

## On imagining what is true (and what is false)

Patricia E. Barres

*Fairleigh Dickinson University, Teaneck, NJ, USA*

P.N. Johnson-Laird

*Princeton University, NJ, USA*

How do people imagine the possibilities in which an assertion would be true and the possibilities in which it would be false? We argue that the mental representation of the meanings of connectives, such as “and”, “or”, and “if”, specify how to construct the true possibilities for simple assertions containing just a single connective. It follows that the false possibilities are constructed by inference from the true possibilities. We report converging evidence supporting this account from four experiments in which the participants had to list the true and the false possibilities for various sorts of assertion. Their systematic errors, response times, and think-aloud protocols corroborated the theory’s predictions. Not even true possibilities are immediately available for complex assertions containing two or more different connectives. The task of listing the true possibilities then becomes difficult, and the task of listing the false possibilities almost impossible.

A fundamental semantic ability is to imagine the circumstances in which an assertion is true. If you understand an assertion, then you should be able to imagine such circumstances; if you cannot imagine any circumstance in which an assertion is true, then you are hardly able to understand it. As Wittgenstein (1922, 4.024) remarked: “To understand a proposition means to know what is the case, if it is true”. Almost as important, however, is the ability to imagine the

---

Address correspondence to Dr P.E. Barres, 13121 N.E. 123rd Street, Apartment 206, Kirkland, WA 98034 USA. Email:pebarres@alumni.princeton.edu

This research was made possible in part by a grant to the second author from the National Science Foundation (Grant BCS 0076287) to study strategies in reasoning. We thank Jonathan Evans, Steve Newstead, and David Over for comments on an earlier draft of this paper. We are grateful to many colleagues for their help and advice: Victoria Bell, Monica Bucciarelli, Ruth Byrne, Zachary Estes, Vittorio Girotto, Yevgeniya Goldvarg, Paolo Legrenzi, Maria Sonino Legrenzi, Juan Madruga, Mary Newsome, Fabien Savary, Walter Schaeken, Patrizia Tabossi, Lisa Torreano, Isabelle Vadeboncoeur, Yingrui Yang, and Lauren Ziskind.

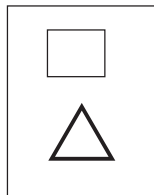
## 2 BARRES AND JOHNSON-LAIRD

circumstances in which an assertion is false. Psychologists have studied the verification of assertions, that is, how individuals decide whether an assertion is true or false in relation to a situation. They have discovered, for example, that it is easier to make such judgements about affirmative sentences than about negative sentences (e.g., Clark & Chase, 1972; Wason, 1959, 1961). They have also studied conditional assertions in the “truth table” task in which individuals have to evaluate whether or not various cases conform to a conditional, and in a task in which they have to construct cases that make it true and cases that make it false (see Evans, Newstead & Byrne, 1993). Strangely, they have neglected other sorts of sentential connectives and assertions containing more than one sentential connective. The aim of the present paper is to make good this neglect and to develop a theory, with supporting evidence, of how individuals imagine the possibilities in which sentences would be true and the possibilities in which they would be false.

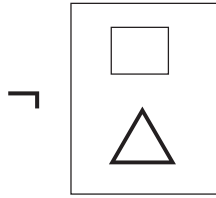
The theory is based on mental models, which were originally postulated in order to explain the comprehension of discourse and deductive reasoning (e.g., Johnson-Laird, 2001; Johnson-Laird & Byrne, 1991). According to this theory, possibilities are central to comprehension. That is, when people understand an assertion, they use its meaning to imagine the possibility or set of possibilities to which it refers, where each possibility is represented in a mental model. Given a simple assertion, such as:

The square is above the triangle

individuals can use their grasp of its meaning to construct a mental model of the spatial arrangement between the two objects. The model captures all the different ways in which the possibility might be realised, because the model specifies what they have in common, namely, the spatial relation exemplified in the following vertical arrangement:



There are many ways in which the assertion would be false, e.g., if the square were below the triangle or on the same level with it. The theory postulates, however, that individuals have the option of using negation as an operator on a model. Hence, they can represent the possibilities in which the assertion is false by constructing a mental model of the form:



where “ $\neg$ ” denotes negation.

Our concern is assertions containing sentential connectives, such as “if”, “or”, and “and”. We refer to these assertions using the standard terminology. Simple assertions that contain neither negation nor connectives, such as, “The teacher owns a house”, are *atomic*, and assertions composed from atomic assertions and sentential connectives, such as, “If the teacher owns a house, then the student rents a room,” are *compound*. The sentences that we focus on are true or false depending on the truth or falsity of the clauses that they interconnect. The following exclusive disjunction, for example:

The battery is dead or else the circuit isn’t connected

is true in the case that the battery is dead and the circuit is connected, and in the case that the battery isn’t dead and the circuit isn’t connected. Otherwise, it is false. If a connective has an interpretation of this sort, then logicians define its meaning using a truth table. Table 1 is a truth table of the preceding exclusive disjunction.

TABLE 1  
The truth table of an exclusive disjunction, “The battery is dead or else the circuit isn’t connected”

<i>The battery is dead</i>	<i>The circuit is connected</i>	<i>The battery is dead or else the circuit isn’t connected</i>
True	True	True
True	False	False
False	True	False
False	False	True

The processing capacity of working memory is limited, and so it seems unlikely that individuals represent assertions in truth tables. Everyday discourse, however, respects Grice’s (1975) cooperative principle, which postulates various maxims governing the conventions of discourse. One such convention is that assertions should be true. Indeed, lies would be ineffective without such a convention. Similarly, truth is more useful than falsity. The fundamental principle of the model theory is accordingly:

#### 4 BARRES AND JOHNSON-LAIRD

The principle of *truth*: each mental model of one or more assertions represents a possibility given the truth of the assertions, and each mental model represents a clause in the assertions only when it is true in that possibility (see, e.g., Johnson-Laird & Savary, 1999).

One consequence of the principle is that the load on working memory is minimised, because mental models—unlike truth tables—do not represent what is false.

The principle of truth can best be explained by way of an example. Consider again the exclusive disjunction:

The battery is dead or else the circuit isn't connected.

It is compatible with two possibilities, and so each of them is represented in a mental model, shown here on separate lines:

dead  
     $\neg$  connected

where “dead” denotes a model of the battery as dead, “ $\neg$ ” denotes negation, and so “ $\neg$  connected” denotes a model of the circuit as not connected. One consequence of the principle is that mental models do not represent possibilities that are false (cf. the truth table for the disjunction in Table 1). Another consequence of the principle is that the first mental model above represents the possibility in which it is true that the battery is dead, but the model does not make explicit that it is false that the circuit is not connected in this possibility. Similarly, the second mental model represents the possibility in which it is true that the circuit is not connected, but the model does not make explicit that it is false that the battery is dead in this possibility. As the example illustrates, mental models represent negative clauses when they are true, but they do not represent *false* clauses, whether they are affirmative or negative. Reasoners make mental “footnotes” to capture the information about what is false, but these footnotes are soon forgotten in normal cases. Johnson-Laird and Byrne (1991) used a special notation to represent such footnotes, but we will forego the notation here.

A conditional 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 the battery is dead then the circuit is connected”, are accordingly:

dead          connected  
    ...

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:

The battery is dead and the circuit is connected

which has only a single model:

dead            connected

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 as long as individuals remember the mental footnotes about what is false, they can use them to construct *fully explicit* models of the true possibilities. Hence, the mental footnotes about what is false allow reasoners to flesh out their models of the exclusive disjunction, “The battery is dead or else the circuit isn’t connected”, to make them fully explicit:

dead            connected  
 $\neg$  dead         $\neg$  connected

The fully explicit models of a conditional, such as: “If the battery is dead then the circuit is connected”, can be constructed from the mental models and the footnote on the implicit model. They are as follows:

dead            connected  
 $\neg$  dead        connected  
 $\neg$  dead         $\neg$  connected

Hence, the conditional can be glossed as: If the battery is dead then the circuit is connected, and if the battery is not dead then the circuit is or isn’t connected.

How do individuals imagine the possibilities in which a compound assertion would be false? In principle, these possibilities could be as available as the true possibilities. If the mind contained truth tables, such as Table 1, individuals would have equal access both to the true possibilities and to the false possibilities. In contrast, the principle of truth implies that individuals have direct access only to true possibilities, and not even to fully explicit models of them. It follows that even in the case of a simple compound assertion with two clauses, such as:

The battery is dead and the circuit is connected



In summary, according to the model theory, individuals have access to the possibilities in which simple compound assertions are true. They construct mental models of such possibilities, albeit these models are incomplete, but to imagine its false possibilities, they need to go a step further and to infer them from their models of the true possibilities. The theory makes three main predictions. First, individuals should be more accurate in imagining true possibilities than in imagining false possibilities, because false possibilities have to be inferred from true possibilities. Second, as a corollary, individuals should be faster to list true possibilities than false possibilities. Third, they should be faster to list false possibilities when they have already listed the true possibilities than when they have to list the false possibilities *ab initio*. In the first case, they have the true possibilities in mind and can derive the false ones from them. But, in the second case, they have to generate the true possibilities before they can derive the false possibilities from them. The aim of our first experiment was to test these three predictions.

## EXPERIMENT 1

Experiment 1 examined the ability of naïve individuals to imagine the possibilities in which assertions would be true and the possibilities in which they would be false. Previous research has shown that material about everyday events can modulate the basic interpretations of compound assertions, both by eliminating otherwise feasible possibilities and by adding temporal and spatial relations (Johnson-Laird & Byrne, 2002). In order to avoid this problem the experiment used propositions concerning the presence of letters and numbers, i.e., propositions such as, “If there is an A then there is a 1”. In these *basic* propositions, there is no relation—either in meaning or in general knowledge—between the two clauses, other than their occurrence in the same conditional. These propositions were combined in assertions based on a single sentential connective, and the experiment examined four sorts of connectives: conjunctions, conditionals, exclusive disjunctions, and inclusive disjunctions.

### Method

*Design and materials.* The participants listed the true possibilities and the false possibilities for four sorts of assertion of the following forms:

Conjunctions: There is an A and there is a 1.

Conditionals: If there is an A then there is a 1.

Exclusive disjunctions: Either there is an A or else there is a 1, but not both.

Inclusive disjunctions: There is an A or there is a 1, or both.

There were eight trials with different contents for each of the four sorts of assertion (32 problems). The participants listed the true possibilities and the false

possibilities in separate blocks of trials, making up a total of 64 trials. They were assigned at random to two groups. One group carried out the block of “true” trials and then the “false” trials (the “True–false” group), and the other group carried out in the two blocks in the opposite order (the “False–true” group). The order of the assertions in the two blocks was the same, but each participant received a different random order. All the assertions had a letter in the first clause and a number in the second clause, and the letters and numbers were assigned at random to the assertions so that no participant encountered the same pair more than once in the experiment. Henceforth, we will state the assertions in an abbreviated form, e.g., “B and 2”.

*Procedure.* The participants were tested individually. They were told that in one block of trials their task was to read an assertion and to write down a list of the possibilities in which it would be true, where each possibility was a pairing of letters and numbers, such as “C 5”. Therefore, participants understood that they were to use different combinations of individual letters and numbers to represent each possibility. In another block of trials, they would receive the same assertions, but now they had to write down a list of the possibilities in which they would be false. A custom-written Hypercard program presented the assertions one-by-one on the screen of a Power Macintosh (7100) computer. The participants listed the possibilities by typing each of them on a separate line in a response box on the computer screen. They were told to use the mouse to click on the response box whenever they were ready to type. When they could think of no more possibilities, they clicked on a button to advance to the next assertion. They were told to concentrate on listing the correct possibilities, and to take as much time as they needed. In fact, the computer recorded the latency from the presentation of each assertion on the screen to the click indicating that the participant was ready to type the first possibility.

*Participants.* A total of 24 Princeton undergraduates, who had no training in formal logic, carried out the experiment. They received course credit for their participation.

## Results and discussion

Participants generated lists of true and false possibilities for the assertions (each list consisting of 1 or more possibilities). For instance, for an exclusive disjunction such as “A or 1, but not both”, participants often listed the two possibilities:

A

1



When they represented negation, they generally did so by writing “not A” or “no A”. But a minority represented negation by using a different letter or number, e.g., instead of “not A 1” they wrote: “E 1”. Appendix A presents four tables summarising the possibilities that the participants listed for each of the four sorts of assertion. The most salient features of these data is that reasoners vary considerably in what they judge to be true and false, and that they tend to make many errors.

What counts as a correct listing? Consider, for example, the exclusive disjunction, Either A or else 1, but not both. The listing of the true fully explicit possibilities:

$$\begin{array}{ll} A & \neg 1 \\ \neg A & 1 \end{array}$$

is obviously correct. But the listing of the true possibilities corresponding to mental models:

$$\begin{array}{ll} A & \\ & 1 \end{array}$$

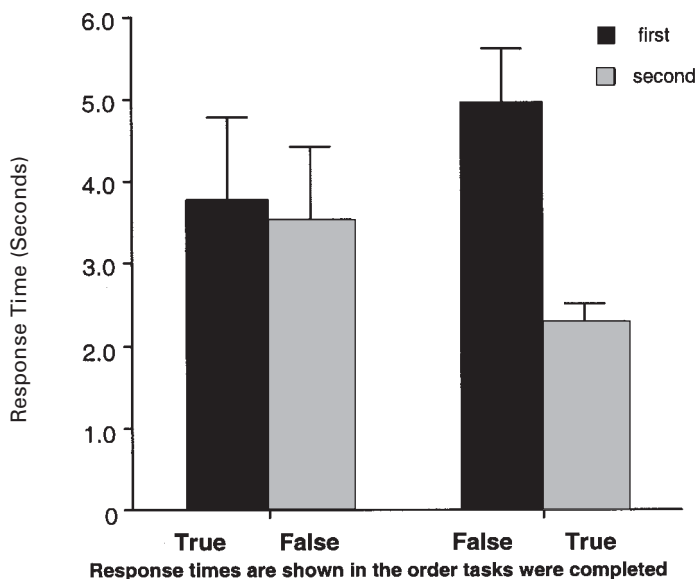
is also correct. Table 3 presents the percentages of both sorts of correct possibilities for each of the four sorts of assertion. As the theory predicts, the participants listed the true possibilities (78% correct) more accurately than the false possibilities (34% correct; Wilcoxon test,  $z = 4.0$ ,  $p < .001$ ). Only conditionals tended not to yield a reliable difference (see Table 3).

TABLE 3  
The percentages of correct responses in listing the true possibilities and the false possibilities for the four sorts of assertions in Experiment 1

	<i>True possibilities</i>	<i>False possibilities</i>
Conjunctions	81 (81)	20
Conditionals	58 (50)	49
Exclusive disjunctions	72 (50)	23
Inclusive disjunctions	76 (60)	30

The percentages in parentheses show the true possibilities corresponding to mental models, and the balance of the percentages corresponded to fully explicit models. The mental model of a conjunction is the same as its fully explicit model. For the conditionals, there were a further 11% of true possibilities and a further 11% of false possibilities corresponding to the fully explicit models for a *biconditional* interpretation.

Figure 1 presents the mean latencies for the two groups of participants. One group listed the true possibilities in the first block of trials and then the false possibilities in the second block (the *True-false* group); and the other group listed the false possibilities in the first block of trials and then the true possibilities in the second block (the *False-true* group). These mean latencies are for all responses, correct and incorrect, because the participants made so many errors. As the theory predicts, the participants were faster to list the true possibilities (mean 3.0 s) than to list the false possibilities (mean 4.3 s; Wilcoxon test,  $z = 3.0$ ,  $p < .02$ ). However, as the theory also predicts, there was a significant advantage in listing the false possibilities for the group who had already listed the true possibilities. The participants who constructed the true possibilities and then the false possibilities did not differ in their latencies of response in the two cases (see the left-hand side of Figure 1). However, the participants who constructed the false possibilities and then the true possibilities took reliably longer to construct the false possibilities than the true ones, but they were also faster to list the true possibilities (see the right-hand side of Figure 1). This interaction was reliable: the difference in response times between the first and second blocks was significantly larger for the False-true group than for the True-false group (Mann-Whitney,  $z = 3.0$ ,  $p < .01$ ). The model theory predicts the interaction:



**Figure 1.** The mean response times (and standard errors) for the true and false cases listed by the True-false group and the False-true group in Experiment 1.

when the participants encounter the false possibilities in the first block, they have to think about the true possibilities in order to recover the false possibilities. This procedure has two consequences. First, the false possibilities take longer to evaluate in the first block. Second, there is a facilitatory effect on the true possibilities in the second block.

The participants made many varied errors in listing the false possibilities. In the case of conjunctions, the majority of errors failed to list all three ways in which the conjunction could be false: 25% of responses listed two correct false possibilities, and 9% of responses listed one correct false possibility. The remaining errors listed erroneous possibilities. In the case of conditionals, 49% of the false possibilities listed were correct for a conditional interpretation, 11% corresponded to a biconditional interpretation, and the remaining 40% of responses were miscellaneous errors. Across assertions, participants were the most accurate in generating the false possibilities for conditionals, and their accuracy scores for the false possibilities were close to those for the true possibilities, although still not as accurate (see Table 3 for a comparison). These results are not surprising, as both the true possibility and the false possibility of a conditional can be represented by a single model (see Appendix A, Table A2). However, response times showed that participants took the same amount of time to generate false possibilities for conditionals as they did the other assertions. Hence, although participants were generating only one possibility for both the true and false cases of conditionals, they still had to work harder to generate the false possibility.

In the case of the two sorts of disjunction, the participants typically listed the possibilities corresponding to mental models for the true possibilities (see above). The most frequent pattern of error in listing the false possibilities was what we refer to as the *mirror-image* error (42% for exclusive disjunctions, and 52% for inclusive disjunctions), that is, the participants constructed the false possibilities merely by negating some or all of the true possibilities. Hence, for the exclusive disjunction: A or 1, but not both, they listed:

False possibilities:  $\neg A$                      $\neg 1$

At first sight, this listing seems very strange. It lists possibilities that correspond to *true* possibilities (see Table 2). The correct and fully explicit response is:

False possibilities:     A            1  
                                  $\neg A$          $\neg 1$

And the first of these possibilities is incompatible with either of the possibilities in the mirror-image error.

Mirror-image errors occurred significantly more often than one would expect by chance. In listing a possibility for an assertion containing two clauses (A, 1) there are nine possible responses: A, not-A, 1, not-1, A 1, A not-1, not-A 1, not-A not-1, and nothing. Hence, the chance probability of the mirror-image error for an exclusive disjunction is  $1/9 \times 1/8 = .01389$ . The probability of the mirror-image error for an inclusive disjunction is  $p = .00002$ . We assume conservatively that the chance probability of a single mirror-image error occurring at least once in 16 trials (for inclusive and exclusive disjunctions) is  $p = .3$ . In fact, 21 out of the 24 participants committed the mirror-image error at least once in the experiment, and the mean number was on eight trials (Binomial test,  $p \ll 1$  in a million). The mirror-image error is, of course, highly diagnostic of a procedure in which individuals imagine the true possibilities as the first step in constructing the false possibilities. In sum, the distribution of errors, the pattern of latencies (see Figure 1), and the occurrence of the mirror-image errors all corroborated the hypothesis that individuals have to infer false possibilities from true possibilities.

## EXPERIMENT 2

When reasoners have to think aloud as they tackle problems, their resulting protocols are revealing about the strategies that they use in reasoning (see, e.g., Bucciarelli & Johnson-Laird, 1999; Schaeken, De Vooght, Vandierendonck, & d'Ydewalle, 2000; Van der Henst, Yang, & Johnson-Laird, 2002). In Experiment 2, we therefore asked the participants to think aloud as they listed true and false possibilities for assertions. They were also given visual cues to remind them to construct fully explicit possibilities.

### Method

*Design and materials.* The design and materials were similar to those of Experiment 1. Each participant listed separately the true and false possibilities for two instances of the four sorts of assertion. The participants were randomly allocated either to a True–false group or to a False–true group. They made their responses by placing cut-out letters and numbers on a board. Within each group, there was a further subdivision. Half of the participants, selected at random, used two separate boards, one board for the letters and numbers that were in present in a possibility and one board for the letters and numbers that were *not* present in a possibility. The other half of the participants used a single board, and they were instructed to represent that a letter, such as “A”, was not present, by using a different letter, such as “B”. Each board was divided into separate lines so that the participants could clearly separate the possibilities.

*Procedure.* The participants were tested individually. They were given the same general instructions as in Experiment 1 with appropriate modifications because the present experiment was not carried out on a computer. They were

also told that they had to “think aloud” at all times. The participants were given the assertions one at a time on separate sheets of paper. They were told to arrange the letters and numbers for each possibility on a separate line on the board, and they were shown how to construct fully explicit cases. The participants in the two-board groups were told that one board was for listing the letter and numbers that were present in a possibility and the other board was for listing those that were absent in that possibility. The participants in the one-board groups were told to represent the absence of a letter or number by using some other number or letter in place of it.

During the experiment, when the participants said that they could think of no more possibilities, the experimenter removed the letters and numbers from the boards and gave them the next assertion. The participants’ responses and thoughts were video-recorded using a camera, which was above them and focused on the boards. They rapidly adapted to the conditions of the experiment. They could take as much time as they needed to for each trial. If they fell silent for more than 3 seconds, the experimenter reminded them to think aloud.

*Participants.* A total of 24 Princeton undergraduates, who had no training in formal logic, carried out the experiment and received course credit for their participation.

## Results and discussion

The patterns of response to the assertions, including the mirror-image error, were very similar to those in Experiment 1. Appendix B contains four tables that summarise the different listings that the participants generated for each type of assertion. As in the previous experiment, we classified the true and false possibilities as correct if they corresponded to the mental models or the fully explicit models of the assertions. Overall accuracy with one and with two boards did not differ significantly, and so we have combined the results from the two groups. As before, the participants were more accurate in listing the true possibilities (95% correct) than in listing the false possibilities (35% correct; Wilcoxon test,  $z = 4.0$ ,  $p < .001$ ). Overall accuracy did not differ between the one- and two-board groups. But, as Appendix B shows, the group with two boards did list more fully explicit possibilities than the group with one board (37% vs 18%; Mann-Whitney,  $z = 2.0$ ,  $p < .05$ ). This effect occurred most strongly with disjunctions, but the participants tended to list more fully explicit possibilities with the other sorts of assertion.

What the participants said as they thought aloud was revealing. The protocols for the true possibilities were brief, because these possibilities required little mental work. But the false cases seemed, in the words of a participant, “Counterintuitive—you need to start out with true and think in the reverse.” Indeed, the protocols for the false possibilities provided overwhelming evidence

that the participants inferred them from the true possibilities. Every single participant produced at least one protocol showing that he or she had considered the true possibilities in order to derive the false possibilities. Here are eight typical examples, each from a different participant, illustrating what they had to say about the construction of false possibilities:

- (1) For a conjunction, *A and I*, a participant constructed a single false possibility:  $\neg A$  and  $\neg I$ , and said: “A and I would both have to be there [for it to be true], so if not A and not I it would be false. I go with the true and then find the false.”

- (2) For the same conjunction, a participant constructed two false possibilities:

A  
I

and said: “This [A and I] means both have to be present, so if A is by itself it won’t be true. Just I by itself won’t be true.”

- (3) For the conditional: *If D then 4*, a participant constructed two false possibilities:

D  
4

and said: “If D then 4, so if D with no 4. If 4 then D, so if 4 with no D then it is false.” This participant clearly made a biconditional interpretation of the assertion.

- (4) For the same conditional, a participant constructed the one correct false possibility:

D       $\neg 4$

and said: “If D then 4 ... If D is there, there has to be a 4, so if anything else is there it is false.” This participant clearly treated the subordinate clause as given, and merely falsified the main clause.

- (5) For an exclusive disjunction: *B or 2, but not both*, a participant constructed two false possibilities:

$\neg B$   
 $\neg 2$

and said: “You could have B and 2 by themselves, so I’ll do the opposite.”

- (6) For the same exclusive disjunction, a participant constructed one false possibility:

B            2

and said: “It says not both ... for that to be true they would not be present together, so I’ll reverse that.”

- (7) For the inclusive disjunction, *C or 3, or both*, a participant constructed three false possibilities:

$\neg C$   
 $\neg C$              $\neg 3$   
 $\neg C$              $\neg 3$

and said: “C or 3 or both—you can have C by itself, so no C, 3 by itself so no 3, and C and 3 together so no C and no 3.” This protocol is a perfect illustration of the mirror-image error.

- (8) For the same inclusive disjunction, a participant constructed the single correct false possibility:

$\neg C$              $\neg 3$

and said: “Both can be present [alone] and then present together, so the only thing that can make it false is when they are both not there.”

As some of the examples show, the participants often represented the absence of a number or letter merely by omitting it. Thus, a participant lists D by itself to indicate the possibility in which there is D but not 4. The same phenomenon occurred in Experiment 1. In other words, the absence of a letter is taken to mean that it is not present. This principle is known as the *closed world assumption* in setting up computer data bases (see Brewka, Dix, & Konolige, 1997). It is clear, however, that individuals did not abide by the assumption consistently. For example, individuals who constructed the following false possibilities for the disjunction, *A or else 1*:

$\neg A$   
 $\neg 1$

obviously did not mean that 1 was absent from the first possibility, and that A was absent from the second possibility. In that case, the two possibilities would have been identical. Perhaps, they intended the absence of a symbol *including its negation* to indicate that the symbol and its negation were not present. Hence, the two false possibilities were:

$$\begin{array}{cc} \neg A & 1 \\ A & \neg 1 \end{array}$$

But this interpretation—though it may have been what was intended—is still odder, because the two possibilities are exactly those required for the truth of the disjunction. Hence, the most charitable interpretation is that the absence of a symbol is ambiguous. Sometimes, it means that the symbol is not present; and, sometimes, it means merely that the participants were uncertain about its status. They may have been uncertain because they did not consider the status of both clauses simultaneously; they may not have even noticed that they had neglected one of them. The “think aloud” protocols did not help us to interpret the absence of symbols. The participants did not say anything to elucidate the interpretation of the missing letters or numbers. They did not mention them at all. Instead, their protocols suggested that they were focusing on the most essential information—that is, what was present in a possibility—and were less aware of what was absent. They think about what is true in a possibility, not about what is false. But, as the protocols and the errors showed, there is no direct route to false possibilities. They have to be inferred from the true possibilities.

### THE RECURSIVE CONSTRUCTION OF TRUE POSSIBILITIES

For a simple compound assertion containing two clauses, such as the assertions in Experiments 1 and 2, individuals need to imagine the situation described by each clause and to combine these representations to construct the true possibilities. These combinations depend in turn on the meaning of the sentential connective, and the model theory postulates that this meaning is stored in memory as part of the lexical representation of the sentential connective. Hence, individuals know that an exclusive disjunction of the form: *A or else B*, calls for a set of models of the following sort:

$$\begin{array}{c} A \\ B \end{array}$$

where A and B represent the possibilities described by the two clauses. The lexicon contains such instructions for constructing the mental models for all the connectives in Table 2.



What happens in the case of more complex compound assertions? For example, consider an assertion of the form:

If A or else B but not both, then C.

Although individuals know the meanings of conditionals and the meaning of disjunctions, they do not have the meanings of their combinations stored in memory. The theory postulates that they need to make a recursive construction of each such combination in order to envisage even the *true* possibilities, let alone the false ones. We have implemented the model theory in a computer program to show how, in principle, the process works. The program operates at various levels of expertise, and we will illustrate how it constructs mental models. The parser constructs a representation of the meaning of the first major constituent of the assertion: *A or B but not both*, and this meaning is used to construct mental models of the true possibilities:

A  
     B

The parser then constructs a representation of the meaning of the conditional, *If x then C* (where *x* designates the preceding exclusive disjunction):

x     C  
   ...

and there is a mental footnote on the implicit model indicating that it represents the possibilities in which *x* is false. The previous models of the disjunction are then substituted for *x* to construct the mental models of the assertion as a whole:

A           C  
     B     C  
   ...

and the footnote on the implicit model indicated that it represents possibilities in which *A or else B* is false. At its most advanced level of performance, the program constructs fully explicit models of the assertion:

A   ¬ B   C  
 ¬ A   B   C  
 A    B   C  
 A    B  ¬ C  
 ¬ A  ¬ B   C  
 ¬ A  ¬ B  ¬ C

The assertion is false when the antecedent of the conditional is true, i.e., the disjunction is true, but the consequent of the conditional is false:

$$\begin{array}{r} A \quad \neg B \quad \neg C \\ \neg A \quad B \quad \neg C \end{array}$$

The need to construct true possibilities recursively should increase the difficulty of imagining true possibilities. In turn, the task of imagining false possibilities should be still harder, granted that they are inferred from the true possibilities. Hence, the theory predicts that false possibilities should be harder to construct than true possibilities even for complex compound assertions. Our next experiment was designed to test this prediction.

### EXPERIMENT 3

#### Method

*Design, materials, and procedure.* The participants were randomly allocated either to a True–false group or to a False–true group. They listed separately the true and false possibilities for two instances of eight sorts of assertion. In each assertion, there was a central connective, which was one of the four connectives used in the previous experiments, and it interconnected a compound clause, based on an exclusive disjunction, and an atomic proposition. The compound occurred either first in the sentence or second. Hence, there were eight sorts of assertion of the following forms:

(A or else B) and C.	A and (B or else C).
If A or else B then C.	If A then (B or else C).
(A or else B) or else C.	A or else (B or else C).
(A or else B) or C.	A or (B or else C).

The contents of the assertions concerned letters, number, and shapes, e.g., “If there is a square or an A, but not both, then there is a 2”. The eight different assertions were presented twice with different contents. The computer-controlled procedure and instructions were the same as in Experiment 1.

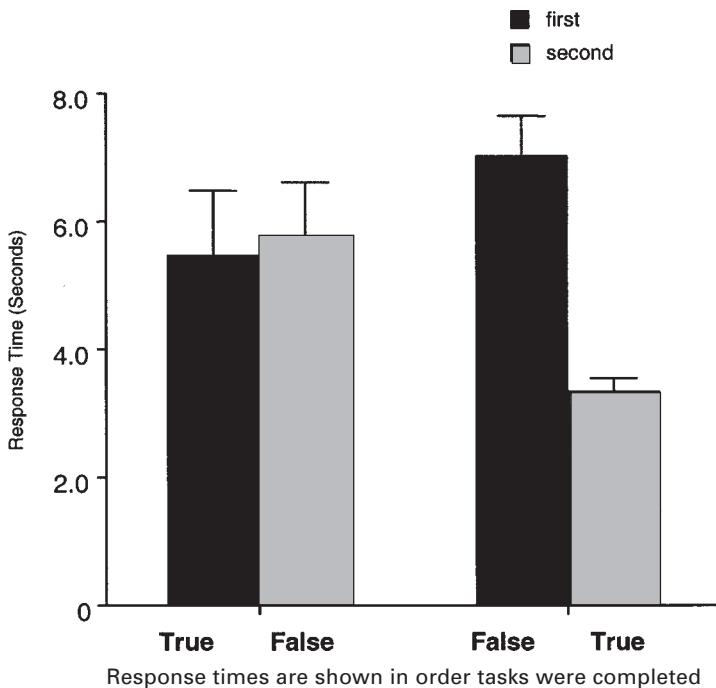
*Participants.* A total of 24 Princeton undergraduates, who had no training in formal logic, carried out the experiment and received course credit.

#### Results and discussion

As in the previous experiments, we classified the true and false possibilities as correct if they corresponded to either the mental models or the fully explicit models of the assertions. Overall accuracy was not affected by whether the single atomic proposition came first or last in the assertions, and so we combined the

data from the two sets of assertions. Appendix C presents these combined data in four tables that summarise the different sorts of responses generated by participants for each type of assertion. Once again, the participants were more accurate in listing the true possibilities (42% correct) than in listing the false possibilities (8% correct; Wilcoxon test,  $z = 4.0$ ,  $p < .001$ ). As we expected, however, performance was much worse overall than in Experiment 1.

Figure 2 presents the latencies of responses for the two groups and the two blocks. The pattern of results replicated those of Experiment 1. As the theory predicts, the participants were faster to list the true possibilities (mean 4.4 s) than to list the false possibilities (mean 6.2 s; Wilcoxon test,  $z = 3.0$ ,  $p < .01$ ). As before, there was a significant advantage in listing the false cases for the group who had already listed the true cases as opposed to the group who had not (see Figure 2). This interaction was reliable: the difference in response times between the first block of trials and the second block was significantly larger for the False–true group than for the True–false group (Mann-Whitney,  $z = 2.0$ ,  $p < .05$ ). Hence, the latencies suggest that the participants were constructing the false possibilities from the true possibilities. It takes them longer to deal with the false possibilities in the first block, but their consideration of the true possibilities in



**Figure 2.** The mean response times (and standard errors) for the true and false cases listed by the True–false group and the False–true group in Experiment 3.



True possibility: A

they constructed a false instance by making the second constituent true too:

False possibility: A C

The same principle transformed:

True possibility: B

into:

False possibility: B C

This process of *recombination* led the participants to construct the false possibilities by forming all possible pairs of separate true possibilities. It transforms the typical:

True possibilities: A  
B  
C

into the following typical listing:

False possibilities: A B  
A C  
A B C

The recombination procedure explains the presence of the error, A B C, and the absence of the genuine false possibility:  $\neg A \neg B \neg C$ . The chance probability of the four responses above is  $p = 1/8 \times 1/7 \times 1/6 \times 1/5$ , and the chance probability of at least one such response in four trials is  $p < .0025$ . Of the 24 participants, 16 made this response at least once in four trials (Binomial test,  $p \ll 1$  in a billion). Likewise, 14 of the 24 participants made the recombination pattern of responses for the false possibilities of the inclusive disjunction (Binomial test,  $p \ll 1$  in a billion).

In summary, Experiment 3 corroborated the predictions of the model theory. It was harder for the participants to imagine the true possibilities, which now depended on a process of recursive construction. The task of listing the false possibilities became still harder. The pattern of latencies and the nature of the errors suggested that they were inferred from the true possibilities.

## EXPERIMENT 4: ILLUSIONS OF TRUTH

Consider the following two assertions, one a conjunction of two conditionals, and the other their inclusive disjunction:

Conjunction:      If A then B, and if C then D.  
 Disjunction:      If A then B, or if C then D, or both.

They differ in meaning, presumably, but how? According to the model theory and its computer implementation, which we described earlier, the two conditionals have the following mental models:

A B  
 ...

and:

C D  
 ...

The conjunction of the two sets of mental models yields:

A	B			[combination of the first explicit model, A B, with the second implicit model]
		C	D	[combination of the second explicit model, C D, with the first implicit model]
A	B	C	D	[combination of the two explicit models]
	...			[combination of the two implicit models]

An inclusive disjunction of two propositions, *X or Y or both*, calls for models of the form:

X  
   Y  
 X Y

Given that *X* is the conditional, *If A then B*, its models are:

A B  
 ...

And given that *Y* is the conditional, *If C then D*, then its models are:

C D  
 ...

and the conjunction of  $X$  and  $Y$ , yields the models shown above for the conjunction of the two conditionals. But these models of the conjunction already include the models for  $X$ , and they already include the models for  $Y$ . Hence, the mental models for the disjunction of the two conditionals are identical to those for the conjunction of the two conditionals. This equivalence is an unforeseen consequence of the theory: the mental models of a conjunction of conditionals are the same as those for a disjunction of conditionals. Moreover, the mental models of an inclusive disjunction of conjunctions:

A and B, or C and D

are as follows:

A	B		
		C	D
A	B	C	D

These mental models differ from the previous ones only in the absence of an implicit model.

Is it really the case that the three sorts of assertion are so similar in meaning, and that there is no difference between a conjunction and a disjunction of conditionals? The answer is shown in Table 4. The three assertions do differ in meaning: they refer to different sets of possibilities. The disjunction of conjunctions is quite distinct from the other two compounds. Likewise, the conjunction of conditionals is compatible with fewer possibilities than the inclusive disjunction of conditionals. In fact, the conjunction validly implies the disjunction, but the converse implication is not valid, because the disjunction can be true in possibilities in which the conjunction would be false. According to the model theory, however, naïve individuals should construct the same true possibilities for the three sorts of assertion. The assertions should give rise to illusions of truth in which individuals radically underestimate the number of true possibilities and list the same sets of true possibilities for the three sorts of assertion. The illusions arise from the principle of truth, which entails that it is impossible to construct the full recursion of true possibilities for complex compound assertions. Experiment 4 was designed to test whether individuals succumbed to the illusions of truth.

## Method

*Design, materials, and procedure.* The participants acted as their own controls and listed the true possibilities and the false possibilities for a series of four assertions. There were three experimental assertions, and one control assertion—a conjunction of conjunctions—for which the model theory predicted

TABLE 4  
The fully explicit models of the three sorts of assertion in Experiment 4

A and B, or C and D				If A then B, and if C then D				If A then B, or if C then D			
A	B	$\neg$ C	$\neg$ D	A	B	$\neg$ C	$\neg$ D	A	B	$\neg$ C	$\neg$ D
$\neg$ A	$\neg$ B	C	D	$\neg$ A	$\neg$ B	C	D	$\neg$ A	$\neg$ B	C	D
A	B	C	D	A	B	C	D	A	B	C	D
A	B	$\neg$ C	D	A	B	$\neg$ C	D	A	B	$\neg$ C	D
$\neg$ A	B	C	D	$\neg$ A	B	C	D	$\neg$ A	B	C	D
A	B	C	$\neg$ D								
A	$\neg$ B	C	D	$\neg$ A	B	$\neg$ C	D	$\neg$ A	B	$\neg$ C	D
				$\neg$ A	B	$\neg$ C	$\neg$ D	$\neg$ A	B	$\neg$ C	$\neg$ D
				$\neg$ A	$\neg$ B	$\neg$ C	D	$\neg$ A	$\neg$ B	$\neg$ C	D
				$\neg$ A	$\neg$ B	$\neg$ C	$\neg$ D	$\neg$ A	$\neg$ B	$\neg$ C	$\neg$ D
								A	B	C	$\neg$ D
								A	$\neg$ B	C	D
								A	$\neg$ B	$\neg$ C	D
								A	$\neg$ B	$\neg$ C	$\neg$ D
								$\neg$ A	B	C	$\neg$ D
								$\neg$ A	$\neg$ B	C	$\neg$ D

that the participants should construct the correct true possibility. The study was carried out in a class in an undergraduate course in which each experimenter was a student and tested separate logically untrained individuals. None of the experimenters knew the predictions of the model theory. The contents of the four assertions were as follows:

- (1) There is a “J” on the board and there is a “9”, or there is a “L” on the board and there is a “7”, and both conjunctions may be true. (The disjunction of conjunctions)
- (2) If there is a “A” on the board then there is a “2”, and if there is a “C” on the board then there is a “3”. (The conjunction of conditionals)
- (3) If there is a “D” on the board then there is a “5”, or if there is a “E” then there is a “6”, and both conditionals may be true. (The inclusive disjunction of conditionals)
- (4) There is a “G” on the board and there is a “8”, and there is an “H” on the board and there is a “4”. (The control conjunction of conjunctions)

For each assertion, the subjects wrote down a list of the possible circumstances in which the assertion would be true, i.e., letters and numbers, such as “A 2”. When they had completed this task, they went through each assertion again, and wrote down a list of the possible circumstances in which the assertion would be false.



*Participants.* A total of 25 Princeton undergraduates, who had no training in formal logic, participated in the experiment.

### Results and discussion

The results corroborated the illusions of truth. Table 5 summarises the most frequent selections for the three experimental assertions and the control assertion. As the Table shows, the participants tended to list the same true possibilities for the three experimental assertions. The numbers of subjects (out of 25) who listed the responses in the order predicted by the computer program were as follows: 16 for the disjunction of conjunctions, 9 for the conjunction of conditionals, and 17 for the disjunction of conditionals. On the assumption that there are 16 possible selections, the likelihood of making by chance these three selections in their given order is  $1/16 \times 1/15 \times 1/14 < 0.0003$ , and so the binomial probability for, say, 9 out of 25 subjects making this selection is at the nano-probability level. All but one of the participants listed the correct true possibility for the control assertion. The participants' performance in listing the false possibilities could hardly have been worse. No-one listed the correct false possibilities for any of the assertions. And no-one accounted for all 16 contingencies in their combined lists of true and false possibilities. We carried out a second experiment with a single experimenter, which corroborated the occurrence of the illusions for the conjunction and the disjunction of conditionals.

## GENERAL DISCUSSION

The ability to imagine the possibilities in which an assertion is true and the possibilities in which it is false is a fundamental semantic ability. Without this ability, it is doubtful whether anyone could be said to understand assertions. Our aim has been to elucidate how people imagine such possibilities for compound

TABLE 5  
The percentages of subjects who listed each of the three most frequent possibilities in Experiment 4 ( $n = 25$ )

<i>The possibilities</i>	<i>The four sorts of assertion</i>			
	<i>A and B, or C and D</i>	<i>If A then B, and if C then D</i>	<i>If A then B, or if C then D</i>	<i>A and B, and C and D</i>
A B	92	84	92	0
C D	88	80	96	0
A B C D	76	80	96	96

assertions based on the connectives “if”, “or”, and “and”. Because the processing capacity of working memory is limited, the model theory postulates a principle of truth: individuals normally represent only true possibilities, and within them they represent the clauses within assertions only when they are true within a possibility. It follows that when people understand simple compound assertions, which contain only a single sentential connective, they have access only to true possibilities, and not even to fully explicit models of them. They have no direct access to the false possibilities, but must infer them from their mental models of the true possibilities. In contrast, if individuals had a mental embodiment of truth tables, then they would have direct access both to true possibilities and to false possibilities.

Experiment 1 corroborated three predictions of the model theory. Individuals are more accurate in imagining true possibilities than in imagining false possibilities. They list true possibilities faster than false possibilities. And when they list the false possibilities in the first block of trials, they have to think about the true possibilities in order to infer the false possibilities. The task accordingly takes longer. But, in the second block of trials, the true possibilities are listed faster. In contrast, when the true possibilities are listed in the first block, their consideration eliminates the additional time needed to cope with the false possibilities in the second block. In other words, the participants in the False–true group have to construct the true possibilities before they can infer the false possibilities in the first block, whereas the participants in the True–false group have the true possibilities in mind to speed up the process of constructing the false possibilities in the second block.

Experiment 2 confirmed that individuals are more accurate in imagining true possibilities than false possibilities. And their reports of what they were thinking showed that they did indeed tend to construct the true possibilities en route to constructing the false possibilities. Both experiments revealed a “mirror-image” error when the participants imagined false possibilities: they often tended to negate separately each of the true possibilities. For instance, a participant in Experiment 2 constructed the following false possibilities for the exclusive disjunction: *B or 2, but not both*:

¬B  
           ¬ 2

and said: “You could have B and 2 by themselves [for the true possibilities], so I’ll do the opposite.”

For compound assertions containing two clauses, such as the preceding example, individuals combine representations of each clause according to the meaning of the sentential connective. The model theory postulates that such meanings are stored in memory as part of the lexical representations of each of

the sentential connectives in Table 2. But, for more complex compound assertions containing several connectives, there is no pre-stored recipe for constructing true possibilities. The theory postulates instead that individuals must make a recursive construction even to imagine the *true* possibilities for an assertion. This process is also modelled in the computer program implementing the theory. It should be more difficult to imagine true possibilities, and, in turn, the task of imagining false possibilities should be still harder. Experiment 3 bore out these predictions with assertions containing two sentential connectives, such as:

If A or else B then C.

The participants listed the true possibilities correctly on 42% of trials, and performance in listing the false possibilities fell to 8% correct. As in Experiment 1, the participants were faster to list the false possibilities when they had already listed the true possibilities. The pattern of errors and the “think-aloud” protocols also corroborated the use of true possibilities as the basis for inferring the false possibilities.

As the recursions needed to imagine true possibilities increase, so the principle of truth begins to affect performance. In particular, the mental models of true possibilities become an increasingly smaller proportion of the fully explicit possibilities. There is also an unexpected consequence, which we discovered from the computer implementation of the theory: assertions that seem superficially different are likely to be interpreted in the same way. Thus, the mental models of a conjunction of conditionals are identical to those of a disjunction of conditionals, and nearly identical to those of a conjunction of disjunctions. Experiment 4 confirmed that naïve individuals tend to list the same true possibilities for each of these assertions. The results suggest that there is a resemblance between conjunctions and conditionals. This resemblance is borne out by the possibilities that children list for simple conditionals of the form, *If A then B*: young children tend to list only the possibility: *A B*, which is the one possibility compatible with a conjunction (Barrouillet & Lecas, 1998). When the recursive construction of true possibilities becomes difficult, not surprisingly it becomes almost impossible to list the correct false possibilities. The illusions in turn can give rise to a variety of illusory inferences. For example, consider the following problem (from Goldvarg & Johnson-Laird, 2000):

Only one of the following assertions is true about a hand of cards:

There is a king or an ace in the hand, or both.

There is a queen or an ace in the hand, or both.

There is a jack or a ten in the hand, or both.

Is it possible that there is an ace in the hand?

The mental models of the premises yield the response, “Yes”, and nearly every participant in the experiment made this response. But it is an illusion. If there were an ace in the hand, then two of the assertions would be true, contrary to the rubric that only one of them is true. Analogous illusions arise in reasoning about necessary conclusions (Johnson-Laird & Savary, 1999), probable conclusions (Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni, 1999), and quantified conclusions (Yang & Johnson-Laird, 2000). They also occur in reasoning about whether sets of assertions are consistent (Johnson-Laird, Legrenzi, Girotto, & Legrenzi, 2000). Underlying all these predicted illusions is the principle of truth. Naïve individuals normally have access only to true possibilities, and they must infer false possibilities from them.

Our experiments used abstract content, and so readers may wonder what effects, if any, the use of real-world content might have had. A study of the effects of such content on the interpretation of conditionals has shown that it can modulate the basic meanings elicited by abstract content (Johnson-Laird & Byrne, 2002). In particular, it can eliminate possibilities that are otherwise feasible. For example, a conditional such as:

If she is in Brazil then she is in Rio

is compatible with just two possibilities (out of the three that are otherwise feasible):

In Brazil	In Rio
¬ In Brazil	¬ In Rio

General knowledge rules out the possibility in which she is not in Brazil but is in Rio. Likewise, participants perform better on the selection task with certain realistic scenarios (see, e.g., Cheng & Holyoak, 1985; Cosmides, 1989). It is conceivable that real-world knowledge can make false possibilities directly available without the need to imagine true possibilities. An analogous case certainly occurs in statements of deontic prohibition. Individuals have rapid access to the case that is *impermissible* in such assertions as, “Being under age prohibits you from drinking alcohol”, i.e., being under age and drinking alcohol. Hence, there may be cases where individuals have rapid access to the case that is false.

We draw two main morals from our investigation. First, it bears out the heuristic value of computer implementations of theories. The program implementing the model theory led us to the illusions of truth. Second, the experiments corroborate the model theory’s account of sentential connectives. It is more difficult to imagine what is false than what is true, because there is no direct access to false possibilities. They have to be inferred from true possibilities. The same prediction applies to other procedures, such as one in

which participants evaluate whether given possibilities are true or false. Mental models are incomplete representations even of true possibilities. The consequences are that in complex recursions even the enumeration of true possibilities begins to break down. There are failures in combining the true cases of the constituent propositions in compound assertions. And the recursion yields illusions about true possibilities, and apparent equivalences between distinct assertions. Alternative theories of reasoning, including those based on formal rules of inference (e.g., Braine & O'Brien, 1998; Rips, 1994), either do not deal with the meaning of connectives or else treat their meanings as equivalent to the rules of inference that govern them. Because the rules for disjunction differ from those for conjunction, which in turn differ from those for conditionals, such theories do not explain why individuals sometimes treat "if" to mean "and", and "and" to mean "or". As Dewey (1910, pp. 5–6) remarked, it is a sign of unusual reflective ability to structure reality in terms of what is false as well as what is true.

Manuscript received 27 September 2001

Revised manuscript received 2 July 2002

## REFERENCES

- Barrouillet, P., & Lecas, J-F. (1998). How can mental models theory account for content effects in conditional reasoning? A developmental perspective. *Cognition*, *67*, 209–253.
- Braine, M.D.S., & O'Brien, D.P. (Eds.). (1998). *Mental logic*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Brewka, G., Dix, J., & Konolige, K. (1997). *Nonmonotonic reasoning: An overview*. Stanford, CA: CSLI Publications.
- Bucciarelli, M., & Johnson-Laird, P.N. (1999). Strategies in syllogistic reasoning. *Cognitive Science*, *23*, 247–303.
- Cheng, P.W., & Holyoak, K.J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*, 391–416.
- Clark, H.H., & Chase, W.G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, *3*, 472–517.
- 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.
- Dewey, J. (1910). *How we think*. Boston, MA: Heath.
- Evans, J.B.T., Newstead, S.E., & Byrne, R.M.J. (1993). *Human reasoning: The psychology of Deduction*. Hove, UK: Lawrence Erlbaum Associates Ltd.
- Goldvarg, Y., & Johnson-Laird, P.N. (2000). Illusions in modal reasoning. *Memory & Cognition*, *28*, 282–294.
- Grice, H.P. (1975). Logic and conversation. In P. Cole & J.L. Morgan (Eds.), *Syntax and semantics. Vol. 3: Speech acts*. New York: Seminar Press.
- Johnson-Laird, P.N. (2001). Mental models and deduction. *Trends in Cognitive Science*, *5*, 434–442.
- Johnson-Laird, P.N., & Byrne, R.M.J. (1991). *Deduction*. Hove, UK: Lawrence Erlbaum Associates Ltd.
- Johnson-Laird, P.N., & Byrne, R.M.J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*, 646–678.

- Johnson-Laird, P.N., Legrenzi, P., Girotto, P., & Legrenzi, M.S. (2000). Illusions in reasoning about consistency. *Science*, 288, 531–532.
- Johnson-Laird, P.N., Legrenzi, P., Girotto, P., Legrenzi, M.S., & Caverni, J-P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review*, 106, 62–88.
- Johnson-Laird, P.N., & Savary, F. (1999). Illusory inferences: A novel class of erroneous deductions. *Cognition*, 71, 191–229.
- Rips, L.J. (1994). *The psychology of proof*. Cambridge, MA: MIT Press.
- Schaeken, W., De Vooght, G., Vandierendonck, A., & d'Ydewalle, G. (Eds.). (2000). *Deductive reasoning and strategies*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Van der Henst, J-B., Yang, Y., & Johnson-Laird, P.N. (2002). Strategies in sentential reasoning. *Cognitive Science*, 26, 425–468.
- Wason, P.C. (1959). The processing of positive and negative information. *Quarterly Journal of Experimental Psychology*, 21, 92–107.
- Wason, P.C. (1961). Response to affirmative and negative binary statements. *British Journal of Psychology*, 52, 133–142.
- Wittgenstein, L. (1922). *Tractatus logico-philosophicus*. London: Routledge & Kegan Paul.
- Yang, Y., & Johnson-Laird, P.N. (2000). Illusions in quantified reasoning: How to make the impossible seem possible, and vice versa. *Memory & Cognition*, 28, 452–465.

## APPENDIX A: EXPERIMENT 1

Each table shows the percentages of the different combinations of true and false possibilities that the participants listed for an assertion of one of the four sorts. The 20 participants carried out two blocks of 4 trials with each sort of assertion, but we have collapsed the data from trials and blocks here. For convenience, we have assumed that each assertion concerned: A and I. The four sorts of assertions were conjunctions (A and I), conditionals (If A then I), exclusive disjunctions (Either A or else I, but not both), and inclusive disjunctions (A or I, or both). The thicker lines enclose the correct responses corresponding to either fully explicit models or mental models.

TABLE A1  
 The percentages of different combinations of true and false cases listed for conjunctions: A and 1

A and 1

		True Cases		
		A 1	misc. errors	totals
False Cases	A ¬ 1 ¬ A 1 ¬ A ¬ 1	15	2	17
	A 1 ¬ A ¬ 1	3	0	3
	¬ A ¬ 1 ¬ A 1	5	2	7
	¬ A ¬ 1	7	2	9
	A ¬ 1 ¬ A 1	4	0	4
	A 1	5	1	6
	¬ A ¬ 1	10	5	15
	misc. errors	32	7	39
	totals	81	19	100

“¬” signifies negation.

TABLE A2  
 The percentages of different combinations of true and false cases  
 listed for conditionals: If A then 1

If A then 1

		True Cases			* False Cases		
		A 1 ¬ A 1 ¬ A ¬ 1	A 1	A 1 ¬ A ¬ 1	A 1 ¬ A 1	misc. errors	totals
False Cases	A ¬ 1	8	29	5	2	5	49
	* A ¬ 1 ¬ A 1	0	7	3	1	0	11
	misc. errors	0	14	3	0	23	40
	totals	8	50	11	3	28	100

¬ signifies negation,

\* signifies sets of cases corresponding to a biconditional interpretation.



TABLE A3  
 The percentages of different combinations of true and false cases listed for exclusive disjunctions: A or 1, but not both

A or 1, but not both

		True Cases				
		A ¬ A	¬ 1 1	A 1	misc. errors	totals
False Cases	A ¬ A	1 ¬ 1	8	9	6	23
	A ¬ A	1 ¬ 1	4	8	3	15
	¬ A ¬ A	¬ 1 ¬ 1	0	3	1	4
	*	¬ A ¬ 1	5	22	6	33
	**	¬ A or ¬ 1	0	3	6	9
		misc. errors	5	5	6	16
		totals	22	50	28	100

¬ signifies negation,

\* signifies mirror-image errors (see text),

\*\* signifies sets that are partial mirror-image errors.

TABLE A4  
 The percentages of different combinations of true and false cases listed for inclusive disjunctions: A or 1, or both

A or 1, but not both

		True Cases				misc. errors	totals
		A	$\neg$ 1	A	1		
False Cases	$\neg$ A	$\neg$ 1	10	17	3	30	
	*	$\neg$ A	$\neg$ 1	0	16	4	20
	**	$\neg$ A	$\neg$ 1	3	18	11	32
	misc. errors		3	9	6	18	
	totals		16	60	24	100	

$\neg$  signifies negation,  
 \* signifies mirror-image errors,  
 \*\* signifies sets that are partial mirror-image errors.

## APPENDIX B: EXPERIMENT 2

Each table shows the percentages of the different combinations of true and false possibilities that the participants listed for an assertion of one of the four sorts. The 24 participants carried out two blocks of two trials with each sort of assertion, but we have collapsed the data from trials, blocks, and the one- and two-board groups (except for the results with disjunctions). For convenience, we have assumed that each assertion concerned A and 1. The thicker lines enclose the correct responses corresponding to either fully explicit models or mental models.

TABLE B1  
 The percentages of different combinations of true and false cases listed for conjunctions: A and 1

A and 1

		True Cases		
		A 1	misc. errors	totals
False Cases	A ¬ 1 ¬ A 1 ¬ A ¬ 1	2	4	6
	A 1 ¬ A ¬ 1	29	2	31
	¬ A ¬ 1 ¬ A 1	15	0	15
	¬ A ¬ 1	15	2	17
	¬ A ¬ 1 ¬ A 1	15	2	17
	A 1	14	0	14
	totals	90	10	100

“¬” signifies negation.

TABLE B2  
 The percentages of different combinations of true and false cases  
 listed for conditionals: If A then 1

If A then 1

		True Cases			* False Cases		
		A 1 ¬ A 1 ¬ A ¬ 1	A 1	A 1 ¬ A ¬ 1	A 1 ¬ A 1	misc. errors	totals
False Cases	A ¬ 1	19	17	10	0	0	46
	* A ¬ 1 ¬ A 1	0	8	8	0	0	16
	misc. errors	0	22	8	4	2	38
	totals	19	47	28	4	2	100

¬ signifies negation,

\* signifies sets of cases corresponding to a biconditional interpretation.

TABLE B3  
 The percentages of different combinations of true and false cases listed for exclusive disjunctions: A or 1, but not both

A or 1, but not both

		True Cases				
False Cases		A	$\neg 1$	A	misc. errors	totals
		$\neg A$	1	1		
	A		8	8	4	20
	$\neg A$		0	<i>21</i>	0	<i>21</i>
			0	21	4	25
	A		4	<i>12</i>	<i>16</i>	<i>32</i>
	$\neg A$		0	0	0	0
			8	5	0	13
*	$\neg A$		17	33	5	55
	$\neg 1$		22	<i>12</i>	0	<i>34</i>
	totals		25	62	13	100
			<i>34</i>	<i>50</i>	<i>16</i>	<i>100</i>

$\neg$  signifies negation.

\* signifies mirror-image errors.

normal numbers signify data from the one-board group.

italicised numbers signify data from the two-board group (see text).

TABLE B4  
 The percentages of different combinations of true and false cases listed for inclusive disjunctions: A or 1, or both

A or 1, but not both

		True Cases			totals
		A $\neg$ A A	$\neg$ 1 1 1	A 1 A	
False Cases			0	21	21
		$\neg$ A $\neg$ 1	25	29	54
	*	$\neg$ A	0	25	25
		$\neg$ A $\neg$ 1	<i>13</i>	29	42
	**	$\neg$ A	4	46	50
		$\neg$ A $\neg$ 1	<i>0</i>	4	4
		misc. errors	0	4	4
	misc. errors	<i>0</i>	<i>0</i>	<i>0</i>	
	totals	4	96	100	
		<i>38</i>	<i>62</i>	<i>100</i>	

$\neg$  signifies negation,  
 \* signifies mirror-image errors,  
 \*\* signifies sets that are partial mirror-image errors,  
 normal numbers signify data from one-board group,  
 italicised numbers signify data from the two-board group (see text).

APPENDIX C: EXPERIMENT 3

Each table shows the percentages of the different combinations of true and false possibilities that the participants listed for four matched pairs of assertions. We have collapsed the data from each sentence in the pair, which differed only in whether the single atomic proposition came first or last in the assertion. For convenience, we have assumed that each assertion concerned A, 1, and a square. Likewise, as shown, the tables are compressed by collapsing percentages of different subsets from the same listing.

TABLE C1  
The percentages of different combinations of true and false cases listed for conjunctions

(A or else 1) and square

		True Cases						
False Cases		A	$\neg$ 1	squ	A	squ	misc. errors	totals
		$\neg$ A	1	squ	1	squ		
**	A 1 squ							
	A 1 A 1 A 1 squ			1		10	1	12
**	A 1 squ							
	A 1 A 1 A 1 squ			2		18	2	22
**	$\neg$ A							
	$\neg$ 1							
	$\neg$ squ							
	$\neg$ A $\neg$ 1							
	$\neg$ A $\neg$ squ							
	$\neg$ 1 $\neg$ squ			0		20	1	21
	misc. errors			6		33	6	45
	totals			9		81	10	100

$\neg$  signifies negation,  
squ abbreviates square,  
\*\* signifies that the responses included only part of the listed set.

TABLE C2  
 The percentages of different combinations of true and false cases  
 listed for conditionals

If (A or else 1) then square

		True Cases						
False Cases		A	$\neg$ 1	squ	A	squ	misc. errors	totals
		$\neg$ A	1	squ	1	squ		
	A $\neg$ squ 1 $\neg$ squ			3		6	0	9
*	A $\neg$ squ 1 $\neg$ squ A 1 squ $\neg$ A $\neg$ 1 squ			0		7	0	7
*	A $\neg$ squ 1 $\neg$ squ A 1 squ			0		5	0	5
**	$\neg$ A $\neg$ 1 $\neg$ squ $\neg$ A $\neg$ 1 $\neg$ A $\neg$ squ $\neg$ 1 $\neg$ squ			0		17	0	17
	misc. errors			7		50	5	62
	totals			10		85	5	100

$\neg$  signifies negation,

squ abbreviates square,

\* signifies a biconditional interpretation,

\*\* signifies that the responses included only part of the listed set.



TABLE C3  
 The percentages of different combinations of true and false cases  
 listed for exclusive disjunctions

(A or else 1) or else square

		True Cases						
False Cases		A	$\neg$ 1	$\neg$ squ	A	1	misc. errors	totals
		$\neg$ A	1	$\neg$ squ	$\neg$ A	$\neg$ 1		
	A 1							
	A squ							
	! A 1				5	27	2	34
**	A 1							
	A squ							
	! A 1				3	14	0	17
	! A 1							
	$\neg$ A $\neg$ 1				1	9	1	11
**	$\neg$ A							
	$\neg$ 1							
	$\neg$ squ							
	$\neg$ A $\neg$ 1							
	$\neg$ A $\neg$ squ				0	11	2	13
	misc. errors				3	21	1	25
	totals				12	82	6	100

$\neg$  signifies negation,  
 squ abbreviates square,  
 \*\* signifies that the responses included only part of the listed set,  
 ! signifies that a true case was incorrectly listed as a false case.

TABLE C4  
 The percentages of different combinations of true and false cases  
 listed for inclusive disjunctions

(A or else 1) or square

True Cases						
False Cases		A $\neg$ 1 $\neg$ squ $\neg$ A 1 $\neg$ squ $\neg$ A $\neg$ 1 squ A $\neg$ 1 squ $\neg$ A 1 squ	A 1 squ A 1 squ			
				misc. errors	totals	
	A 1 $\neg$ A $\neg$ 1 $\neg$ squ		0	9	1	10
	A 1 ! A 1 squ		2	30	4	36
**	$\neg$ A $\neg$ 1 $\neg$ squ $\neg$ A $\neg$ 1 $\neg$ A $\neg$ 1 $\neg$ squ $\neg$ 1 $\neg$ squ		0	19	4	23
	misc. errors		8	14	9	31
	totals		10	72	18	100

$\neg$  signifies negation,

squ abbreviates square,

\*\* signifies that the responses included only part of the listed set,

! signifies that a true case was incorrectly listed as a false case.