Mental models and human reasoning

Philip N. Johnson-Laird[1]

Department of Psychology, Princeton University, Princeton, NJ 08540

To be rational is to be able to reason. Thirty years ago psychologists believed that human reasoning depended on formal rules of inference akin to those of a logical calculus. This hypothesis ran into difficulties, which led to an alternative view: reasoning depends on envisaging the possibilities consistent with the starting point—a perception of the world, a set of assertions, a memory, or some mixture of them. We construct mental models of each distinct possibility and derive a conclusion from them. The theory predicts systematic errors in our reasoning, and the evidence corroborates this prediction. Yet, our ability to use counterexamples to refute invalid inferences provides a foundation for rationality. On this account, reasoning is a simulation of the world fleshed out with our knowledge, not a formal rearrangement of the logical skeletons of sentences.

abduction | deduction | induction | logic | rationality

Human beings are rational—a view often attributed to Aristotle—and a major component of rationality is the ability to reason (1). A task for cognitive scientists is accordingly to analyze what inferences are rational, how mental processes make these inferences, and how these processes are implemented in the brain. Scientific attempts to answer these questions go back to the origins of experimental psychology in the 19th century. Since then, cognitive scientists have established three robust facts about human reasoning. First, individuals with no training in logic are able to make logical deductions, and they can do so about materials remote from daily life. Indeed, many people enjoy pure deduction, as shown by the world-wide popularity of Sudoku problems (2). Second, large differences in the ability to reason occur from one individual to another, and they correlate with measures of academic achievement, serving as proxies for measures of intelligence (3). Third, almost all sorts of reasoning, from 2D spatial inferences (4) to reasoning based on sentential connectives, such as *if* and *or*, are computationally intractable (5). As the number of distinct elementary propositions in inferences increases, reasoning soon demands a processing capacity exceeding any finite computational device, no matter how large, including the human brain.

Until about 30 y ago, the consensus in psychology was that our ability to reason depends on a tacit mental logic, consisting of formal rules of inference akin to those in a logical calculus (e.g., refs. 6–9). This view has more recent adherents (10, 11). However, the aim of this article is to describe an alternative theory and some of the evidence corroborating it. To set the scene, it outlines accounts based on mental logic and sketches the difficulties that they ran into, leading to the development of the alternative theory. Unlike mental logic, it predicts that systematic errors in reasoning should occur, and the article describes some of these frailties in our reasoning. Yet, human beings do have the seeds of rationality within them—the development of mathematics, science, and even logic itself, would have been impossible otherwise; and as the article shows, some of the strengths of human reasoning surpass our current ability to understand them.

Mental Logic

The hypothesis that reasoning depends on a mental logic postulates two main steps in making a deductive inference. We recover the logical form of the premises; and we use formal rules to prove a conclusion (10, 11). As an example, consider the premises:

> Either the market performs better or else I won't be able to retire.
> I will be able to retire.

The first premise is an exclusive disjunction: either one clause or the other is true, but not both. Logical form has to match the formal rules in psychological theories, and so because the theories have no rules for exclusive disjunctions, the first premise is assigned a logical form that conjoins an inclusive disjunction, which allows that both clauses can be true, with a negation of this case:

> A or not-B & not (A & not-B).
> B.

The formal rules then yield a proof of A, corresponding to the conclusion: *the market performs better*. The number of steps in a proof should predict the difficulty of an inference; and some evidence has corroborated this prediction (10, 11). Yet, problems exist for mental logic.

A major problem is the initial step of recovering the logical form of assertions. Consider, for instance, Mr. Micawber's famous advice (in Dickens's novel, *David Copperfield*):

> Annual income twenty pounds, annual expenditure nineteen pounds nineteen and six, result happiness. Annual income twenty pounds, annual expenditure twenty pounds ought and six, result misery.

What is its logical form? A logician can work it out, but no algorithm exists that can recover the logical form of all everyday assertions. The difficulty is that the logical form needed for mental logic is not just a matter of the syntax of sentences. It depends on knowledge, such as that nineteen pounds nineteen (shillings) and six (pence) is less than twenty pounds, and that happiness and misery are inconsistent properties. It can even depend on knowledge of the context in which a sentence is uttered if, say, its speaker points to things in the world. Thus, some logicians doubt whether logical form is pertinent to everyday reasoning (12). Its extraction can depend in turn on reasoning itself (13) with the consequent danger of an infinite regress.

Other difficulties exist for mental logic. Given the premises in my example about retirement, what conclusion should you draw? Logic yields only one constraint. The conclusion should be *valid*; that is, if the premises are true, the conclusion should be true too, and so the conclusion should hold in every possibility in which the premises hold (14). Logic yields infinitely many valid inferences from any set of premises. It follows validly from my premises that: *I will be able to retire and I will be able to retire, and I will be able to retire*. Most of these valid conclusions are silly, and silliness is hardly rational. Yet, it is compatible with logic, and so logic alone cannot alone characterize rational reasoning (15). Theories of mental logic take pains to prevent silly inferences but then have difficulty in explaining how we recognize that the silly inference above is valid.

Another difficulty for mental logic is that we withdraw valid deductions when brute facts collide with them. Logic can establish such inconsistencies—indeed, one method of logic exploits them to yield valid inferences: you negate the conclusion to be proved, add it to the premises, and show that the resulting set of sentences is inconsistent (14). In orthodox logic, however, any conclusions whatsoever follow from a contradiction, and so it is never necessary to withdraw a conclusion. Logic is accordingly monotonic: as more premises are added so the number of valid conclusions that can be drawn increases monotonically. Humans do not reason in this way. They tend to withdraw conclusions that conflict with a brute fact, and this propensity is rational—to the point that theorists have devised various systems of reasoning that are not monotonic (16). A further problem for mental logic is that manipulations of content affect individuals' choices of which cases refute a general hypothesis in a problem known as Wason's "selection" task (17). Performance in the task is open to various interpretations (18, 19), including the view that individuals aim not to reason deductively but to optimize the amount of information they are likely to gain from evidence (20). Nevertheless, these problems led to an alternative conception of human reasoning.

Mental Models

When humans perceive the world, vision yields a mental model of what things are where in the scene in front of them (21). Likewise, when they understand a description of the world, they can construct a similar, albeit less rich, representation—a mental model of the world based on the meaning of the description and on their knowledge (22). The current theory of mental models (the "model" theory, for short) makes three main assumptions (23). First, each mental model represents what is common to a distinct set of possibilities. So, you have two mental models based on Micawber's advice: one in which you spend less than your income, and the other in which you spend more. (What happens when your expenditure equals your income is a matter that Micawber did not address.) Second, mental models are *iconic* insofar as they can be. This concept, which is due to the 19th century logician Peirce (24), means that the structure of a representation corresponds to the structure of what it represents. Third, mental models of descriptions represent what is true at the expense of what is false (25). This principle of *truth* reduces the load that models place on our working memory, in which we hold our thoughts while we reflect on them (26). The principle seems sensible, but it has an unexpected consequence. It leads, as we will see, to systematic errors in deduction.

When we reason, we aim for conclusions that are true, or at least probable, given the premises. However, we also aim for conclusions that are novel, parsimonious, and that maintain information (15). So, we do not draw a conclusion that only repeats a premise, or that is a conjunction of the premises, or that adds an alternative to the possibilities to which the premises refer—even though each of these sorts of conclusion is valid. Reasoning based on models delivers such conclusions. We search for a relation or property that was not explicitly asserted in the premises. Depending on whether it holds in all, most, or some of the models, we draw a conclusion of its necessity, probability, or possibility (23, 27). When we assess the deductive validity of an inference, we search for counterexamples to the conclusion (i.e., a model of a possibility consistent with the premises but not with the conclusion). To illustrate reasoning based on mental models, consider again the premises:

Either the market performs better or else I won't be able to retire.
I will be able to retire.

The principal data in the construction of mental models are the meanings of premises. The first premise elicits two models: one of the market performing better and the other of my not being able to retire. The second premise eliminates this second model, and so the first model holds: the market does perform better.

Icons

A visual image is iconic, but icons can also represent states of affairs that cannot be visualized, for example, the 3D spatial representations of congenitally blind individuals, or the abstract relations between sets that we all represent. One great advantage of an iconic representation is that it yields relations that were not asserted in the premises (24, 28, 29). Suppose, for example, you learn the spatial relations among five objects, such as that A is to the left of B, B is to the left of C, D is in front of A, and E is in front of C, and you are asked, "What is the relation between D and E?" You could use formal rules to infer this relation, given an axiom capturing the transitivity of "is to the left of." You would infer from the first two premises that A is to the left of C, and then using some complex axioms concerning two dimensions, you would infer that D is to the left of E. A variant on the problem should make your formal proof easier: the final premise asserts instead that E is in front of B. Now, the transitive inference is no longer necessary: you have only to make the 2D inference. So, mental logic predicts that this problem should be easier. In fact, experiments yield no reliable difference between them (28). Both problems, however, have a single iconic model. For example, the first problem elicits a spatial model of this sort of layout:

A   B   C
D   E

The second problem differs only in that E is in front of B. In contrast, inferences about the relation between D and E become much harder when a description is consistent with two different layouts, which call either for two models, or at least some way to keep track of the spatial indeterminacy (30–32). Yet, such descriptions can avoid the need for an initial transitive inference, and so mental logic fails to make the correct prediction. Analogous results bear out the use of iconic representations in temporal reasoning, whether it is based on relations such as "before" and "after" (33) or on manipulations in the tense and aspect of verbs (34).

Many transitive inferences are child's play. That is, even children are capable of making them (35, 36). However, some transitive inferences present a challenge even to adults (37). Consider this problem:

Al is a blood relative of Ben.
Ben is a blood relative of Cath.
Is Al a blood relative of Cath?

Many people respond, "Yes." They make an intuitive inference based on a single model of typical lineal descendants or filial relations (38). A clue to a counterexample helps to prevent the erroneous inference: people can be related by marriage. Hence, Al and Cath could be Ben's parents, and not blood relatives of one another.

Visual images are iconic (39, 40), and so you might suppose that they underlie reasoning. That is feasible but not necessary (41). Individuals distinguish between relations that elicit visual images, such as *dirtier than*, those that elicit spatial relations, such as *in front of*, and those that are abstract, such as *smarter than* (42). Their reasoning is slowest from relations that are easy to visualize. Images impede reasoning, almost certainly because they call for the processing of irrelevant visual detail. A study using functional magnetic resonance imaging (fMRI) showed that visual imagery is not the same as building a mental model (43). Deductions that evoked visual images again took longer for participants to make, and as Fig. 1 shows, these inferences, unlike those based on spatial or abstract relations, elicited extra activity in an area of visual cortex. Visual imagery is not necessary for reasoning.
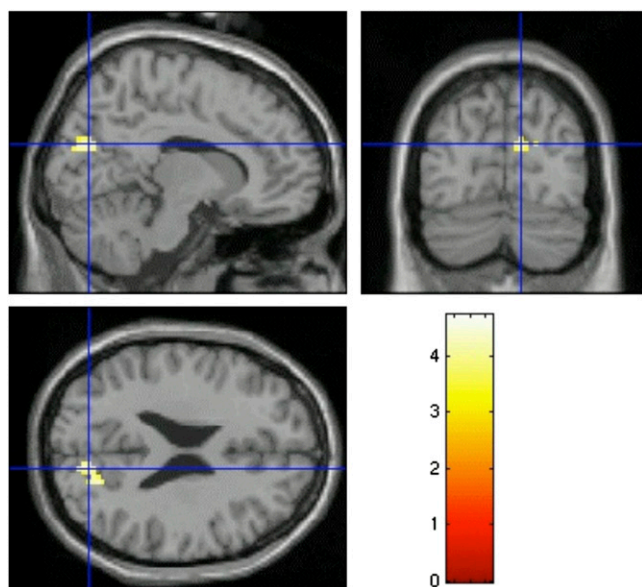
Fig. 1. Reasoning about relations that are easy to visualize but hard to envisage spatially activated areas in the secondary visual cortex, V2 (43). The figure shows the significant activation from the contrast between such problems and abstract problems (in terms of the color scale of z values of the normal distribution) with the crosshairs at the local peak voxel. *Upper Left:* Sagittal section, showing the rear of the brain to the left. *Lower Left:* Horizontal section, showing the rear of the brain to the left. *Upper Right:* Coronal section, with the right of the brain to the right. [Reproduced with permission from ref. 43 (Copyright 2003, MIT Press Journals).]

Frailties of Human Reasoning

The processing capacity of human working memory is limited (26). Our intuitive system of reasoning, which is often known as "system 1," makes no use of it to hold intermediate conclusions. Hence, system 1 can construct a single explicit mental model of premises but can neither amend that model recursively nor search for alternatives to it (22). To illustrate the limitation, consider my definition of an optimist: an optimist $=_{def}$ a person who believes that optimists exist. It is plausible, because you are not much of an optimist if you do not believe that optimists exist. Yet, the definition has an interesting property. Both Presidents Obama and Bush have declared on television that they are optimists. Granted that they were telling the truth, you now know that optimists exist, and so according to my recursive definition you too are an optimist. The definition spreads optimism like a virus. However, you do not immediately grasp this consequence, or even perhaps that my definition is hopeless, because pessimists too can believe that optimists exist. When you deliberate about the definition using "system 2," as it is known, you can use recursion and you grasp these consequences.

Experiments have demonstrated analogous limitations in reasoning (44, 45), including the difficulty of holding in mind alternative models of disjunctions (46). Consider these premises about a particular group of individuals:

Anne loves Beth.
Everyone loves anyone who loves someone.

Does it follow that everyone loves Anne? Most people realize that it does (45). Does it follow that Charles loves Diana? Most people say, "No." They have not been told anything about them, and so they are not part of their model of the situation. In fact, it does follow. Everyone loves Anne, and so, using the second premise again, it follows that everyone loves everyone, and that entails that Charles loves Diana—assuming, as the

question presupposes, that they are both in the group. The difficulty of the inference is that it calls for a recursive use of the second premise, first to establish that everyone loves Anne, and then to establish that everyone loves everyone. Not surprisingly, it is even harder to infer that Charles does not love Diana in case the first premise is changed to "Anne does not love Beth" (45).

Number of Mental Models

The model theory predicts that the more models that we need to take into account to make an inference, the harder the inference should be—we should take longer and make more errors. Many studies have corroborated this prediction, and no reliable results exist to the contrary—one apparent counter-example (10) turned out not to be (46). So, how many models can system 2 cope with? The specific number is likely to vary from person to person, but the general answer is that more than three models causes trouble.

A typical demonstration made use of pairs of disjunctive premises, for example:

Raphael is in Tacoma or else Julia is in Atlanta, but not both.
Julia is in Atlanta or else Paul is in Philadelphia, but not both.
What follows?

Each premise has two mental models, but they have to be combined in a consistent way. In fact, they yield only two possibilities: Raphael is in Tacoma and Paul is in Philadelphia, or else Julia is in Atlanta. This inference was the easiest among a set of problems (47). The study also included pairs of *inclusive* disjunctions otherwise akin to those above. In this case, each premise has three models, and their conjunction yields five. Participants from the general public made inferences from both sorts of pair (47). The percentage of accurate conclusions fell drastically from exclusive disjunctions (just over 20%) to inclusive disjunctions (less than 5%). For Princeton students, it fell from approximately 75% for exclusive disjunctions to less than 30% for inclusive disjunctions (48). This latter study also showed that diagrams can improve reasoning, provided that they are iconic. Perhaps the most striking result was that the most frequent conclusions were those that were consistent with just a single model of the premises. Working memory is indeed limited, and we all prefer to think about just one possibility.

Effects of Content

Content affects all aspects of reasoning: the interpretation of premises, the process itself, and the formulation of conclusions. Consider this problem:

All of the Frenchmen in the restaurant are gourmets.
Some of the gourmets in the restaurant are wine-drinkers.
What, if anything, follows?

Most of the participants in an experiment (49) spontaneously inferred:

Therefore, some of the Frenchmen in the restaurant are wine-drinkers.

Other participants in the study received a slightly different version of the problem, but with the same logical form:

All of the Frenchmen in the restaurant are gourmets.
Some of the gourmets in the restaurant are Italians.
What, if anything, follows?

Only a very small proportion of participants drew the conclusion:

Some of the Frenchmen are Italians.

The majority realized that no definite conclusion followed about the relation between the Frenchmen and the Italians.

The results are difficult to reconcile with mental logic, because the two inferences have the same logical form. However, the model theory predicts the phenomenon. In the first case, individuals envisage a situation consistent with both premises, such as the following model of three individuals in the restaurant:

Frenchman	gourmet	wine-drinker
Frenchman	gourmet	wine-drinker
Frenchman	gourmet	

This model yields the conclusion that some of the Frenchmen are wine-drinkers, and this conclusion is highly credible—the experimenters knew that it was, because they had already asked a panel of judges to rate the believability of the putative conclusions in the study. So, at this point—having reached a credible conclusion—the reasoners were satisfied and announced their conclusion. The second example yields an initial model of the same sort:

Frenchman	gourmet	Italian
Frenchman	gourmet	Italian
Frenchman	gourmet	

It yields the conclusion that some of the Frenchmen are Italians. However, for most people, this conclusion is preposterous—again as revealed by the ratings of the panel of judges. Hence, the participants searched more assiduously for an alternative model of the premises, which they tended to find:

Frenchman	gourmet
Frenchman	gourmet
Frenchman	gourmet
gourmet	Italian
gourmet	Italian

In this model, none of the Frenchman is an Italian, and so the model is a counterexample to the initial conclusion. The two models together fail to support any strong conclusion about the Frenchmen and the Italians, and so the participants responded that nothing follows. Thus, content can affect the process of reasoning, and in this case motivate a search for a counterexample (see also ref. 50). Content, however, has other effects on reasoning. It can modulate the interpretation of connectives, such as "if" and "or" (51). It can affect inferences that depend on a single model, perhaps because individuals have difficulty in constructing models of implausible situations (52).

Illusory Inferences

The principle of truth postulates that mental models represent what is true and not what is false. To understand the principle, consider an exclusive disjunction, such as:

Either there is a circle on the board or else there isn't a triangle.

It refers to two possibilities, which mental models can represent, depicted here on separate lines, and where "¬" denotes a mental symbol for negation:

$$\text{o}$$
$$\neg \Delta$$

These mental models do not represent those cases in which the disjunction would be false (e.g., when there is a circle and not a triangle). However, a subtler saving also occurs. The first model does not represent that it is false that there is not a triangle (i.e., there is a triangle). The second model does not represent that it is false there is a circle, (i.e., there is not a circle). In simple tasks, however, individuals are able to represent this information in *fully explicit* models:

$$\text{o} \qquad \Delta$$
$$\neg \text{o} \qquad \neg \Delta$$

Other sentential connectives have analogous mental models and fully explicit models.

A computer program implementing the principle of truth led to a discovery. The program predicted that for certain premises individuals should make systematic fallacies. The occurrence of these fallacies has been corroborated in various sorts of reasoning (25, 53–55). They tend to be compelling and to elicit judgments of high confidence in their conclusions, and so they have the character of cognitive illusions. This, in turn, makes

their correct analysis quite difficult to explain. However, this illusion is easy to understand, at least in retrospect:

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 ten, or both.
Is it possible that there is an ace in the hand?

In an experiment, 95% of the participants responded, "Yes" (55). The mental models of the premises support this conclusion, for example, the first premise allows that in two different possibilities an ace occurs, either by itself or with a king. Yet, the response is wrong. To see why, suppose that there *is* an ace in the hand. In that case, both the first and the second premises are true. Yet, the rubric to the problem states that only one of the premises is true. So, there cannot be an ace in the hand. The framing of the task is not the source of the difficulty. The participants were confident in their conclusions and highly accurate with the control problems.

The key to a correct solution is to overcome the principle of truth and to envisage fully explicit models. So, when you think about the truth of the first premise, you also think about the concomitant falsity of the other two premises. The falsity of the second premise in this case establishes that there is not a queen and there is not an ace. Likewise, the truth of the second premise establishes that the first premise is false, and so there is not a king and there is not an ace; and the truth of the third premise establishes that both the first and second premises are false. Hence, it is impossible for an ace to be in the hand. Experiments examined a set of such illusory inferences, and they yielded only 15% correct responses, whereas a contrasting set of control problems, for which the principle of truth predicts correct responses, yielded 91% correct responses (55).

Perhaps the most compelling illusion of all is this one:

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

My colleagues and I have given this inference to more than 2,000 individuals, including expert reasoners (23, 25), and almost all of them drew the conclusion:

There is an ace in the hand.

To grasp why this conclusion is invalid, you need to know the meaning of "if" and "or else" in daily life. Logically untrained individuals correctly consider a conditional assertion—one based on "if" —to be false when its if-clause is true and its then-clause is false (15, 56). They understand "or else" to mean that one clause is true and the other clause is false (25). For the inference above, suppose that the first conditional in the disjunctive premise is false. One possibility is then that there is a queen but not an ace. This possibility suffices to show that even granted that both premises are true, no guarantee exists that there is an ace in the hand. You may say: perhaps the participants in the experiments took "or else" to allow that both clauses could be true. The preceding argument still holds: one clause could have been false, as the disjunction still allows, and one way in which the first clause could have been false is that there was a queen without an ace. To be sure, however, a replication replaced "or else" with the rubric, "Only one of the following assertions is true" (25).

Conditional assertions vary in their meaning,* in part because they can refer to situations that did not occur (65), and so it would be risky to base all claims about illusory inferences on

---

*A large literature exists on reasoning from conditionals. The model theory postulates that the meaning of clauses and knowledge can modulate the interpretation of connectives so that that they diverge from logical interpretations (57). Evidence bears out the occurrence of such modulations (58), and independent experimental results corroborate the model theory of conditionals (59–62). Critics, however, have yet to be convinced (63, but cf. 64).

them. The first illusion above did not make use of a conditional, and a recent study has shown a still simpler case of an illusion based on a disjunction:

Suppose only one of the following assertions is true:
1. You have the bread.
2. You have the soup or the salad, but not both.
Also, suppose you have the bread. Is it possible that you also have both the soup and the salad?

The premises yield mental models of the three things that you can have: the bread, the soup, the salad. They predict that you should respond, "No"; and most people make this response (66). The fully explicit models of the premises represent both what is true and what is false. They allow that when premise 1 is true, premise 2 is false, and that one way in which it can be false is that you have both the soup and the salad. So, you can have all three dishes. The occurrence of the illusion, together with other disjunctive illusions (67), corroborates the principle of truth.

Strengths of Human Reasoning

An account of the frailties of human reasoning creates an impression that individuals are incapable of valid deductions except on rare occasions; and some cognitive scientists have argued for this skeptical assessment of rationality. Reasoners, they claim, have no proper conception of validity and instead draw conclusions based on the verbal "atmosphere" of the premises, so that if, say, one premise contains the quantifier "some," as in the earlier example about the Frenchmen, individuals are biased to draw a conclusion containing "some" (68). Alternatively, skeptics say, individuals are rational but draw conclusions on the basis of probability rather than deductive validity. The appeal to probability fits a current turn toward probabilistic theories of cognition (69). Probabilities can influence our reasoning, but theories should not abandon deductive validity. Otherwise, they could offer no account of how logically untrained individuals cope with Sudoku puzzles (2), let alone enjoy them. Likewise, the origins of logic, mathematics, and science, would be inexplicable if no one previously could make deductions. Our reasoning in everyday life has, in fact, some remarkable strengths, to which we now turn.

Strategies in Reasoning.
If people tackle a batch of reasoning problems, even quite simple ones, they do not solve them in a single deterministic way. They may flail around at first, but they soon find a strategy for coping with the inferences. Different people spontaneously develop different strategies (70). Some strategies are more efficient than others, but none of them is immune to the number of mental models that an inference requires (71). Reasoners seem to assemble their strategies as they explore problems using their existing inferential tactics, such as the ability to add information to a model of a possibility. Once they have developed a strategy for a particular sort of problem, it tends to control their reasoning.

Studies have investigated the development of strategies with a procedure in which participants can use paper and pencil and think aloud as they are reasoning.[†] The efficacy of the main strategy was demonstrated in a computer program that parsed an input based on what the participants had to say as they thought aloud, and then followed the same strategy to make the same inference. Here is an example of a typical problem (71):

Either there is a blue marble in the box or else there is a brown marble in the box, but not both. Either there is a brown marble in the box or else there is white marble in the box, but not both. There is a white marble in the box if and only if there is a red marble in the

box. Does it follow that: If there is a blue marble in the box then there is a red marble in the box?

The participants' most frequent strategy was to draw a diagram to represent the different possibilities and to update it in the light of each subsequent premise. One sort of diagram for the present problem was as follows, where each row represented a different possibility:

blue   white   red
brown

The diagram establishes that the conclusion holds. Other individuals, however, converted each disjunction into a conditional, constructing a coreferential chain of them:

If blue then not brown. If not brown then white. If white then red.

The chain yielded the conclusion. Another strategy was to make a supposition of one of the clauses in the conditional conclusion and to follow it up step by step from the premises. This strategy, in fact, yielded the right answer for the wrong reasons when some individuals made a supposition of the then-clause in the conclusion.

In a few cases, some participants spontaneously used counterexamples. For instance, given the problem above, one participant said:

Assuming there's a blue marble and *not* a red marble.
Then, there's not a white.
Then, a brown and not a blue marble.
No, it is impossible to get from a blue to not a blue.
So, if there's a blue marble there is a red.

You might wonder whether individuals use these strategies if they do not have to think aloud. However, the relative difficulty of the inferences did remain the same in this condition (71). The development of strategies may itself depend on metacognition, that is, on thinking about your own thinking and about how you might improve it (23). This metacognitive ability seems likely to have made possible Aristotle's invention of logic and the subsequent development of formal systems of logic.

Counterexamples.
Counterexamples are crucial to rationality. A counterexample to an assertion shows that it is false. However, a counterexample to an inference is a possibility that is consistent with the premises but not with the conclusion, and so it shows that the inference is invalid. Intuitive inferences based on system 1 do not elicit counterexamples. Indeed, an alternative version of the mental model theory does without them too (76). Existing theories of mental logic make no use of them, either (10, 11). So, what is the truth about counterexamples? Is it just a rare individual who uses them to refute putative inferences, or are we all able to use them?

Experiments have answered both these questions, and they show that most people do use counterexamples (77). We have already seen that when reasoners infer unbelievable conclusions, they tend to look for counterexamples. A comparable effect occurs when they are presented with a conclusion to evaluate. In fact, there are two sorts of invalid conclusion. One sort is a flat-out contradiction of the premises. Their models are disjoint, as in this sort of inference:

A or B, or both.
Therefore, neither A nor B.

The other sort of invalidity should be harder to detect. Its conclusion is consistent with the premises, that is, it holds in at least one model of them, but it does not follow from them, because it fails to hold in at least one other model, for example:

A or B, or both.
Therefore, A and B.

In a study of both sorts of invalid inference, the participants wrote their justifications for rejecting conclusions (78). They were more accurate in recognizing that a conclusion was invalid when it was inconsistent with the premises (92% correct) than when it was consistent with the premises (74%). However, they

---

[†]The task of thinking aloud is relevant to a recent controversy about whether moral judgments call for reasoning (72, 73). The protocols from participants who have to resolve moral dilemmas suggest that they do reason, although they are also affected by the emotions that the dilemmas elicit (53, 74, 75).

used counterexamples more often in the consistent cases (51% of inferences) than in the inconsistent cases (21% of inferences). When the conclusion was inconsistent with the premises, they tended instead to point out the contradiction. They used other strategies too. One individual, for example, never mentioned counterexamples but instead reported that a necessary piece of information was missing from the premises. The use of counterexamples, however, did correlate with accuracy in the evaluation of inferences. A further experiment compared performance between two groups of participants. One group wrote justifications, and the other group did not. This manipulation had no reliable effect on the accuracy or the speed of their evaluations of inferences. Moreover, both groups corroborated the model theory's prediction that invalidity was harder to detect with conclusions consistent with the premises than with conclusions inconsistent with them.

Other studies have corroborated the use of counterexamples (79), and participants spontaneously drew diagrams that served as counterexamples when they evaluated inferences (80), such as:

> More than half of the people at this conference speak French.
> More than half of the people at this conference speak English.
> Therefore, more than half of the people at this conference speak both French and English.

Inferences based on the quantifier "more than half" cannot be captured in the standard logic of the first-order predicate calculus, which is based on the quantifiers "any" and "some" (14). They can be handled only in the second-order calculus, which allows quantification over sets as well as individuals, and so they are beyond the scope of current theories of mental logic.

An fMRI study examined the use of counterexamples (81). It contrasted reasoning and mental arithmetic from the same premises. The participants viewed a problem statement and three premises and then either a conclusion or a mathematical formula. They had to evaluate whether the conclusion followed from the premises, or else to solve a mathematical formula based on the numbers of individuals referred to in the premises. The study examined easy inferences that followed immediately from a single premise and hard inferences that led individuals to search for counterexamples, as in this example:

> There are five students in a room.
> Three or more of these students are joggers.
> Three or more of these students are writers.
> Three or more of these students are dancers.
> Does it follow that at least one of the students in the room is all three: a jogger, a writer, and a dancer?

You may care to tackle this inference for yourself. You are likely to think first of a possibility in which the conclusion holds. You may then search for a counterexample, and you may succeed in finding one, such as this possibility: students 1 and 2 are joggers and writers, students 3 and 4 are writers and dancers, and student 5 is a jogger and dancer. Hence, the conclusion does not follow from the premises.

As the participants read the premises, the language areas of their brains were active (Broca's and Wernicke's area), but then nonlanguage areas carried out the solution to the problems, and none of the language areas remained active. Regions in right prefrontal cortex and inferior parietal lobe were more active for reasoning than for calculation, whereas regions in left prefrontal cortex and superior parietal lobe were more active for calculation than for reasoning. Fig. 2 shows that only the hard logical inferences calling for a search for counterexamples elicited activation in right prefrontal cortex (the right frontal pole). Similarly, as the complexity of relations increases in problems, the problems become more difficult (37, 82, 83), and they too activate prefrontal cortex (36, 84). The anterior frontal lobes evolved most recently, they take longest to mature, and their maturation relates to measured intelligence (85). Other studies of reasoning, however, have not found activation in right frontal pole (86, 87), perhaps
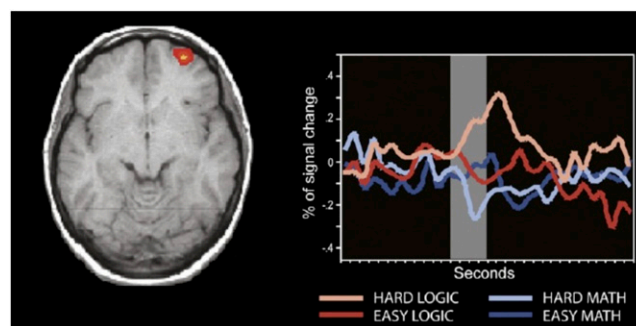

Fig. 2. Interaction between the type of problem and level of difficulty in right frontal pole, Brodmann's area 10 (81). After the 8-s window of the presentation of the problems (shown in gray, allowing for the hemodynamic lag), the inferences that called for a search for counterexamples activated this region more than the easy inferences did. There was no difference in activation between the hard and easy mathematical problems, and only the counterexample inferences showed activity above baseline. [Reproduced with permission from ref. 81 (Copyright 2008, Elsevier).]

because these studies did not include inferences calling for counterexamples. Indeed, a coherent picture of how the different regions of the brain contribute to reasoning has yet to emerge.

Abduction. Human reasoners have an inductive skill that far surpasses any known algorithm, whether based on mental models or formal rules. It is the ability to formulate explanations. Unlike valid deductions, inductions increase information, because they use knowledge to go beyond the strict content of the premises, for example:

> A poodle's jaw is strong enough to bite through wire.
> Therefore, a German shepherd's jaw is strong enough to bite through wire.

We are inclined to accept this induction, because we know that German shepherds are bigger and likely to be stronger than poodles (88). However, even if the premises are true, no guarantee exists that an inductive conclusion is true, precisely because it goes beyond the information in the premises. This principle applies *a fortiori* to those inductions that yield putative explanations—a process often referred to as *abduction* (24).

A typical problem from a study of abduction is:

> If a pilot falls from a plane without a parachute then the pilot dies.
> This pilot didn't die. Why not?

Some individuals respond to such problems with a valid deduction: the pilot did not fall from a plane without a parachute (89). However, other individuals infer explanations, such as:

> The plane was on the ground & he [sic] didn't fall far.
> The pilot fell into deep snow and so wasn't hurt.
> The pilot was already dead.

A study that inadvertently illustrated the power of human abduction used pairs of sentences selected at random from pairs of stories, which were also selected at random from a set prepared for a different study. In one condition, one or two words in the second sentence of each pair were modified so that the sentence referred back to the same individual as the first sentence, for example:

> Celia made her way to a shop that sold TVs.
> She had just had her ears pierced.

The participants were asked: what is going on? On the majority of trials, they were able to create explanations, such as:

> She's getting reception in her earrings and wanted the shop to investigate.
> She wanted to see herself wearing earrings on close-circuit TV.
> She won a bet by having her ears pierced, using money to buy a new TV.

The experimenters had supposed that the task would be nearly impossible with the original sentences referring to different

individuals. In fact, the participants did almost as well with them (23).

Suppose that you are waiting outside a café for a friend to pick you up in a car. You know that he has gone to fetch the car, and that if so, he should return in it in about 10 min—you walked with him from the car park. Ten minutes go by with no sign of your friend, and then another 10 min. There is an inconsistency between what you validly inferred—he will be back in 10 min—and the facts. Logic can tell you that there is an inconsistency. Likewise, you can establish an inconsistency by being unable to construct a model in which all of the assertions are true, although you are likely to succumb to illusory inferences in this task too (90, 91). However, logic cannot tell you what to think. It cannot even tell you that you should withdraw the conclusion of your valid inference. Like other theories (16, 92), however, the model theory allows you to withdraw a conclusion and to revise your beliefs (23, 89). What you really need, however, is an *explanation* of what has happened to your friend. It will help you to decide what to do.

The machinery for creating explanations of events in daily life is based on knowledge of causal relations. A study used simple inconsistencies akin to the one about your friend, for example:

If someone pulled the trigger, then the gun fired.
Someone pulled the trigger, but the gun did not fire.
Why not?

The participants typically responded with causal explanations that resolved the inconsistency, such as: someone emptied the gun and there were no bullets in its chamber (89). In a further study, more than 500 of the smartest high school graduates in Italy assessed the probability of various putative explanations for these inconsistencies. The explanations were based on those that the participants had created in the earlier experiment. On each trial, the participants assessed the probabilities of a cause and its effect, the cause alone, the effect alone, and various control assertions. The results showed unequivocally that the participants ranked as the most probable explanation, a cause and its effect, such as: someone emptied the gun and there were no bullets in its chamber. Because these conjunctions were ranked as more probable than their individual constituent propositions, the assessments violated the probability calculus—they are an instance of the so-called "conjunction" fallacy in which a conjunction is considered as more probable than either of its constituents (93). Like other results (94), they are also contrary to a common view—going back to William James (95)—that we accommodate an inconsistent fact with a minimal change to our existing beliefs (92, 96).

Conclusions

Human reasoning is not simple, neat, and impeccable. It is not akin to a proof in logic. Instead, it draws no clear distinction between deduction, induction, and abduction, because it tends to exploit what we know. Reasoning is more a simulation of the world fleshed out with all our relevant knowledge than a formal manipulation of the logical skeletons of sentences. We build mental models, which represent distinct possibilities, or that unfold in time in a kinematic sequence, and we base our conclusions on them.

When we make decisions, we use heuristics (93, 97), and some psychologists have argued that we can make better decisions when we rely more on intuition than on deliberation (98). In reasoning, our intuitions make no use of working memory (in system 1) and yield a single model. They too can be rapid—many of the inferences discussed in this article take no more than a second or two. However, intuition is not always enough for rationality: a single mental model may be the wrong one. Studies of disasters illustrate this failure time and time again (99). The capsizing of *The Herald of Free Enterprise* ferry is a classic example. The vessel was a "roll on, roll off" car ferry: cars drove into it through the open bow doors, a member of the crew closed the bow doors, and the ship put out to sea. On March 6, 1987, the *Herald* left Zebrugge in Belgium bound for England. It sailed out of the harbor into the North Sea with its bow doors wide open. In the crew's model of the situation, the bow doors had been closed; it is hard to imagine any other possibility. Yet such a possibility did occur, and 188 people drowned as a result.

In reasoning, the heart of human rationality may be the ability to grasp that an inference is no good because a counterexample refutes it. The exercise of this principle, however, calls for working memory—it depends on a deliberative and recursive process of reasoning (system 2). In addition, as I have shown elsewhere, deliberative reasoning was crucial to success in the Wright brother's invention of a controllable heavier-than-air craft, in the Allies breaking of the Nazi's Enigma code, and in John Snow's discovery of how cholera was communicated from one person to another (23). Even when we deliberate, however, we are not immune to error. Our emotions may affect our reasoning (100), although when we reason about their source, the evidence suggests that our reasoning is better than about topics that do not engage us in this way (101) —a phenomenon that even occurs when emotions arise from psychological illnesses (102). A more serious problem may be our focus on truth at the expense of falsity. The discovery of this bias corroborated the model theory, which—uniquely—predicted its occurrence.

1. Baron J (2008) *Thinking and Deciding* (Cambridge Univ Press, New York), 4th Ed.
2. Lee NYL, Goodwin GP, Johnson-Laird PN (2008) The psychological problem of Sudoku. *Think Reasoning* 14:342–364.
3. Stanovich KE (1999) *Who Is Rational? Studies of Individual Differences in Reasoning* (Erlbaum, Mahwah, NJ).
4. Ragni M (2003) An arrangement calculus, its complexity and algorithmic properties. *Lect Notes Comput Sci* 2821:580–590.
5. Garey M, Johnson D (1979) *Computers and Intractability: A Guide to the Theory of NP-Completeness* (Freeman, San Francisco).
6. Beth EW, Piaget J (1966) *Mathematical Epistemology and Psychology* (Reidel, Dordrecht, The Netherlands).
7. Osherson DN (1974–6) *Logical Abilities in Children* (Erlbaum, Hillsdale, NJ), Vol 1–4, .
8. Braine MDS (1978) On the relation between the natural logic of reasoning and standard logic. *Psychol Rev* 85:1–21.
9. Rips LJ (1983) Cognitive processes in propositional reasoning. *Psychol Rev* 90:38–71.
10. Rips LJ (1994) *The Psychology of Proof* (MIT Press, Cambridge, MA).
11. Braine MDS, O'Brien DP, eds (1998) *Mental Logic* (Erlbaum, Mahwah, NJ).
12. Barwise J (1989) *The Situation in Logic* (CSLI Publications, Stanford University, Palo Alto, CA).
13. Stenning K, van Lambalgen MS (2008) *Human Reasoning and Cognitive Science* (MIT Press, Cambridge, MA).
14. Jeffrey R (1981) *Formal Logic: Its Scope and Limits* (McGraw-Hill, New York), 2nd Ed.
15. Johnson-Laird PN, Byrne RMJ (1991) *Deduction* (Erlbaum, Hillsdale, NJ).
16. Brewka G, Dix J, Konolige K (1997) *Nonmonotonic Reasoning: An Overview* (CLSI Publications, Stanford University, Palo Alto, CA).
17. Wason PC (1966) Reasoning. *New Horizons in Psychology*, ed Foss BM (Penguin, Harmondsworth, Middlesex, United Kingdom).
18. Sperber D, Cara F, Girotto V (1995) Relevance theory explains the selection task. *Cognition* 52:3–39.
19. Nickerson RS (1996) Hempel's paradox and Wason's selection task: Logical and psychological problems of confirmation. *Think Reasoning* 2:1–31.
20. Oaksford M, Chater N (1996) Rational explanation of the selection task. *Psychol Rev* 103:381–391.
21. Marr D (1982) *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (Freeman, San Francisco).
22. Johnson-Laird PN (1983) *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness* (Harvard Univ Press, Cambridge, MA).
23. Johnson-Laird PN (2006) *How We Reason* (Oxford Univ Press, New York).
24. Peirce CS (1931–1958) *Collected Papers of Charles Sanders Peirce*, eds Hartshorne C, Weiss P, Burks A (Harvard Univ Press, Cambridge, MA).
25. Johnson-Laird PN, Savary F (1999) Illusory inferences: A novel class of erroneous deductions. *Cognition* 71:191–229.

26. Baddeley A (2003) Working memory: Looking back and looking forward. *Nat Rev Neurosci* 4:829–839.
27. Johnson-Laird PN, Legrenzi P, Girotto V, Legrenzi MS, Caverni J-P (1999) Naive probability: a mental model theory of extensional reasoning. *Psychol Rev* 106:62–88.
28. Byrne RMJ, Johnson-Laird PN (1989) Spatial reasoning. *J Mem Lang* 28:564–575.
29. Taylor H, Tversky B (1992) Spatial mental models derived from survey and route descriptions. *J Mem Lang* 31:261–292.
30. Carreiras M, Santamaría C (1997) Reasoning about relations: Spatial and nonspatial problems. *Think Reasoning* 3:191–208.
31. Vandierendonck A, Dierckx V, De Vooght G (2004) Mental model construction in linear reasoning: evidence for the construction of initial annotated models. *Q J Exp Psychol A* 57:1369–1391.
32. Jahn G, Knauff M, Johnson-Laird PN (2007) Preferred mental models in reasoning about spatial relations. *Mem Cognit* 35:2075–2087.
33. Schaeken WS, Johnson-Laird PN, d'Ydewalle G (1996) Mental models and temporal reasoning. *Cognition* 60:205–234.
34. Schaeken WS, Johnson-Laird PN, d'Ydewalle G (1996) Tense, aspect, and temporal reasoning. *Think Reasoning* 2:309–327.
35. Bryant PE, Trabasso T (1971) Transitive inferences and memory in young children. *Nature* 232:456–458.
36. Waltz JA, et al. (1999) A system for relational reasoning in human prefrontal cortex. *Psychol Sci* 10:119–125.
37. Goodwin GP, Johnson-Laird PN (2005) Reasoning about relations. *Psychol Rev* 112:468–493.
38. Goodwin GP, Johnson-Laird PN (2008) Transitive and pseudo-transitive inferences. *Cognition* 108:320–352.
39. Shepard RN, Metzler J (1971) Mental rotation of three-dimensional objects. *Science* 171:701–703.
40. Kosslyn SM, Ball TM, Reiser BJ (1978) Visual images preserve metric spatial information: Evidence from studies of image scanning. *J Exp Psychol Hum Percept Perform* 4:47–60.
41. Pylyshyn Z (2003) Return of the mental image: Are there really pictures in the brain? *Trends Cogn Sci* 7:113–118.
42. Knauff M, Johnson-Laird PN (2002) Visual imagery can impede reasoning. *Mem Cognit* 30:363–371.
43. Knauff M, Fangmeier T, Ruff CC, Johnson-Laird PN (2003) Reasoning, models, and images: Behavioral measures and cortical activity. *J Cogn Neurosci* 15:559–573.
44. Schroyens W, Schaeken W, Handley S (2003) In search of counter-examples: Deductive rationality in human reasoning. *Q J Exp Psychol A* 56:1129–1145.
45. Cherubini P, Johnson-Laird PN (2004) Does everyone love everyone? The psychology of iterative reasoning. *Think Reasoning* 10:31–53.
46. García-Madruga JA, Moreno S, Carriedo N, Gutiérrez F, Johnson-Laird PN (2001) Are conjunctive inferences easier than disjunctive inferences? A comparison of rules and models. *Q J Exp Psychol A* 54:613–632.
47. Johnson-Laird PN, Byrne RMJ, Schaeken WS (1992) Propositional reasoning by model. *Psychol Rev* 99:418–439.
48. Bauer MI, Johnson-Laird PN (1993) How diagrams can improve reasoning. *Psychol Sci* 4:372–378.
49. Oakhill J, Johnson-Laird PN, Garnham A (1989) Believability and syllogistic reasoning. *Cognition* 31:117–140.
50. Evans JS, Barston JL, Pollard P (1983) On the conflict between logic and belief in syllogistic reasoning. *Mem Cognit* 11:295–306.
51. Johnson-Laird PN, Byrne RMJ (2002) Conditionals: A theory of meaning, pragmatics, and inference. *Psychol Rev* 109:646–678.
52. Klauer KC, Musch J, Naumer B (2000) On belief bias in syllogistic reasoning. *Psychol Rev* 107:852–884.
53. Bucciarelli M, Johnson-Laird PN (2005) Naïve deontics: A theory of meaning, representation, and reasoning. *Cognit Psychol* 50:159–193.
54. Yang Y, Johnson-Laird PN (2000) Illusions in quantified reasoning: How to make the impossible seem possible, and vice versa. *Mem Cognit* 28:452–465.
55. Goldvarg Y, Johnson-Laird PN (2000) Illusions in modal reasoning. *Mem Cognit* 28:282–294.
56. Evans JS, Over DE (2004) *If* (Oxford Univ Press, Oxford).
57. Byrne RMJ, Johnson-Laird PN (2009) 'If' and the problems of conditional reasoning. *Trends Cogn Sci* 13:282–287.
58. Quelhas AC, Johnson-Laird PN, Juhos C (2010) The modulation of conditional assertions and its effects on reasoning. *Q J Exp Psychol*, in press.
59. Schroyens W, Schaeken W, d'Ydewalle G (2001) The processing of negations in conditional reasoning: A meta-analytical case study in mental models and/or mental logic theory. *Think Reasoning* 7:121–172.
60. Markovits H, Barrouillet P (2002) The development of conditional reasoning: A mental model account. *Dev Rev* 22:5–36.
61. Oberauer K (2006) Reasoning with conditionals: A test of formal models of four theories. *Cognit Psychol* 53:238–283.
62. Barrouillet P, Gauffroy C, Lecas J-F (2008) Mental models and the suppositional account of conditionals. *Psychol Rev* 115:760–771, discussion 771–772.
63. Evans JS, Over DE (2010) Conditional truth: Comment on Byrne and Johnson-Laird. *Trends Cogn Sci* 14:5; author reply 6.
64. Byrne RMJ, Johnson-Laird PN (2010) Models redux: Reply to Evans and Over. *Trends Cogn Sci* 14:6.
65. Byrne RMJ (2005) *The Rational Imagination: How People Create Alternatives to Reality* (MIT Press, Cambridge, MA).
66. Khemlani S, Johnson-Laird PN (2009) Disjunctive illusory inferences and how to eliminate them. *Mem Cognit* 37:615–623.
67. Walsh CR, Johnson-Laird PN (2004) Co-reference and reasoning. *Mem Cognit* 32:96–106.
68. Gilhooly KJ, Logie RH, Wetherick NE, Wynn V (1993) Working memory and strategies in syllogistic-reasoning tasks. *Mem Cognit* 21:115–124.
69. Oaksford M, Chater N (2007) *Bayesian Rationality: The Probabilistic Approach to Human Reasoning* (Oxford Univ Press, Oxford).
70. Schaeken W, De Vooght G, Vandierendonck A, d'Ydewalle G (2000) *Deductive Reasoning and Strategies* (Erlbaum, Mahwah, NJ).
71. Van der Henst JB, Yang Y, Johnson-Laird PN (2002) Strategies in sentential reasoning. *Cogn Sci* 26:425–468.
72. Haidt J (2001) The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychol Rev* 108:814–834.
73. Gazzaniga M, et al. (2010) *Does Moral Action Depend on Reasoning?* (John Templeton Foundation, West Conshohocken, PA).
74. Bucciarelli M, Khemlani S, Johnson-Laird PN (2008) The psychology of moral reasoning. *Judgm Decis Mak* 3:121–139.
75. Green DW, McClelland AM, Muckli L, Simmons C (1999) Arguments and deontic decisions. *Acta Psychol (Amst)* 101:27–47.
76. Polk TA, Newell A (1995) Deduction as verbal reasoning. *Psychol Rev* 102:533–566.
77. Byrne RMJ, Espino O, Santamaría C (1999) Counterexamples and the suppression of inferences. *J Mem Lang* 40:347–373.
78. Johnson-Laird PN, Hasson U (2003) Counterexamples in sentential reasoning. *Mem Cognit* 31:1105–1113.
79. Bucciarelli M, Johnson-Laird PN (1999) Strategies in syllogistic reasoning. *Cogn Sci* 23:247–303.
80. Neth H, Johnson-Laird PN (1999) The search for counterexamples in human reasoning. *Proceedings of the Twenty-first Annual Conference of the Cognitive Science* (Erlbaum, Mahwah, NJ), p 806.
81. Kroger JK, Nystrom LE, Cohen JD, Johnson-Laird PN (2008) Distinct neural substrates for deductive and mathematical processing. *Brain Res* 1243:86–103.
82. Halford GS, Wilson WH, Phillips S (1998) Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behav Brain Sci* 21:803–831; discussion 831–864.
83. Kroger J, Holyoak KJ (2004) Varieties of sameness: The impact of relational complexity on perceptual comparisons. *Cogn Sci* 28:335–358.
84. Kroger JK, et al. (2002) Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: A parametric study of relational complexity. *Cereb Cortex* 12:477–485.
85. Shaw P, et al. (2006) Intellectual ability and cortical development in children and adolescents. *Nature* 440:676–679.
86. Goel V, Dolan RJ (2004) Differential involvement of left prefrontal cortex in inductive and deductive reasoning. *Cognition* 93:B109–B121.
87. Monti MM, Osherson DN, Martinez MJ, Parsons LM (2007) Functional neuroanatomy of deductive inference: A language-independent distributed network. *Neuroimage* 37:1005–1016.
88. Smith EE, Shafir E, Osherson D (1993) Similarity, plausibility, and judgments of probability. *Cognition* 49:67–96.
89. Johnson-Laird PN, Girotto V, Legrenzi P (2004) Reasoning from inconsistency to consistency. *Psychol Rev* 111:640–661.
90. Johnson-Laird PN, Legrenzi P, Girotto V, Legrenzi MS (2000) Illusions in reasoning about consistency. *Science* 288:531–532.
91. Legrenzi P, Girotto V, Johnson-Laird PN (2003) Models of consistency. *Psychol Sci* 14:131–137.
92. Alchourrón C, Gärdenfors P, Makinson D (1985) On the logic of theory change: Partial meet contraction functions and their associated revision functions. *J Symbolic Logic* 50:510–530.
93. Tversky A, Kahneman D (1983) Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychol Rev* 90:292–315.
94. Walsh CR, Johnson-Laird PN (2009) Changing your mind. *Mem Cognit* 37:624–631.
95. James W (1907) *Pragmatism* (Longmans Green, New York).
96. Lombrozo T (2007) Simplicity and probability in causal explanation. *Cognit Psychol* 55:232–257.
97. Kahneman D, Frederick S (2005) A model of heuristic judgment. *The Cambridge Handbook of Thinking and Reasoning*, eds Holyoak KJ, Morrison RG (Cambridge Univ Press, Cambridge, UK).
98. Gigerenzer G, Goldstein DG (1996) Reasoning the fast and frugal way: Models of bounded rationality. *Psychol Rev* 103:650–669.
99. Perrow C (1984) *Normal Accidents: Living with High-risk Technologies* (Basic Books, New York).
100. Oaksford M, Morris F, Grainger R, Williams JMG (1996) Mood, reasoning, and central executive processes. *J Exp Psychol Learn Mem Cogn* 22:477–493.
101. Blanchette I, Richards A, Melnyk L, Lavda A (2007) Reasoning about emotional contents following shocking terrorist attacks: A tale of three cities. *J Exp Psychol Appl* 13:47–56.
102. Johnson-Laird PN, Mancini F, Gangemi A (2006) A hyper-emotion theory of psychological illnesses. *Psychol Rev* 113:822–841.