importance of affect in guiding judgments and decisions. As used here, affect means the specific quality of "goodness" or "badness" (1) experienced as a feeling state (with or without consciousness) and (2) demarcating a positive or negative quality of a stimulus. Affective responses occur rapidly and automatically--note how quickly you sense the feelings associated with the stimulus word before or into. We argue that reliance on such feelings can be characterized as the affect heuristic. In this chapter, we trace the development of the affect heuristic across a variety of research paths followed by ourselves and many others. We also discuss some of the important practical implications resulting from ways that this heuristic impacts our daily lives.

BACKGROUND

Although affect has long played a key role in many behavioral theories, it has rarely been recognized as an important component of human judgment and decision making. Perhaps belittling its rationalistic origins, the main focus of descriptive decision research has been cognitive, rather than affective. When principles of utility maximization appeared to be descriptively inadequate, Simon (1956) oriented the field toward problem-solving and information-processing models based on bounded rationality. The work of Tversky and Kahneman (1974) and Kahneman, Slovic, and Tversky (1982) demonstrated how boundedly rational individuals use such heuristics as availability, representativeness, and anchoring and adjustment to make judgments, and how they use simplified strategies such as "elimination by aspects" to make choices (Tversky, 1972). Other investigators elaborated the cognitive strategies underlying judgment and choice through modes of constructed preferences (Slovic, 1985; Payne, Bettman, & Johnson, 1988); dominance structuring (Montgomery, 1983), and comparative advantages (Shafir, Osherson, & Smith, 1985). In 1993, the entire volume of the journal Cognition was dedicated to the topic Reason-Based Choice, in which it was argued that "Decisions ... are often reached by focusing

Financial support for the writing of this chapter was provided by the National Science Foundation under grant 505 897 6387.