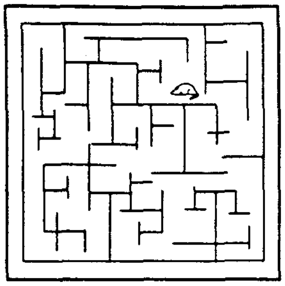

Reasoning and Natural Selection

LEDA COSMIDES, JOHN TOOBY, *Center for Advanced Study in the Behavioral Sciences*

Glossary

Adaptation Aspect of an organism that was created by the process of natural selection because it served an adaptive function

Adaptive Contributing to the eventual reproduction of an organism or its relatives

Bayes's theorem Specifies the probability that a hypothesis is true, given new data; $P(H|D) = P(H)P(D|H)/P(D)$, where H is the hypothesis and D is the new data

Cognitive psychology Study of how humans and other animals process information

Natural selection Evolutionary process responsible for constructing, over successive generations, the complex functional organization found in organisms, through the recurring cycle of mutation and subsequent increased reproduction of the better design

Normative theory Theory specifying a standard for how something *ought* to be done (as opposed to how it actually *is* done)

Valid argument Argument that is logically derived from premises; a conclusion may be valid, yet false, if it is logically derived from false premises

THE STUDY OF REASONING is an important component of the study of the biology of behavior. To survive and reproduce, animals must use data to make decisions, and these decisions are controlled, in part, by processes that psychologists label "inference" or "reasoning." To avoid predators, for example, a monkey must infer from a rustle in the grass and a glimpse of fur that a leopard is nearby and use information about its proximity to decide whether to take evasive action or continue eating. Because almost all action requires inferences to regulate it, the mechanisms controlling reasoning participate in almost every kind of behavior that humans, or other animals, engage in. Human reasoning has traditionally been studied without asking what kind of reasoning procedures our ancestors would have needed to survive and reproduce in the environment in which they evolved. In recent years, however, an increasing number of researchers have been using an evolutionary framework.

I. What Is Reasoning and How Is It Studied?

When psychologists study how humans reason, they are trying to discover what rules people use to make inferences about the world. They investigate whether there are general principles that can describe what people will conclude from a set of data.

One way of studying reasoning is to ask "If one were trying to write a computer program that could simulate human reasoning, what kind of program would have to be written? What kind of information-processing procedures (rules or algorithms) would the programmer have to give this program, and what kind of data structures (representations) would those procedures operate on?"

Of course, the human brain was not designed by an engineer with foresight and purposes; it was "designed" by the process of natural selection. Natural selection is the only natural process known that is

Encyclopedia of Human Biology, Volume 6. Copyright © 1991 by Academic Press, Inc. All rights of reproduction in any form reserved.

capable of creating complex and organized biological structures, such as the human brain. Contrary to widespread belief, natural selection is not "chance"; it is a powerful positive feedback process fueled by differential reproduction. If a change in an organism's design allows it to outreproduce other members of its species, that design change will become more common in the population—it will be *selected for*. Over many generations that design change will spread through the population until all members of the species have it. Design changes that enhance reproduction can be selected for; those that hinder reproduction are selected against.

When evolutionary biologists study how humans reason, they are asking, "What kind of cognitive programs was natural selection likely to have designed, and is there any evidence that humans have such programs?"

A. Mind Versus Brain

At present, researchers find it useful to study the brain on different descriptive and explanatory levels. Neuroscientists describe the brain on a physiological level—as the interaction of neurons, hormones, neurotransmitters, and other organic aspects. Cognitive psychologists, on the other hand, study the brain as an information-processing system—that is, as a collection of programs that process information—without worrying about exactly how neurophysiological processes perform these tasks. The study of cognition is the study of how humans and other animals process information.

For example, ethologists have traditionally studied very simple cognitive programs: A newborn herring gull, for instance, has a cognitive program that defines a red dot on the end of a beak as salient information from the environment, and that causes the chick to peck at the red dot upon perceiving it. Its mother has a cognitive program that defines pecking at her red dot as salient information from her environment, and that causes her to regurgitate food into the newborn's mouth when she perceives its pecks. This simple program adaptively regulates how the herring gull feeds its offspring. (If there is a flaw anywhere in the program—if the mother or chick fails to recognize the signal or to respond appropriately—the chick starves. If the flaw has a genetic basis, it will not be passed on to future generations. Thus natural selection controls the design of cognitive programs.)

These descriptions of the herring gull's cognitive programs are entirely in terms of the functional relationships among different pieces of information; they describe two simple information-processing systems. Of course, these programs are embodied in the herring gull's neurological "hardware." Knowledge of this hardware, however, would add little to our understanding of these programs as information-processing systems. Presumably, one could build a silicon-based robot, using hardware completely different from what is present in the gull's brain, that would produce the same behavioral output (pecking at red dot) in response to the same informational input (seeing red dot). The robot's cognitive programs would maintain the same functional relationships among pieces of information and would therefore be, in an important sense, identical to the cognitive programs of the herring gull. But the robot's neural hardware would be totally different.

The specification of a cognitive program constitutes a complete description of an important level of causation, independent of any knowledge of the physiological hardware the program runs on. Cognitive psychologists call this position "functionalism," and they use it because it provides a precise language for describing complex information-processing architectures, without being limited to studying those few processes that neurophysiologists presently understand. (Eventually, of course, one wants to understand the neurophysiological processes that give rise to a cognitive program as well.) Cognitive scientists use the term "mind" solely to refer to an information-processing description of the functioning of the brain, and not in any colloquial sense.

II. The Mind as Scientist: General-Purpose Theories of Human Reasoning

Traditionally, cognitive psychologists have acknowledged that the mind (i.e., the information-processing structure of the brain) is the product of evolution, but their research framework was more strongly shaped by a different premise: that the mind was a general-purpose computer. They thought the function of this computer was self-evident: to discover the truth about whatever situation or problem it encountered. In other words, they started from the reasonable assumption that the procedures that governed human reasoning were there because they functioned to produce valid knowledge in nearly any context a person was likely to encounter.

They reasoned that if the function of the human mind is to discover truth, then the reasoning procedures of the human mind should reflect the methods by which truth can be discovered. Because science is the attempt to discover valid knowledge about the world, psychologists turned to the philosophy of science for *normative theories*—i.e., for theories specifying how one *ought* to reason if one is to produce valid knowledge. Their approach was to use the normative theories of what constitutes good scientific reasoning as a standard against which to compare actual human reasoning performance. The premise was that humans should be reasoning like idealized scientists about whatever situation they encountered, and the research question became: To what extent is the typical person's reasoning like an ideal scientist's?

The normative theories of how scientists—and hence the human mind—should be reasoning fall broadly into two categories: inductive reasoning and deductive reasoning. Inductive reasoning is reasoning from specific observations to general principles; deductive reasoning is reasoning from general principles to specific conclusions.

Ever since Hume, induction has carried a heavy load in psychology while taking a sound philosophical beating. In psychology, it has been the learning theory of choice since the British Empiricists argued that the experience of spatially and temporally contiguous events is what allows us to jump from the particular to the general, from sensations to objects, from objects to concepts. Many strands of psychology, including Pavlovian reflexology, Watsonian and Skinnerian behaviorism, and the sensory-motor parts of Piagetian structuralism, have been elaborations on the inductive psychology of the British Empiricists. Yet when Hume, a proponent of inductive inference as a psychological learning theory, donned his philosopher's hat, he demonstrated that induction could never justify a universal statement. To use a familiar example, no matter how many white swans you might see, you could never be justified in concluding "All swans are white," because it is always possible that the next swan you see will be black. Thus Hume argued that the inductive process whereby people were presumed to learn about the world could not ensure that the generalizations it produced would be valid.

Only recently, with the publication in 1935 of Karl Popper's *The Logic of Scientific Discovery*, has a logical foundation for psychology's favorite learning theory been provided. Popper argued that although a universal statement of science can never be proved true, it deductively implies particular assertions about the world, called hypotheses, and particular assertions can be proved false. Although no number of observed white swans can prove that "All swans are white" is true, just one black swan can prove it false. Generalizations cannot be confirmed, but they can be falsified, so inductions tested via deductions coupled with observations are on firmer philosophical ground than knowledge produced through induction alone.

This view had broad consequences for psychologists interested in learning. Psychologists who assumed that the purpose of human learning is to produce valid generalizations about the world reasoned that learning must be some form of Popperian hypothesis testing. Inductive reasoning must be used to generate hypotheses, and deductive reasoning coupled with observation must be used to try to falsify them. Furthermore, these reasoning procedures should be general-purpose: They should be able to yield valid inferences about any subject that one is interested in.

A broad array of cognitive psychologists such as Jean Piaget, Jerome Bruner, and Peter Wason, adopted a version of hypothesis testing—often an explicitly Popperian version—as their model of human learning. They used it to set the agenda of cognitive psychology in the 1950s and 1960s, and this view remains popular today. Some psychologists investigated inductive reasoning, by seeing whether people reason in accordance with the normative theories of inferential statistics; others investigated deductive reasoning, by seeing whether people reason in accordance with the rules of inference of the propositional calculus (formal propositional logic.

A. Deductive Reasoning

Psychologists became interested in whether the human mind included a "deductive component": mental rules that are the same as the rules of inference of the propositional calculus. They performed a wide variety of experiments to see whether people were able (1) to *recognize* the difference between a valid deductive inference and an invalid one, or (2) to *generate* valid conclusions from a set of premises. If people have a "deductive component," then they should be good at tasks like these. For example, in reasoning about conditional statements, one can make two valid inferences and two invalid inferences (see Fig. 1).

One of the most systematic bodies of work exploring the idea that people have reasoning procedures that embody the rules of inference of the propositional calculus was produced by Peter Wason and P. N. Johnson-Laird, together with their students and colleagues. Their research provides strong evidence that people do not reason according to the canons of formal propositional logic. For example:

(1) *Recognition of an argument as valid.* To see whether people are good at recognizing an argument as valid, psychologists gave them arguments like the ones in Fig. 1; for example, a subject might be asked to judge the validity of the the following argument: "If the object is rectangular, then it is blue; the object is rectangular; therefore the object is blue." In some of the experiments, unfamiliar conditionals were used; in others, familiar ones were used. These experiments indicated that people are good at recognizing the validity of a *modus ponens* inference, but they frequently think *modus tollens* is an invalid inference and that the two invalid inferences in Figure 1 are valid. Furthermore, they frequently view logically distinct conditionals as implying each other, and they have a pronounced tendency to judge an inference valid when they agree with the conclusion and invalid when they do not agree with the conclusion, regardless of its true validity.

(2) *Generating valid conclusions from a set of premises.* In other experiments, psychologists gave people sets of premises and asked them to draw conclusions from them. Many of the problems requiring the use of *modus ponens* were done incorrectly, and most of those requiring the use of *modus tollens* were done incorrectly. It may at first seem puzzling that people who are good at recognizing a *modus ponens* argument as valid would have trouble using *modus ponens* to generate a conclusion from premises. However, analogous experiences are common in everyday life: sometimes one cannot recall a person's name but can recognize it on a list. If humans all had rules of reasoning that mapped on to *modus ponens*, then they would be able both to generate a valid *modus ponens* inference and to recognize one. The fact that people cannot do both indicates that they lack this rule of reasoning. They may simply be able to recognize a contradiction when they see one, even though they cannot reliably generate valid inferences.

Valid inferences		Invalid inferences	
Modus ponens	Modus tollens	Affirming the Consequent	Denying the Antecedent
If P then Q P	If P then Q not-Q	If P then Q Q	If P then Q not-P
Therefore Q	Therefore not-P	Therefore P	Therefore not-Q

FIGURE 1 "P" and "Q" can stand for any proposition; for example, if "P" stands for "it rained" and "Q" stands for "the grass is wet," then the *modus ponens* inference above would read "If it rained, then the grass is wet; it rained, therefore the grass is wet." *Affirming the consequent* and *denying the antecedent* are invalid because the conditional "If P then Q" does not claim that P is the only possible antecedent of Q. If it did *not* rain, the grass could still be wet—the lawn could have been watered with a sprinkler, for example.

Perhaps the most intriguing and widely used experimental paradigm for exploring deductive reasoning has been the Wason selection task (see Fig. 2a). Peter Wason was interested in Popper's view that the structure of science was hypothetico-deductive. He wondered if learning were really hypothesis testing—i.e., the search for evidence that contradicts a hypothesis. Wason devised his selection task because he wanted to see whether people really do test a hypothesis by looking for evidence that could potentially falsify it. In the Wason selection task, a subject is asked to see whether a conditional hypothesis of the form "If P then Q" has been violated by any one of four instances, represented by cards.

A hypothesis of the form "If P then Q" is violated only when "P" is true but "Q" is false—the rule in Figure 2a, for example, can be violated only by a card that has a D on one side and a number other than 3 on the other side. Thus, one would have to turn over the "P" card (to see if it has a "not-Q" on the back) and the "not-Q" card (to see if it has a "P" on the back)—i.e., D and 7 for the rule in Figure 1a. The logically correct response, then, is always "P and not-Q."

Wason expected that people would be good at this. Nevertheless, he and many other psychologists have found that few people actually give this logically correct answer (<25% for rules expressing unfamiliar relations). Most people choose either the "P" card alone or "P and Q." Few people choose the "not-Q" card, even though a "P" on the other side of it would falsify the rule.

A wide variety of conditional rules that describe some aspect of the world ("descriptive" rules) have been tested; some of these have expressed relatively familiar relations, such as "If a person goes

a. Abstract Problem (AP)

Part of your new clerical job at the local high school is to make sure that student documents have been processed correctly. Your job is to make sure the documents conform to the following alphanumeric rule:

"If a person has a 'D' rating, then his documents must be marked code '3'."
(If P then Q)*

You suspect the secretary you replaced did not categorize the students' documents correctly. The cards below have information about the documents of four people who are enrolled at this high school. Each card represents one person. One side of a card tells a person's letter rating and the other side of the card tells that person's number code.

Indicate only those card(s) you definitely need to turn over to see if the documents of any of these people violate this rule.

D	F	3	7
(P)	(not-P)	(Q)	(not-Q)

b. Drinking Age Problem (DAP; adapted from Griggs & Cox, 1982)

In its crackdown against drunk drivers, Massachusetts law enforcement officials are revoking liquor licenses left and right. You are a bouncer in a Boston bar, and you'll lose your job unless you enforce the following law:

"If a person is drinking beer, then he must be over 20 years old."
(If P then Q)

The cards below have information about four people sitting at a table in your bar. Each card represents one person. One side of a card tells what a person is drinking and the other side of the card tells that person's age.

Indicate only those card(s) you definitely need to turn over to see if any of these people are breaking this law.

drinking beer	drinking coke	25 years old	16 years old
(P)	(not-P)	(Q)	(not-Q)

c. Structure of Social Contract (SC) Problems

It is your job to enforce the following law:

Rule 1 — Standard Social Contract (STD-SC): "If you take the benefit, then you pay the cost."
(If P then Q)

Rule 2 — Switched Social Contract (SWC-SC): "If you pay the cost, then you take the benefit."
(If P then Q)

The cards below have information about four people. Each card represents one person. One side of a card tells whether a person accepted the benefit and the other side of the card tells whether that person paid the cost.

Indicate only those card(s) you definitely need to turn over to see if any of these people are breaking this law.

	Benefit Accepted	Benefit NOT Accepted	Cost Paid	Cost NOT Paid
Rule 1 — STD-SC:	(P)	(not-P)	(Q)	(not-Q)
Rule 2 — SWC-SC:	(Q)	(not-Q)	(P)	(not-P)

FIGURE 2 Content effects on the Wason selection task. The logical structures of these three Wason selection tasks are identical; they differ only in propositional content. Regardless of content, the logical solution to all three problems is the same: To see if the rule has been violated, choose the "P" card (to see if it has a "not-Q" on the back) and choose the "not-Q" card (to see if it has a "P" on the back). Fewer than 25% of college students choose "P & not-Q" for the Abstract Problem (*a*), whereas about 75% choose both these cards for the Drinking Age Problem (*b*)—a familiar social contract. *c* shows the abstract structure of a social contract problem. A "look for cheaters" procedure would cause one to choose the "benefit accepted" card and the "cost NOT paid" card, regardless of which logical categories they represent. For Rule 1, these cards represent the values "P & not-Q," but for Rule 2 they represent the values "Q & not-P." Consequently, a person who was looking for cheaters would appear to be reasoning logically in response to Rule 1 but illogically in response to Rule 2. * The logical categories (P and Q) marked on the rules and cards here are only for the reader's benefit; they never appear on problems given to subjects in experiments.

to Boston, then he takes the subway" or "If a person eats hot chili peppers, then he will drink a cold beer." Others have expressed unfamiliar relations, such as "If you eat duiker meat, then you have found an ostrich eggshell" or "If there is an 'A' on one side of a card, then there is a '3' on the other side." In many experiments, performance on familiar descriptive rules is just as low as it is on unfamiliar ones; some familiar rules, however, do elicit a higher percentage of logically correct responses than unfamiliar ones. Even so, familiar descriptive rules typically elicit the logically correct response in fewer than half of the people tested. Recently, rules expressing causal relations have been tested; the pattern of results is essentially the same as for descriptive rules.

It is particularly significant that performance on the Wason selection task is so poor when the descriptive or causal rule tested is unfamiliar. If the function of our reasoning procedures is to allow us to discover new things about the world, then they must be able to function in novel—i.e., unfamiliar—situations. If they cannot be used in unfamiliar situations, then they cannot be used to learn anything new. Thus, the view that the purpose of human reasoning is to learn about the world is particularly undermined by the finding that people are not good at looking for violations of descriptive and causal rules, especially when they are unfamiliar.

B. Inductive Reasoning

The hypotheses that scientists test do not appear from thin air. Some of them are derived from theories; others come from observations of the world. For example, although no number of observations of white swans can prove that all swans are white, a person who has seen hundreds of white swans and no black ones may be more likely to think this hypothesis is worthy of investigation than a person who has seen only one white swan. The process of inferring hypotheses from observations is called inductive inference.

Using probability theory, mathematicians have developed a number of different normative theories of inductive inference, such as Bayes's theorem, null hypothesis testing, and Neyman–Pearsonian decision theory. These theories specify how scientists should make inferences from data to hypotheses. They are collectively known as inferential statistics.

A number of psychologists have studied the extent to which people's inductive reasoning conforms to the normative theories of inferential statistics and probability theory. One of the most extensive research efforts of this kind was spearheaded by Amos Tversky and Daniel Kahneman, along with their students and colleagues. They tested people's inductive reasoning by giving them problems in which they were asked to judge the probability of uncertain events. For example, a subject might be asked to reason about a diagnostic medical test: "If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming you know nothing about the person's symptoms or signs?" If the subject's answer is different from what a theory of statistical inference says it should be, then the experimenters conclude that our inductive-reasoning procedures do not embody the rules of that normative statistical theory.

The consensus among many psychologists is that this body of research demonstrates that (1) the human mind does not calculate the probability of events in accordance with normative probability theories, and (2) the human mind does not include information-processing procedures that embody the normative theories of inferential statistics. In other words, they conclude that the human mind is not innately equipped to do college-level statistics. Instead, these psychologists believe that people make inductive inferences using heuristics—cognitive shortcuts or rules of thumb. These heuristics frequently lead to the correct answer but can also lead to error precisely because they do not embody the formulas and calculational procedures of the appropriate normative theory. These psychologists also believe that humans suffer from systematic biases in their reasoning, which consistently lead to errors in inference.

Recently, however, a powerful critique by Gerd Gigerenzer, David Murray, and their colleagues has called this consensus into serious doubt. Their critique is both theoretical and empirical. Gigerenzer and Murray point out that the Tversky and Kahneman research program is based on the assumption that a statistical problem has only one correct answer; when the subject's response deviates from that answer, the experimenter infers that the subject is not reasoning in accordance with a normative statistical theory. However, Gigerenzer and Murray show that the problems subjects are typically asked to solve do not have only one correct answer. There are several reasons why this is true.

1. Statistics Does Not Speak With One Voice

There are a number of different statistical theories, and not all of them give the same answer to a problem. For example, although subjects' answers to certain problems have been claimed to be incorrect from the point of view of Bayes's theorem (but see below), their answers can be shown to be correct from the point of view of Neyman–Pearsonian decision theory. These subjects may be very good "intuitive statisticians," but simply applying a different normative theory than the experimenter is.

2. Concepts Must Match Exactly

For a particular statistical theory to be applicable, the concepts of the theory must match up precisely with the concepts in the problem. Suppose, for example, that you have some notion of how likely it is that a green cab or a blue cab would be involved in a hit-and-run accident at night. You are then told that there was a hit-and-run accident last night, and that a witness who is correct 80% of the time reported that it was a green cab. Bayes's theorem allows you to revise your prior probability estimate when you receive new information, in this case, the witness's testimony.

But what should your prior probability estimate (i.e., the estimate that you would make if you did not have the witness's testimony) be based on? It could, for example, be based on (1) the relative number of green and blue cabs in the city, (2) the relative number of reckless driving arrests for green versus blue cab drivers, (3) the relative number of drivers who have alcohol problems, or (4) the relative number of hit-and-run accidents they get into at night.

There is no normative theory for deciding which of these four kinds of information is the most relevant. Yet Bayes's theorem will generate different answers, depending on which you use. If subjects and experimenter differ in which kind of information they believe is most relevant, they will give different answers, even if each is correctly applying Bayes's theorem. Indeed, experimental data suggest that this happens. If one assumes that the subjects in these experiments were making certain very reasonable assumptions, then they *were* answering these questions correctly.

3. Structural Assumptions of the Theory Must Hold for the Problem

Assume that nature had selected for statistical rules; then it also should have selected for an assumption-checking program. For a particular statistical theory to be applicable, the assumptions of the theory must hold for the problem. For instance, a frequent assumption for applying Bayes's theorem is that a sample was randomly drawn. But in the real world there are many situations in which events are not randomly sampled: Diagnostic medical tests, for example, are rarely given to a random sample of people—instead, they are given only to those who already have symptoms of the disease. By their content, certain problems tested invited the inference that the random sampling assumption was violated; given this assumption, the "incorrect" answers subjects were giving were, in fact, correct. Indeed, in an elegant series of experiments, Gigerenzer and his colleagues showed that if one makes the random sampling assumption explicit to subjects, they do appear to reason in accordance with Bayes's theorem.

These experiments and theoretical critiques cast serious doubt on the conclusion that people are not good "intuitive statisticians." Evidence suggests that people are very good at statistical reasoning if the problem is about a real-world situation in which the structural assumptions of the theory hold. Their apparent errors may be because they are making assumptions about the problem that are different from the experimenters', or because they are consistently applying one set of statistical principles in one context and other sets in different contexts. What is clear from the research on inductive reasoning, however, is that the content of the problem matters, a theme we will return to in Section III.

C. Did We Evolve to Be Good Intuitive Scientists?

Good design is the hallmark of adaptation: To demonstrate that human reasoning evolved to fulfill a particular function, one must show that our reasoning procedures are well designed to fulfill that function. If the human mind was designed by natural selection to generate logically valid, scientifically justifiable knowledge about the world, then we ought to be good at drawing correct inductive and deductive inferences. Moreover, this ability ought to be context-independent, to allow us to learn about new, unfamiliar domains. After all, everything is initially unfamiliar.

But the data on deductive reasoning indicate that our minds do not include rules of inference that conform to the canons of deductive logic. The data on inductive reasoning indicate that we do not have inductive-reasoning procedures that operate independently of content and context. We may have inductive-reasoning procedures that conform to normative theories of statistical inference, but if we do, their application in any particular instance is extremely context-dependent, as the issues of conceptual and structural matching show. The evidence therefore suggests that we do not have formal, content-independent reasoning procedures. This indicates that the hypothesis that the adaptive function of human reasoning is to generate logically valid knowledge about the world is false.

III. The Mind as a Collection of Adaptations: Evolutionary Approaches to Human Reasoning

Differential reproduction is the engine that drives natural selection: If having a particular mental structure, such as a rule of inference, allows an animal to outreproduce other members of its species, then that mental structure will be selected for. Over many generations it will spread through the population until it becomes a universal, species-typical trait.

Consequently, alternative phenotypic traits are selected for not because they allow the organism to more perfectly apprehend universal truths, but because they allow the organism to outreproduce others of its species. Truth-seeking can be selected for only to the extent that it promotes reproduction. Although it might seem paradoxical to think that reasoning procedures that sometimes produce logically incorrect inferences might be more adaptive than reasoning that always leads to the truth, this will frequently be the case. Among other reasons, organisms usually must act before they have enough information to make valid inferences. In evolutionary terms, the design of an organism is like a system of betting: What matters is not each individual outcome, but the statistical average of outcomes over many generations. A reasoning procedure that sometimes leads to error, but that usually allows one to come to an adaptive conclusion (even when there is not enough information to justify it logically), may perform better than one that waits until it has sufficient information to derive a valid truth without error. Therefore, factors such as the cost of acquiring new information, asymmetries in the payoffs of alternative decisions (believing that a predator is in the shadow when it is not versus believing a predator is not in the shadow when it is), and trade-offs in the allocation of limited attention may lead to the evolution of reasoning procedures whose design is sharply at variance with scientific and logical methods for discovering truth.

Although organisms do not need to discover universal truths or scientifically valid generalizations to reproduce successfully, they do need to be very good at reasoning about important adaptive problems and at acquiring the kinds of information that will allow them to make adaptive choices in their natural environment. Natural selection favors mental rules that will enhance an animal's reproduction, whether they lead to truth or not. For example, rules of inference that posit features of the world that are usually (but not always) true may provide an adequate basis for adaptive decision-making. Some of these rules may be general-purpose: For example, the heuristics and biases proposed by Tversky and Kahneman are rules of thumb that will get the job done under the most commonly encountered circumstances. Their availability heuristic, for instance, is general-purpose insofar as it is thought to operate across domains: One uses it whether one is judging the frequency of murders in one's town or of words in the English language beginning with the letter "k." However, there are powerful reasons for thinking that many of these evolved rules will be special-purpose.

Traditionally, cognitive psychologists have assumed that the human mind includes only general-purpose rules of reasoning and that these rules are few in number. But natural selection is also likely to produce many mental rules that are specialized for reasoning about various evolutionarily important domains, such as cooperation, aggressive threat, parenting, disease avoidance, predator avoidance, and the colors, shapes, and trajectories of objects. This is because different adaptive problems frequently have different optimal solutions. For example, vervet monkeys have three major predators: leopards, eagles, and snakes. Each of these predators requires different evasive action: climbing a tree (leopard), looking up in the air or diving straight into the bushes (eagle), or standing on hind legs and looking into the grass (snake). Accordingly, vervets have a different alarm call for each of these three predators. A single, general-purpose alarm call would be less efficient because the monkeys would not know which of the three different evasive actions to take.

When two adaptive problems have different optimal solutions, a single general solution will be inferior to two specialized solutions. In such cases, a jack of all trades is necessarily a master of none, because generality can be achieved only by sacrificing efficiency.

The same principle applies to adaptive problems that require reasoning: There are cases where the rules for reasoning adaptively about one domain will lead one into serious error if applied to a different domain. Such problems cannot, in principle, be solved by a single, general-purpose reasoning procedure. They are best solved by different, special-purpose reasoning procedures. We will consider some examples of this below.

A. Internalized Knowledge and Implicit Theories

Certain facts about the world have been true for all of our species' evolutionary history and are critical to our ability to function in the world: The sun rises every 24 hours; space is locally three-dimensional; rigid objects thrown through space obey certain laws of kinematic geometry. Roger Shephard has argued that a human who had to learn these facts through the slow process of "trial and possibly fatal error" would be at a severe selective disadvantage compared to a human whose perceptual and cognitive system was designed in such a way that it already assumed that such facts were true. In an elegant series of experiments, Shepard showed that our perceptual–cognitive system has indeed internalized laws of kinematic geometry, which specify the ways in which objects move in three-dimensional Euclidean space. Our perceptual system seems to expect objects to move in the curvilinear paths of kinematic geometry so strongly that we see these paths even when they do not exist, as in the phenomenon of visual apparent motion. This powerful form of inference is specific to the motion of objects; it would not, for example, help you to infer whether a friend is likely to help you when you are in trouble.

Learning a relation via an inductive process that is truly general-purpose is not only slow, it is impossible in principle. There are an infinite number of dimensions along which one can categorize the world, and therefore an infinite number of possible hypotheses to test ("If my elbow itches, then the sun will rise tomorrow", "If a blade of grass grows in the flower pot, then a man will walk in the door"; i.e., "If P then Q," "If R then Q," "If S then Q" *ad infinitum*). The best a truly unconstrained inductive machine could do would be to randomly generate each of an infinite number of inductive hypotheses and deductively test each in turn.

Those who have considered the issue recognize that an organism could learn nothing this way. If any learning is to occur, then one cannot entertain all possible hypotheses. There must be constraints on which hypotheses one entertains, so that one entertains only those that are most likely to be true. This insight led Susan Carey and a number of other developmental psychologists to suggest that children are innately endowed with mental models of various evolutionarily important domains. Carey and her colleagues call these mental models *implicit theories*, to reflect their belief that all children start out with the same set of theories about the world, embodied in their thought processes.

These implicit theories specify how the world works in a given domain; they lead the child to test hypotheses that are consistent with the implicit theory, and therefore likely to be true (or at least useful). Implicit theories constrain the hypothesis space so that it is no longer infinite, while still allowing the child to acquire new information about a domain. Implicit theories are thought to be domain-specific because what is true of one domain is not necessarily true of another. For example, an implicit theory that allows one to predict a person's behavior if one knows that person's beliefs and desires will not allow one to predict the behavior of falling rocks, which have no beliefs and desires. The implicit-theory researchers have begun to study children's implicit theories about the properties of organisms, the properties of physical objects and motion, the use of tools, and the minds of others.

B. Reasoning about Prescriptive Social Conduct

The reasoning procedures discussed so far function to help people figure out what the world is like and how it works. They allow one to acquire knowledge that specifies what kind of situation one is facing from one moment to the next. For example, a rule such as "If it rained last night, then the grass will be wet this morning" purports to describe the way the world is. Accordingly, it has a truth value: A descriptive rule can be either true or false. In contrast, a rule such as "If a person is drinking beer, then that person must be over 21 years old" does not describe the way things are. It does not even describe the way existing people behave. It *prescribes*: It communicates the way some people want other people to behave. One cannot assign a truth value to it.

From an evolutionary perspective, knowledge about the world is just a means to an end, and that end is behaving adaptively. Once an organism knows what situation it is in, it has to know how to act, so reasoning about the facts of the world should be paired with reasoning about appropriate conduct. For this reason, the mind should have evolved rules of reasoning that specify what one ought to do in various situations—rules that prescribe behavior. Because different kinds of situations call for different kinds of behavior, these rules should be situation-specific. For example, the rules for reasoning about cooperation should differ from those for reasoning about aggressive threat, and both should differ from the rules for reasoning about the physical world. Recent research by Cosmides & Tooby, Manktelow & Over, and others has explored such rules.

Social exchange, for example, is cooperation between two or more people for mutual benefit, such as the exchange of favors between friends. Humans in all cultures engage in social exchange, and the paleoanthropological record indicates that such cooperation has probably been a part of human evolutionary history for almost 2 million years. Game-theoretic analyses by researchers such as Robert Trivers, Robert Axelrod, and W. D. Hamilton have shown that cooperation cannot evolve unless people are good at detecting "cheaters" (people who accept favors or benefits without reciprocating). Given a social contract of the form "If you take the benefit, then you pay the cost," a cheater is someone who took the benefit but did not pay the required cost (see Fig. 2c). Detecting cheaters is an important adaptive problem: A person who was consistently cheated would be incurring reproductive costs, but receiving no compensating benefits. Such individuals would dwindle in number, and eventually be selected out of the population.

Rules for reasoning about descriptive relations would lead one into serious error if applied to social contract relations. In the previous discussion of the Wason selection task, we saw that the logically correct answer to a descriptive rule is "P and not-Q," no matter what "P" and "Q" stand for (i.e., no matter what the rule is about). But this definition of violation differs from the definition of cheating on a social contract. A social contract rule has been violated whenever a person has taken the benefit without paying the required cost, *no matter what logical category these actions correspond to.* For the social contract expressed in Rule 1 of Figure 2c, a person who was looking for cheaters would, by coincidence, produce the logically correct answer. This is because the "benefit accepted" card and the "cost NOT paid" card correspond to the logical values "P" and "not-Q," respectively, for Rule 1. But for the social contract expressed in Rule 2, these two cards correspond to the logical values "Q" and "not-P"—a logically incorrect answer. The logically correct answer to Rule 2 is to choose the "cost paid" card, "P," and the "benefit NOT accepted" card, "not-Q." Yet a person who has paid the cost cannot possibly have cheated, nor can a person who has not accepted the benefit.

Thus, for the social contract in Rule 2, the adaptively correct answer is logically incorrect, and the logically correct answer is adaptively incorrect. If the only reasoning procedures that our minds contained were the general-purpose rules of inference of the propositional calculus, then we could not, in principle, reliably detect cheating on social contracts. This adaptive problem can be solved only by inferential procedures that are specialized for reasoning about social exchange.

The Wason selection task research discussed previously showed that we have no general-purpose ability to detect violations of conditional rules—unfamiliar descriptive and causal rules elicit the logically correct response from <25% of subjects. But when a conditional rule expresses a social contract, people are very good at detecting cheaters. Approximately 75% of subjects choose the "benefit accepted" card and the "cost NOT paid" card, regardless of which logical category they correspond to and regardless of how unfamiliar the social contract rule is. This research indicates that the human mind contains reasoning procedures that are specialized for detecting cheaters on social contracts. Recently, the same experimental procedures have been used to investigate reasoning about aggressive threat. Although there is only one way to violate the terms of a social contract, there are two ways of violating the terms of a threat: Either the person making the threat can be bluffing (i.e., he does not carry out the threat, even though the victim refuses to comply), or he can be planning to double-cross the person he is threatening (i.e., the victim complies with his demand, but the threatener punishes him anyway). The evidence indicates that people are good at detecting both bluffing and double-crossing. Similarly, two British researchers, Kenenth Manktelow and David Over, have found that people are very good at detecting violations of "precaution rules." Precaution rules specify what precautions should be taken to avoid danger in hazardous situations.

Situations involving social contracts, threats, and precaution rules have recurred throughout human evolutionary history, and coping with them successfully constituted powerful selection pressures. An individual who cannot cooperate, cannot avoid danger, or cannot understand a threat is at a powerful selective disadvantage in comparison to those who can. More important, what counts as a violation differs for a social contract rule, a threat, and a precaution rule. Because of this difference, the same reasoning procedure cannot be successfully applied to all three situations. As a result, there cannot be a general-purpose reasoning procedure that works for all of them. If these problems are to be solved at all, they must be solved by specialized reasoning procedures. Significantly, humans do reason successfully about these problems, suggesting that natural selection has equipped the human mind with a battery of functionally specialized reasoning procedures, designed to solve specific, recurrent adaptive problems.

IV. Summary

Reasoning procedures are an important part of how organisms adapt. Adaptive behavior depends on adaptive inferences to regulate decisions. Although initial approaches within psychology to the study of human reasoning uncovered many interesting phenomena, the search for a few, general rules of reasoning that would account for human-reasoning performance and explain how humans cope with the world was largely unsuccessful. The recent emergence of an evolutionary perspective within cognitive psychology has led to a different view of how inference in the human mind is organized. Instead of viewing the mind as a general-purpose computer, employing a few general principles that are applied uniformly in all contexts, an evolutionary perspective suggests that the mind consists of a larger collection of functionally specialized mechanisms, each consisting of a set of reasoning procedures designed to efficiently solve particular families of important adaptive problems. In the last decade, a growing body of research results has validated this approach, indicating that humans have specialized procedures for reasoning about such things as the motion of objects, the properties of living things, cooperation, threat, and avoiding danger.

Bibliography

Carey, S. (1985). "Conceptual Change in Childood." MIT Press, Cambridge, Massachusetts.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition* **31,** 187–276.

Cosmides, L., and Tooby, J. (1989). Evolutionary psychology and the generation of culture, part II. Case study: A computational theory of social exchange. *Ethol. Sociobiol.* **10,** 51–97.

Dawkins, R. (1976). "The Selfish Gene." Oxford University Press, Oxford.

Dawkins, R. (1986). "The Blind Watchmaker." Norton, New York.

Evans, J. St. B. T. (ed.) "Thinking and Reasoning." Routledge and Kegan Paul, London.

Gigerenzer, G., Hell, W., and Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *J. Exp. Psych.: Hum. Percep. Perform.* **14,** 513–525.

Gigerenzer, G., and Murray, D. (1987). "Cognition as Intuitive Statistics." LEA, Hillsdale, New Jersey.

Johnson-Laird, P. N. (1982). Thinking as a skill. *Q. J. Exp. Psych.* **34A,** 1–29.

Kahneman, D., Slovic, P., and Tversky, A. (1982). "Judgment Under Uncertainty: Heuristics and Biases." Cambridge University Press, Cambridge.

Keil, F. (1989). "Concepts, Kinds, and Cognitive Development." MIT Press, Cambridge, Massachusetts.

Manktelow, K. I., and Over, D. E. (1987). Reasoning and rationality. *Mind Lang.* **2,** 199–219.

Manktelow, K. I., and Over, D. E. (1990). Deontic thought and the selection task. *In* "Lines of Thinking," Vol. 1, (K. J. Gilhooly, M. T. G. Keane, R. H. Logie, and G. Erdos, eds.). Wiley, New York.

Shepard, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychol. Rev.* **91,** 417–447.

Wason, P., and Johnson-Laird, P. N. (1972). "Psychology of Reasoning: Structure and Content." Harvard University Press, Cambridge, Massachusetts.

Williams, G. C. (1966). "Adaptation and Natural Selection." Princeton University Press, Princeton.