Annu. Rev. Psychol. 1999. 50:109–35

DEDUCTIVE REASONING

P. N. Johnson-Laird

Department of Psychology, Princeton University, Princeton, New Jersey 08544;
e-mail: phil@clarity.princeton.edu

KEY WORDS: deduction, logic, rules of inference, mental models, thinking

ABSTRACT

This chapter describes the main accounts of deductive competence, which explain *what* is computed in carrying out deductions. It argues that people have a modicum of competence, which is useful in daily life and a prerequisite for acquiring logical expertise. It outlines the three main sorts of theory of deductive performance, which explain *how* people make deductions: They rely on factual knowledge, formal rules, or mental models. It reviews recent experimental studies of deductive reasoning in order to help readers to assess these theories of performance.

CONTENTS

0084-6570/99/0201-0109$08.00

INTRODUCTION

Reasoning is a process of thought that yields a conclusion from percepts, thoughts, or assertions. The process may be one of which reasoners are painfully aware or of which they are almost unconscious. But it is a systematic process if it is reasoning, as opposed to, say, daydreaming.

This chapter is concerned with one sort of reasoning, deduction. By definition, deduction yields *valid* conclusions, which must be true given that their premises are true, e.g.:

> If the test is to continue, then the turbine must be rotating fast enough.
> The turbine is not rotating fast enough.
> Therefore the test is *not* to continue.

Some deductions are difficult, and the failure to draw this particular valid conclusion probably contributed to the Chernobyl disaster. Despite such mistakes, the business of life depends on the ability to make deductions. Individuals differ in this ability, and those who are better at it—at least as measured by intelligence tests—appear to be more successful. If so, it is not surprising. A person who is poor at reasoning is liable to blunder in daily life. Conversely, without deduction, there would be no logic, no mathematics, and no Annual Review articles.

Psychologists have studied reasoning for a century. Not until Piaget, however, did anyone purport to explain how people were able to make deductions. In his account of the genesis of knowledge, he argued that children spontaneously recapitulate the history of mathematics and arrive at formal reasoning in early adolescence (e.g. Beth & Piaget 1966). By the mid-1970s, researchers assumed that even though Piaget's theory might not be viable in detail, it was right on the grand scale. People were equipped with a mental logic. The task for psychologists—so they thought—was to delineate its principles. This approach ignored an unsettling discovery made by Wason (1966). Intelligent adults in his "selection" task regularly committed a logical error. He laid out four cards in front of them:

> A B 2 3

They knew that each card had a letter on one side and a number on the other side. He showed them a conditional rule: If a card has the letter A on one side, then it has the number 2 on the other side. He then asked them to select those cards that had to be turned over to discover whether the rule was true or false about the four cards. Most people selected the A card alone, or the A and 2 cards. What was puzzling was their failure to select the 3 card: If it has an A on its other side, the rule is false. Indeed, nearly everyone judges it to be false in that case. Yet when Wason changed the content to a sensible everyday generalization, many people made the correct selections (Wason & Johnson-Laird 1972). Insofar as the task is deductive, mental logic is stymied by these effects of content, which have no bearing on its logic.

So much for the historical introduction to the present review. Its plan is simple. It begins with accounts of what the mind is computing when it makes deductions, that is, accounts of *deductive competence*. It then describes theories of how the mind carries out these computations, that is, theories of *deductive performance*. The controversy among these theories is hot, so the chapter reviews recent experiments to enable readers to make up their own minds about deduction.

RATIONALITY AND DEDUCTIVE COMPETENCE

Naive individuals, who have no training in logic, may err in tests of deductive reasoning yet achieve their goals in daily life. This discrepancy is the fundamental paradox of rationality. Psychologists react to it in several different ways, each of which yields a different account of logical competence (for a philosophical analysis, see Engel 1991). This section reviews these accounts, which are couched at the "computational" level—they characterize *what* is computed but not *how* the process is carried out.

One reaction to the paradox is that people are wholly rational but the psychological tests do not reflect their competence. "I have never found errors," Henle (1978) wrote, "which could unambiguously be attributed to faulty reasoning." The philosopher LJ Cohen (1981) has concurred that in such cases there is a glitch in an information-processing mechanism. The strength of this view is that it explains how it was possible for humans to invent logic. Its weakness is that it makes little sense of genuine inferential blunders (for catalogs of the irrational, see Sutherland 1992, Piattelli-Palmarini 1994). Deduction is not tractable: As the number of premises increases, any system of reasoning will eventually run out of time and memory before it reaches a conclusion (Cook 1971). Unless the brain somehow bypasses computational constraints, reasoning is bounded (Simon 1982). Perfect rationality is for the angels.

A different reaction to the paradox is that logic is the wrong normative theory (pace Piaget, Devlin 1997). It permits inferences that no sane individual (other than a logician) is liable to draw, e.g.:

Ann is here
Therefore Ann is here or Ben is here, or both.

Inferences of this sort are valid, but naive individuals balk at them, presumably because their conclusions are less informative than their premises (Johnson-Laird & Byrne 1991). Likewise, in logic, if a conclusion follows from premises, then no subsequent premise can invalidate it. But in human reasoning, subsequent information can undermine a deduction. Philosophers and artificial intelligencers formulate such systems of "defeasible" or "nonmonotonic" reasoning (e.g. Harman 1986, Brewka et al 1997), but psychologists do not know how people reason in this way (Chater & Oaksford 1993).

If logic is the wrong normative theory, what should replace it? One strategy, advocated by Anderson (1990), is to make a "rational analysis" of the domain. The analysis calls for a model of the environment in which the cognitive system operates, because, for Anderson, rationality *is* an optimal adaptation to the environment. Evolutionary psychologists assume that natural selection is the main engine of mental adaptation. They abandon general deductive competence in favor of specialized inferential modules, such as a module for "checking for cheaters," which are supposed to have evolved because they conferred a selective advantage on our hunter-gatherer ancestors (Cosmides 1989). But psychologists cannot go back to the Stone Age to test natural selection at work on shaping the mind. The best they can do is to use ethological and paleontological evidence to speculate about what was adaptive for our ancestors (Cummins 1996, Mithen 1996). The strength of the approach is that it explains success in the laboratory where a task taps into a hypothesized module and failure where it does not (Gigerenzer & Hug 1992). But what do evolutionary theorists do if a test fails to corroborate a module (cf Girotto et al 1989, Pollard 1990)? They can allow, like Pinker (1997), that not all modules are adaptations, or they can rewrite their "just so" story to fit the facts (Simon 1991). No result can jeopardize evolutionary psychology. It may be a useful heuristic for generating hypotheses, but it is irrefutable.

Another reaction to the paradox is that people have two sorts of rationality (Evans & Over 1996): rationality_1 is a tacit competence for coping with life's problems, and rationality_2 is a conscious mechanism for normative reasoning. The commentaries on Evans & Over (1997) are a cross section of the views of leading researchers. Some accept the dichotomy (e.g. Ball et al 1997, Santamaria 1997), and some reject it and argue for a unitary competence based to a first approximation either on rationality_1 (e.g. Cummins 1997, Hertwig et al 1997) or on rationality_2 (e.g. Noveck 1997, Ormerod 1997). There is a history of similar dichotomies, particularly between tacit and conscious reasoning, and "gut reactions" and deliberation. Sloman (1996) links the two sorts of reasoning to associative and rule-based thinking. The strong point of the dichotomy is that it makes sense of both competence and incompetence in life and the laboratory. Its weakness is that it may accommodate too much.

The various approaches to competence each have their strengths and weaknesses. Perhaps the following conception distills only their strong points. Naive individuals have a modicum of deductive competence (Johnson-Laird & Byrne 1991). The persuasiveness of an inference depends on the credibility of its premises and on the proportion of situations that satisfy both them and the conclusion. The process of reasoning can lead reasoners to abandon a premise or assumption or to abandon an inference as invalid. Conclusions, however, do not throw semantic information away by adding disjunctive alternatives to the premises. They are more parsimonious than their premises, do not repeat those premises that can be taken for granted, and assert a relevant proposition that was not stated explicitly in the premises. When no conclusion meets these constraints, then naive individuals tend to respond, "nothing follows." Logically, the response is wrong, because infinitely many valid conclusions follow from any premises.

The modicum of rationality is important to achieving the goals of everyday life and crucial for technical expertise in logic. But it is only a modicum, so it explains why people make mistakes in reasoning, particularly in the laboratory, but sometimes in life. And it explains why tests of reasoning predict academic success, and how it was possible for humanity to invent logic. The challenge to theories of performance is accordingly to accommodate a modicum of competence.

THEORIES OF DEDUCTIVE PERFORMANCE

This section outlines the three main schools of thought about deductive performance. The first school bases performance on factual knowledge. The second school bases it on a system of formal rules of inference akin to those of a logical calculus. The third school bases it on mental models akin to those of the semantic theory of a logical calculus. In short, deduction is controversial: It may rely on knowledge, formal rules, or mental models, or on some mixture of them (Falmagne & Gonsalves 1995).

Deduction as a Process Based on Factual Knowledge

Psychologists propose that the mind uses content-specific conditional rules to make inferences from general knowledge. The proposal is part of two seminal theories of cognitive architecture: Anderson's (1993) ACT theory and Newell's (1990) SOAR theory, in which the rules, or productions, as they are known, are triggered by the current contents of working memory and then carry out various actions. These actions may, in turn, add new information to working memory and in this way yield a chain of inferences.

Knowledge plays its most specific role in the theory that reasoning is based on memories of previous inferences (e.g. Riesbeck & Schank 1989, Kolodner 1993). Indeed, according to this theory of "case-based" reasoning, human thinking has nothing to do with logic. What happens is that one inference calls to mind another—a procedure that is useful in artificial intelligence (Smyth & Keane 1995). When an activity has been repeated often enough, however, it begins to function like a content-specific rule (cf Eisenstadt & Simon 1997).

Knowledge certainly enters into everyday deductions, but whether it is represented by rules or specific cases is an open question. It might, after all, be represented by *assertions* in a mental language rather than by rules. It might even have a distributed representation that has no explicit structure (Shastri & Ajjanagadde 1993), though it is not known whether distributed systems can acquire the full compositional semantics needed for the representation of assertions.

One drawback with all knowledge-based theories is that they offer no immediate explanation of the ability to reason about the unknown. Even if you know nothing about atonality, you can make the following deduction:

If it's atonal then it's nondeterministic.
It's atonal.
Therefore it's nondeterministic.

This abstract deductive competence is necessary for logic and mathematics. Knowledge-based theories do not explain it, so we turn to alternative accounts.

Deduction as a Formal, Syntactic Process

The idea that deductive performance depends on formal rules of inference is remarkably pervasive. It goes back to the ancient doctrine that the laws of logic are the laws of thought. It was championed by Piaget (e.g. Beth & Piaget 1966), and it underlies several current psychological theories (Nisbett 1993). Rips (1994) and the late Martin Braine (1998) and his colleagues (e.g. Braine & O'Brien 1991) proposed that reasoners extract the logical forms of premises and use rules to derive conclusions. Rips's theory is implemented in a computer program (known as PSYCOP). Both theories have rules for sentential connectives, such as "if" and "or," and for quantifiers, such as "all" and "some." Both are based on a method in logic known as natural deduction, so they have rules for introducing, and for eliminating, sentential connectives. The following rule of *modus ponens*, for instance, eliminates "if":

If A then B
A
Therefore B.

A key feature of natural deduction is the use of suppositions, which are assumptions made for the sake of argument. One way to use a supposition is to show that together with the premises it leads to a contradiction and must therefore be false (*reductio ad absurdum*). Thus, consider the following proof:

1. If the test is to continue then the turbine is rotating.
2. The turbine isn't rotating.
3. The test is to continue. (*A supposition*)
4. The turbine is rotating. (*Modus ponens* applied to 1 and 3)

There is a contradiction between a premise (The turbine isn't rotating) and a conclusion derived from the supposition (The turbine is rotating). The rule of *reductio ad absurdum* negates the supposition:

5. Therefore the test is *not* to continue.

Theorists could adopt a single rule for this *modus tollens* inference, but it is more difficult than *modus ponens*, so they assume instead that it depends on the chain of inferential steps. The choice of rules in a theory is thus an empirical matter.

The two main problems in developing a formal rule theory are to ensure that it is computationally viable and that it fits the empirical data. An example of a computational difficulty is that the rule for introducing "and" can run wild, e.g.

A
B
Therefore A and B
Therefore A and (A and B)
Therefore A and [A and (A and B)]

and so on ad infinitum. Two sorts of rules are dangerous: those that introduce a connective and those that make suppositions. One way to curb a rule is to build its effects into another rule that is safe to use (Braine 1994). Alternatively, the dangerous rule can be restricted to chains of inference working backward from a given conclusion to the premises. Rips embodies this idea in PSYCOP, which has forward rules, backward rules, and rules used in either direction.

Rips (1994) advances a Deduction-System hypothesis: Formal rules of inference underlie not just deduction but also the cognitive architecture of the mind. His rules can be used as a programming language in which to implement, for example, a production system. The problem with the Deduction-System hypothesis is illustrated by Eisenstadt & Simon's (1997) counterarguments that a production system underlies cognitive architecture (see also Newell 1990) and that formal rules can be derived from them. The critical question is, What do formal rules contribute to cognitive architecture, as opposed to, say, productions? This issue must be distinguished from the predictions made from the particular use of the rules in "programming" thinking, because what can be programmed by using rules can also be programmed by using productions, the lambda calculus, or any other universal basis for computation. Until this question is answered, it is difficult to test different theories of cognitive architecture.

Deduction as a Semantic Process Based on Mental Models

Human beings can understand the meaning of assertions, envisage the corresponding situations, and ascertain whether a conclusion holds in them. The theory of mental models accordingly postulates that reasoning is based not on syntactic derivations from logical forms but on manipulations of mental models representing situations (Johnson-Laird & Byrne 1991, Polk & Newell 1995). Models can represent the world (e.g. Glasgow 1993), simulate a process (e.g. Hegarty 1992), and yield inductive or deductive inferences (for reviews, see Rogers et al 1992, Oakhill & Garnham 1996).

If deduction depends on models, then the process is semantic because their construction from discourse depends on meaning and knowledge (Glenberg & Langston 1992, Stevenson 1993, Garnham & Oakhill 1996). Each mental model represents a *possibility*, and its structure and content capture what is common to the different ways in which the possibility might occur. For example, when individuals understand a conjunction such as "There is triangle and there is circle," they represent its meaning, from which they can construct a mental model of the situation to which it refers:

o   Δ,

in which o represents a circle, and Δ represents a triangle. The model captures what is common to any situation in which a triangle and a circle are present. Such a model may, or may not, give rise to an image, but models are distinct from images. They can contain abstract elements, such as negation, that cannot be visualized (Johnson-Laird & Byrne 1991).

The theory gives a unified account of deductions about what is possible, probable, and necessary. Given the truth of the premises, a conclusion is *possible* if it holds in at least one model of the premises; its *probability* depends on the proportion of models in which it holds; and it is *necessary* if it holds in all the models of the premises. The model theory gives an account of connectives and quantifiers, including such quantifiers as "most" and "few," and a variety of other sorts of constructions, such as spatial, temporal, and causal relations, and counterfactual conditionals (Johnson-Laird & Byrne 1991, Byrne 1997).

A fundamental assumption of the theory is *the principle of truth*:

> Individuals minimize the load on working memory by tending to construct mental models that represent explicitly only what is true, and not what is false.

The principle applies at two levels. First, individuals represent only true possibilities. Second, for each true possibility, they represent only those literal propositions in the premises—affirmative or negative—that are true (Barres & Johnson-Laird 1997). For example, an inclusive disjunction, such as "There is a circle and/or there is *not* a triangle," elicits three alternative models to represent the three true possibilities:

o
¬Δ
o ¬Δ,

where each row denotes a model of a separate possibility and ¬ represents negation (see Newell 1990, for a defense of such representations). Each model represents only what is true in a particular possibility. Hence, the model in the first row represents that there is a circle, but it does not represent explicitly that it is false that there is not a triangle in this case. Similarly, the second model does not represent explicitly that it is false that there is a circle in this case. Reasoners have only a modicum of competence—they try to remember what is false, but these "mental footnotes" on models are soon forgotten.

In contrast to mental models, *fully explicit* models represent the false components in each possibility. The fully explicit models of the inclusive disjunction above are as follows:

o Δ
¬o ¬Δ
o ¬Δ

In fully explicit models, false affirmatives are represented by true negations, and false negatives are represented by true affirmatives. The three fully explicit models match the three rows that are true in the truth table of the inclusive disjunction.

The mental models for a conditional

If there is a circle then there is a triangle

include a model of the salient case (the presence of a circle and a triangle). People realize that there are possibilities in which the antecedent of the conditional (there is a circle) is false. They do not normally represent them explicitly; instead they represent them in a single implicit model that has no content, symbolized here by an ellipsis:

o Δ
. . .

Readers will notice the similarity to the model of the conjunction. The difference is that the conditional has an implicit model allowing for possibilities in which its antecedent is false. A biconditional

If, and only if, there is a circle, then there is a triangle

has the same mental models as a conditional, but its implicit model corresponds to the possibility in which both the antecedent and the consequent of the conditional are false. If reasoners remember the "mental footnotes" about what is false in the implicit model, then they can construct a set of fully explicit models for a conditional or biconditional. But they soon forget these mental footnotes, especially with assertions that contain more than one connective.

THE PHENOMENA OF DEDUCTIVE REASONING

The previous section reviewed the three principal approaches to deductive performance—the use of knowledge, formal rules, and mental models. The present section turns to the evidence bearing on these theories. The study of deductive performance has burgeoned in recent years (see Evans et al 1993, Garnham & Oakhill 1994, and the new journal *Thinking & Reasoning*). To help readers make sense of the plethora of studies, and to reach their own judgments about the theories of performance, the review is organized in terms of the different branches of deduction.

Reasoning with Sentential Connectives

Most evidence in favor of formal rule theories comes from studies of sentential connectives. According to these theories, the greater the number of steps in a derivation, the harder the inference should be, though other factors cause difficulty, such as the accessibility and complexity of rules. Braine et al (1984) asked naive reasoners to rate the difficulty of each inference in a large set. They used the data to estimate the accessibility of each of their postulated rules and then combined these estimated parameters to predict the ratings for each inference in the set. The fit was satisfactory. Johnson-Laird et al (1992a) used a program to count the overall number of mental models required for each inference, and these parameter-free predictions correlated with the ratings just as well as the post hoc estimates of Braine et al. Bonatti (1994) argued that counting models was inconsistent with the normal use of models. In fact, the theory is not usually used in correlational studies, and it is not clear how else it could yield predictions about the data.

Rips (1994) reported several experiments with sentential connectives, including a study of a sample of inferences in which the participants evaluated given conclusions. His theory fitted the results satisfactorily. It also accounted for the times that the participants took to understand explicit proofs and for their memory of proofs: They remember sentences in the same domain as the premises better than those in a subdomain based on a supposition. Other studies show that individuals automatically make logical inferences in interpreting text (Lea et al 1990, Lea 1995).

There are robust differences in the difficulty of reasoning with the various connectives. Conjunctions are easier than conditionals, which in turn are easier than disjunctions (Johnson-Laird et al 1992a, Schaeken et al 1995). Likewise, exclusive disjunctions (two mental models) are easier than inclusive disjunctions (three mental models; Johnson-Laird et al 1992a). Certain diagrams—those that make alternative possibilities more explicit—both speed up and improve reasoning with disjunctions, but exclusive disjunctions remain easier than inclusive disjunctions (Bauer & Johnson-Laird 1993). These results are predicted by the model theory from the number of mental models required for the respective connectives (for a global experimental confirmation, see Klauer & Oberauer 1995). Formal rule theories contain no machinery for *predicting* differences from one connective to another.

One result at first sight refutes the model theory. Rips (1994) found that an inference of the following form based on a conjunction,

A and B
If A then C
If B then C
Therefore C,

was no easier to evaluate than one based on a disjunction,

A or B
If A then C
If B then C
Therefore C.

Contrary to the model theory, PSYCOP predicts that the conjunctive inference should be more difficult than the disjunctive inference, but Rips (1994:368) claims that the disjunctive rule may be “harder for subjects to apply.” Madruga et al (1998) argue that Rips’s null result is a ceiling effect. They corroborated the model theory’s predictions when reasoners drew their own conclusions from these premises. They also showed that when the premises are presented with the conditionals first, or one at a time on a computer screen, the results corroborate the model theory even in the evaluation task. Reasoners’ strategies may differ among the different tasks. They develop various high-level strategies in tackling inferential problems, and these strategic principles are just as important as the underlying reasoning mechanisms (Byrne & Handley 1997; M Bucciarelli & PN Johnson-Laird, submitted). The model theory allows for a variety of strategies (Byrne et al 1995), but rule theories rely on a single deterministic strategy.

Conditional Reasoning

Studies of conditionals continue to focus on their four main inferences (e.g. Dugan & Revlin 1990, Evans et al 1995):

If A then B, A, therefore B. (*modus ponens*)
If A then B, not B, therefore not A. (*modus tollens*)
If A then B, B, therefore A. (*affirming the consequent premise*)
If A then B, not A, therefore not B. (*denying the antecedent premise*)

Affirming the consequent and denying the antecedent are valid only if the conditional premise is interpreted as a biconditional; otherwise, they are fallacies. Formal rule theorists argue that there are no rules for the fallacies, because they can be suppressed by a second conditional premise, e.g. the premises

If he went fishing then he had fish for supper.
If he bought some fish then he had fish for supper.
He had fish for supper.

suppress the tendency to infer "He went fishing." Byrne (1989), however, showed that an additional conditional premise can also suppress *valid* inferences, e.g. the premises

If he went fishing then he had fish for supper.
If he caught some fish then he had fish for supper.
He went fishing.

suppress the tendency to infer "He had fish for supper." By parity of argument, Byrne claimed, there is no mental rule for *modus ponens*. Politzer & Braine (1991) replied that the additional conditional falsifies the first conditional. But Byrne (1991) demonstrated suppression without affecting its believability (see also RMJ Byrne, O Espino, and C Santamaria, submitted). The tendency to make conditional inferences is influenced by knowledge undermining the sufficiency, or the necessity, of the antecedent to bring about the consequent (Thompson 1994, 1995). Markovits and his colleagues demonstrated subtle effects of this sort, asking the reasoners themselves to generate their own instances of these conditions (Markovits & Vachon 1990; see also Cummins et al 1991). Stevenson & Over (1995) showed that lowering the credibility of a conditional suppresses inferences from it, and they argued that the sufficiency of an antecedent to bring about a consequence is represented by the proportion of models in which the antecedent holds but the consequent does not.

Modus ponens is easier than *modus tollens* (cf the Chernobyl inference in the Introduction). Formal rules and mental models both account for the difference. But Girotto et al (1997) used the model theory to predict a surprising result. *Modus tollens* is easier when the categorical premise is presented first than when it is presented second. Presented first, it provides an initial negative model

$\neg$ B

so that reasoners should be more likely to flesh out their models of the conditional, If A then B, to include the cases where the antecedent is false. RMJ

Byrne & A Tasso (submitted) used the model theory to predict another surprising result. *Modus tollens* is drawn more often from counterfactual than from factual conditionals. A counterfactual of the form "If A had happened then B would have happened" calls for models of both the counterfactual situation (A B) and the factual situation (¬A ¬B) that the conditional presupposes. Formal rule theories predict neither phenomena.

Evans et al (1996; see also Evans 1993) proposed a useful correction to the model theory. It had postulated that a negative assertion was likely to elicit models of both the assertion and its corresponding unnegated proposition (Johnson-Laird & Byrne 1991). Thus, a conditional of the form "If not A then B" was supposed to elicit the models

¬ A   B
  A

The assumption was made to account for "matching" bias, which is the tendency to ignore negatives in conditionals in matching them to categorical premises or situations (Evans 1998; cf Oaksford & Stenning 1992). Evans and his colleagues discovered that the key to the phenomenon is the ease of grasping that one proposition refutes another. The task is easier if one proposition explicitly negates the other than if the relation is merely implicit, e.g. "The number is 4" refutes "The number is 9." Hence, a conditional of the form, If not A then B, should elicit the models,

¬ A   B
    . . .

The model theory can thus revert to the general principle of truth (cf Evans et al 1998).

One intriguing developmental trend is that very young children treat conditionals as conjunctions; slightly older children treat them as biconditionals, and adolescents and adults are able to treat them as one-way conditionals (see e.g. Taplin et al 1974). If this result is reliable (cf Russell 1987, Markovits 1993), it is a nice confirmation of the model theory: Conjunctions have one fully explicit model, biconditionals have two, and one-way conditionals have three. VM Sloutsky & BJ Morris (submitted) observed that children tend to ignore the second clause of a compound premise, so the premise calls for exactly one model. They are most likely to ignore the second clause of a tautology or contradiction, less likely to ignore it in a disjunction, and least likely to ignore it in a conjunction. Even adults treat conditionals as conjunctions in complex premises. They consider an assertion such as "If A then 2 or if B then 3" to be true in just those cases that they consider the assertion "A and 2, or B and 3" to be true (Johnson-Laird & Barres 1994). The mental models for the two sorts of assertion are identical.

Ormerod and his colleagues have pioneered the study of immediate inferences from one sort of conditional to another, and from conditionals to disjunctions, and vice versa (Ormerod et al 1993, Richardson & Ormerod 1997). They argue that their results can best be explained in terms of a revised model theory. Reasoners construct a minimal set of models of a premise to infer the required paraphrase. Minimal sets of models reduce the load on working memory, and they may underlie reasoning in general.

Reasoning About Relations

Some simple deductions depend on relations, e.g.:

Ann is better than Beth.
Cath is worse than Beth.
Who is worst?

Early studies of these inferences suggested that reasoners construct mental models (De Soto et al 1965, Huttenlocher 1968), though they are also affected by linguistic variables (Clark 1969). Recent studies show that deductions depending on one model are easier than those depending on multiple models. Consider the following spatial problem (from Byrne & Johnson-Laird 1989), which in the original concerned common objects such as knives, forks, and spoons:

A is on the right of B.
C is on the left of B.
D is in front of C.
E is in front of B.
What is the relation between D and E?

The model theory predicts that reasoners should construct a single two-dimensional model of the premises:

```
C  B  A
D  E
```

The model supports the conclusion that D is on the left of E, which no model of the premises refutes. In contrast, the problem

B is on the right of A.
C is on the left of B.
D is in front of C.
E is in front of B.
What is the relation between D and E?

calls for two distinct models because the premises do not relate A and C:

C	A	B		A	C	B
D		E			D	E

Both models yield the same answer: D is on the left of E, so it is valid. But the inference depends on more than one model, so it should be harder than the first one. Formal rule theories predict that there will be no difference between the two problems: Their first premise is irrelevant, and the derivations depend solely on the remaining premises, which are identical in the two problems. In fact, reasoners drew a greater percentage of correct conclusions to the one-model problems than to the multiple-model problems. Rips (1994:415) counters that the instructions may bias people to use images and that the irrelevant premises may sidetrack them. Why they should be sidetracked more by the second problem than the first is unclear. Braine (1994:245) concedes that "much reasoning does use mental models."

Schaeken et al (1996a) replicated the difference between one-model and multiple-model problems in inferences about temporal relations, such as "John takes a shower before he drinks coffee." The difference also occurred when temporal order was implicit only in the use of tense and aspect (Schaeken et al 1996b) and when neither sort of problem contained an irrelevant premise (Schaeken et al 1998). Vandierendonck & De Vooght (1996) confirmed the difference in temporal and spatial problems. They also showed that a task that preoccupies visuo-spatial memory interferes with these inferences (Vandierendonck & De Vooght 1997). This finding contrasts with earlier results (Gilhooly et al 1993, Toms et al 1993) but corroborates an investigation of sentential and spatial reasoning (Klauer et al 1997).

Syllogisms and Reasoning with Quantifiers

Quantifiers underlie many deductions, including syllogisms, such as

Some actuaries are businessmen.
All businessmen are conformists.
Therefore some actuaries are conformists.

One controversy about syllogisms is whether individuals reason or merely select a conclusion that matches the superficial form (or "mood") of a premise (Wetherick & Gilhooly 1990; N Chater & M Oaksford, submitted). Another controversy is whether, if they reason, they use formal rules, Euler circles, or mental models. Stenning & Yule (1997) framed equivalent normative theories of the three sorts, so the real issue is to account for errors.

Rips (1994) fitted his rules to experimental data. He also showed how the theory might account for the results of a study in which the participants drew their own conclusions (Johnson-Laird & Bara 1984): PSYCOP implies that they guessed tentative conclusions to work backward from them. Yang (1997) and Yang et al (1998) fitted Braine's (1998) rules to the rated difficulty of inferences depending on quantifiers and connectives. Yang (personal communication), however, showed that the numbers of tokens in the mental models for these inferences give just as good a fit, even though they do not require the estimate of parameters from the data.

Euler circles represent a premise of the form "All A are B" with two separate diagrams: In one diagram, a circle A lies wholly within a circle B to represent that set A is properly included within set B; in the other diagram, the two circles coincide to represent that the two sets are coextensive. Analogous topological relations between the two circles represent the other sorts of premises. The traditional method calls for the construction of all the different diagrams for each premise and all their different combinations for a pair of premises—a demand that leads to a combinatorial explosion. Stenning and his colleagues, however, devised a way to use Euler circles that obviates the explosion (e.g. Stenning & Yule 1997). Ford (1995) postulates a similar procedure: Reasoners use the verbal premises as reminders of which areas within the circles cannot be empty. Individuals who use Euler circles can thus avoid the traditional method in favor of one resembling the use of mental models (cf Cardaci et al 1996). One puzzle is whether people who have never seen circles used to represent sets spontaneously invent Euler circles. The main disadvantage of the method is that it does not generalize to relational inferences, such as

> All horses are animals.
> Therefore all horses' heads are animals' heads.

The model theory of quantified reasoning accommodates relations. It postulates that individuals build an initial model from the premises, formulate a conclusion from this model, and, if they are prudent, search for counterexamples to the conclusion (Johnson-Laird & Byrne 1991). The theory predicts the majority of erroneous conclusions, which correspond to the initial models of multiple-model syllogisms (Bara et al 1995). It therefore provides an alternative explanation for the so-called atmosphere effects (Shaw & Johnson-Laird 1998), i.e. the alleged tendency to draw conclusions that merely match the superficial form of a premise.

The model theory has been criticized on several grounds. Ford (1995) has argued vigorously that some reasoners use formal rules, others use Euler circles, but no one uses mental models (for different results, see M Bucciarelli & PN Johnson-Laird, submitted). Ford formulated a set of formal rules for verbal substitutions, but they are equivalent to a model-based procedure proposed by Johnson-Laird & Bara (1984). The outward signs of substitutions are silent on whether their inward occurrence is in sentences or models.

Polk & Newell (1995) claimed that the search for counterexamples appeared to underlie few predictions: Their own model theory gave a better account of individual differences when it dropped this component. With hindsight, syllogisms are not ideal for demonstrating a search for counterexamples. *Modal* reasoning is better, because the model theory predicts a key interaction hinging on a search for counterexamples: It should be easier to determine that a situation is possible (one model of the premises suffices as an example) than not possible (all the models of the premises must be tested), whereas it should be easier to determine that a situation is not necessary (one model serving as a counterexample suffices) than that it is necessary (all models must be tested). The interaction has been corroborated in reasoning with sentential connectives (Bell & Johnson-Laird 1998) and with quantifiers (JStBT Evans, SE Handley, CNJ Harper & PN Johnson-Laird, submitted). Hence, in tasks where counterexamples are useful, reasoners appear to search for them. Indeed, Barwise (1993) emphasized that the only way to *know* that a conclusion is invalid is by constructing a model of the premises that is a counterexample to it.

Another criticism is that the model theory fails to specify how reasoners search for counterexamples (Martín-Cordero & González-Labra 1994). Various computer programs implementing the theory contain a search procedure (cf Hardman 1996, for a critique of an earlier program), so the problem was to obtain relevant evidence. M Bucciarelli & PN Johnson-Laird (submitted) therefore devised a technique to externalize thinking: The participants had to construct *external* models using cut-out shapes. They were able to construct counterexamples, that is, models of the premises that refuted invalid conclusions. In drawing their own valid conclusions, they constructed more models for multiple-model syllogisms than for one-model syllogisms. In both tasks, they used the same search operations as the program, but it seriously underestimated the variety of different reasoning strategies and the variety of different interpretations of the premises (see also Langford & Hunting 1994). The construction of external models—let alone counterexamples—is beyond the scope of current formal rule theories.

Inferences can depend on relations and multiple quantifiers, e.g.:

> Some of the Avon letters are in the same place as all the Bury letters.
> All the Bury letters are in the same place as all the Caton letters.
> Therefore some of the Avon letters are in the same place as all the Caton letters.

Johnson-Laird et al (1989) showed that one-model problems of this sort yield a greater proportion of correct conclusions than multiple-model problems. Greene (1992) pointed out a possible confounding: The multiple-model problems called for conclusions of the form "None of the Avon letters are in the same place as *some* of the Caton letters," which are difficult for people to formulate. He argued that there was therefore no need to postulate the use of mental models. However, the multiple-model problems also support straightforward conclusions, such as "Some of the Caton letters are not in the same place as any of the Avon letters." The model theory also predicted the participants' erroneous conclusions: They mainly corresponded to the initial models of the premises (Johnson-Laird et al 1992b).

The Effects of Content on Deduction

If deduction depends on formal rules, the content of inferences cannot affect the *process* of reasoning, which is purely syntactic. But if deduction depends on models, then content can affect the process, because reasoners can stop searching for alternative models of the premises if their initial models yield believable conclusions. Most investigations of the effects of believability concern syllogisms, such as

All of the Frenchmen are wine drinkers.
Some of the wine drinkers are gourmets.
Therefore some of the Frenchmen are gourmets.

Many early studies were methodologically flawed, but recent research has established three main phenomena (Evans 1989, Newstead & Evans 1993). First, reasoners accept more valid conclusions than invalid conclusions. Second, they accept more believable than unbelievable conclusions. Many people, for example, draw the conclusion in the syllogism above (Oakhill et al 1990). It is invalid, but highly believable. Third, the effects of believability are greater on invalid inferences than on valid ones. But when invalid conclusions are not consistent with the premises, then the interaction disappears (Newstead et al 1992).

There are several possible accounts for the phenomena. Reasoners may rely on mental models. They may not realize that a conclusion that is merely consistent with the premises does not thereby follow from them (the *misinterpreted necessity* hypothesis). They may accept believable conclusions and only assess the validity of unbelievable ones (the *selective scrutiny* hypothesis). As Quayle & Ball (1997) showed, however, the model theory can be reconciled with misinterpreted necessity. Necessary conclusions are supported by models, and impossible conclusions are refuted by them. But if a conclusion is merely possible, then reasoners may be uncertain and so be more likely to accept believable conclusions and to reject unbelievable conclusions (cf Hardman & Payne 1995). This account is supported by reasoners' confidence ratings: They are highly confident in valid conclusions, which is a good index of syllogistic competence.

One study has examined the effects of believability on conditional reasoning (Santamaria et al 1998). A pair of believable conditionals can yield a believable or an unbelievable conclusion, as in

> If Marta is hungry, then she takes an afternoon snack.
> If Marta takes an afternoon snack, then she has a light dinner.
> Therefore if Marta is hungry, then she has a light dinner.

The conclusion lacks the causal link (eating the snack), so it violates the normal relation between hunger and dinner. Reasoners were more likely, and quicker, to draw a valid conclusion when it was believable than when it was unbelievable. Hence, there may be a process of "filtering out" valid conclusions that are not believable (Oakhill et al 1990).

The Selection Task

The selection task (see the Introduction) has launched a thousand studies, but the literature has grown faster than knowledge. Selections are sensitive to the probability of encountering a falsifying instance, and the introduction of such Bayesian considerations has revitalized research (Oaksford & Chater 1994, 1996; Kirby 1994; Nickerson 1996). This normative approach—inspired by Anderson's (1990) "rational analysis"—accounts for many phenomena, but there is, as yet, no corresponding theory of the mental processes underlying performance. Naive individuals are unlikely to be explicitly calculating, say, the expected gain in information from selecting a card.

A change in the content of the selection task can yield a striking improvement in performance, particularly with *deontic* conditionals concerning what is permissible (Cheng & Holyoak 1985, Kroger et al 1993). Manktelow & Over (1991) used the deontic conditional, "If you tidy your room then you may go out to play," and the participants' selections depended on whose point of view they were asked to take. The mother's concern is that her child does not cheat, and those with her point of view tended to select the cards did-not-tidy and played. Her son's concern is that his mother does not renege on the deal, and those with his point of view tended to select the cards tidied and did-not-play. Even children are sensitive to point of view (Light et al 1990), and adults with a neutral point of view tend to select all four cards (Politzer & Nguyen-Xuan 1992).

Theorists debate the causes of these effects (see e.g. Holyoak & Cheng 1995 and the commentaries thereon). Cheng & Holyoak (1985) argued that a deontic conditional maps onto a "pragmatic reasoning schema" such as

> If the action (e.g. playing) is to be taken, then the precondition must be satisfied (e.g. tidying the room).

Cosmides (1989) proposed that there is an innate module for "checking for cheaters" (see also Gigerenzer & Hug 1992). But Manktelow & Over (1995) argued against both of these positions and in favor of mental models. However, the model theory, they say, underestimates the importance of probabilities, preferences, and pragmatics.

Because the selection task is sensitive to Bayesian considerations, one might conclude that it does not depend on deductive reasoning. The claim is premature. On the one hand, individuals of higher cognitive ability tend to make deductively correct selections in versions of the task with an abstract content (Stanovich 1998). On the other hand, deductions can yield probabilistic conclusions, and such deductions can be accounted for by the model theory: The probability of an event depends on the proportion of models in which it occurs (Johnson-Laird & Savary 1996, Johnson-Laird et al 1998). The model theory predicts that any manipulation that emphasizes what would falsify the rule should improve performance in the selection task. Such effects do occur, even in tasks that are neither deontic nor concern checking for cheaters. Instructions to check for *violations* improved performance in the abstract task (Platt & Griggs 1993, Griggs 1995, Dominowski 1995). Green (1995) showed that instructions to envisage counterexamples also improved performance (see also Green & Larking 1995). In work that brings together the model theory and Bayesian considerations, Green et al (1997) showed that reasoners' assessments of how likely they were to encounter a counterexample (in four stacks of cards) predicted their selections (see also Green 1997). Sperber et al (1995) used a more indirect procedure to render counterexamples more *relevant*—in the sense of Sperber & Wilson (1995)—and thereby improved performance. Love & Kessler (1995) used a context that suggested the possibility of counterexamples, and Liberman & Klar (1996) demonstrated that apparent effects of "checking for cheaters" are better explained in terms of the participants' grasp of appropriate counterexamples and of the relevance of looking for them.

Systematic Fallacies in Reasoning

The controversy between rules and models has been going on for a long time. Roberts (1993) suggested that it was neither fruitful nor resolvable. Yet most criticisms of formal rules concern their power to explain empirical results, whereas most criticisms of mental models concern alleged shortcomings in the formulation of the theory—it is unfalsifiable, or it is obviously false; it relies on images too much, or it does not rely on images enough; it relies on formal rules or it needs to rely on formal rules. These mutually inconsistent criticisms are strange, especially for a theory systematically implemented in computer programs. They may reflect the difficulty of grasping a different and nondeterministic paradigm.

Evidence in favor of rules and against models is hard to find. The best case rests on studies of hemispherical differences. The lack of right-hemisphere involvement in reasoning about easily visualized content has been taken to jeopardize the model theory (see Wharton & Grafman 1998 for a review). But as these authors point out, the studies are preliminary, and the results count more against the claim that models are images—a claim that Johnson-Laird & Byrne (1991) repudiate—rather than the theory that reasoning depends on models.

Is there any evidence in favor of models and against rules? Preceding sections discussed several cases, but we turn now to a phenomenon that may be decisive. The principle of truth, which underlies the model theory, has a surprising consequence. It implies that human beings reason in a systematically fallacious way. Consider the following problem:

> Only one of the following premises is true about a particular hand of cards:
> There is a king in the hand or there is an ace, or both.
> There is a queen in the hand or there is an ace, or both.
> There is a jack in the hand or there is a 10, or both.
> Is it possible that there is an ace in the hand?

The mental models of the first premise are

King	
	Ace
King	Ace

They support the conclusion that an ace is possible, and most people draw this conclusion (Johnson-Laird & Goldvarg 1997). They fail to take into account that when one premise is true, the others are false. Hence, if the first premise is true, the second premise is false, so there cannot be an ace. Indeed, if there were an ace in the hand, then *two* of the premises would be true, contrary to the rubric that only one of them is true.

Reasoners are vulnerable to a variety of fallacies in modal and probabilistic reasoning (Johnson-Laird & Savary 1996). The rubric "Only one of the premises is true" is equivalent to an exclusive disjunction, and a compelling illusion occurs in the following inference about a particular hand of cards:

> If there is a king in the hand then there is an ace in the hand, or else if there isn't a king in the hand then there is an ace in the hand.
> There is a king in the hand.
> What, if anything, follows?

Nearly everyone, expert and novice alike, infers that there is an ace in the hand. It follows from the mental models of the premises. Yet it is a fallacy granted a disjunction, exclusive or inclusive, between the two conditionals. Hence, one or other of the conditionals could be false. If, say, the first conditional is false, then there is no guarantee that there is an ace in the hand even though there is a king. Granted that the fallacies arise from a failure to reason about what is false, any manipulation that emphasizes falsity should alleviate them. An experiment that used the rubric "Only one of the following two premises is false" did reliably reduce their occurrence (Tabossi et al 1998).

No current formal rule theory predicts the fallacies: These theories rely solely on valid inferential procedures (see e.g. Rips 1997). Perhaps a simple procedural assumption can at least accommodate the fallacies post hoc. If not, then they may resolve the controversy. They do not imply, however, that all possible formal rule theories are wrong. No result could do so. As the Deductive-System hypothesis shows, formal rules can be used to program any computable theory, including the mental model theory itself.

CONCLUSIONS

Deduction is under intense investigation. The two dominant accounts of its underlying mechanisms are based on rules and on models—a distinction that echoes the contrast in logic between "proof" theory and "model" theory. Rule theorists are impressed by the automatic ease with which we make certain inferences. They formulate rules that correspond to these elementary deductions, and they assume that more difficult inferences call for chains of elementary deductions. In contrast, what strikes model theorists is that reasoning is just the continuation of comprehension by other means. They notice that arguments are seldom laid out as proofs and that public reasoning is often dialectical. We are all better critics of other people's inferences than of our own. We recognize the force of counterexamples, but we more readily construct models that reflect our own views than find refutations of them (Baron 1990, Legrenzi et al 1993).

Both theoretical approaches have a penumbra of protective stratagems. Yet both have testable consequences. Psychology is difficult, but it is not impossible. The evidence suggests that naive reasoners have a modicum of deductive competence based on mental models. In principle, rules and models are not incompatible. Indeed, advanced reasoners may learn to construct formal rules for themselves—a process that ultimately leads to the discipline of logic.

Literature Cited

Anderson JR. 1990. *The Adaptive Character of Thought*. Hillsdale, NJ: Erlbaum

Anderson JR. 1993. *Rules of the Mind*. Hillsdale, NJ: Erlbaum

Ball LJ, Sutherland EJ, Quayle JD. 1997. Situating deductive competence within a models framework: some further considerations. *Curr. Psychol. Cogn.* 16:52–62

Bara B, Bucciarelli M, Johnson-Laird PN. 1995. The development of syllogistic reasoning. *Am. J. Psychol.* 108:157–93

Baron J. 1990. *Thinking and Deciding*. New York: Cambridge Univ. Press. 2nd ed.

Barres PE, Johnson-Laird PN. 1997. Why is it hard to imagine what is false? In *Proc. 19th Annu. Conf. Cogn. Sci. Soc.,* pp. 475–78. Mahwah, NJ: Erlbaum

Barwise J. 1993. Everyday reasoning and logical inference. *Behav. Brain Sci.* 16:337–38

Bauer MI, Johnson-Laird PN. 1993. How diagrams can improve reasoning. *Psychol. Sci*. 4:372–78

Bell V, Johnson-Laird PN. 1998. A model theory of modal reasoning. *Cogn. Sci.* In press

Beth EW, Piaget J. 1966. *Mathematical Epistemology and Psychology.* Dordrecht: Reidel

Bonatti L. 1994. Propositional reasoning by model? *Psychol. Rev.* 101:725–33

Braine MDS. 1994. Mental logic and how to discover it. In *The Logical Foundations of Cognition*, ed. J Macnamara, GE Reyes, pp. 241–63. New York: Oxford Univ. Press

Braine MDS. 1998. Steps towards a mental predicate logic. See Braine & O'Brien 1998, pp. 273–331

Braine MDS, O'Brien DP. 1991. A theory of If: a lexical entry, reasoning program and pragmatic principles. *Psychol. Rev.* 98: 182–203

Braine MDS, O'Brien DP, eds. 1998. *Mental Logic.* Mahwah, NJ: Erlbaum

Braine MDS, Reiser BJ, Rumain B. 1984. Some empirical justification for a theory of natural propositional logic. In *The Psychology of Learning and Motivation,* 18:313–71. New York: Academic

Brewka G, Dix J, Konolige K. 1997. *Nonmonotonic Reasoning: An Overview.* Stanford, CA: CLSI Publ., Stanford Univ.

Byrne RMJ. 1989. Suppressing valid inferences with conditionals. *Cognition* 31: 61–83

Byrne RMJ. 1991. Can valid inferences be suppressed? *Cognition* 39:71–78

Byrne RMJ. 1997. Cognitive processes in counterfactual thinking about what might have been. In *The Psychology of Learning and Motivation, Advances in Research and Theory*, ed. DL Medin, 37:105–54. San Diego, CA: Academic

Byrne RMJ, Handley SJ. 1997. Reasoning strategies for suppositional deductions. *Cognition* 62:1–49

Byrne RMJ, Handley SJ, Johnson-Laird PN. 1995. Reasoning from suppositions. *Q. J. Exp. Psychol.* 48A:915–44

Byrne RMJ, Johnson-Laird PN. 1989. Spatial reasoning. *J. Mem. Lang.* 28:564–75

Cardaci M, Gangemi A, Pendolino G, Di Nuovo S. 1996. Mental models vs. integrated models: explanations of syllogistic reasoning. *Percept. Mot. Skills* 82: 1377–78

Chater N, Oaksford M. 1993. Logicism, mental models and everyday reasoning. *Mind Lang.* 8:72–89

Cheng PN, Holyoak KJ. 1985. Pragmatic reasoning schemas. *Cogn. Psychol.* 17: 391–416

Clark HH. 1969. Linguistic processes in deductive reasoning. *Psychol. Rev.* 76: 387–404

Cohen LJ. 1981. Can human irrationality be experimentally demonstrated? *Behav. Brain Sci.* 4:317–70

Cook SA. 1971. The complexity of theorem proving procedures. In *Proc. 3rd Annu. ACM Symp. Theory Computing*, pp. 151–58. Arlington, VA: Assoc. Comput. Mach.

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

Cummins DD. 1996. Evidence for the innateness of deontic reasoning. *Mind Lang.* 11: 160–90

Cummins DD. 1997. Rationality: biological, psychological, and normative theories. *Curr. Psychol. Cogn.* 16:78–86

Cummins DD, Lubart T, Alksnis O, Rist R. 1991. Conditional reasoning and causation. *Mem. Cogn.* 19:274–82

De Soto LB, London M, Handel LS. 1965. Social reasoning and spatial paralogic. *J. Pers. Soc. Psychol.* 2:513–21

Devlin K. 1997. *Goodbye Descartes: The End of Logic and the Search for a New Cosmology of the Mind*. New York: Wiley

Dominowski RL. 1995. Content effects in Wason's selection task. See Newstead & Evans 1995, pp. 41–65

Dugan CM, Revlin R. 1990. Response options and presentation format as contributors to conditional reasoning. *Q. J. Exp. Psychol.* 42A:829–48

Eisenstadt SA, Simon HA. 1997. Logic and thought. *Minds Mach.* 7:365–85

Engel P. 1991 (1989). *The Norm of Truth: An Introduction to the Philosophy of Logic.* Toronto: Univ. Toronto Press

Evans JStBT. 1989. *Bias in Human Reasoning: Causes and Consequences*. Hillsdale, NJ: Erlbaum

Evans JStBT. 1993. The mental model theory of conditional reasoning: critical appraisal and revision. *Cognition* 48:1–20

Evans JStBT. 1998. Matching bias in conditional reasoning: Do we understand it after 25 years? *Think. Reasoning* 4:45–82

Evans JStBT, Clibbens J, Rood B. 1995. Bias in conditional inference: implications for mental models and mental logic. *Q. J. Exp. Psychol.* 48A:644–70

Evans JStBT, Clibbens J, Rood B. 1996. The role of implicit and explicit negation in conditional reasoning bias. *J. Mem. Lang.* 35:392–404

Evans JStBT, Legrenzi P, Girotto V. 1998. The influence of linguistic form on reasoning: the case of matching bias. *Q. J. Exp. Psychol.* In press

Evans JStBT, Newstead SE, Byrne RMJ. 1993. *Human Reasoning: The Psychology of Deduction.* Mahwah, NJ: Erlbaum

Evans JStBT, Over DE. 1996. *Rationality and Reasoning*. Hove, East Sussex, UK: Psychology Press

Evans JStBT, Over DE. 1997. Rationality in reasoning: the problem of deductive competence. *Curr. Psychol. Cogn.* 16:3–38

Falmagne RJ, Gonsalves J. 1995. Deductive inference. *Annu. Rev. Psychol.* 46:525–59

Ford M. 1995. Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition* 54:1–71

Garnham A, Oakhill JV. 1994. *Thinking and Reasoning*. Oxford: Blackwell

Garnham A, Oakhill JV. 1996. The mental models theory of language comprehension. In *Models of Understanding Text*, ed. BK Britton, AC Graesser, pp. 313–39. Hillsdale, NJ: Erlbaum

Gigerenzer G, Hug K. 1992. Domain specific reasoning: social contracts, cheating, and perspective change. *Cognition* 43:127–71

Gilhooly KJ, Keane MTG, Logie RH, Erdos G, eds. 1990. *Lines of Thinking,* Vol. 1. New York: Wiley

Gilhooly KJ, Logie RH, Wetherick NE, Wynn V. 1993. Working memory and strategies in syllogistic reasoning tasks. *Mem. Cogn.* 21:115–24

Girotto V, Blaye A, Farioli F. 1989. A reason to reason: pragmatic basis of children's search for counterexamples. *Eur. Bull. Cogn. Psychol.* 9:297–321

Girotto V, Mazzocco A, Tasso A. 1997. The effect of premise order in conditional reasoning: a test of the mental model theory. *Cognition* 63:1–28

Glasgow JI. 1993. Representation of spatial models for geographic information systems. In *Proc. ACM Workshop on Advances in Geographic Information Systems*, ed. N Pissinou, pp. 112–17. Arlington, VA: Assoc. Comput. Mach.

Glenberg AM, Langston WE. 1992. Comprehension of text: pictures help to build mental models. *J. Mem. Lang.* 31:129–51

Green DW. 1995. Externalization, counterexamples and the abstract selection task. *Q. J. Exp. Psychol.* 48A:424–46

Green DW. 1997. Hypothetical thinking in the selection task: amplifying a model-based approach. *Curr. Psychol. Cogn.* 16:93–102

Green DW, Larking R. 1995. The locus of facilitation in the abstract selection task. *Think. Reasoning* 1:183–99

Green DW, Over D, Pyne R. 1997. Probability and choice in the selection task. *Think. Reasoning* 3:209–35

Greene S. 1992. Multiple explanations for multiply quantified sentences: Are multiple models necessary? *Psychol. Rev.* 99: 184–87

Griggs RA. 1995. The effects of rule clarification, decision justification, and selection instruction on Wason's abstract selection task. See Newstead & Evans 1995, pp. 17–39

Hardman DK. 1996. Mental models: The revised theory brings new problems. *Behav. Brain Sci.* 19:542–43

Hardman DK, Payne SJ. 1995. Problem difficulty and response format in syllogistic reasoning. *Q. J. Exp. Psychol.* 48A: 945–75

Harman G. 1986. *Change in View: Principles of Reasoning*. Cambridge, MA: Bradford Books/MIT Press

Hegarty M. 1992. Mental animation: inferring motion from static displays of mechanical systems. *J. Exp. Psychol.: Learn. Mem. Cogn.* 18:1084–102

Henle M. 1978. Foreword to *Human Reasoning*, ed. R Revlin, RE Mayer. Washington, DC: Winston

Hertwig R, Ortmann A, Gigerenzer G. 1997. Deductive competence: a desert devoid of content and context. *Curr. Psychol. Cogn.* 16:102–7

Holyoak KJ, Cheng PW. 1995. Pragmatic reasoning with a point of view. *Think. Reasoning* 1:289–313

Huttenlocher J. 1968. Constructing spatial images: a strategy in reasoning. *Psychol. Rev.* 75:286–98

Johnson-Laird PN, Bara B. 1984. Syllogistic inference. *Cognition* 16:1–61

Johnson-Laird PN, Barres PE. 1994. When 'or' means 'and': a study in mental models. *Proc. 16th Annu. Conf. Cogn. Sci. Soc.,* pp. 475–78. Hillsdale, NJ: Erlbaum

Johnson-Laird PN, Byrne RMJ. 1991. *Deduction*. Hillsdale, NJ: Erlbaum

Johnson-Laird PN, Byrne RMJ, Schaeken WS. 1992a. Propositional reasoning by model. *Psychol. Rev.* 99:418–39

Johnson-Laird PN, Byrne RMJ, Tabossi P. 1989. Reasoning by model: the case of multiple quantification. *Psychol. Rev.* 96: 658–73

Johnson-Laird PN, Byrne RMJ, Tabossi P. 1992b. In defense of reasoning: a reply to Greene. *Psychol. Rev.* 99:188–90

Johnson-Laird PN, Goldvarg Y. 1997. How to make the impossible seem possible. In *Proc. 19th Annu. Conf. Cogn. Sci. Soc.,* pp. 354–57. Mahwah, NJ: Erlbaum

Johnson-Laird PN, Legrenzi P, Girotto V, Legrenzi MS, Caverni J-P. 1998. Naive probability: a mental model theory of extensional reasoning. *Psychol. Rev.* In press

Johnson-Laird PN, Savary F. 1996. Illusory inferences about probabilities. *Acta Psychol.* 93:69–90

Kirby KN. 1994. Probabilities and utilities of fictional outcomes in Wason's four-card selection task. *Cognition* 51:1–28

Klauer KC, Oberauer K. 1995. Testing the mental model theory of propositional reasoning. *Q. J. Exp. Psychol.* 48A:671–87

Klauer KC, Stegmaier R, Meiser T. 1997. Working memory involvement in propositional and spatial reasoning. *Think. Reasoning* 3:9–47

Kolodner J. 1993. *Case-Based Reasoning*. San Mateo, CA: Morgan Kaufman

Kroger K, Cheng PW, Holyoak KJ. 1993. Evoking the permission schema: the impact of explicit negation and a violation-checking context. *Q. J. Exp. Psychol.* 46A: 615–35

Langford PE, Hunting R. 1994. A representational communication approach to the development of inductive and deductive logic. In *Intelligence, Mind, and Reasoning,* ed. A Demetriou, A Efklides, pp. 191–209. New York: Elsevier

Lea RB. 1995. On-line evidence for elaborative inferences in text. *J. Exp. Psychol. Learn. Mem. Cogn*. 21:1469–82

Lea RB, O'Brien DP, Fisch SM, Noveck IA, Braine MDS. 1990. Predicting propositional-like inferences in text-comprehension. *J. Mem. Lang.* 29:361–87

Legrenzi P, Girotto V, Johnson-Laird PN. 1993. Focussing in reasoning and decision making. *Cognition* 49:37–66

Liberman N, Klar Y. 1996. Hypothesis testing in Wason's selection task: social exchange cheating detection or task understanding. *Cognition* 58:127–56

Light PH, Girotto V, Legrenzi P. 1990. Children's reasoning on conditional promises and permissions. *Cogn. Dev.* 5:369–83

Love R, Kessler C. 1995. Focussing in Wason's selection task: content and instruction effects. *Think. Reasoning* 1:153–82

Madruga JAG, Moreno S, Carriedo N, Gutiérrez F. 1998. Task, premise order and strategies in Rips's conjunction-disjunction and conditionals problems. In *Deductive Reasoning and Strategies*, ed. W Schaeken. In press

Manktelow KI, Over DE. 1991. Social rules and utilities in reasoning with deontic conditionals. *Cognition* 39:85–105

Manktelow KI, Over DE. 1995. Deontic reasoning. See Newstead & Evans 1995, pp. 91–114

Markovits H. 1993. The development of conditional reasoning: a Piagetian reformulation of mental models theory. *Merrill-Palmer Q.* 39:133–60

Markovits H, Vachon R. 1990. Conditional reasoning, representation, and level of abstraction. *Dev. Psychol.* 26:942–51

Martín-Cordero J, González-Labra MJ. 1994. Amnesic mental models do not completely spill the beans of deductive reasoning. *Behav. Brain Sci.* 17:773–74

Mithen S. 1996. *The Prehistory of Mind: The Cognitive Origins of Art, Religion and Science*. London: Thames & Hudson

Newell A. 1990. *Unified Theories of Cognition*. Cambridge. MA: Harvard Univ. Press

Newstead SE, Evans JStBT. 1993. Mental models as an explanation of belief bias effects in syllogistic reasoning. *Cognition* 46:93–97

Newstead SE, Evans JStBT, eds. 1995. *Perspectives on Thinking and Reasoning: Essays in Honour of Peter Wason.* Hillsdale, NJ: Erlbaum

Newstead SE, Pollard P, Evans JStBT, Allen J. 1992. The source of belief bias in syllogistic reasoning. *Cognition* 45:257–84

Nickerson R. 1996. Hempel's paradox and Wason's selection task: logical and psychological puzzles of confirmation. *Think. Reasoning* 2:1–31

Nisbett RE, ed. 1993. *Rules for Reasoning*. Hillsdale, NJ: Erlbaum

Noveck IA. 1997. Deductive competence need not be problematic: some comments on Evans and Over's proposal concerning human rationality. *Curr. Psychol. Cogn.* 16: 162–72

Oakhill J, Garnham A, eds. 1996. *Mental Models in Cognitive Science*. Hove, East Sussex, UK: Psychology Press

Oakhill J, Garnham A, Johnson-Laird PN. 1990. Belief bias effects in syllogistic reasoning. See Gilhooly et al 1990, pp. 125–38

Oaksford M, Chater N. 1994. A rational analysis of the selection task as optimal data selection. *Psychol. Rev.* 101:608–31

Oaksford M, Chater N. 1996. Rational explanation of the selection task. *Psychol. Rev.* 103:381–91

Oaksford M, Stenning K. 1992. Reasoning with conditionals containing negated constituents. *J. Exp. Psychol. Learn. Mem. Cogn.* 18:835–54

Ormerod TC. 1997. Rationalities 1 and 2: dual processes or different task demands? *Curr. Psychol. Cogn.* 16:181–89

Ormerod TC, Manktelow KI, Jones GV. 1993. Reasoning with three types of conditional: biases and mental models. *Q. J. Exp. Psychol.* 46A:653–78

Piattelli-Palmarini M. 1994. *Inevitable Illusions: How Mistakes of Reason Rule Our Minds.* New York: Wiley

Pinker S. 1997. *How the Mind Works*. New York: Norton

Platt RD, Griggs RA. 1993. Facilitation in the abstract selection task: the effects of attentional and instructional factors. *Q. J. Exp. Psychol.* 46A:591–613

Politzer G, Braine MDS. 1991. Responses to inconsistent premises cannot count as suppression of valid inferences. *Cognition* 38: 103–8

Politzer G, Nguyen-Xuan A. 1992. Reasoning about conditional promises and warnings: Darwinian algorithms, mental models, relevance judgements or pragmatic schemas? *Q. J. Exp. Psychol.* 44:401–12

Polk TA, Newell A. 1995. Deduction as verbal reasoning. *Psychol. Rev.* 102:533–66

Pollard P. 1990. Natural selection for the selection task: limits to social exchange theory. *Cognition* 36:195–204

Quayle JD, Ball LJ. 1997. Subjective confidence and the belief bias effect in syllogistic reasoning. In *Proc. 19th Annu. Conf. Cogn. Sci. Soc.*, pp. 626–31. Mahwah, NJ: Erlbaum

Richardson J, Ormerod TC. 1997. Rephrasing between disjunctives and conditionals: mental models and the effects of thematic content. *Q. J. Exp. Psychol*. 50A:358–85

Riesbeck CK, Schank RC. 1989. *Inside Case-Based Reasoning.* Hillsdale, NJ: Erlbaum

Rips LJ. 1994. *The Psychology of Proof*. Cambridge, MA: MIT Press

Rips LJ. 1997. Goals for a theory of deduction: reply to Johnson-Laird. *Minds Mach.* 7: 409–24

Roberts MJ. 1993. Human reasoning: deduction rules or mental models? *Q. J. Exp. Psychol*. 46A:569–89

Rogers Y, Rutherford A, Bibby PA, eds. 1992. *Models in the Mind: Theory, Perspective and Application.* London: Academic

Russell J. 1987. Rule-following, mental models, and the developmental view. In *Meaning and the Growth of Understanding,* ed. M Chapman, RA Dixon, pp. 23–48. New York: Springer-Verlag

Santamaria C. 1997. On implicit reasoning and mental models. *Curr. Psychol. Cogn.* 16:205–10

Santamaria C, Madruga JAG, Johnson-Laird PN. 1998. Reasoning from double conditionals: the effects of logical structure and believability. *Think. Reasoning.* In press

Schaeken W, Girotto V, Johnson-Laird PN. 1998. The effect of an irrelevant premise on temporal and spatial reasoning. *Kognitionswisschenschaft.* In press

Schaeken W, Johnson-Laird PN, Byrne RMJ, d'Ydewalle G. 1995. A comparison of conditional and disjunctive inferences: a case study of the mental model theory of reasoning. *Psychol. Belg.* 35:57–70

Schaeken WS, Johnson-Laird PN, d'Ydewalle G. 1996a. Mental models and temporal reasoning. *Cognition* 60:205–34

Schaeken WS, Johnson-Laird PN, d'Ydewalle G. 1996b. Tense, aspect, and temporal reasoning. *Think. Reasoning* 2:309–27

Shastri L, Ajjanagadde V. 1993. From simple associations to systematic reasoning: a connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behav. Brain Sci.* 16: 417–94

Shaw VF, Johnson-Laird PN. 1998. Dispelling the 'atmosphere' effect in reasoning. *Anal. Psychol.* In press

Simon HA. 1982. *Models of Bounded Rationality,* Vols. 1, 2. Cambridge, MA: MIT Press

Simon HA. 1991. Cognitive architectures and rational analysis: comment. In *Architectures for Intelligence: The 22nd Carnegie Mellon Symposium on Cognition*, ed. K VanLehn. Hillsdale, NJ: Erlbaum

Sloman SA. 1996. The empirical case for two systems of reasoning. *Psychol. Bull.* 119: 3–22

Smyth B, Keane MT. 1995. Remembering to forget: a competence- preserving deletion policy for CBR systems. In *14th Int. Jt. Conf. Artificial Intelligence*. Los Altos, CA: Morgan Kaufman

Sperber D, Cara F, Girotto V. 1995. Relevance theory explains the selection task. *Cognition* 52:3–39

Sperber D, Wilson D. 1995. *Relevance: Communication and Cognition*. Oxford: Blackwell. Rev. ed.

Stanovich KE. 1998. *Who Is Rational? Studies of Individual Differences in Reasoning.* Mahwah, NJ: Erlbaum. In press

Stenning K, Yule P. 1997. Image and language in human reasoning: a syllogistic illustration. *Cogn. Psychol.* 34:109–59

Stevenson RJ. 1993. *Language, Thought and Representation*. New York: Wiley

Stevenson RJ, Over DE. 1995. Deduction from uncertain premises *Q. J. Exp. Psychol*. 48A:613–43

Sutherland S. 1992. *Irrationality: The Enemy Within*. London: Constable

Tabossi P, Bell VA, Johnson-Laird PN. 1998. Mental models in deductive, modal, and probabilistic reasoning. In *Mental Models in Discourse Processing and Reasoning*, ed. C Habel, G Rickheit. Berlin: Benjamins. In press

Taplin JE, Staudenmayer H, Taddonio JL. 1974. Developmental changes in conditional reasoning: linguistic or logical? *J. Exp. Child Psychol*. 17:360–73

Thompson VA. 1994. Interpretational factors in conditional reasoning. *Mem. Cogn.* 22: 742–58

Thompson VA. 1995. Conditional reasoning: the necessary and sufficient conditions. *Can. J. Exp. Psychol.* 49:1–60

Toms M, Morris N, Ward D. 1993. Working memory and conditional reasoning *Q. J. Exp. Psychol*. 46A:679–99

Vandierendonck A, De Vooght G. 1996. Evidence for mental-model-based reasoning: a comparison of reasoning with time and space concepts. *Think. Reasoning* 2: 249–72

Vandierendonck A, De Vooght G. 1997. Working memory constraints on linear reasoning with spatial and temporal contents. *Q. J. Exp. Psychol*. 50A:803–20

Wason PC. 1966. Reasoning. In *New Horizons in Psychology*, ed. BM Foss, pp. 135–51. Harmondsworth, Middx, UK: Penguin

Wason PC, Johnson-Laird PN. 1972. *The Psychology of Deduction: Structure and Content*. Cambridge, MA: Harvard Univ. Press/London: Batsford

Wetherick NE, Gilhooly KJ. 1990. Syllogistic reasoning: effects of premise order. See Gilhooly et al 1990, pp. 99–108

Wharton CM, Grafman J. 1998. Deductive reasoning and the brain. *Trends Cogn. Sci*. 2:54–59

Yang Y. 1997. *Predicting the difficulty of predicate-logic inference problems: an empirical examination of a predicate-logic model.* PhD thesis. New York Univ., New York

Yang Y, Braine MDS, O'Brien DP. 1998. Some empirical justifications of the mental-predicate-logic model. See Braine & O'Brien 1998, pp. 333–65

Annual Review of Psychology
Volume 50, 1999

CONTENTS