Cognitive Science 26 (2002) 425–468

COGNITIVE SCIENCE

http://www.elsevier.com/locate/cogsci

Strategies in sentential reasoning

Jean-Baptiste Van der Henst [a], Yingrui Yang [b], P.N. Johnson-Laird [c,*]

[a] *University of Leuven, Leuven, Belgium*
[b] *Rensselaer Polytechnic Institute, Troy, NY, USA*
[c] *Department of Psychology, Princeton University, Green Hall, Princeton, NJ 08544, USA*

Received 6 September 2001; received in revised form 7 April 2002; accepted 24 April 2002

Abstract

Four experiments examined the strategies that individuals develop in sentential reasoning. They led to the discovery of five different strategies. According to the theory proposed in the paper, each of the strategies depends on component tactics, which all normal adults possess, and which are based on mental models. Reasoners vary their use of tactics in ways that have no deterministic account. This variation leads different individuals to assemble different strategies, which include the construction of incremental diagrams corresponding to mental models, and the pursuit of the consequences of a single model step by step. Moreover, the difficulty of a problem (i.e., the number of mental models required by the premises) predisposes reasoners towards certain strategies. Likewise, the sentential connectives in the premises also bias reasoners towards certain strategies, e.g., conditional premises tend to elicit reasoning step by step whereas disjunctive premises tend to elicit incremental diagrams.

Keywords: Strategies; Reasoning; Mental models

1. Introduction

Logical reasoning is central both to the development of science and mathematics and to the solution of problems in daily life. Naïve individuals can grasp that a set of propositions logically implies a conclusion. The term *naïve* here refers to individuals who have no explicit mastery of formal logic or any other cognate discipline. It does not impugn their intelligence. What underlies their logical ability, however, is controversial. Theorists have proposed that

* Corresponding author. Tel.: +1-609-258-4432; fax: +1-609-258-1113.
E-mail address: phil@princeton.edu (P.N. Johnson-Laird).

PII: S0364-0213(02)00074-5

it depends on a memory for previous inferences (e.g., Kolodner, 1993), on conditional rules that capture general knowledge (e.g., Newell, 1990), on "neural nets" representing concepts (e.g., Shastri & Ajjanagadde, 1993), or on specialized innate modules for matters that were important to our hunter–gatherer ancestors (Cosmides, 1989). But, none of these accounts readily explains the ability to reason about matters for which you have no general knowledge. Suppose, for instance, that you know nothing about computers but you are given the following premises:

If the software is right and the cable is correct then the printer works.
The software is right, but the printer does not work.

You are able to infer the conclusion:

The cable is not correct.

This inference is *valid*, that is, its conclusion must be true granted that its premises are true. It is an example of a major class of logical deductions, *sentential* inferences, which hinge on negation and sentential connectives, such as "if," "or," and "and," and which are captured in an idealized way in the branch of logic known as the "sentential" or "propositional" calculus.

Theorists have two alternative views about how naïve individuals make sentential inferences (see Baron, 1994, for a review). Originally, they thought that naïve individuals rely on formal rules of inference akin to those of logic (e.g., Braine & O'Brien, 1998; Inhelder & Piaget, 1958; Rips, 1994). The discovery that content influences reasoning (see Wason & Johnson-Laird, 1972), coupled with the need to account for the mental representation of discourse, led to a different view of reasoning. Individuals grasp the meaning of premises, and they use this meaning to construct *mental models* of the possibilities that the premises describe. They evaluate an inference as valid if its conclusion holds in all their mental models of the premises (Johnson-Laird & Byrne, 1991; Polk & Newell, 1995).

The controversy between the two competing views has been fruitful. It has led to the development of explicit computer models of reasoning, and to more stringent experiments. But, it has focused on simple inferences. Most studies of sentential reasoning have examined inferences based on no more than two premises (for reviews, see Evans, Newstead, & Byrne, 1993; Garnham & Oakhill, 1994). They have aimed to reveal the hidden nature of inferential mechanisms: do they rely on formal inference rules or on mental models? In this paper, however, we want to go beyond the usual investigation of the basic inferential mechanisms and to examine an aspect of reasoning that has often been neglected—the strategies that individuals develop to make complex inferences (see also Schaeken, De Vooght, Vandierendonck, & d'Ydewalle, 2000).

We propose the following working definition:

A *strategy* in reasoning is a systematic sequence of elementary mental steps that an individual follows in making an inference.

We refer to each of these mental steps as a *tactic*, and so a strategy is a sequence of tactics that an individual uses to make an inference. We illustrate this terminology with the following problem, which you are invited to solve:

There is a white pill in the box if and only if there is a green pill.
Either there is a green pill in the box or else there is a red pill, but not both.
There is a red pill in the box if and only if there is a blue pill.
Does it follow that:
If there isn't a white pill in the box then there is a blue pill in the box?

Like most people, you probably responded correctly to this problem:

Yes, if there isn't a white pill in the box, then there is a blue pill in the box.

But, how did you solve the problem? What kind of tactical steps did you carry out and how were they organized? These are the questions that we want to address in this paper.

One possible strategy is to use a *supposition*, i.e., an assumption for the sake of argument. Thus, you might have said to yourself:

Suppose that there isn't a white pill in the box. It follows from the first premise that the box does not contain a green pill, either. It then follows from the second premise that there *is* a red pill in the box. And it follows from the third premise that there is a blue pill too. So, if there isn't a white pill in the box, then it follows that there is a blue pill in the box. The conclusion in the question is correct.

Hence, in this strategy, you made a supposition corresponding to a single possibility and followed up its consequences step by step. One tactic in the *strategy* is to make a supposition, and another is to draw a conclusion from the supposition and the first premise. *Strategies* are therefore the molar units of analysis, tactics are the molecular units, and the inferential *mechanisms* underlying tactics are the atomic units. Whereas the nature of inferential mechanisms is the topic of several hundred papers in the literature, strategies and tactics have not yet been investigated for sentential reasoning.

We use the term "strategy" in much the same sense as Bruner, Goodnow, and Austin (1956), who described strategies in concept attainment. Like them, we do not imply that a strategy is necessarily conscious. It may become conscious as reasoners try to develop a way to cope with a problem. But, as we will see, individuals describe each tactical step they take in an inference rather than their high-level strategy. We can infer their strategies from these descriptions (cf. Bruner et al., 1956, p. 55). Miller, Galanter, and Pribram (1960, p. 16) defined a "plan" as "any hierarchical process in the organism that can control the order in which a sequence of operations is to be performed." We could have used the term "plan" instead of "strategy," but it has misleading connotations, particularly in artificial intelligence. It suggests that people first plan how they will make an inference and then carry out the plan. Reasoners, however, do not seem to proceed in this way. Instead, they start reasoning at once, and one tactical step leads to another, and so on, as their strategy unfolds spontaneously. Tactics are also akin to the component processes that Sternberg postulated in his analysis of analogies, numerical series problems, and syllogisms (see e.g., Sternberg, 1977, 1983, 1984). We likewise follow in the tradition of analyzing cognitive tasks into their component processes, or mechanisms, within the information-processing methodology (Hunt, 1999).

As Fig. 1 shows, our account distinguishes four levels in a hierarchy of thinking and reasoning. Each level in the hierarchy is a level of organization, but it depends on the level below for its implementation, much as a programming language is a level of organization that depends for its implementation on the lower level of machine language. In other words, the levels are

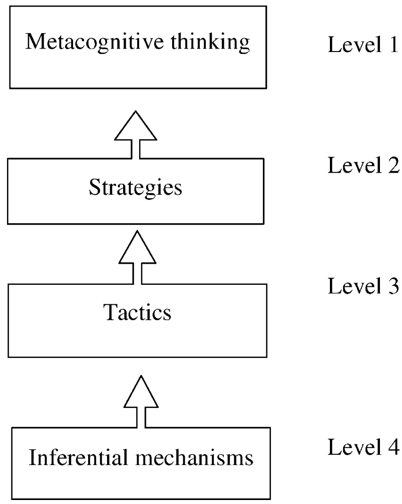

Fig. 1. The four levels in the hierarchy of thinking. As the arrow denotes, the units at a lower level make up the constituents of the units at the next level up, e.g., strategies are made up of a sequence of tactics.

not independent of one another, and the organization at one level has to be underpinned by what happens at a lower level. At the highest level, there is *metacognitive thinking*, i.e., thinking about thinking, usually in order to develop a solution to a problem. Thinking at this level often occurs in solving complex problems such as the tower of Hanoi, and in "game-theoretic" situations such as the prisoner's dilemma (Berardi-Coletta, Buyer, Dominowski, & Rellinger, 1995). When it is not obvious how to solve a problem, individuals may think about thinking in this self-conscious way. But, as we will see, they do not seem to develop reasoning strategies at the metacognitive level. At the second level are the *strategies* in thinking. They unfold in a series of actions without the individual necessarily having a conscious awareness of an overall strategy. At the third level are each of the *tactics* in a strategy, such as making a supposition, or combining it with a premise to make an inference. At the lowest level are the cognitive *mechanisms* that underlie the tactical steps, e.g., the construction of a mental model if the model theory is correct, or the application of a rule of inference if formal rule theories are correct. Our aim in the present paper is to delineate the nature of the strategies and tactics underlying sentential reasoning.

The levels of thought are not independent of one another: strategies depend on tactics, and tactics depend on inferential mechanisms. Hence, we need an account of inferential mechanisms in order to explain tactics. We propose that inferential mechanisms are based on mental models (Johnson-Laird, 2001), and we will try to show how such mechanisms at the lowest level can compose the tactics one level up, which in turn make up strategies at the second level (see Section 5). It is conceivable that inferential mechanisms are based on formal rules of inference instead of mental models, and we return to this possibility in Section 7.

The phenomena occurring at the strategic and tactical levels are comparable to the control procedures and the logical components in the implementation of a theorem prover in artificial intelligence. Logic itself is not enough (Stenning & Oaksford, 1993). It cannot specify the procedure for proving theorems. Hence, nonlogical decisions are necessary to obtain a practical implementation of a theorem prover. Typically, these procedures are designed to avoid a combinatorial explosion in demands on resources. For instance, PROLOG, a programming language based on the analogy between programs and proofs, uses heuristic tools such as "backward-chaining" to implement an effective theorem prover. Similarly, inferential mechanisms at the lowest level of thought do not determine how human reasoning proceeds. They are the basic tools. The ways they are put to use is a matter of strategies and tactics in reasoning.

How can we find out what reasoning strategies individuals develop? In our view, the first steps are to observe what they say as they think-aloud while they reason, and to give them paper and pencil and to see what they write down and what they draw. These data, however, are a controversial source of evidence. One problem is their validity. People can be unable to describe how they reached a certain decision or even be mistaken about why they acted as they did (see e.g., Greenwald & Banaji, 1995; Nisbett & Wilson, 1977). The need to think-aloud and to use paper and pencil may also change the nature of the thought process, and even impair it. Schooler, Ohlsson, and Brooks (1993) interrupted people who were trying to solve insight problems. After the interruption, those who had to make a retrospective report on their thinking solved fewer problems than those who carried out an unrelated task (a crossword puzzle). Other studies, however, report that thinking aloud enhanced performance (e.g., Berardi-Coletta et al., 1995). It can slow people down, but it often appears to have no other major effects (cf. Russo, Johnson, & Stephens, 1989). In general, it can be a reliable guide to the sequence of a person's thoughts (Ericsson & Simon, 1980; Newell & Simon, 1972). Our view is that the use of "think-aloud" protocols and drawings and writings is indeed only a first step in the analysis of strategies. Bell (1999) has compared reasoning when individuals think-aloud and when individuals think to themselves in the usual way. The patterns of results were similar in the two conditions. The present paper makes no such comparisons, because its main goal is to delineate the variety of strategies that reasoners develop. Even if these strategies were unique to thinking aloud with pencil and paper, a major goal of psychology should be to give an account of them and of how individuals develop them.

The paper begins with an account of how psychologists have thought about reasoning strategies in the past (Section 2). It then presents a taxonomy of strategies in sentential reasoning based on experiments in which the participants thought aloud as they either evaluated given conclusions or drew their own conclusions from premises (Section 3). It outlines the core principles of the theory of mental models (Section 4), which it uses to formulate a theory of strategies in sentential reasoning (Section 5). It reports three experiments corroborating the theory's account of how people develop strategies (Section 6). And it concludes with an appraisal of strategic and tactical thinking in reasoning (Section 7).

2. Previous studies of strategies in logical reasoning

The pioneering studies of reasoners' strategies investigated relational "series" problems (Hunter, 1957; Huttenlocher, 1968; Piaget, 1921), such as:

John is taller than Pete.
Pete is smaller than Bob.
Who is the tallest?

When participants carry out problems based on five premises, they rapidly develop various "short cut" strategies (Wood, 1969; Wood, Shotter, & Godden, 1974). With premises that each contain the same relation, say, "taller than," they look to see whether a term occurs only on the left-hand side of a single premise—in which case, it denotes the tallest entity. The result of this strategy is that the participants can answer the question posed in the problem, but are less likely to be able to answer a second unexpected question about other items in the series (Wood, 1969; Wood et al., 1974). Another strategy is to use the question as a guide to select relevant premises in order to construct a chain linking the terms occurring in the question (see also Ormrod, 1979). Quinton and Fellows (1975) asked their participants to talk about the strategies that they had developed. After repeated experience with problems sharing the same formal properties, the participants tended to identify invariants (e.g., an extreme term is mentioned only once, and the middle term is never the answer of the question), and to use them to solve the problems with minimal effort. Quinton and Fellows described five different "perceptual strategies," such as one in which the participants try to answer the question solely from the information in the first premise. If they obtain an answer, e.g., "John" to the problem above, and this term does not occur in the second premise, then they do not need to represent the second premise: the answer is correct. The strategy works only for determinate problems, which yield an order for all three individuals.

From a study of metareasoning in so-called "knight-and-knave" problems, Rips (1989) argued that reasoners rely on a single deterministic strategy based on categorical premises or, if there are none, on suppositions. An example of such a problem is:

> There are only two sorts of people: knights, who always tell the truth, and knaves, who always lie.
> Arthur says, "Lancelot is a knight and Gawain is a knave."
> Lancelot says, "Arthur is a knave."
> Gawain says, "Arthur is a knave."
> What are Arthur, Lancelot, and Gawain?

Rips proposed that reasoners solve these problems by making a supposition:

> Suppose Arthur is a knight (i.e., tells the truth).
> It follows that Lancelot is a knight.
> But Lancelot asserts that Arthur is a knave (i.e., tells lies).
> Hence, Arthur cannot be a knight.
> And so on.

Likewise, Rips's (1994) PSYCOP computer program for sentential and quantified reasoning follows a single deterministic strategy relying on formal rules of inference at the lowest level of thinking. Braine and O'Brien (1998) appear to make a similar case for a single deterministic strategy. But, as the design of theorem provers shows (see Wos, 1988), the use of formal rules does not necessitate a single strategy.

In contrast, Johnson-Laird and Byrne (1990) argued that naive reasoners use various strategies for "knight-and-knave" problems. Consider the problem above. Some people do indeed use suppositions. But, others report that they solved the problem when they noticed that both Lancelot and Gawain assert the same proposition, and so either they are both knights or else they are both knaves. Hence, Arthur's assertion cannot be true, because he says that one of them is a knight and one of them is a knave. So, he must be a knave, and both Lancelot and Gawain must be knights. Johnson-Laird and Byrne (1990) developed a computer program modeling five distinct strategies for knight-and-knave problems. Subsequently, Byrne and Handley (1997) showed that reasoners do develop a variety of strategies for them.

In general, the model theory has always been compatible with a diversity of reasoning strategies: Johnson-Laird (1983) discussed individual differences in reasoning, and Johnson-Laird and Bara (1984) described two alternative strategies. Bucciarelli and Johnson-Laird (1999) have investigated the strategies that naïve reasoners use in syllogistic reasoning. The participants were video-taped as they used cut-out shapes to make syllogistic inferences, such as:

Some of the chefs are musicians.
None of the musicians are painters.
∴ None of the chefs are painters.

The most striking aspect of the results was the diversity in the participants' strategies. They sometimes began by constructing an external model of the first premise to which they added the information from the second premise; they sometimes proceeded in the opposite order. They sometimes built an initial model that satisfied the conclusion, and modified it to refute the conclusion; they sometimes built an initial model that immediately refuted the conclusion.

There has been a dearth of studies of strategies in sentential reasoning. We suspect that experiments have used too simple premises for strategies to differ, and that the data have failed to reveal reasoner's strategies. Indeed, in the field of sentential reasoning, the data considered by researchers are responses and their latencies. These sorts of results can be compared to a known strategy, but they do not reveal unknown strategies.

3. A taxonomy of strategies in sentential reasoning

How can experimenters best observe the strategies that reasoners use in sentential reasoning? In our view, there are three desiderata. First, the inferential problems should be sufficiently time-consuming to force the participants to think, but not so difficult that they make many errors. Second, the experimental procedure needs to externalize strategies as much as possible. Third, the content of inferences should be neutral and unlikely to trigger general knowledge. Those materials that do engage general knowledge tend to bias logical reasoning (see e.g., Evans et al., 1993), and particularly to eliminate possibilities that are normally compatible with the interpretation of sentential connectives. For example, Johnson-Laird and Byrne (2002) have shown that manipulations of content can yield 12 distinct interpretations of conditionals. Such a diversity of interpretations makes reasoners' strategies hard to discern. Neutral materials are indeed commonly used in studies of logical reasoning (e.g., Braine & O'Brien, 1998; Rips, 1994), but they risk violating ecological validity and thereby leading experimental participants to adopt wholly artificial ways of thinking. In our view, this risk is small, and it is accordingly appropriate to begin the study of strategies in sentential reasoning with materials that are sensible but unlikely to trigger general knowledge. In sum, if the participants have to think-aloud in reasoning about time-consuming problems with neutral materials, and are allowed to use paper and pencil, then their video-taped protocols might be revealing about their strategies.

3.1. Experiment 1: strategies in evaluating given conclusions

Our first experiment was designed to find out whether the "think-aloud" procedure would reveal reasoning strategies. We needed a set of problems that would fit our desiderata for investigating strategies—they should be easy but time-consuming. We accordingly used sentential problems based on three premises, but each set of premises was compatible with only two alternative possibilities. The task was to evaluate a given conclusion. Here is a typical problem:

> 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 premises are compatible with the following two models of the possible contents of the box, shown here on separate horizontal lines:

blue	white	red
brown		

The first premise calls for two possibilities (blue or brown) and the subsequent premises add further information to the first of them. Thus, the integration of the three premises gives rise to two possibilities, and the conclusion follows from them. We will explain in more detail in Section 4 the process of reasoning on the assumption that each mental model represents a possibility.

3.1.1. Method

The participants carried out 12 inferences, which each had a conclusion to be evaluated. These problems are stated in an abbreviated form in Table 1. We use the following abbreviations: "iff" for biconditionals of the form "if and only if," "ore" for exclusive disjunctions of the form "either _ or else _, but not both," and "or" for inclusive disjunctions of the form "_ or _, or both". For half of the problems the correct answer was "yes" (i.e., the given conclusion was valid), and for the other half of the problems the correct answer was "no" (i.e., the given conclusion was invalid). Eight problems were based on three premises, and four problems were based on four premises. Two of the problems (4 and 6) had a redundant first premise, and two of the problems (11 and 12) were stated in a discontinuous order, i.e., the first two premises did not contain any proposition in common. The premises were mainly biconditionals and exclusive disjunctions, and the conclusions were conditionals except for two problems (3 and 5), which had exclusive disjunctions as conclusions.

The contents of the problems concerned different colored marbles. The color terms were eight frequent one-syllable English words: black, blue, brown, gray, green, pink, red, and white. We made two different random assignments of the color terms to the problems, ensuring that no two problems in an assignment had the same subset of words. Half the participants were

Table 1
The form of the 12 problems in Experiment 1

Valid problems		Invalid problems	
1.	A ore B B ore C C iff D If A then D?	2.	A iff B B ore C C iff D If A then D?
3.	A iff B If B then C C ore D A ore D?	4.	A ore B B ore C C iff D If B then D?
5.	A iff B B iff C C ore D A ore D?	6.	A iff B B ore C C iff D If B then D?
7.	A iff B B ore C C iff D If not-A then D?	8.	A ore B B iff C C ore D If not-A then D?
9.	A iff B B iff C C ore D D ore E If A then E?	10.	A iff B B ore C C iff D D iff E If A then E?
11.	A iff B C ore D B iff C D ore E If A then E?	12.	A iff B C iff D B ore C D iff E If A then E?

"Iff" denotes "if and only if," "or" denotes inclusive disjunction, and "ore" denotes an exclusive disjunction. The question at the end of each problem is the conclusion to be evaluated. A, B, C, . . . stand for different propositions.

tested with one assignment, and half the participants were tested with the other assignment. The problems were presented in a different random order to each participant.

The participants were told that the aim of the experiment was to try to understand how people reasoned. They were encouraged to use the pencil and paper, and they were told to think-aloud as they tackled each problem. We video-recorded what they said and what they wrote down and drew. The camera was above them and focused on the paper on which they wrote, and they rapidly adapted to the conditions of the experiment. They could take as much time as they needed to make each inference. Each problem was available to them throughout the period of solution. They had to try to maintain a running commentary on their thoughts. If they fell silent for more than 3 s, the experimenter reminded them to think-aloud. They were given one illustrative problem and four practice problems to which they drew their own conclusions. The aim of these problems was to familiarize the participants with the task, and to get them used to thinking aloud and to being video-recorded. We tested eight Princeton students, who had no training in logic, and who had not previously participated in any experiment on reasoning.

3.1.2. Results

The participants often floundered for one or two practice problems, but the 12 experimental problems were easy. None of the participants made any errors in evaluating the given conclusions, though they were not always correct for the correct reasons. We transcribed the tapes *verbatim* apart from repetitions of words, filled pauses, and hesitations. These protocols also included anything that the participants wrote down and a record, step by step, of any drawings or diagrams. The transcription was labor intensive, but we were able to make sense of almost everything that the participants said, wrote, and drew.

The protocols reflected intelligent individuals thinking aloud and revealing the sequences of their tactical steps. Most participants used two or more distinct strategies, but two of them stuck to the same strategy throughout the experiment. What the protocols did not reveal were the mechanisms underlying the tactical steps (the lowest level in the hierarchy of thinking, see Fig. 1). We were able, however, to categorize all the protocols from all the participants into one of the strategies in the taxonomy below.

3.2. The taxonomy of strategies

The taxonomy is based on the protocols from Experiment 1, but it also takes into account the results from Experiments 2 and 3 below. Its aim is to capture the main strategies with which the reasoners tackled the problems. Unless two protocols for a problem are identical in every step, one could argue that they represent two distinct strategies. Our view, however, is that the same strategy can occur in distinct protocols. For example, one protocol might show that a reasoner made a supposition based on the *antecedent* of a conditional, and then combined it with a premise in order to infer an intermediate conclusion, and continued in this way step by step. Another protocol might show that a reasoner made a supposition based on the *consequent* of the conditional, and then continued in a step by step way. Despite the superficial differences between the protocols, what constrains them is the same strategy realized in slightly different ways. As far as possible, the taxonomy is based on the assumption that a strategy should be applicable to any sort of problem in sentential reasoning. Hence, we have tried to describe strategies in ways that do not depend on the specifics of problems.

The taxonomy distinguishes five main strategies. It may be necessary to add further strategies: no one can ever know when the classification is complete. As we will see, however, it is possible to advance a theory of the "space" of humanly possible strategies. We begin with an informal description of each of the five strategies illustrated with examples of *verbatim* protocols. Readers can find additional protocols in the appendix on the Website www.cognitivesciencesociety.org.

3.2.1. The incremental diagram strategy

This strategy depends on drawing a single diagram that keeps track of all the possibilities compatible with the premises. The diagram represents one or more possibilities, depending on the number of possibilities implied by the premises. It corresponds to a set of models, typically

Problem 11:
Yellow or black.
Black iff green.
Green ore red.

Verbal protocol:	Diagrams:			
Yellow or black.		Yellow	Black	Black Yellow
Black iff green.		Yellow	Black Green	Black Yellow Green
Green ore red	Yellow Red	Yellow	Black Green	Black Yellow Green

Conclusion: (Yellow and red) or (yellow) or (black and green) or (black, yellow and green)

General comment: The participant represents the first premise by drawing three possibilities in vertical columns, and then adds the contents of each premise incrementally to the initial diagram. With the last premise, the participant infers wrongly that the red marble should both be added, and not added, to the yellow marble, and so he adds another possibility to the diagram in which the yellow and the red marbles are together. In fact, yellow cannot occur by itself.

Fig. 2. The incremental diagram strategy in generating conclusions (a verbatim protocol from participant 24 in Experiment 3). "Iff" denotes a biconditional, "or" denotes an inclusive disjunction, and "ore" denotes an exclusive disjunction.

mental models, but sometimes fully explicit models (see Section 4). Fig. 2 presents an example of a protocol from an individual who uses this strategy to generate a conclusion. Participants using this strategy tended to work through the premises in an order that allowed them to increment their diagrams. As Fig. 2 illustrates, the incremental diagram strategy can yield a set of models that naïve individuals have difficulty in condensing into a succinct conclusion. They draw instead a conclusion in so-called "disjunctive normal form," i.e., each possibility is described in a conjunction, and these conjunctions are combined with disjunctions, e.g., *Red and green, or blue and yellow*.

A telltale sign of the strategy is a single diagram representing the possibilities compatible with the premises. Another telltale sign is the representation in the diagram of premises that are irrelevant to evaluating the conclusions. The diagrams sometimes have additional annotations, which represent the connectives in the premises. These annotations are frequent when a participant is in the process of developing this strategy. Indeed, a precursor to the strategy is often to draw separate diagrams for each of the premises.

3.2.2. The step strategy

Reasoners pursue step by step the consequences of either a categorical proposition or a supposition. They accordingly infer a sequence of what logicians refer to as "literals," where a *literal* is a proposition that does not contain any sentential connectives: it may be an atomic proposition, *A*, or its negation, *not-A*. Fig. 3 is a protocol of a participant using the step strategy based on a supposition. A supposition is a tactic, which reasoners use to derive an intermediate conclusion from a premise. They then use this intermediate conclusion to derive another conclusion from another premise, and so on, until they derive the other literal (or its negation) in the conditional conclusion.

The sign of a supposition is a phrase, such as, "Supposing there were . . . " or "Assuming we have . . . ." Another tactic, which is observed when reasoners draw their own conclusion (as in Experiment 3), consists in integrating the supposition and the intermediate conclusion into a complex conditional conclusion. Its antecedent contains one or two of the literals in the premises and its consequent contains others literals that follow when the antecedent is satisfied, e.g., *If A then B, not-C, and D.* Individuals may prefer to formulate complex conditionals rather than simple ones, such as: *If A then D*, because complex conditionals convey more semantic information than simple ones, and because they reflect intermediate steps. If reasoners have

Problem:
Pink iff black.
Black ore gray.
Gray iff blue.
If not pink then blue?

Verbal Protocol:	Drawing:	Comments
Pink iff black. Black ore gray. Gray iff blue. If not pink then blue.		Repetition of the premises
Assuming we have no pink:		Supposition
There is no pink.	Crosses out "pink" in premise.	
So there is no black.	Crosses out "black" in both premises.	Inference
There is gray.	Circles "gray" in second premise.	Inference
There is blue.		Inference
Yes.		
Not pink and blue.		Conclusion
Yes.		

General comment: The participant starts by making a supposition that matches the antecedent of the conclusion to evaluate, and derives a set of categorical conclusions.

Fig. 3. The step strategy with a supposition in evaluating a conclusion (a verbatim protocol from participant 8 in Experiment 1).

to evaluate complex conditionals, then presumably they would be able to do so. On some occasions, however, a reasoner constructed several complex conditional conclusions based on different suppositions from the same premises, e.g.:

If only white then pink and gray.
If only red then not pink or gray.
If red and white then pink and gray.

The participants in our experiments used suppositions in a variety of ways, not all of which were logically correct. They made suppositions both to evaluate given conclusions and to create their own conclusions. They made suppositions to derive disjunctive conclusions. They made suppositions of literals common to two premises. They made suppositions to draw modal conclusions about possibilities. They sometimes made suppositions even when there was a categorical premise. If a given conclusion is interpreted as a biconditional, i.e., *A iff C*, then the suppositional strategy needs to be used twice, e.g., once to show that the supposition of *A* yields the consequent, *C*, and once to show that the supposition of *C* yields the antecedent, *A*. The participants did not know these subtleties. They did not realize that the proof of a biconditional conclusion or an exclusive disjunction calls for two suppositional inferences. They also assumed wrongly that a conditional conclusion can be proved from a supposition of its consequent. Hence, many of their inferences, strictly speaking, were invalid. In general, naïve reasoners are not fastidious about the suppositions that they make—given, that is, that they are prepared to make suppositions. Some reasoners never made any suppositions.

One variant of the step strategy was highly sophisticated. A few participants made a supposition of a *counterexample* to a conclusion, and then used the step strategy to pursue its consequences. For instance, given a problem of the form:

A ore B.
B ore C.
C iff D.
A iff D?

a participant (17 in Experiment 2) reasoned as follows:

Assuming A and not-D.	(a counterexample to the conclusion)
Then not-C.	(from the supposition and the third premise)
Then not-A.	(from the previous step and the second and first premises)
No, it is impossible to get from A to not-A.	

The telltale sign of the step strategy is a sequence of inferences, which starts with a categorical premise or a supposition, and continues with a sequence of literal conclusions, which may be incorporated within a single complex conditional.

3.2.3. The compound strategy

Reasoners draw a compound conclusion from two compound assertions, i.e., assertions containing a sentential connective. For instance, an example of a compound inference is the following:

A ore B
B iff C
∴ A ore C

In a sequence of such compound inferences, reasoners derive an ultimate conclusion, either one to be evaluated or one that they draw for themselves. The source of the premises for a compound inference may be the stated premises or previous compound conclusions. Likewise, the source may be a sentence or a diagram that a participant has drawn, and the conclusion may be expressed verbally or in the form of another diagram, or sometimes both. Fig. 4 shows a complete protocol in which the reasoner combines the first two premises to yield an intermediate compound conclusion, and then combines this conclusion with the third premise to draw the final compound conclusion. Strictly speaking, the participant erred. His conclusion is correct, but to prove an exclusive disjunction, it is necessary to establish not just the conditional here,

Problem 9:
White iff blue.
If blue then pink.
Pink ore brown.
White ore brown? [given conclusion]

Verbal protocol:	Drawing:	Comments:
White iff blue.	Blue → White	
If blue then pink.		
White from blue.	Points to diagram.	
If blue then pink.		
If blue then pink.	Writes: "if Blue, then Pink."	
If pink then white.	Pink → White	Invalid compound inference.
Pink ore brown.		
Pink and white.	Points to previous diagram.	
If brown then not white.	Brown, ~~white~~	Invalid compound inference.
White ore brown.		
Yes.		Accepts conclusion.

General comments: The participant starts by stating and reformulating the first two premises and making an invalid compound inference: If blue then pink; white from blue; therefore, If pink then white. He draws another invalid compound inference: if pink then white; pink ore brown; therefore, if brown then not white. Finally, he takes this conditional to imply the conclusion.

Fig. 4. The compound strategy in evaluating a conclusion (a verbatim protocol from participant 1 in Experiment 1).

but also its converse. Another example of a compound inference is shown here:

If B then not-C.	The participant points at the diagram representing an exclusive disjunction: b × c
D for C.	The participant points at a diagram
∴ It isn't the case that if B then D.	representing a biconditional premise: d → c

The telltale sign of the strategy is a sequence of compound conclusions.

3.2.4. *The chain strategy*

Reasoners construct a chain of *conditionals* leading from one constituent of a compound conclusion to its other constituent. The conclusion may be one that reasoners construct for themselves or one that they are evaluating. Fig. 5 shows a protocol of the chain strategy in generating a conclusion. The strategy's telltale signs are three-fold. First, reasoners do not draw a sequence of conclusions in the form of literals, but rather a sequence of conditionals. Second, they make an immediate inference from any premise that is not a conditional, i.e., a disjunctive or a biconditional premise, to convert it into an appropriate conditional. Third, the consequent of one conditional matches the antecedent of the next conditional in the chain. The strategy is valid provided that reasoners construct a chain leading from the antecedent of a conditional conclusion to its consequent. However, reasoners often worked invalidly in the converse direction. Likewise, the valid use of the strategy to prove a biconditional or exclusive disjunction calls for two chains, but reasoners usually rely on just a single chain.

3.2.5. *The concatenation strategy*

This strategy is subtle, and we did not observe it in Experiments 1 and 3, but a few reasoners resorted to it in Experiment 2. They used the tactic of concatenating two or more premises in

---

Problem 5
Green ore blue.
Blue iff black.
Black iff red.

Protocol:	Comments:
Green ore blue.	
Blue iff black.	Repetition and reformulation
Black iff red.	of the premises
Black if red.	
If red then black.	Immediate inference from premise 3
If black then blue.	Immediate inference from premise 2
If blue then no green.	Immediate inference from premise 1

Conclusion:
If red then no green.

General comment: The participant constructs a chain from premise 3 to premise 1.

Fig. 5. The *chain strategy* in generating a conclusion (a verbatim protocol from participant 5 in Experiment 3).

order to form an intermediate conclusion. They usually went on to use some other strategy, such as a supposition and a step. In some protocols, however, reasoners formed a conclusion by concatenating all the premises, and this conclusion was then used as the premise for an immediate inference yielding the required conclusion. For example, one participant (14 in Experiment 2) argued from the premises:

A and B.
B iff C.
C iff D.

to the concatenation:

A and (B iff C iff D).

and then made an immediate inference to the required conclusion:

A and D.

The strategy accordingly depends on concatenating at least two premises into a single conclusion, and then either drawing such a conclusion, or else evaluating a given conclusion, if necessary by an immediate inference. Its telltale sign is the concatenation of the premises and their connectives within a single conclusion.

Table 2 presents a taxonomy of the five sorts of strategy, and their underlying tactics. It is designed to enable investigators to categorize strategies. The initial tactic is highly diagnostic of a strategy, but there are exceptions. Sometimes, reasoners use one initial tactic as a precursor to another initial tactic and thence to a different strategy. For example, a reasoner may start with a supposition, but then use it to initiate the incremental diagram strategy.

The 12 problems in Experiment 1 appear to be typical of those within the competence of logically-untrained individuals, as shown by the fact that they got them all correct, though not always for the correct reasons. Hence, we calculated the total number of times each strategy occurred in the protocols, and expressed these numbers as percentages of the total number of occurrences of strategies. The results were as follows:

Incremental diagram strategy:	34% of overall use.
Supposition and step strategy:	21% of overall use.
Compound strategy:	19% of overall use.
Chain strategy:	25% of overall use.

The most salient feature of the protocols was that different participants used different strategies. On a few occasions, they changed from one strategy to another during a single problem. More often, they switched from one strategy to another from one problem to another. They sometimes reverted to a strategy that they had used earlier in the experiment.

4. Reasoning with mental models

The taxonomy in Table 2 describes the strategies, but it does not explain them or their tactical steps. Our aim is to formulate a theory of strategies and tactics, and we proceed by first

Table 2
A taxonomy of strategies

	The five strategies				
	Incremental diagram	Step	Compound	Chain	Concatenation
First tactic	Start a diagram based on first premise.	Make a supposition or find a literal in a categorical premise.	Make a compound inference from compound premises.	Find a premise containing an end literal.	Concatenate two premises.
Iterated tactics	Increment the diagram with information from next premise.	Infer a literal from the previous literal and a premise (and add to the consequent of a conditional conclusion).	Make a compound inference from the previous conclusion and a compound premise.	Find a premise containing current literal; if necessary, make an immediate inference to form a conditional. Its consequent becomes the current literal.	Concatenate the next premise.
Final tactic for evaluating given conclusion	Evaluate in relation to the diagram.	Evaluate in relation to initial literal and final literal.	Evaluate in relation to the final compound conclusion.	Evaluate in relation to the final literal in the chain.	Evaluate in relation to an immediate inference from the final concatenation.
Final tactic for formulating a conclusion	Describe each possibility in the diagram.	State final complex conditional.	State the final compound conclusion.	State a conditional conclusion: if initial literal then final literal.	State the final concatenation.

The first tactic is not wholly diagnostics of a strategy (see text).

accounting for inferential tactics, and then for how they are integrated within strategies. Tactics include reading a single premise, writing it down, and drawing a diagram to represent it. Our concern, however, is with *inferential* tactics and with the mechanisms that underlie them. So we turn to the mental model theory, which concerns this lowest level of thinking, and we will show how models can underlie tactics.

A mental model represents a possibility, and so a set of mental models is akin to a truth table (Johnson-Laird & Byrne, 1991). However, a crucial difference arises from the theory's two-fold principle of *truth* (for an extensive discussion of the principle of truth, see Johnson-Laird & Savary, 1999):

> First, the mental models of a set of assertions represent only those situations that are possible given the truth of the assertions. Second, each model represents the literals in the premises (affirmative or negative) only when they are true within the possibility.

As an example of the principle, consider an exclusive disjunction based on two literals (not-A, B):

Not-A ore B.

The principle of truth implies that individuals envisage only the two true possibilities. In one model, they represent the truth of *not-A*; in the other, they represent the truth of *B*. They therefore construct the following two mental models (shown on separate rows):

¬a	
	b

where "¬" denotes negation, and "a" and "b" denote mental models of the corresponding propositions. The principle of truth has a further, less obvious, consequence. When people think about the first possibility, they tend to neglect the fact that *B* is false in this case. Likewise, when they think about the second possibility, they tend to neglect the fact that *not-A* is false in this case. Some commentators have argued that the principle of truth is a misnomer, because individuals merely represent those propositions that are mentioned in the premises. This view is mistaken, however. The same propositions can be mentioned in, say, a conjunction and a disjunction, but the mental models of these assertions are very different. Mental models correspond to those rows that are true in a truth table of the conjunction or the disjunction, and they represent each clause in these assertions, affirmative or negative, only when it is true in the row.

According to the principle of truth, reasoners normally represent what is true. The principle does not imply, however, that they never represent what is false. Indeed, the theory proposes that they represent what is false in "mental footnotes," but that these footnotes are ephemeral (Johnson-Laird & Byrne, 1991). As long as they are remembered, they can be used to construct *fully explicit* models, which represent true possibilities in a fully explicit way. Hence, the mental footnotes about what is false allow reasoners to flesh out their models of the preceding exclusive disjunction, *not-A ore B*, to make them fully explicit:

¬a	¬b
a	b

These explicit models correspond to those for a biconditional of the form: *A iff B*. Yet, most people are surprised to discover that the exclusive disjunction is equivalent to this biconditional. They normally consider mental models, not fully explicit models.

According to the theory, a conditional assertion has two mental models. One model represents the salient possibility in which both the antecedent and the consequent are true. The other model has no explicit content, but is a "place holder" that allows for the possibilities in which the antecedent is false. The mental models for a conditional of the form, *If A then B*, are accordingly:

a b
. . .

where the ellipsis denotes the place holder, which is a wholly implicit model, with a footnote indicating that the antecedent is false in the possibilities that it represents. It is the implicit model that distinguishes the models of a conditional from the model of a conjunction, such as:

A and B

which has only a single model:

a b

The fully explicit models of the conditional can be constructed from the mental models and the footnote on the implicit model. They are as follows:

a b
¬a b
¬a ¬b

Thus, a conditional can be glossed as: *If A then B, and if not-A then B or not-B*. A biconditional has the same mental models as a conditional, but the mental footnote indicates that the implicit model represents the possibility in which both the antecedent and the consequent are false. Hence, the fully explicit models of the biconditional are:

a b
¬a ¬b

The specific meanings of clauses, and general knowledge, can add further information to models, but they can also block the construction of models, giving rise, for example, to other interpretations of conditionals (Johnson-Laird & Byrne, 2002) and disjunctions (Ormerod & Johnson-Laird, 2002).

How can inferences be made with mental models? The next example illustrates a simple method of the sort underlying the step strategy:

A or B.
Not A.
What follows?

The inclusive disjunction yields the mental models:

a	
	b
a	b

The categorical premise has the mental model:

¬a

This model eliminates the first and third models of the disjunction, but it is consistent with the second model, which yields the conclusion: B. This conclusion is valid, i.e., it is necessarily true given the truth of the premises, because it holds in all the models—in this case, the single model—consistent with the premises.

Experimental evidence has corroborated the model theory (see e.g., Johnson-Laird & Byrne, 1991). Inferences based on one model are easier than inferences based on multiple models. Reasoners tend to overlook models and so their systematic errors correspond to a proper subset of the models, typically just a single model. The model theory also predicts the occurrence of illusory inferences. These are compelling, but invalid, inferences that arise from the failure to represent what is false (see Goldvarg & Johnson-Laird, 2000; Johnson-Laird, Legrenzi, Girotto, & Legrenzi, 2000; Johnson-Laird & Savary, 1999; Yang & Johnson-Laird, 2000).

5. The theory of reasoning strategies

In this part of the paper, we develop a theory of strategies and tactics. It derives from the theory of mental models, and from its application to earlier work on strategies in other sorts of reasoning (Bucciarelli & Johnson-Laird, 1999; Johnson-Laird & Byrne, 1990). We formulate the theory in terms of three main assumptions. Following Harman (1973), our first assumption is that reasoning is not a deterministic process that unwinds like clockwork:

1. The principle of *nondeterminism*: thinking in general and sentential reasoning in particular is governed by constraints, but there is seldom just a single path it must follow. It varies in a way that can be captured only in a nondeterministic account.

A deterministic process is one in which each step depends solely on the current state of the process and whatever input it may have (Hopcroft & Ullman, 1979). Psychological theories, however, cannot treat reasoning deterministically, because of the impossibility of predicting precisely what will happen next in any given situation. Our theory of inferential mechanisms constrains the process, but it cannot predict the precise sequence of tactical steps. The correct interpretation of nondeterminism, however, is unknown. On the one hand, the brain might be genuinely nondeterministic. On the other hand, its apparent nondeterminism might merely reflect ignorance: if psychologists had a better understanding of the brain, then they might discover that it was deterministic. Experiment 1 corroborated the principle of nondeterminism, and it did so at two levels. At a high level, the participants developed diverse strategies. At a low level, there was tactical variation within strategies, e.g., individuals differed in which proposition they used as a supposition.

Reasoners are equipped with a variety of inferential tactics, such as making a supposition, and combining compound premises. As they reason about problems, the natural variation in their use of tactics, which is captured in the principle of nondeterminism, leads them to assemble sequences of tactics in novel ways. The result can sometimes be a new reasoning strategy. Our second assumption is accordingly:

2. The principle of *strategic assembly*: naïve reasoners assemble reasoning strategies bottom-up as they explore problems using their existing inferential tactics. Once they have developed a strategy bottom-up, it can control their reasoning in a top-down way.

A corollary of the principle of assembly is that individuals should not develop a reasoning strategy working "top-down" from a high-level specification. This procedure may be possible for experts who think in a self-conscious way about a branch of logic. But, naïve individuals tackling problems spontaneously work "bottom-up" from their existing inferential tactics, trying out different sequences of them. Once they have developed a strategy, and mastered its use in a number of problems, then the strategy itself can unfold in a top-down way.

Granted the principle of strategic assembly, it follows that the space of possible strategies is defined by the different ways in which inferential tactics can be sequenced in order to make inferences. Hence, if we can enumerate tactics exhaustively, then we have specified the recursive basis for all humanly feasible strategies.

Where do the tactics themselves come from? If the mechanism underlying reasoning depends on mental models, then each inferential tactic must be based on models. We therefore make a third assumption:

3. The principle of *model-based tactics*: inferential tactics are based on models. The mechanisms for constructing models are, in turn, constrained by the nature of the human mind, which reflects innate constraints and individual experiences.

The first test of the three principles is to show that mental models can underlie all the strategies and tactics in our taxonomy. A variety of tactics concern reading premises, making diagrams to represent them, and so on, but our concern is inferential tactics, i.e., those tactics that play an essential role in inference. In the following account, we show how each of the five strategies depends on a sequence of model-based tactics. We italicize tactics below, and summarize their role in the strategies in Table 3.

1. The incremental diagram strategy is isomorphic to the cumulative construction of a single set of models based on the premises. The strategy corresponds closely to the processes in the original computer program implementing the model theory (Johnson-Laird & Byrne, 1991). The strategy *finds* either the first premise in the list or a premise containing an end literal, i.e., a literal in a given conclusion, or one that occurs only once in the set of premises. The next step is to *construct* the mental models of the premise. For example, given the problem:

   Blue ore brown.
   Brown ore white.
   White iff red.
   If blue then red?

Table 3
The model-based tactics underlying each of the five strategies: + indicates the use of a tactic, and (+) indicates its optional use

Tactics	The five strategies				
	Incremental diagram	Step	Compound	Chain	Concatenation
Find a premise	+	+	+	+	
Make a supposition	(+)	+			
Concatenate premises or intermediate conclusion		(+)	(+)		+
Construct models	+	+	+	+	+
Update models	+	+	+		
Immediate inference from models	(+)		(+)	+	+
Formulate intermediate conclusion from models		+	+		
Evaluate or formulate a conclusion from models	+	+	+	+	+

the strategy can construct the models for the first premise (shown here on separate rows):

blue
brown

Thereafter, the strategy finds a premise containing a literal already represented in the set of models, and uses the premise to *update* the models. Hence, it uses the second premise above to update the models:

blue	white
brown	

It iterates the process for the third premise:

blue	white	red
brown		

When there are no further premises to be used in incrementing the models, the strategy evaluates the given conclusion in relation to the final set of models. If there is no conclusion to be evaluated, the strategy can use the models to *formulate* a conclusion in disjunctive normal form. Thus, the preceding models can yield the conclusion:

Blue, white, and red, or else brown.

Keeping track of all possibilities compatible with the premises places a heavy load on working memory, though this load can be reduced by the use of an external diagram.

2. The step strategy starts either by *finding* a categorical premise or by *making* a supposition. Consider the problem:

Pink iff black.
Black ore gray.

Gray iff blue.
If not pink then blue?

Because there is no categorical premise, the strategy starts with a supposition corresponding, say, to the antecedent of the conclusion:

Suppose that there is not pink.

The next step is to find a premise containing the same literal or its negation, i.e., pink iff black, and then to *construct* its models, to *update* them with the model of the literal, and to *formulate* an intermediate conclusion based on the result. In the present case, an updating of the fully explicit models of the biconditional:

pink	black
¬pink	¬black

with the model of the supposition, ¬pink, yields the model:

¬pink ¬black

The resulting conclusion is the literal:

Not black.

In cases where the result is more than one model, the conclusion has a modal qualification, e.g., “possibly, there isn’t a black marble,” and any subsequent conclusions are themselves modal in the same sense. The strategy iterates, until it constructs a model containing the other end literal. Hence, the iteration with the second premise above yields:

Gray.

And its iteration with the third premise yields the other end literal:

Blue.

Since there is a given conclusion, it is *evaluated* in relation to this result. The supposition of not pink has led to the conclusion blue, and this relation matches the putative conclusion:

If not pink then blue.

Hence, the conclusion follows from the premises.

If there is no given conclusion, the strategy *formulates* a conclusion. The supposition can be integrated into a complex conditional. If the supposition corresponds to a conjunction, then the antecedent takes the form of a conjunction. If it corresponds to a single literal, then the antecedent is the single literal. The inferred literals are concatenated in the consequent of the conditional. The strategy places a minimal load on working memory because each step pursues the consequences of a single mental model. However, it does not follow that all problems are of the same difficulty if one uses this strategy: the mental models of each premise still need to be built, and the difficulty of a problem increases with the number of models compatible with a premise.

3. The compound strategy relies on a series of compound inferences in which pairs of premises or intermediate conclusions yield compound conclusions, e.g.:

   A ore B.
   B ore C.
   ∴ A iff C (participant 5, problem 19, Experiment 2).

Such inferences are straightforward for the model theory. One premise is used to *construct* models, and the other premise is used to *update* them. In the preceding example, for instance, the premises have two models:

a		c
	b	

which can be used to *formulate* the biconditional conclusion. The combination of two compound premises can put a heavy load on working memory, especially when both premises have multiple models (see Johnson-Laird, Byrne, & Schaeken, 1992). The model theory provides the mechanism required for the compound inferences that underlie the strategy.

The strategy proceeds by *finding* a pair of premises containing an end literal and another literal in common. It *constructs* models of the first premise and uses the second premise to *update* them. It *formulates* a compound conclusion based on the resulting models omitting the literal in common to the two premises. It iterates this procedure until it constructs models containing the two end literals from the premises. If there is a given conclusion, it *evaluates* it in relation to these models. Otherwise, these models are used to *formulate* a compound conclusion.

4. The chain strategy depends on the construction of a chain of conditionals. A chain begins with *finding* an end literal in a given conclusion or a premise. Hence, with the following problem:

   Gray iff red.
   Red ore white.
   White iff blue.
   If not gray then blue?

The first tactic yields the literal:

Not gray.

The chain itself may be based on the premises or on individual diagrams representing them. The next step is to *find* a premise that contains the literal or its negation and that has not been used in the chain:

Gray iff red.

The procedure iterates if this premise is conditional with the literal as its antecedent. Otherwise, as in this case, the premise is used to *construct* a set of models:

gray	red
¬gray	¬red

These models allow an *immediate inference* to a conditional with the literal as an antecedent:

If not gray then not red.

Richardson and Ormerod (1997) have studied how such immediate inferences occur and they have argued that a process of constructing minimal models gives a good account of them. The procedure iterates with the literal in the consequent of the conditional, not red, and makes an immediate inference from the second premise:

If not red then white.

and then an immediate inference from the third premise:

If white then blue.

The chain ends with a conditional containing the other end literal or its negation, or else it is abandoned for want of an appropriate premise. In the present case, the chain leads from one literal in the conditional conclusion to the other, and so the conclusion is *evaluated* as following from the premises. If the given conclusion is not a conditional, then an *immediate inference* converts it into a conditional in which its antecedent matches the initial literal in the chain. If the final consequent in the chain matches the other end literal in the conclusion then the procedure responds that the conclusion follows; otherwise, it responds that the conclusion does not follow. If a conclusion has to be drawn, then a conditional is *formulated* with the first end literal as its antecedent and the final end literal in the chain as its consequent.

5. The concatenation strategy appears at first sight to rely on purely syntactic operations, and therefore to violate the principle of model-based tactics. In fact, the strategy provides a striking vindication of mental models, because it depends critically on them. Given premises of the form:

A iff B.
B ore C.
C iff D.

there are five possible concatenated conclusions depending on the parentheses, e.g., (A iff (B ore (C iff D))). The reader is invited to determine which of them is valid. In fact, none of them is valid. Yet, four participants (10, 11, 13, and 17) in Experiment 2 spontaneously constructed this conclusion:

(A iff B) ore (C iff D).

It is the only concatenation out of the five possibilities that has the same mental models as those of the premises:

a	b		
		c	d

But, the conclusion is an illusory inference, because its fully explicit models do not correspond to those of the premises. Ten participants in Experiment 2 used the tactic of concatenating a conclusion on one or more occasions. On 82% of occasions, the resulting conclusions were compatible with the mental models of the premises, and nine of the ten participants concatenated more conclusions of this sort than not (Sign test, $p < .02$). We conclude that concatenation is not blindly syntactic. Instead, it reflects intuitions about the plausibility of the results, which tend to be accepted only if they yield the same mental models as the premises.

The strategy depends on the tactic of *concatenating* a conclusion. Once an intermediate conclusion is formed in this way, its mental models are *constructed* and compared with those of the premises. The *evaluation* of the conclusion depends on whether the two sets of mental models are the same. The process continues until there are no further premises to be concatenated. The *evaluation* of a given conclusion depends on an *immediate inference* from the concatenation to the given conclusion. The participants did not use the strategy to draw their own conclusions in any of our experiments, though such a use seems feasible.

We conclude that all the strategies that we have observed can be based on tactics that manipulate mental models. Table 3 shows each of the inferential tactics and the role that they play in the five strategies. All the tactics occur in more than one strategy.

Granted the need for nondeterministic theories, we need an exact way to express them so that they can be compared with think-aloud protocols. We propose a methodology that depends on the following steps. First, the different possibilities allowed by the theory are captured in a *grammar*. In the case of a reasoning strategy, we need a grammar in which each step in the strategy calls on a tactic selected from the set of possible tactics. Second, in implementing the strategy in a computer program, each tactic must be modeled in an explicit mechanism that carries out the appropriate inferential process. Third, the computer program includes a parser that uses the grammar to parse think-aloud protocols. Hence, as it uses the grammar to parse a protocol, the program carries out the actual inferential tactics that the theory attributes to reasoners following the strategy. The grammar is thus a parsimonious representation of all the ways in which a strategy can unfold as a sequence of tactical steps. A grammar of a language embodies a theory of all the possible syntactic structures in the language. Likewise, a grammar of a strategy embodies a theory of all the possible tactical sequences in the strategy.

We assume a so-called *regular* grammar of strategies, which corresponds to a finite-state automaton. Finite-state automata do not require any working memory for intermediate results, and so they are the least powerful computational device capable of generating infinitely many sequences (Hopcroft & Ullman, 1979). Of course, the program as a whole makes use of working memory as do human reasoners: our assumption of a regular grammar concerns only the identification of tactical steps in parsing a protocol.

In a grammar of a strategy, each rule corresponds to a tactical step in the strategy. It specifies the state of the system by a numerical label such as S0, the next tactical step, and the resulting state of the system after this step is taken, e.g.:

S0 → read-premise S1
S1 → paraphrase S3
S3 → draw-diagram

Fig. 6 shows such a grammar and an equivalent finite-state automaton. They both implement reading a premise, paraphrasing or making an immediate inference from it, and drawing a diagram based on its meaning. As the figure shows, the system starts in state S0 and then has a choice of different routes. It reads a premise and may stay in the same state (S0)—so that it can

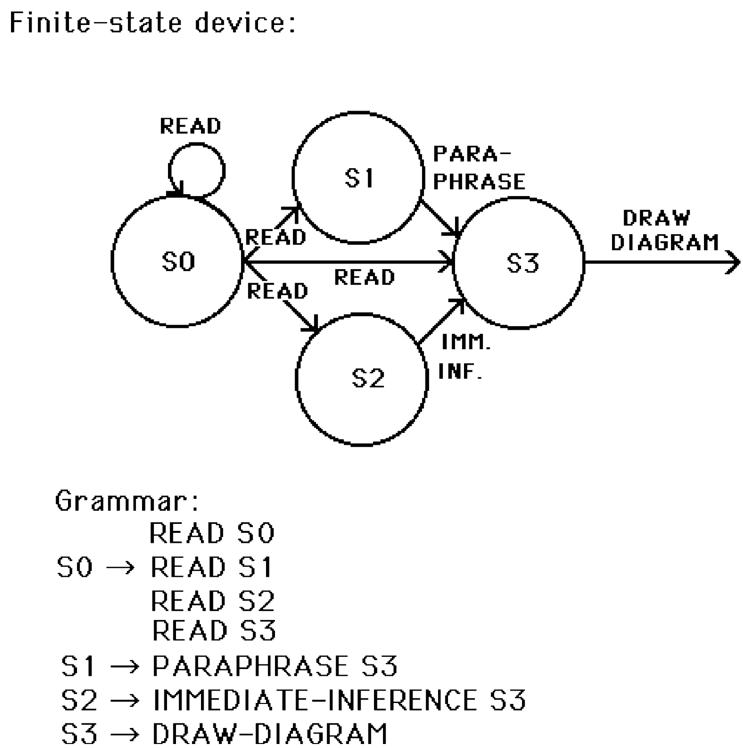

Fig. 6. A nondeterministic finite-state device, and its corresponding grammar, for reading a premise and drawing a diagram to represent it.

read the premise repeatedly, or it jumps to a state (S1) where its next action is to paraphrase the premise, or to a state (S2) where its next action is to make an immediate inference from the premise, or to a state (S3) where its next action is to draw a diagram of the premise. Nothing in the automaton or grammar determines which of these routes is taken. That is why the procedure is nondeterministic.

We can illustrate the method with an example of the program parsing the chain strategy. As the program parses a protocol, it examines each item in the protocol to determine its tactical status. It simulates the mental processes that the theory attributes to reasoners, carrying out all the required tactical steps—drawing diagrams of individual premises, making immediate inferences, and adding conditionals to the chain—as it proceeds through the protocol. Indeed, its ability to carry out these steps provides a check on the accuracy of its tactical assignments to each step in the protocol. Its output recreates both the protocol and its underlying inferential processes according to the mental model theory. The result is that the program makes the same inference as the original participant, and each step in the protocol is annotated to show the postulated mental process. Fig. 7 shows a complete think-aloud protocol of a participant who is using the chain strategy. We have substituted "a," "b," "c," and "d" for the propositions referring to the different colors. We have added comments that label the tactical steps, and we have shown the diagrams drawn by the participant in the notation used by the program. Fig. 8

Problem:
(a ore b)
(b ore c)
(c iff d)
(if a then d) Conclusion

Verbal Protocol:	Diagrams:	Tactical steps:
Drawing initial diagrams:-		
(a ore b)		Read-premise
(a or b)		Immediate-inference
	(draw a / b)	Draw
(b ore c)		Read-premise
	(draw b / c)	Draw
(iff d then c)		Immediate-inference
(if d then c)		Immediate-inference
	(draw d -> c)	Draw
(if d then c)		Immediate-inference
	(show d -> c)	Show-diagram
Checking previous steps, showing diagrams:-		
(a or b)	(show a / b)	Show-diagram
(b or c)	(show b / c)	Show-diagram
(b or c)	(show b / c)	Show-diagram
The chain strategy from d in conclusion:-		
(if d then c)	(show d -> c)	Show-diagram
(if c then not b)		Immediate-inference
	(draw c -> - b)	Draw
(if c then not b)	(show c -> - b)	Check previous two steps
(if not b then a)		Immediate-inference
	(draw - b -> a)	Draw
(if not b then a)	(show - b -> a)	Check previous two steps
(if a then d)		Read-assess-conclusion
(yes)		Assert conclusion

General comment: The participant starts by making a diagrammatic representation of the premises, but subsequently uses the chain strategy to evaluate the conclusion.

Fig. 7. An annotated protocol of the chain strategy in the format used by the computer program modeling the strategy. Paraphrases have been subsumed under the more general tactic of making an immediate inference.

shows the actual output of the program as it parses this protocol and carries out the appropriate inferential tactics.

6. The development of strategies

How do reasoning strategies develop? The process might be idiosyncratic, but the evidence supports the occurrence of robust differences from one individual to another. People are likely

Problem:
(a ore b)
(b ore c)
(c iff d)
(if a then d) Conclusion

The program's output:
Drawing initial diagrams:-

(Read-premise (a ore b))
(Immediate-inference (a or b) from (a ore b))
(Diagram (draw a / b) from (a ore b))
(Read-premises (b ore c))
(Diagram (draw b/c) from (b ore c))
(Immediate-inference (iff d then c) from (c iff d))
(Immediate-inference (if d then c) from (c iff d))
(Diagram (draw d →c) from (c iff d))

. . . Checking previous steps, showing diagrams

The chain strategy from d in conclusion:-
(Diagram (show d →c)from (c iff d))
(Immediate-inference (if c then not b) from (b ore c))
(Diagram (draw c →- b) from (b ore c))
. . . checks two previous steps
(Immediate-inference (if not b then a) from (a ore b))
(Diagram (draw - b →a) from (a ore b))
. . . checks two previous steps Chain is complete.
(Read-conclusion (if a then d))
(Asserts-conclusion (Yes))

The parse was successful.

Fig. 8. A computer parse of the protocol in Fig. 7. We have omitted the program's output for the repetitions of certain steps. Everything within parentheses is an output of the program; our comments are on separate lines without parentheses.

to differ in their reasoning experiences, in the capacity of their working memories, and in their ability to employ complex inferential mechanisms. They are therefore likely to develop different strategies that reflect these differences. Yet, the principle of model-based tactics implies that everyone at the lowest level of thought has the mechanisms for manipulating mental models. It follows that they should be able to acquire any strategy. Some evidence corroborates this hypothesis: Bell (1999) taught naïve participants how to use both the incremental diagram strategy and the step strategy based on suppositions. They all acquired these strategies, and used them with much better than chance accuracy.

The model theory predicts that the nature of the inferential problems given to reasoners should influence the development of strategies. According to the principle of strategic assembly, the characteristics of particular problems should trigger certain strategies "bottom-up." It follows that any element of problems affecting the manipulation of mental models should influence tactics and therefore the development of strategies. One instance of this prediction concerns the effects of number of models. Problems yielding a single explicit mental model should tend to elicit the step and the chain strategy, because these strategies follow up the consequences of a single mental model. Thus, the step strategy follows up the consequences of a categorical premise or supposition. The chain strategy similarly depends on constructing a single explicit mental model corresponding to a chain of conditionals leading from one clause in a conclusion to another. In contrast, problems yielding multiple models should tend to elicit the diagrammatic strategy, because it keeps track of all the possibilities compatible with the premises. In principle, the strategy places a larger load on working memory, but the load is mitigated if reasoners can rely on a diagram as an external memory aid. But, with no categorical premises, and no premises offering a single mental model as a starting point, the step and chain strategies are harder to apply. They call for the construction of fully explicit models. In the case of a one-model problem, such as:

A and B.
B iff C.
C or else D.
Does it follow that A and not-D?

it is easy to generate the consequences of the model of *B*, step by step, and to draw the conclusion. But, in the case of a two-model problem, such as:

A or else B.
B iff C.
C or else D.

the step strategy is harder to apply. Reasoners have to make a supposition—a tactic that some individuals are reluctant to use—and they also need to consider the fully explicit models of the premises in order to draw the final conclusion *D* from the initial supposition *A*. It follows that the greater the number of models called for by an inference, the more likely reasoners should be to use the diagrammatic strategy and the less likely they should be to use the other strategies.

Experiment 2 tested this prediction. The model theory also predicts that the greater the number of models for a problem, the greater the number of errors—a prediction that has been observed in many experiments on logical reasoning (see e.g., Johnson-Laird, 2001). This prediction is independent of reasoners' strategies, because it depends on the process of interpreting premises at the lowest level of thought, e.g., the comprehension of a premise of the form, *A or B or both*, calls for the construction of three models.

6.1. *Experiment 2: number of models and the development of strategies*

6.1.1. *Method*

The participants acted as their own controls and evaluated given conclusions to 36 problems presented in three blocks: 12 one-model inferences, 12 two-model inferences, and 12 three-model inferences. Each participant was assigned at random to one of six groups in order to control for the order of presentation of the three blocks. Within each block, the problems were presented in counterbalanced orders to the participants. Sixteen problems had valid conclusions, and 20 problems had invalid conclusions.

A typical one-model problem was of the form:

A and B.
B ore C.
C iff D.
A and not-D?

where the letters denote propositions about different colored marbles in a box. The set of premises yield just one model:

a	b	¬c	¬d

Each of the 12 problems had premises consisting of one conjunction, either first or last. The other two premises were biconditionals and exclusive disjunctions. The putative conclusion was a conjunction, and for some problems one of its clauses was negative.

A typical two-model problem was of the form:

A iff B.
B ore C.
C iff D.
A iff not-D?

Its premises yield two models:

a	b		
		c	d

Each of these 12 problems had premises consisting of either two biconditionals and one exclusive disjunction or else one biconditional and two exclusive disjunctions. The putative conclusions were either biconditionals or exclusive disjunctions.

A typical three-model problem was of the form:

A iff B.
B iff C.
C or D.
A or D?

Its premises yield three models:

a	b	c	
a	b	c	d
			d

These 12 problems had premises consisting of one inclusive disjunction, either first or last. The other two premises were biconditionals and exclusive disjunctions. The putative conclusions were either conditionals or inclusive disjunctions. As the examples illustrate, all three sorts of problems were laid out so that the clauses in the premises followed one another in the continuous arrangement: A–B, B–C, C–D.

The participants' task was to read the premises and conclusion for each problem and then to decide whether or not the conclusion followed from the premises, i.e., the conclusion must be true given that the premises were true. We used the same think-aloud and video-recording procedure as before. The participants were free to take as much time as they wanted for each problem, but they were not allowed to return to an earlier problem. We tested individually 20 undergraduates from Princeton University, who participated for course credit.

6.1.2. Results

As the model theory predicts, errors increased with the number of models: there were 8% of errors with one-model problems, 15% of errors with two-model problems, and 20% of errors with three-model problems (Page's $L = 251.5$, $p < .05$; this nonparametric test is for a predicted trend over related data and is accordingly one-tailed, see Siegel & Castellan, 1988). The result is in line with the existing literature on reasoning. Even though there was no feedback, the participants showed a marginal tendency to increase in accuracy during the course of the experiment (Page's $L = 243$, $p < .1$). We were able to determine the participants' strategies for 95% of the protocols. Table 4 presents the percentages of the different strategies for the one-, two-, and three-model problems. The only strategies in frequent use were the incremental diagram, step, and compound ones. The participants relied increasingly on the incremental diagram strategy as the problems required a greater number of models (Page's $L = 254.5$, $p < .05$). Concomitantly, they tended to use the step strategy with one-model problems, but its use declined with an increasing number of models. Hence, the results corroborated the principle of strategic assembly: reasoners develop strategies "bottom-up" depending on the sorts of problem that they encounter. With one-model problems, the strategy of choice is to follow up the consequences of a single possibility (based on the conjunctive premise) step by step. As the number of models increases, however, the use of this strategy declines in favor of the diagrammatic strategy. It becomes harder because a supposition has to be made, and the strategy subsequently calls for fleshing out mental models to make them fully explicit. The diagrammatic strategy, however, tracks the multiple possibilities, and much of its memory load is externalized by the use of a diagram.

Table 4
The percentages of the different strategies for the three sorts of problems in Experiment 2 in which the participants evaluated given conclusions

	The strategies			
	Incremental diagram	Step	Compound	Miscellaneous
One-model premises	21	69	0.5	2.5
Two-model premises	26	56	12	3
Three-model premises	49	45	1.5	0.5
Total	32	56	5	2

The chain and concatenation strategies are classified as miscellaneous. The balances of the percentages (5% overall) were uncategorizable strategies.

Although the chain strategy occurred in Experiment 1, we did not observe it in Experiment 2. In Experiment 1, however, there were only two-model problems and for 10 of the 12 problems the conclusions to evaluate were conditionals. In contrast, in Experiment 2, there were no conditional conclusions to evaluate for two-model problems. Conditional conclusions may induce the process of converting the premises into conditionals. There were such conclusions for three-model problems, but the conversion of premises into conditionals may be harder for inclusive disjunctions than for the exclusive disjunctions in Experiment 1 (Richardson & Ormerod, 1997). The absence of conditional conclusions and the presence of inclusive disjunctions may have inhibited the development of the chain strategy. Indeed, the model theory predicts that linguistic cues should elicit certain tactics and hence the development of strategies. We examine this prediction in Experiment 4.

6.2. Experiment 3: strategies in formulating conclusions

The main aims of this experiment were to examine the strategies that reasoners develop when they draw their own conclusions rather than evaluate given conclusions (as in the previous experiments) and to investigate the effects of strategies on the sorts of conclusions that reasoners draw. As in Experiment 2, however, the present experiment also manipulated the number of models.

6.2.1. Method

The participants acted as their own controls and carried out four one-model inferences, four two-model inferences, and four three-model inferences. The inferences were similar to those of Experiment 2, except that there were no conclusions to evaluate. The problems were presented in a different random order to each participant, with the constraint that those with the same number of models never occurred consecutively. For each problem, the premises and the question, “What, if anything, follows?” were printed on a sheet of paper with plenty of space for the participants to write or draw on. We used the same video-recording procedure as before. Participants were instructed that if they thought that no valid conclusion followed from the premises, they had to write down “nothing follows.” There was one training problem of the form:

A iff B.
B iff C.
C iff D.

The participants were free to take as much time as they wanted for each problem, but they were not allowed to return to an earlier problem. We tested individually 24 undergraduates from Princeton University, who participated for course credit.

6.2.2. Results

Once again, the participants developed diverse strategies, and the realization of any particular strategy varied from trial to trial even for the same participant. Table 5 presents the percentages of the different sorts of conclusions for the three sorts of problem: invalid conclusions, “nothing follows” responses, modal conclusions, and incomplete conclusions. As the model theory

Table 5
The percentages of invalid conclusions, "nothing follows" responses, modal conclusions, and incomplete conclusions in Experiment 3

	Sorts of conclusion			
	Invalid	"Nothing follows"	Modal	Incomplete
One-model premise	11	6	5	3
Two-model premises	16	6	10	13
Three-model premises	42	20	19	27

The complement of the invalid responses were valid, nonmodal, and complete. The invalid responses include the "nothing follows" responses and some of the modal and incomplete responses.

predicts, the percentage of invalid conclusions increased with the number of models (Page's $L = 311.5$, $z = 3.39$, $p < .0005$). Some of the conclusions contained modal terms, such as "may" or "might," "can" or "could," or "possibly," e.g.:

> If there is a blue marble then there is a white marble, a red marble and possibly a pink one.

Some of these modal conclusions were valid, but to conclude that there might be, say, a white marble when in fact there *is* a white marble is to draw a conclusion that is weaker and less informative than it need be. Other modal conclusions were invalid. There should be more modal conclusions from multiple-model premises, because uncertainty will increase with the number of possibilities. As Table 5 shows, there was a trend in the predicted direction, but it was not significant (Page's $L = 298$, $z = 1.44$, n.s. $p < .08$).

Some participants drew *incomplete* conclusions that were based on reasoning that failed to take into account all the premises. For example, given problem 12 of the form:

> A or B.
> B ore C.
> C iff D.

one participant drew the conclusion:

> If D then not-B.

This conclusion is valid but it is incomplete because it fails to take into account the first premise. A complete valid conclusion is:

> If D then A.

As Table 5 shows, the occurrence of incomplete conclusions increased with the number of models (Page's $L = 309$, $z = 3.03$, $p < .002$). The phenomenon is easily explained by the difficulty of problems, which in turn is attributable to the number of models. As the difficulty of problems increases, reasoners settle for the less costly effort of drawing a conclusion from two premises rather than from three. Likewise, they show an increasing trend to respond that nothing follows from the premises.

Table 6
The percentages of the different strategies for the three sorts of problems in Experiment 3

	The strategies			
	Incremental diagram	Step	Compound	Chain
One-model premises	14	76	5	3
Two-model premises	33	22	20	9
Three-model premises	36	25	13	7
Overall	28	41	13	7

The balances of the percentages (11% overall) were uncategorizable strategies.

Table 6 presents the percentages of the different strategies in the experiment. As in the previous experiment, the use of the incremental diagram strategy increased with the number of mental models required by the premises. With one-model problems, the participants were likely to use the step strategy, but there was an increase in the use of the incremental diagram strategy with multiple-model inferences. This trend was reliable (Page's $L = 299.5$, $z = 1.66$, $p < .05$). Hence, the principle of strategic assembly is borne out by the present experiment too.

Strategies should influence the form of the conclusions that reasoners draw. In particular, they should tend to draw conclusions in disjunctive form with the incremental diagram strategy. It is difficult for reasoners to see what is common to a number of alternative possibilities, and so they should tend to describe each possibility separately and to combine these descriptions in a disjunction. The other strategies, however, are unlikely to yield conclusions of this sort. These strategies focus on a single possibility, such as a supposition. We examined this prediction by dividing the participants in Experiment 3 into two *post hoc* groups. In the *diagram* group (nine participants), more than half of the participants' identifiable strategies yielding conclusions were cases of the incremental diagram strategy. In the *nondiagram* group (15 participants), more than half of the participants' identifiable strategies yielding conclusions were some other sort. For the diagram group, 63% of the problems solved with the diagrammatic strategy had a conclusion that was a disjunction of possibilities, but for the nondiagram group only 11% of the problems solved with a nondiagrammatic strategy had such a conclusion (Mann–Whitney test, $z = 2.87$, $p < .005$ one-tailed). Different strategies do tend to yield different sorts of conclusion.

6.3. *Experiment 4: strategies and the nature of the premises*

The principle of strategic assembly implies that a way to elicit the incremental diagram strategy is to use premises that are naturally represented as *sets* of possibilities. Disjunctive premises are the obvious candidates, particularly inclusive disjunctions because they have three mental models of possibilities. Hence, problems containing a high proportion of disjunctive premises should predispose reasoners to adopt the diagrammatic strategy. Moreover, disjunctions are less likely to elicit the step and chain strategies, because these strategies require an immediate inference to convert disjunctions into conditionals and these immediate inferences are even harder from inclusive than from exclusive disjunctions (Richardson & Ormerod, 1997).

The simplest way to convert a disjunction (*A or B*) into a conditional is to envisage what follows from one of the disjuncts (e.g., *A*). In the case of an exclusive disjunction (*A or else B*) the result is the negation of the other disjunct (*If A then not-B*). But, in the case of an inclusive disjunction (*A or B or both*), nothing follows. It is necessary to envisage the negation of a disjunct (*Not-A*) in order to infer the other disjunct (*If not-A then B*). In contrast, conditional and biconditional premises yield only one explicit mental model, and so they should be less likely to elicit the diagrammatic strategy, and more likely to elicit the other strategies, including the step and chain strategies, which are based on a single explicit mental model. The aim of this experiment was to test these predictions.

6.3.1. Method

The participants acted as their own controls and drew their own conclusions to two sets of problems: four *disjunctive* problems and four logically equivalent *conditional* problems. The disjunctive problems had as premises one inclusive disjunction and two exclusive disjunctions. The inclusive disjunction was either the first or the third premise, and the two exclusive disjunctions were either of two affirmative literals or one affirmative and one negative literal, e.g.:

A or B.
B ore not-C.
C ore not-D.

The conditional problems were constructed from the disjunctive problems. We transformed each inclusive disjunction into a logically equivalent conditional with a negated antecedent, and each exclusive disjunction into a logically equivalent biconditional with a negated consequent. The preceding problem, for instance, yielded the conditional problem:

If not-A then B.
B if and only if C.
C if and only if D.

The two versions of each problem are logically equivalent, that is, they are compatible with the same set of possibilities. Half the participants received the four disjunctive problems in a random order followed by the four conditional problems in a random order; and half the participants received the two blocks of problems in the opposite order. As in the previous experiments, the participants used pencil and paper, and they had to think-aloud as they tackled the problems. Their protocols were video-recorded. The instructions were the same as those in Experiment 3. We tested 20 undergraduate students from Princeton University, who participated for a course credit.

6.3.2. Results and discussion

One participant could not follow the instructions and was replaced by another prior to the analysis of the data. Table 7 presents the percentages of the different strategies for the two sorts of problems, and it presents the data separately for the two blocks of trials. As predicted, the participants were more likely to use the incremental diagram strategy (56%) for the disjunctive problems than for the conditional problems (23%; Wilcoxon test $T = 66$, $n = 11$, $p < .0005$).

Table 7
The percentages of the different strategies (incremental diagram vs. step, compound, and chain strategies) for (a) the disjunctive problems and (b) the conditional problems in Experiment 4

	The strategies	
	Incremental diagram	Step, compound, and chain
(a) Disjunctive problems		
Presented in first block	58	35
Presented in second block	55	35
Overall	56	35
(b) Conditional problems		
Presented in first block	10	90
Presented in second block	35	60
Overall	23	75

The balances of the percentages are trials with erroneous responses or unclassifiable strategies.

The table shows that the participants who received the conditional problems in the first block rarely developed the incremental diagram strategy (10% of these problems), but their use of the strategy increased reliably for the disjunctive problems in the second block (55% of problems, with seven participants increasing their use, and three participants who never used the strategy in the entire experiment, Sign test, $p < .02$, two-tailed). In contrast, those who received the disjunctive problems in the first block frequently developed the incremental diagram strategy (58% of problems), and did not reliably reduce its use with the conditional problems in the second block (35% of problems, with four participants reducing their use, three maintaining their use, and three never used the strategy in the entire experiment). This difference between the two groups was reliable (Mann–Whitney $U = 21$, $p < .05$, two-tailed). An explanation for the differential transfer is that the incremental diagram strategy is simpler to use with any sort of sentential connective, whereas the step and chain strategies call for additional immediate inferences to convert disjunctive premises into conditionals.

The experiment corroborated the principle of strategic assembly. The nature of the sentential connectives in the premises biases reasoners to adopt particular strategies. Disjunctive premises tend to elicit the incremental diagram strategy, whereas conditional premises tend to elicit the step and the chain strategies. However, the results slightly qualify the prediction of a "top-down" residual effect of a strategy. The incremental diagram strategy increases in use when the problems switch from conditional to disjunctive premises, but does not reliably decline in use when the problems switch from disjunctive to conditional premises. This strategy is more flexible than the other strategies, which are more finely tuned to conditional premises.

7. General discussion

Current accounts of sentential reasoning have neglected strategies. Our aim has been to remedy the neglect and to advance a theory of strategies in sentential reasoning. The theory depends on three assumptions:

1. *Nondeterminism*: thinking in general and sentential reasoning in particular are governed by constraints, but vary in ways that can be captured only in a nondeterministic account.
2. *Strategic assembly*: naïve reasoners assemble reasoning strategies bottom-up as they explore problems using their existing inferential tactics. Once they have developed a strategy, it can control their reasoning in a top-down way.
3. *Model-based tactics*: reasoners' inferential tactics are based on mechanisms that make use of models.

In other words, naïve reasoners are equipped with a set of inferential tactics. As they reason about problems, the variation in their performance leads them to assemble these tactics in novel ways so that they yield a reasoning strategy. As a result, reasoners can develop diverse strategies. All the strategies, however, depend on tactics that can rely on mental models, and so, depending on the properties of problems such as the number of models that they elicit, it is possible to influence which particular strategies reasoners are likely to develop.

Experiment 1 examined the strategies that the participants developed spontaneously to deal with two-model problems, i.e., the premises were compatible with two distinct possibilities. The participants thought aloud as they used pencil and paper to evaluate given conclusions. Their video-taped protocols revealed their strategies, showing that they did indeed develop various strategies, and that within any strategy, the sequence of their tactics differed from trial to trial. For instance, the participants varied in whether they read a premise once or more than once, in whether they drew a diagram of a premise, in whether the proposition they chose as a supposition was the antecedent or the consequent of a conditional conclusion, and so on. There were no fixed sequences of steps that anyone invariably followed.

Naïve reasoners use at least five distinct strategies. Each strategy is built from tactical steps that could rely on the manipulation of models (see Table 3). The incremental diagram strategy keeps track of all the models of possibilities compatible with the premises. The step strategy pursues the step by step consequences of a single model—either one derived from a categorical assertion in a premise or one created by a supposition. The compound strategy combines the models of compound premises to infer what is necessary or possible. The chain strategy pursues a single model in a sequence of conditionals leading from one proposition in a conclusion to another. It calls for model-based inferences that convert premises into the appropriate conditionals for the chain. The concatenation strategy forms a conclusion by concatenating the premises, but normally only if the resulting conclusion has the same mental models as the premises. Because it relies on mental models, it can give rise to illusory inferences, i.e., conclusions that seem highly plausible, but that are invalid (Johnson-Laird & Savary, 1999).

Some of these strategies could rely on formal rules of inference as conceived in current theories (Braine & O'Brien, 1998; Rips, 1994). Thus, the step strategy and the compound strategy could certainly be based on formal rules. The incremental diagram strategy, however, is beyond the scope of formal rule theories, because possibilities play no direct role in these theories. One might argue that each possibility could be treated as a supposition, but this notion runs into severe problems. Consider, for instance, an inference based on an initial disjunction:

Yellow or black, or both.

The reasoner would have to make three suppositions corresponding to the three possibilities, but such a tactic is impossible according to the formal rules postulated in current formal rule theories. Rips (1994) allows suppositions to be made only working backwards from a putative conclusion; Braine and O'Brien (1998) allow suppositions to be made only after direct rules of inference have been exhausted. Both theories demand that suppositions be discharged either by drawing a conditional conclusion, or by negating those suppositions that yield contradictions. No conditional conclusions emerge in the incremental diagram strategy corresponding to each of the possibilities. It is moot point whether the chain and concatenation strategies could be based on formal rules. In any case, the decisive problem for current formal rule theories is that they postulate only a single deterministic strategy, and our results show that this assumption is false.

Experiments 2 and 3 corroborated the strategies with a greater range of problems. These experiments also confirmed that logical accuracy declines with an increased number of mental models. Previous studies had shown such effects for simple inferences based on a single sentential connective. The new experiments, however, generalized the results to inferences based on three connectives. Moreover, Experiment 3, in which the participants had to draw their own conclusions, showed that with a greater number of models, the participants were more likely to draw conclusions that failed to take into account all the premises, and to make responses of the form, "nothing follows."

The theory explains how people develop reasoning strategies. They assemble strategies from their existing tactics, but according to the principle of strategic assembly various properties of inferential problems should trigger the use of particular strategies. The problems in Experiments 2 and 3 called for one, two, or three models. As the principle predicts, this variable had a reliable effect on the development of strategies. The participants tended to use the conjunction in one-model problems as the starting point for the step strategy. But, this strategy calls for more complicated processes with multiple-model problems, and, in particular, for fleshing out mental models to make them fully explicit. Hence, the use of this strategy declines with problems that call for multiple models. With such problems, reasoners are instead more likely to use the incremental diagram strategy, which keeps track of all the different possibilities compatible with the premises. In principle, this strategy places a greater load on working memory, but the load is mitigated by the use of an external diagram. The theory predicts that without such a diagram the use of the strategy is likely to be less common, and less effective, with multiple-model problems, because reasoners would have to hold the alternative models in working memory.

Experiment 4 also bore out the principle of strategic assembly. Different sentential connectives in logically equivalent problems biased reasoners to develop different strategies in a predictable way. Disjunctive premises call for multiple mental models and so they tended to elicit the incremental diagram strategy, whereas conditional premises call for only a single explicit mental model and so they tended to elicit the step, chain, and compound strategies. The participants increased their use of the incremental diagram strategy on switching to disjunctive premises, but they did not decrease its use reliably on switching to conditional problems.

Incremental diagrams are more flexible than those strategies that are geared to conditional premises. Bell (1999) has corroborated this claim in a pedagogical study (cf. Nickerson, 1994). She taught naïve reasoners to make an explicit use of the incremental diagram strategy. The teaching procedure took only a few minutes, and it led to a striking improvement in both the speed and accuracy of sentential reasoning—as much as a 30% improvement in accuracy. It even improved performance when the participants were denied the use of pencil and paper.

What results would have refuted the model theory? At the lowest level of reasoning, the level of inferential mechanisms, it would have been refuted if multiple-model problems had not led to an increase in difficulty. But, Experiments 2 and 3 showed that difficulty did increase in the predicted way, e.g., in an increased number of errors. This phenomenon has been observed in previous studies, but not before in inferences based on three sentential connectives.

At the tactical level, the model theory would have been refuted if reasoners used tactics incompatible with manipulations of models. Suppose, for example, that concatenation had not been sensitive to the mental models of the premises, and therefore had not led reasoners to make systematic illusory inferences. In that case, a tactic would have been controlled purely by syntactic considerations, and it would have been inconsistent with the theory. Moreover, if the principle of model-based tactics is correct, then certain logically valid inferential tactics cannot be part of human competence. For example, consider an inference based on inclusive disjunctions:

A or B.
Not-B or C.
∴ A or C.

This inference is valid, and many systems of automated theorem-proving rely on a corresponding formal rule of inference, which is known as the *resolution* rule. No such inference, however, should be an inferential tactic, because its mental models would place too great a load on the processing capacity of working memory. The first premise has the mental models:

a	
	b
a	b

The second premise has the mental models:

¬b	
	c
¬b	c

The correct conjunction of these two sets requires reasoners to take into account information about what is false. A program that we have implemented performs at this level of competence and it yields the following set of four models:

a	¬b	
a	¬b	c
	b	c
a	b	c

These models support the valid conclusion: *A or C*. But, the need to construct four models will prevent naïve reasoners from using this inferential tactic spontaneously.

At the strategic level, the model theory could have been refuted in at least two ways. One way would have been if reasoners had developed strategies based on tactics that do not depend on models, e.g., the resolution rule. Hence, reasoners should not develop a resolution strategy, in which they convert premises into disjunctions, and construct a chain of disjunctions analogous to the conditionals in the chain strategy. Thus, given the problem:

If A then B.
C iff B.
D or not-C.
∴ If A then D.

naïve reasoners should not develop a strategy that constructs a chain of disjunctions:

Not-A or B.	(immediate inference from the first premise)
Not-B or C.	(immediate inference from the second premise)
Not-C or D.	(paraphrase of third premise)
∴ Not-A or D.	(immediate inference from the conclusion)

Such strategies are commonly used in artificial intelligence (Wos, 1988), but they should be psychologically impossible because they violate the principles of strategic assembly and model-based tactics. Another way in which the model theory would have been refuted is if reasoners had uniformly developed a single deterministic strategy (cf. Rips, 1994). In fact, the evidence suggests that any intelligent adult is able to acquire any strategy that relies on model-based tactics (Bell, 1999). It is impossible to predict precisely which strategy an individual will spontaneously develop in tackling a set of problems. Some people seem to be set in their ways, and then suddenly change their strategy; other people do not even settle down to a consistent choice. At this level of theorizing, we must settle for nondeterminism. Even at the level of tactics, however, when a person reads a premise aloud, for example, they may do so once and then proceed to the next premise, or they may read the premise and then read it again, and even again. We cannot predict precisely what they will do on a given occasion. Theories of reasoning must accordingly be nondeterministic from the highest to the lowest level.

At the highest level, the model theory would have been refuted if the participants had developed strategies top-down from a metacognitive specification. But the manifest signs of such an approach were totally lacking in all the protocols from our experiments. No one ever remarked, for example, "The way I should solve these problems is to construct a chain of conditionals." None of our participants ever described an insightful strategy. According to the principle of strategic assembly, however, such metacognitive remarks should be observable if naïve reasoners have first been able to build strategies bottom-up, and only then are asked to describe them.

All our results come from studies in which highly intelligent Princeton undergraduates think-aloud as they reason about problems concerning colored marbles. Reasoning by other sorts of individual about other sorts of materials without the need to think-aloud might yield very different strategies. Such a skeptical view is possible, but improbable. Our principal result is that different people develop different strategies. If our narrow sample of individuals working on highly constrained materials corroborated this claim, then more diverse people working on more diverse materials are unlikely to overturn it. Similarly, it is unlikely that the sorts of thinking that occur in silence are wholly different from the sorts of thinking that occur in our studies. "Think-aloud" data, as we argued in Section 1, can be a reliable guide to the sequence of a person's thoughts. Indeed, one striking communality is that the number of models called for by the premises reliably predicts the difficulty of inferences whether or not reasoners have to think-aloud. This variable also affects highly intelligent individuals and those from the population at large (see e.g., Johnson-Laird, 1983, p. 117–121), though the former tend to be better reasoners than the latter (Stanovich & West, 2001). And number of models also affects inferences based on everyday contents (e.g., Ormerod & Johnson-Laird, 2002). Nevertheless, a sensible task for the future is to examine how intellectual ability, materials from daily life, and various experimental procedures, affect the development of strategies in sentential reasoning. The manipulation of these variables may modify the frequency of usage of the various strategies; they may lead to the discovery of new strategies; they may show that reasoners' goals influence the strategies that they develop. But, in our view, at the root of sentential strategies is an everyday understanding of negation and connectives such as "if," "or," and "and." Hence, the effects of variables such as intelligence, or everyday contents, are unlikely to overturn our basic findings.

Reasoning depends on strategies that call for a nondeterministic account. The strategies develop as a result of reasoners trying out various tactical steps, but these manipulations are sensitive to the properties of problems, both the nature of their premises and the number of models that they elicit. The tactical steps rely in turn on unconscious inferential mechanisms that manipulate mental models. Unlike other domains such as arithmetic (Lemaire & Siegler, 1995), the study of strategies in reasoning has barely begun. Future studies need to delineate the effectiveness and efficiency of the various strategies. They need to account for the sequences of strategies that reasoners pass through as they gain experience and expertise. Logic, one could say, is the ultimate strategy that some highly gifted individuals attain.

Acknowledgments

This research was made possible in part by grants to the first and the third authors, respectively from the European Commission (Marie Curie Fellowship) and the National Science Foundation (Grant BCS-0076287) to study strategies in reasoning. We thank Fabien Savary for carrying out Experiment 1 and for transcribing its protocols. We are also grateful for the help of many colleagues, including Bruno Bara, Patricia Barres, Victoria Bell, Monica Bucciarelli, Ruth Byrne, Wim DeNeys, Kristien Dieussaert, Zachary Estes, Vittorio Girotto, Yevgeniya Goldvarg, Uri Hasson, Patrick Lemaire, Paolo Legrenzi, Maria Legrenzi, Juan Madruga, Hansjoerg Neth, Mary Newsome, Ira Noveck, Guy Politzer, Walter Schaeken, Walter Schroyens, Vladimir Sloutsky, Dan Sperber, Patrizia Tabossi, and Isabelle Vadeboncoeur. Some of the research was presented to the Brussels Workshop on Deductive Reasoning and Strategies, 1997; and we thank the participants, and the organizers Gery d'Ydewalle, André Vandierendonck, G. De Vooght, and Walter Schaeken, for their constructive remarks. We thank Jon Baron, Nick

Chater, Mike Oaksford, and an anonymous reviewer for their constructive criticisms of an earlier version of the paper.

References

Baron, J. (1994). *Thinking and deciding* (2nd ed.). New York: Cambridge University Press.

Bell, V. (1999). *The model method*. An unpublished Ph.D. dissertation. Department of Psychology, Princeton University.

Berardi-Coletta, B., Buyer, L., Dominowski, R., & Rellinger, E. (1995). Metacognition and problem solving: A process-oriented approach. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*, 205–223.

Braine, M. D. S., & O'Brien, D. P. (Eds.). (1998). *Mental logic*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bruner, J. S., Goodnow, J. A., & Austin, G. A. (1956). *A study of thinking*. New York: Wiley.

Bucciarelli, M., & Johnson-Laird, P. N. (1999). Strategies in syllogistic reasoning. *Cognitive Science*, *23*, 247–303.

Byrne, R. M. J., & Handley, S. J. (1997). Reasoning strategies for suppositional deductions. *Cognition*, *62*, 1–49.

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.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*, 215–251.

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

Garnham, A., & Oakhill, J. V. (1994). *Thinking and reasoning*. Oxford: Basil Blackwell.

Goldvarg, Y., & Johnson-Laird, P. N. (2000). Illusions in modal reasoning, memory. *Memory* & *Cognition*, *28*, 282–294.

Greenwald, A., & Banaji, M. (1995). Implicit social cognition: Attitudes, self-esteem and stereotypes. *Psychological Review*, *102*, 4–27.

Harman, G. (1973). *Thought*. Princeton, NJ: Princeton University Press.

Hopcroft, J. E., & Ullman, J. D. (1979). *Formal languages and their relation to automata*. Reading, MA: Addison-Wesley.

Hunt, E. B. (1999). What is a theory of thought? In R. J. Sternberg (Ed.), *The nature of cognition* (pp. 3–49). Cambridge, MA: MIT Press.

Hunter, I. M. L. (1957). The solving of three terms series problems. *British Journal of Psychology*, *48*, 286–298.

Huttenlocher, J. (1968). Constructing spatial images: A strategy in reasoning. *Psychological Review*, *75*, 550–560.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. London: Routledge & Kegan Paul.

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

Johnson-Laird, P. N. (2001). Mental models and deductive reasoning. *Trends in Cognitive Science*, *5*, 434–442.

Johnson-Laird, P. N., & Bara, B. G. (1984). Syllogistic inference. *Cognition*, *16*, 1–61.

Johnson-Laird, P. N., & Byrne, R. M. J. (1990). Metalogical problems: Knights, knaves, and rips. *Cognition*, *36*, 69–81.

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

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review* (in press).

Johnson-Laird, P. N., & Savary, F. (1999). Illusory inferences: A novel class of erroneous deductions. *Cognition*, *71*, 191–229.

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

Johnson-Laird, P. N., Legrenzi, P., Girotto, V., & Legrenzi, M. S. (2000). Illusions in reasoning about consistency. *Science*, *288*, 531–532.

Kolodner, J. (1993). *Case-based reasoning*. San Mateo, CA: Morgan Kaufman.

Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General*, *124*, 83–97.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Henry Holt.

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

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs: Prentice-Hall.

Nickerson, R. S. (1994). The teaching of thinking and problem solving. In R. J. Sternberg (Ed.), *Thinking and problem solving* (pp. 409–449). San Diego: Academic Press.

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

Ormrod, J. E. (1979). Cognitive processes in the solution of three-term series problems. *American Journal of Psychology*, *92*, 235–255.

Ormerod, T. C., & Johnson-Laird, P. N. (2002). *How pragmatics modulates the interpretation of sentential connectives* (submitted for publication).

Piaget, J. (1921). Une forme verbale de la comparaison chez l'enfant. *Archives de Psychologie*, 141–172.

Polk, T. A., & Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review*, *102*, 533–566.

Quinton, G., & Fellows, B. J. (1975). "Perceptual" strategies in the solving of three-term series problems. *British Journal of Psychology*, *66*, 69–78.

Richardson, J., & Ormerod, T. C. (1997). Rephrasing between disjunctives and conditionals: Mental models and the effects of thematic content. *Quarterly Journal of Experimental Psychology*, *50A*, 358–385.

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

Rips, L. J. (1994). *The psychology of proof*. Cambridge, MA: MIT Press.

Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols. *Memory* & *Cognition*, *17*, 759–769.

Schaeken, W., De Vooght, G., Vandierendonck, A., & d'Ydewalle, G. (2000). *Deductive reasoning and strategies*. Mahwah, NJ: Lawrence Erlbaum Associates.

Schooler, J., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General*, *122*, 166–183.

Shastri, L., & Ajjanagadde, V. (1993). From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behavioral and Brain Sciences*, *16*, 417–494.

Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw Hill.

Stanovich, K. E., & West, R. F. (2001). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*, 645–726.

Stenning, K., & Oaksford, M. R. (1993). Rational reasoning and human implementations of logic. In K. I. Manktelow & D. E. Over (Eds.), *Rationality: Psychological and philosophical perspectives*. Florence, KY: Taylor and Francis/Routledge.

Sternberg, R. J. (1977). *Intelligence, information-processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J. (1983). Components of human intelligence. *Cognition*, *15*, 1–48.

Sternberg, R. J. (Ed.). (1984). *Human abilities: An information-processing approach*. San Francisco: Freeman.

Wason, P. C., & Johnson-Laird, P. N. (1972). *The psychology of deduction: Structure and content*. Cambridge, MA: Harvard University Press, London: Batsford.

Wood, D. J. (1969). Approach to the study of human reasoning. *Nature*, *223*, 102–103.

Wood, D. J., Shotter, J. D., & Godden, D. (1974). An investigation of the relationships between problem solving strategies, representation and memory. *Quarterly Journal of Experimental Psychology*, *26*, 252–257.

Wos, L. (1988). *Automated reasoning: 33 basic research problems*. Englewood Cliffs, NJ: Prentice-Hall.

Yang, Y., & Johnson-Laird, P. N. (2000). Illusory inferences with quantified assertions: How to make the impossible seem possible, and vice versa. *Memory* & *Cognition*, *28*, 452–465.