Research Article

MODELS OF CONSISTENCY

Paolo Legrenzi,[1] Vittorio Girotto,[1,2] and P.N. Johnson-Laird[3]

[1]*University of Architecture, Venice, Italy;* [2]*University of Provence, Aix-en-Provence, France; and* [3]*Princeton University*

Abstract—*This article presents a theory of how individuals detect whether descriptions of an entity are consistent or inconsistent. The theory postulates that individuals try to construct a* mental model *of the entity in which all the propositions are true. If they succeed, they infer that the description is consistent; otherwise, they infer that it is inconsistent. We report three experiments that corroborated the theory. Experiment 1 confirmed that evaluating consistency is easier when an initial model suffices than when reasoners have to find an alternative model. Experiment 2 established the occurrence of illusory inferences about the properties of entities. Experiment 3 showed that the illusions correspond to mental models of the assertions, even when these models are wrong because they fail to represent what is false.*

A hallmark of rationality is to maintain consistent beliefs: If your beliefs are inconsistent, then the rational response is to modify them to restore consistency. The detection of inconsistency, however, is not trivial. One might think of inconsistency as a conflict between two propositions, but it can occur in a set of propositions in which any proper subset is consistent, for example:

Nassau Street runs east to west.
Route 1 runs north to south.
Nassau Street is parallel to Route 1.

Each pair of assertions in this set is consistent, but the three together are inconsistent. In general, the detection of inconsistency is intractable: As the number of clauses in a description increases, so the task makes exponential demands on time and memory (Cook, 1971). One antidote is to maintain separate sets of beliefs insulated one from another (Klein, 1998). That may be why some individuals accept both scientific explanations and unjustified belief systems (Jahoda, 1969). It may also explain why people succumb to self-deception (e.g., Elster, 1986; Mele, 1997).

How do logically untrained individuals determine whether or not descriptions are consistent? The present article aims to answer this question. One possibility is that they rely on formal rules of inference (e.g., Braine & O'Brien, 1998; Rips, 1994). Theories based on this idea aim to explain how individuals draw conclusions from premises. The evaluation of consistency is a very different task. A possible method is for reasoners to select one of the assertions in the description, and to try to prove its negation from the remaining assertions. If they succeed, the set of assertions is inconsistent; otherwise, it is consistent. This procedure seems psychologically implausible, and no one has proposed it. An alternative theory, however, can be based on the following assumption:

The principle of *models of consistency*: Individuals evaluate the consistency of a set of assertions by searching for a mental model of a possibility in which all the assertions are true. If there is such a model, the assertions are consistent; otherwise, they are inconsistent (Johnson-Laird, Legrenzi, Girotto, & Legrenzi, 2000).

Address correspondence to P.N. Johnson-Laird, Department of Psychology, Princeton University, Green Hall, Princeton, NJ 08544; e-mail: phil@princeton.edu.

Possibilities are central to this *model theory*, and a mental model, by definition, represents a possibility (e.g., Johnson-Laird & Byrne, 1991; Polk & Newell, 1995). Another assumption of the model theory is the following:

The principle of *truth*: People represent assertions by constructing mental models. Each model represents a true possibility, and each clause in the assertions (affirmative or negative) is represented in a model only if it is true in that possibility (Johnson-Laird & Savary, 1999).

For instance, consider the following *exclusive* disjunction:

There is not a circle or else there is a triangle.

The mental models of this disjunction are

¬ o	
	Δ

where each horizontal line represents a model of a separate possibility, "¬" denotes negation, "o" denotes a model of the presence of a circle, and "Δ" denotes a model of the presence of a triangle. Following the principle of truth, the first model does not represent explicitly that it is false that there is a triangle in this possibility. Reasoners make "mental footnotes" to keep track of the information about what is false, but they soon forget them. With these footnotes, however, they can convert mental models into *fully explicit* models. In this case, the fully explicit models are as follows:

¬ o	¬ Δ
o	Δ

A conditional, such as *If there is a circle then there is a triangle*, has the following mental models:

o	Δ
. . .	

The first model represents the possibility in which the antecedent is true, and the possibilities in which it is false are represented only implicitly by a model with no content other than a footnote that the antecedent is false. With this footnote, individuals can construct fully explicit models:

o	Δ
¬ o	Δ
¬ o	¬ Δ

The mental models of a biconditional, *There is a circle if and only if there is a triangle*, are identical to those for the conditional except that the footnote indicates that both the antecedent and consequent are false in the possibilities represented by the implicit model. The interpretation of assertions is also affected by pragmatic factors, including general knowledge (Evans & Over, 1996; Garnham & Oakhill, 1994; Johnson-Laird & Byrne, 2002).

The evaluation of consistency differs from deduction because a single model suffices for consistency whereas a deductive conclusion must hold in all the models of the premises. The model theory makes two main predictions about consistency. First, consistency should be harder to establish when a model of the initial assertion is inconsistent with a subsequent assertion than when a model of the initial assertion is consistent with a subsequent assertion. In the first case, individuals have to find an alternative model of the initial assertion consistent with the subsequent assertion; in the second case, this step is not necessary. Second, as a computer program implementing the theory predicted, compelling illusions should occur (cf. Johnson-Laird et al., 2000; Johnson-Laird & Savary, 1999). In the present research, we aimed to test both predictions, and thereby to show that the model theory extends to consistency.

EXPERIMENT 1

In this experiment, the participants evaluated sets of assertions: If they judged that a set was consistent, they had to describe the properties of the corresponding entity; otherwise, they had to respond that no description was possible. This procedure yields more information than one in which the participants merely evaluate consistency. Consider the following description:

1. The chair is saleable if and only if it is elegant.
   The chair is elegant if and only if it is stable.
   The chair is saleable or it is stable, or both.

To evaluate the consistency of this set, individuals should begin by constructing a mental model of the chair satisfying the first assertion:

saleable   elegant

They should extend this possibility according to the second assertion:

saleable   elegant   stable

They should judge that this model is compatible with the third assertion, and so they should describe the chair as *saleable, elegant, and stable*. In contrast, consider the following description:

1′. The chair is unsaleable if and only if it is inelegant.
The chair is inelegant if and only if it is unstable.
The chair is saleable or it is stable, or both.

Reasoners should construct a model satisfying the first two assertions:

unsaleable   inelegant   unstable

Because this model conflicts with the third assertion, they have to consider an alternative model of the first two assertions. The first biconditional allows the following two properties:

saleable   elegant

and the second biconditional allows the additional property:

saleable   elegant   stable

This model is consistent with the third assertion. But Problem 1′ should be harder than Problem 1 because of the need to reject the initial model and to construct an alternative. Of course, this particular problem might be harder because it contains implicit negatives, and so their occurrence was counterbalanced in the one-model and alternative-model problems that we used in this experiment.

Method

Participants

The participants were Italian applicants to Scuola Sant'Anna (a highly selective university in Pisa, Italy), and their mean age was about 19 years. We tested 20, who acted as their own controls and evaluated five pairs of problems. Each pair consisted of a one-model problem and a matching alternative-model problem, as shown in Table 1. The first four pairs of problems have consistent assertions. Problems 5 and 5′ have inconsistent assertions, but the principle of models of consistency predicts that Problem 5 should be easier than Problem 5′. We made two random assignments of contents to the 10 problems, and half the participants were tested with one assignment and half the participants were tested with the other assignment. Each participant received a booklet with the 10 problems in a different random order.

Materials and procedure

The problems concerned the properties of everyday entities. These properties and their implicit negations corresponded to common Italian adjectives, for example, "inelegante" (inelegant), "invendible" (unsaleable), and "instabile" (unstable). The key instructions were, "Your task is to describe with a single phrase a person or object. . . . you should use all the sentences in the description. . . . In some cases, it is not possible to find a set of properties of the person or object on the basis of the description. . . . In this case, you should check the box with the phrase: It is not possible to describe a set of properties for the person or object based on the description."

Results and Discussion

Table 1 presents the percentages of correct responses to the problems; to be scored as correct, descriptions for the consistent problems had to include all three correct properties. (An analysis in which responses for each property were scored separately yielded the same pattern of results.) The participants were more accurate with the one-model problems (97% correct) than with the alternative-model problems (39% correct): Responses of 19 participants fit the prediction, and the responses of 1 participant were tied (binomial test, $p = .5^{19}$). Responses to all five pairs of problems fit the prediction (binomial test, $p < .05$). The problems with implicit negatives in their first two assertions were not reliably harder than those with affirmatives in the first two assertions.

The results showed that the evaluation of consistency was harder when reasoners had to consider an alternative model than when the initial model sufficed for the task. The process of rejecting an initial model calls for subsequent premises to negate earlier premises. This negation is not an alternative explanation of the phenomenon, but merely another way of describing it. In the alternative-model problems, however, some individual clauses in the early assertions were contradicted in the final premise. This mismatch was not present in the one-model problems. Could this factor

Table 1. *Problems in Experiment 1, their correct descriptions, and percentages of correct responses*

One-model problems			Alternative-model problems		
Problem	Description	% correct responses	Problem	Description	% correct responses
1. A if and only if B. B if and only if C. A or C, or both.	A, B, C	100	1′. ¬ A if and only if ¬ B. ¬ B if and only if ¬ C. A or C, or both.	A, B, C	40
2. ¬ A if and only if ¬ B. ¬ B if and only if ¬ C. ¬ A or ¬ C, or both.	¬ A, ¬ B, ¬ C	100	2′. A if and only if B. B if and only if C. ¬ A or ¬ C, or both.	¬ A, ¬ B, ¬ C	35
3. A or else B, but not both. B or else C, but not both. A or C, or both.	A, ¬ B, C	95	3′. ¬ A or else ¬ B, but not both. ¬ B or else ¬ C, but not both. A or C, or both.	A, ¬ B, C	45
4. ¬ A or else ¬ B, but not both. ¬ B or else ¬ C, but not both. ¬ A or ¬ C, or both.	¬ A, B, ¬ C	90	4′. A or else B, but not both. B or else C, but not both. ¬ A or ¬ C, or both.	¬ A, B, ¬ C	50
5. A if and only if B. B if and only if C. A if and only if ¬ C.	None	100	5′. A or else ¬ B, but not both. ¬ B or else C, but not both. A or else C, but not both.	None	25

Note. The symbol for negation, ¬, followed by a letter denotes an implicitly negative property, such as "inelegant."

account for our results? Experiment 2 examined this question and a new prediction of the model theory.

EXPERIMENT 2

The principle of truth predicts the occurrence of *illusory* inferences, that is, compelling but erroneous inferences. There should be illusions of consistency in which reasoners infer that a description is consistent when, in fact, it is inconsistent, and there should be illusions of inconsistency in which reasoners infer that a description is inconsistent when, in fact, it is consistent. In Experiment 2, we used four sorts of problem based on the simplest assertions that should give rise to illusions (Johnson-Laird et al., 2000). Consider a pair of assertions of the form

A or else not both B and C.
A and not B.

The first assertion has the mental models

a		
	¬ b	c
	b	¬ c
	¬ b	¬ c

Absence from a model is taken to be consistent with negation; for example, the first model is consistent with *A and not-B*, and so individuals should respond that the two assertions are consistent. This response is an illusion. The mental models fail to take into account that when one of the clauses in the disjunction is true, the other clause is false. In contrast, the fully explicit models of the disjunction, which take falsity into account, are

a	b	c
¬ a	¬ b	c
¬ a	b	¬ c
¬ a	¬ b	¬ c

These models show that the two assertions are inconsistent. The problem should therefore yield an illusion of consistency. It is an example of the first sort of problem, which we designate as a *C/I* problem, to indicate that "consistency" (C) is the predicted response, and "inconsistency" (I) is the correct response. A control problem pairs the same disjunction with a different conjunction:

Not-B and not-C.

This conjunction corresponds to the fourth mental model, and so reasoners should respond that the two assertions are consistent. As the fully explicit models show, this response is correct. We designate this second sort of problem as *C/C*.

The theory also predicts illusions of inconsistency. An example of this third sort of problem, an *I/C* problem, pairs the same disjunction in the previous problem type with the assertion

A, B, and C.

The mental model of this conjunction does not occur among the mental models for the disjunction, and so reasoners should respond that the two assertions are inconsistent. Yet, as the fully explicit models of the disjunction show, this response reflects an I/C illusion. Finally, the fourth sort of problem is an *I/I* control. For the preceding illusion, this sort of control problem pairs the previous disjunction with the conjunction

Not-A, B, and C

which matches neither the mental models nor the fully explicit models.

Table 2. *Problems of Experiment 2 and percentages of correct evaluations*

Problem		% correct responses	Mental models of the first assertion	Fully explicit models of the first assertion
A or else not both B and C.			a	a b c
			¬ b c	¬ a ¬ b c
			b ¬ c	¬ a b ¬ c
			¬ b ¬ c	¬ a ¬ b ¬ c
1. A and not-B.	C/I	9		
2. A, not-B, and not-C.	C/I	9		
3. Not-B, and not-C.	C/C	92		
4. Not-A, B, and not-C.	C/C	85		
5. A, B, and C.	I/C	37		
6. Not-A, B, and C.	I/I	88		
Not both A and B, or else not both B and C.			a ¬ b	¬ a b c
			¬ a b	a b ¬ c
			¬ a ¬ b	
			b ¬ c	
			¬ b c	
			¬ b ¬ c	
7. A and not-B.	C/I	10		
8. A, not-B, and not-C.	C/I	9		
9. Not-A and B.	C/C	91		
10. Not-A, B, and C.	C/C	44		
11. A, B, and not-C.	I/C	41		
12. A, B, and C.	I/I	96		

Note. Each problem consisted of an initial disjunction followed by a conjunction. The table shows the mental models and the fully explicit models for each of the two sorts of disjunction. Problems expected to yield an illusion of consistency are denoted C/I; their control problems are denoted C/C. Problems expected to yield an illusion of inconsistency are denoted I/C; their control problems are denoted I/I.

Method

A new sample of 489 Pisan applicants acted as their own controls and responded to 12 problems, which are shown in Table 2, together with their mental models and their fully explicit models. We included two trials with different contents of Problems 5, 6, 11, and 12, so that there would be four trials for each of the four sorts of problem: illusions of consistency (C/I) and their controls (C/C), plus illusions of inconsistency (I/C) and their controls (I/I). The contents concerned everyday objects or people, and we made four different assignments of contents to the forms of the problems, testing approximately equal numbers of the participants with each of the different assignments. Each participant received the 16 trials in one of four different random orders. The materials and procedure were similar to those of Experiment 1, except that the participants had to decide whether or not the pairs of assertions could both be true at the same time.

Results and Discussion

Table 2 presents the percentage of correct responses to each problem, and Table 3 summarizes the results for the four sorts of problem. The participants performed better with the control problems than with the illusory problems (responses of 459 participants fit this pattern, responses of 11 participants went against it, and responses of 19 were tied; binomial test, $p < 5$ times in 10^{99}). Participants also performed better with problems having a predicted answer of "inconsistent" than with those having a predicted answer of "consistent" (responses of 427 participants fit this pattern, responses of 15 participants went against it, and the remainder were ties; binomial test, $p < 3$ times in 10^{105}). One reason for this result is that the illusions of consistency were more seductive than the illusions of inconsistency (responses of 269 participants fit this pattern, responses of 23 participants went against it, and the remainder were ties; binomial test, $p < 2$ times in 10^{54}). This effect probably arose because a judgment of "consistent" calls only for a single model satisfying the assertions, whereas a judgment of "inconsistent" calls for a check to ensure that no possible model satisfies the assertions.

In discussing Experiment 1, we described the alternative hypothesis that reasoners respond "inconsistent" whenever an individual clause and its negation occur in the set of assertions. The present results decisively refute

Table 3. *Percentages of correct evaluations of consistency and inconsistency for the four sorts of problems in Experiment 2*

Expected response	Illusions	Controls
"Consistent"	C/I: 9	C/C: 78
"Inconsistent"	I/C: 45	I/I: 94

Note. Problems expected to yield an illusion of consistency are denoted C/I; their control problems are denoted C/C. Problems expected to yield an illusion of inconsistency are denoted I/C; their control problems are denoted I/I.

this hypothesis. Problem 4, for example, has two such contradictions, but most participants correctly responded that the assertions were consistent.

The procedure in this experiment called for participants to evaluate consistency. It did not reveal the possibilities that the participants envisaged when they succumbed to illusions. Also, the experiment relied on Italian speakers to interpret "oppure" as an exclusive disjunction (see Johnson-Laird et al., 2000). Do reasoners really construct mental models of the sort proposed by the theory? And do the illusions occur when the exclusive disjunction is unequivocal? In Experiment 3, we aimed to answer both these questions.

EXPERIMENT 3

On each trial in this experiment, the participants had to describe a possibility satisfying two assertions. As before, we used four sorts of problems (C/I, C/C, I/C, I/I). A typical problem from Experiment 3 is

> Only one of the following assertions is true:
>
> The tray is heavy or elegant, or both.
> The tray is elegant and portable.
>
> The following assertion is definitely true:
>
> The tray is elegant and portable.
>
> Write a description of the tray. ________________

The rubric that only one of the two assertions is true expresses an unequivocal exclusive disjunction. The program implementing the theory yields the following mental models of the disjunction of the first two assertions

heavy		
	elegant	
heavy	elegant	
	elegant	portable

The conjunction that is definitely true corresponds to the last of these models. Hence, reasoners should formulate a description based on this last model. Because "heavy" is missing from the model, they should tend to describe the tray with its implicit negation: *light, elegant, and portable*.

The fully explicit models of the disjunction of the first two assertions are

heavy	¬ elegant	¬ portable
¬ heavy	elegant	¬ portable
heavy	elegant	¬ portable
heavy	¬ elegant	portable

The conjunction is inconsistent with these models, and so the expected response should reflect an illusion of consistency (C/I). Its C/C control problem pairs the same initial disjunction with a different conjunction that is definitely true:

> The tray is heavy and elegant.

The model for this conjunction is consistent with the third mental model of the disjunction, and so reasoners should respond with the description: *heavy, elegant, nonportable* ("intransportabile" in Italian). As the fully explicit models show, this response is correct.

An illusory I/C problem consists of the disjunction paired with the definitely true conjunction:

> The tray is inelegant and portable.

The model for this conjunction is not consistent with the mental models of the disjunction of assertions, but, as the fully explicit models show, the assertions in the problem are consistent. Finally, a control I/I problem pairs the disjunction with a conjunction of two implicit negations:

> The tray is light and inelegant.

We expected that participants would succumb to the illusions, perform well with control problems, and produce descriptions for illusory C/I and control C/C problems that correspond to mental models of the assertions.

Method

A new sample of 20 Pisan applicants acted as their own controls, and worked on four instances of each of the four sorts of problem (C/I, C/C, I/C, I/I). Table 4 presents the 16 problems. The contents concerned everyday objects. We made four different assignments of contents to the forms of the problems, testing equal numbers of the participants with each of the different assignments. The instructions and procedure were similar to those of Experiment 1.

Results

When the participants produced descriptions containing only two properties, these properties were those in the true conjunctive premises on 76% of such trials: Responses of 11 participants showed this trend, responses of 2 went against it, and the remainder were ties (binomial test, $p < .02$). Table 4 therefore presents frequencies of the predicted descriptions only for descriptions with three properties. Overwhelmingly, these descriptions were those that matched the predicted mental models of the problems (17 participants produced more of these predicted descriptions than unpredicted descriptions, 1 produced fewer of them, and 2 produced equal numbers of predicted and unpredicted descriptions; binomial test, $p < .0001$). Seven out of the eight relevant problems yielded more predicted descriptions than unpredicted descriptions, and for the remaining problem predicted and unpredicted descriptions were tied (binomial test, $p < .01$). The descriptions for the illusions of consistency, of course, are of nonexistent possibilities (i.e., the assertions in the problems are inconsistent). The descriptions for Problem 6, a C/C problem, were striking. Nearly all of them corresponded, as predicted, to a mental model, but it is a mental model that is impossible (cf. the fully explicit models in Table 4).

Table 5 shows the overall percentages of correct responses to the four sorts of problem. Every participant produced more correct responses to the control problems than to the illusions (binomial test, $p = .5^{20}$). W. Schaeken (personal communication, March 20, 2002) has pointed out that this experiment elicited more illusions of inconsistency than the previous one (13% vs. 45% correct answers). We suspect that the difference arose because Experiment 2 offered only

Table 4. *Problems of Experiment 3, their models, and the frequency of descriptions containing three properties that match the mental models*

Problem		Mental models of the disjunction	Fully explicit models of the disjunction	Frequencies of responses predicted by mental models[a]	
Only one is true:		a	a b ¬c		
A and/or B.		b	a ¬b c		
B and C.		a b	a ¬b ¬c		
		b c	¬a b ¬c		
Definitely true:					
1. B and C.	C/I			¬A B C:	8 out of 10
2. A and B.	C/C			A B ¬C:	12 out of 13
3. Not-B and C.	I/C			None:	18 out of 20
4. Not-A and not-B.	I/I			None:	20 out of 20
Only one is true:		a b	¬a ¬b c		
If A then B.		b c	a ¬b ¬c		
If C then B.		. . .			
Definitely true:					
5. A and B.	C/I			A B ¬C:	11 out of 13
6. A and not-C.	C/C			A B ¬C:	15 out of 16
7. A and not-B.	I/C			None:	19 out of 20
8. Not-A and not-C.	I/I			None:	20 out of 20
Only one is true:		a b	a b ¬c		
If A then B.		c	¬a b ¬c		
C.		. . .	¬a ¬b ¬c		
			a ¬b c		
Definitely true:					
9. Not-A and C.	C/I			¬A ¬B C:	9 out of 18
10. A and B.	C/C			A B ¬C:	14 out of 14
11. A and C.	I/C			None:	18 out of 20
12. B and C.	I/I			None:	18 out of 20
Only one is true:		a b	a b ¬c		
If A then B.		b c	¬a b ¬c		
B and C.		. . .	¬a ¬b c		
			¬a ¬b ¬c		
Definitely true:					
13. B and C.	C/I			¬A B C:	9 out of 10
14. A and B.	C/C			A B ¬C:	9 out of 10
15. Not-A and not-B.	I/C			None:	16 out of 20
16. A and not-B.	I/I			None:	19 out of 20

Note. Each problem consisted of a disjunction of two assertions followed by a conjunction. The table shows the mental models and the fully explicit models for each of the four sorts of initial disjunction. Problems expected to yield an illusion of consistency are denoted C/I; their control problems are denoted C/C. Problems expected to yield an illusion of inconsistency are denoted I/C; their control problems are denoted I/I. The symbol for negation, ¬, followed by a letter denotes an implicitly negative property, such as "inelegant."
[a]$N = 20$. "None" indicates the participants' judgment that there was no description of the assertions. For C/I and C/C problems, the table shows the number of responses that matched the prediction (out of the number of responses that included three properties).

two possible responses ("yes" vs. "no"), and so some participants could have guessed.

GENERAL DISCUSSION

Our results support the use of models of consistency. Experiment 1 showed that reasoners describe the properties of consistent assertions more accurately when a single model suffices than when it is necessary to reject an initial model and to consider an alternative. The principle of truth predicts the occurrence of illusions of consistency and inconsistency, which Experiment 2 corroborated. It also ruled out an alternative explanation according to which reasoners respond "inconsistent" whenever a clause and its negation occur among the set of assertions. Experiment 3 showed that individuals were indeed relying on mental models in order to perform this task, and that their illusory inferences did not depend on difficulties in interpreting "oppure" in Italian as an exclusive disjunction. When they succumbed to an illusion of consistency, they described the properties of entities that matched a mental model of the premises. These results corroborate the extension of the model theory to the evaluation of consistency.

Table 5. *Percentages of correct responses to the problems in Experiment 3*

Expected response	Illusions	Controls
"Consistent"	C/I: 8	C/C: 93
"Inconsistent"	I/C: 13	I/I: 96

Note. Problems expected to yield an illusion of consistency are denoted C/I; their control problems are denoted C/C. Problems expected to yield an illusion of inconsistency are denoted I/C; their control problems are denoted I/I.

Individuals can detect inconsistencies when in a self-consciously critical frame of mind, as when they participate in experiments or, say, review manuscripts submitted for publication. But, in a more relaxed setting in daily life, they can also notice inconsistencies in sets of assertions. Could the inconsistencies of daily life be simpler in structure than those in our experiments, and therefore not elicit a model-based strategy? Could ordinary reasoning be based instead on formal rules of inference (Braine & O'Brien, 1998; Rips, 1994)? The difficulties with such a proposal are twofold. First, in daily life, people do seem to think about possibilities, and this assumption predicted performance in our experiments. The participants' ability to think about possibilities yielded their excellent performance with the one-model problems in Experiment 1 and the control problems in Experiments 2 and 3, which were often identical in form to the illusory problems (see Tables 2 and 4). Second, illusory inferences do occur in daily life (see Johnson-Laird & Savary, 1999), but it is difficult to see how any theory using valid formal rules could explain their occurrence. Valid rules cannot yield systematically erroneous evaluations. It may be tempting to introduce invalid rules of inference to accommodate the phenomena, but a theory based on invalid rules would lead to wholesale irrationality far beyond illusory inferences.

Reasoners appear to model the consistency of assertions by trying to envisage a compatible possibility. If they construct such a model, they can use it to describe an entity consistent with the assertions; if they cannot construct such a model, then they can declare that the assertions are inconsistent. The difficulty of the task depends on whether an initial model suffices to formulate a description. Individuals also succumb to illusory inferences about the properties of entities. And these illusions correspond to mental models of assertions, even when these models are wrong because they fail to represent what is false.

Acknowledgments—This research was funded by Grant ex-40% (1999) from the Italian Ministry of Education and University and by National Science Foundation Grant BCS-0076287 to P.N.J-L. to study strategies in reasoning. We thank Riccardo Varaldo, Director, Scuola Sant'Anna of Pisa, for allowing us to conduct experiments there; Alberto Mazzocco for enabling us to conduct research in the Department of Psychology, Padua University; Maria Legrenzi and Stefania Pizzini for assistance with all stages of the research; and many colleagues, too numerous to name, for advice. Finally, we are grateful to Daniel Reisberg, Walter Schaeken, and an anonymous reviewer for helpful criticisms.

REFERENCES

Braine, M.D.S., & O'Brien, D.P. (Eds.). (1998) *Mental logic*. Mahwah, NJ: Erlbaum.
Cook, S.A. (1971). The complexity of theorem proving procedures. *Proceedings of the Third Annual Association of Computing Machinery Symposium on the Theory of Computing*, *3*, 151–158.
Elster, J. (1986). *The multiple self*. Cambridge, England: Cambridge University Press.
Evans, J. St.B.T., & Over, D. (1996). *Rationality and reasoning*. Hove, England: Psychology Press.
Garnham, A., & Oakhill, J.V. (1994). *Thinking and reasoning*. Oxford, England: Basil Blackwell.
Jahoda, G. (1969). *The psychology of superstition.* Harmondsworth, Middlesex, England: Penguin Books.
Johnson-Laird, P.N., & Byrne, R.M.J. (1991). *Deduction*. Hillsdale, NJ: Erlbaum.
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., & Savary, F. (1999). Illusory inferences: A novel class of erroneous deductions. *Cognition*, *71*, 191–229.
Klein, G. (1998). *Sources of power: How people make decisions.* Cambridge, MA: MIT Press.
Mele, A.R. (1997). Real self-deception. *Behavioral and Brain Sciences*, *20*, 91–136.
Polk, T.A., & Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review*, *102*, 533–566.
Rips, L. (1994). *The psychology of proof.* Cambridge, MA: MIT Press.

(Received 11/17/01; Revision accepted 4/23/02)