Psychologica Belgica
1997, 37-4, 219-247.

BELGIAN CORNER
INVITED ADDRESS

META-PROPOSITIONAL REASONING ABOUT THE TRUTH OR FALSITY OF PROPOSITIONS

Walter SCHROYENS
University of Leuven

A combination of premises such as "Person A asserts that if and only if he's a liar, then person B is a liar" and "Person B asserts that person A is a truth-teller", with the question what the status of persons A and B might be (truth-teller or liar), elicits meta-propositional reasoning, that is, reasoning about the truth and falsity of propositions. Both an inference rule and a mental models approach have been proposed to explain for meta-propositional reasoning. These proposals are compared to one another with respect to the strategies that people use, and this suggests that the proposed strategies are to a large extent ad hoc amendments to either theory. A review of a series of studies (Schroyens, 1995), controlling some confounding aspects in previous research, gives little evidence for a specific short-cut strategy that hinges on making backwards inferences. This counters the original proposal of Johnson-Laird and Byrne (1990) and recent corroborations of this strategy (Byrne & Handley, 1997; Schroyens, Schaeken, & d'Ydewalle, 1996). Other findings of our studies, however, indicate that, relative to the starting hypothesis by which one enters the formal structure of the problems, reasoners do not always engage in an exhaustive search strategy. This counters Rips (1989, 1990a) rule-based model, but is in accordance with the mental model theory's principle that a validating search for counter-examples is carried out by default but is not invariably complete.

Introduction

Rips (1989) introduced a novel type of reasoning problems. An example of these problems is given in (1).

(1) Person A asserts: Person A is a liar or person B is a truth-teller, or both.
Person B asserts: Person A is a truth-teller.

In 1996, Walter Schroyens won the Award for the Best License Thesis of the Belgian Psychological Society. The present invited address is a part of the award. It presents more recent work performed by Walter Schroyens.

I am deeply endebted to Walter Schaeken and Géry d'Ydewalle for their guidance in conducting and reporting the research for my licence thesis, of which the present invited address offers a summary overview. I also thank the anonymous reviewers who offered helpful suggestions to improve the present manuscript. I gladly thank Jeremy Pacht for reading the manuscript and help on improving my English. The Flanders Fund for Scientific Research (F.W.O-Vlaanderen, Belgium) currently supports my research on the psychology of reasoning.

Correspondence concerning this article should be addressed to Walter Schroyens, Laboratory of Experimental Psychology, Department of Psychology, University of Leuven, Tiensestraat 102, 3000 Leuven. Electronic mail may be sent to Walter.Schroyens@psy.kuleuven.ac.be.

These problems are phrased in a context in which one is to imagine two sorts of people that can only be distinguished by their propensity to tell the truth. On the one hand there are truth-tellers: persons who always tell the truth. On the other hand there are liars: persons who always tell a lie. The question becomes to determine, if possible, the truth-status of both assertors. Problems like (1) are also known as "knight-knave" problems because they were originally phrased in the context of "knights" and "knaves" (see Smullyan, 1979). I refer to reasoning about these problems as *meta-propositional reasoning* because the accounts that have been proposed rely heavily on the models for propositional reasoning or reasoning about sentences containing connectives such as "if", "and", "or", and "not". Meta-propositional reasoning is a meta-deductive type of reasoning because it calls for considerations about the truth and falsity of propositions.

Byrne and Handley (1997) recently pointed out some of the advantages of studying these meta-propositional problems. First, they allow us to examine the strategies that people use when being confronted with complex problems. This aspect of reasoning is not specific to meta-propositional reasoning, though it is sometimes considered as such. The complexity of the problems results in multiple routes being possible to reach their solution. Some of the questions that must be taken into account are: Where and how does one start tackling these problems; if there's a "fork" in the formal structure of these problems, which route does one take; and at what point does one stop?

Second, meta-propositional problems require an assessment of the truth or falsity of assertions. This is contrasted with the fact that participants in experiments on deductive reasoning are always asked to assume the truth of the premises. As such it can be argued that psychological theories on deductive reasoning scrutinize the cognitive processes that enable people to bridge the truth of the premises and the conclusion. However, determining the truth-value of an assertion itself constitutes a problem in our daily interactions. We practically never have absolute certainty about the truth-value of the premises on which we base our arguments. Indeed, logical systems have been proposed in which more than just two truth-values are used. In these systems a proposition is not either true or false, but can have an intermediate truth-value (see, e.g., Post, 1921). In our reflections and inferences about the truth-value or "verisimilitude" (Popper, 1963) of propositions, both coherence and correspondence will play a role. On the one hand, people can look for the correspondence between the state of affairs that is described by the assertion and a state of affairs in reality. Not all our knowledge, though, is directly verifiable against externally observable facts. On the other hand, people can check whether a particular proposition is consistent with other propositions. One checks the plausibility of a proposition relative to generally accepted, supposedly true background knowledge. In most of our daily problems a recursive integration occurs between the logical reasoning mechanisms and the meta-logical reflections about the propositions on the basis of which we make logical inferences.

There is another potential advantage of studying these problems. They may enable us to confront alternative theories of the higher-level control processes that guide the basic deductive machinery in a sequence of inferences. This topic is closely linked to the more descriptive topic of the strategies that people employ during the process of meta-propositional reasoning. However, extant theories of deductive reasoning fail to specify strategies that follow directly from their basic processing principles (see, e.g., Evans & Over, 1997; Roberts, 1993). The models that have been presented within the general framework of one or the other theory seem to be largely ad hoc concerning the strategical component. Obviously, it remains perfectly debatable which theory does better in accommodating a strategy.

I have been referring to alternative theories about the psychology of reasoning. With respect to peoples' competence in making deductive inferences, at present there are two main opposing theories: rule theories and the mental model theory (for a review see, e.g, Evans, 1989; Evans, Newstead, & Byrne, 1993). The two approaches have been extended to the topic of meta-propositional reasoning. First, Rips (1989, 1990a) gave an account of meta-propositional reasoning by mental inference rules. He proposed an extension to his earlier ANDS model (A Natural Deduction System; Rips, 1983; see also Rips, 1994) by adding some task-specific rules to the basic set of mental inference rules for propositional logic. The inference rules were embedded in a single exhaustive proof procedure. Second, in response to Rips' (1989) challenge to come up with an alternative model for meta-propositional reasoning, Johnson-Laird and Byrne (1990, 1991) elaborated the mental model theory to provide such an alternative account. In contrast with Rips' single exhaustive proof procedure, Johnson-Laird proposed multiple strategies.

In the following I will first present and discuss the rule theoretic and model theoretic accounts of meta-propositional reasoning, pointing out some flaws in each account. Next I present a meta-analytic survey of a series of studies that tried to unconfound research supporting the mental model theory in its supposition that people use short-cut strategies which hinge on making backwards inferences. The evidence in favour of a short-cut strategy proposed by the mental model theory is shown not to be conclusive. In a third section it is argued that in order to fully understand short-cut strategies one should also take into consideration the formal structure of the meta-propositional reasoning problems (i.e., where in the reasoning process an assertor can be assigned a truth-status: late or early, definitely or not). Several experiments are reviewed in which the applicability of specific short-cut strategies was disentangled from the formal structure of the problems. The results indicate that the formal structure of the problems, and more particularly how reasoners enter the formal structure, narrows down the possible points at which they can satisfactorily abandon a line of reasoning. In the final section I discuss the implications of the results with respect to rule-based and model-based approaches to reasoning.

Psychological Theories of Meta-Propositional Reasoning

Meta-Propositional Reasoning by the Application of Mental Rules of Inference

Rule-theoretic proposals have been the major approach in the psychology of reasoning (e.g., Braine, 1978; Braine, Reiser, & Rumain, 1984; Johnson-Laird, 1975; Osherson, 1975; Rips, 1983, 1988, 1990b, 1994). Rips (1983) most aptly expresses the general idea of these rule theories:

> Deductive reasoning consists in the application of mental inference rules to the premises and conclusion of an argument. The sequence of applied rules forms a mental proof or derivation of the conclusion from the premises, where these implicit proofs are analogous to the explicit proofs of elementary logic. (p. 40)

Rule theories suppose that when a mental representation of the premises is constructed especially the logical form is represented (Rips, 1990b). These propositional representations are abstracted from natural language and trigger the application of mental inferences rules. Applying a particular inference rule results in making the inference. The following inference schema responsible for a so-called Modus Ponens argument, for instance, is incorporated in all rule based models:

IF p, THEN q
p
——
q

The working principle of these inference rules, of which typically less than twenty are incorporated in a psychological model, can best be illustrated in terms of production rules: If a certain condition is met, then execute the action. For instance, if working memory contains the propositions *IF p THEN q* and *p*, then add the proposition *q* to working memory. Rule theories are a formal approach to reasoning: Application of the inference rules is held to occur irrespective of the specific content of the propositions and is determined by the structure of the sentences.

As mentioned, Rips (1989) extended his model for propositional reasoning (Rips, 1983, 1994). He added four task specific meta-inference rules in order to handle meta-propositional problems.

i. If truth-teller(x) and asserts(x,p) then p
ii. If liar(x) and asserts(x,p) then NOT(p)
iii. If NOT(truth-teller(x)) then liar(x); If liar(x) then NOT(truth-teller(x))
iv. If NOT(liar(x)) then truth-teller(x); If truth-teller(x) then NOT(liar(x))

These inference rules allow people to infer that a truth-teller tells the truth (i), that a liar does not make true assertions (ii), that someone who is no truth-teller is a liar (iii.), and that someone who is no liar is a truth-teller (iv). These meta-inference rules and the standard set of propositional inference rules are embedded within a single and exhaustive proof procedure. One generates and tests both hypotheses about the first assertor (and other assertors when their status remains undetermined during the process of following up a supposition about the first assertor). Reasoners first start with the hypothesis that the first assertor is a truth-teller. A hypothesis that results in a contradiction is rejected and a person is assigned a particular truth-status if this is consistent across all sets of hypotheses.

Table 1 illustrates how Rips' rule-theoretic proof procedure works with respect to problem (1), which for simplicity is presented here again.

Table 1

An Example of Meta-Propositional Reasoning by Application of Mental Rules of Inferences in the Frame of an Exhaustive Proof Procedure

1.	Person A asserts: Person A is a liar or person B is a truth-teller, or both.	
2.	Person B asserts: Person A is a truth-teller.	
1.	asserts (A, liar(A) OR truth-teller(B))	
2.	asserts (B, truth-teller(A))	
3.	truth-teller(A)	Hypothesis generating rule
4.	liar(A) OR truth-teller(B)	Meta-inference (i) on 1.
5.	NOT(liar(A))	Meta-inference (iv) on 3.
6.	truth-teller(B)	OR-elimination on 4 and 5
7.	truth-teller(A)	Meta-inference (i) on 2
		No further inferences rules applicable
3'	liar(A)	Hypothesis generating rule
4'	NOT(liar(A) OR truth-teller(B))	Meta-inference (ii) on 1.
5'	NOT(liar(A)) AND NOT(truth-teller(B))	De Morgan on 4= and 5'.
6a'	NOT(liar(A))	AND-Elimination on 5'
b	NOT(truth-teller(B))	
7'.	NOT(liar(A))	NOT-intr.:Contradiction 3' and 7'
8.	truth-teller(A)	Meta-inference 4 on 7'
	Consistent set of assignments: truth-teller(a) and truth-teller(b)	

(1) Person A asserts: Person A is a liar or person B is a truth-teller, or both.
Person B asserts: Person A is a truth-teller.

First, a hypothesis generating rule is responsible for positing the hypothesis that A is a truth-teller. As such, meta-inference Rule (i) is applicable: A's assertion must be true. The hypothesis and the inclusive disjunction jointly trigger a rule for a disjunctive inference. That is, the proposition that B is a truth-teller is added to working memory. This proposition yields the meta-propositional inferences that person b's assertion must be true. That is, the proposition that A is a truth-teller is added to working memory. At this point no further rules are applicable, but a consistent set of assignments is delineated: Both A and B are truth-tellers. However, when reasoners are following the exhaustive proof procedure, they will also generate and test the alternative hypothesis that A is a liar. Given this hypothesis, A's assertion should not be true. The negation of A's assertion is worked out by a form of the so-called "De Morgan" inference rule, yielding the conjunction of both negated constituent clauses. Elimination of a conjunction is accomplished by placing both conjuncts in separate nodes of working memory. As such the meta-inference Rule 4 can be applied to the proposition that A is not a liar: the proposition that A is a truth-teller is added. But this proposition contradicts the starting assumption that A is a liar and so the hypothesis is rejected: A is not a liar. In sum, both A and B are truth-tellers, because this assignment is consistent across all sets of hypotheses.

Generating and testing hypotheses is one of the twin pillars of Rips's (1983) ANDS model that has recently had been elaborated as PSYCOP (the acronym for "Psychology of Proof"; Rips, 1994). However, there are some flaws in the way ANDS handles hypotheses, as Johnson-Laird (in press) argues with respect to PSYCOP. Rips (1989) has to contend with the fact that none of the standard propositional inference rules is triggered by the premises of a meta-propositional problem. This seems to result in the thesis that "*to make any headway*, the program has to try out some assumptions, and the first assumption it makes is *knight(A)* [truth-teller(a)]" (Rips, 1989, p. 92, first italics added). That is, generating this Gricean hypothesis seems to be an ad hoc postulate, since it falls beyond the margins of both ANDS' and PSYCOP's competence in relation to the problem format of meta-propositional problems. Concerning the status of backward rules responsible for generating hypotheses in ANDS, Rips (1983) reports that "the only inferences it can produce are those of its forward rules" (p. 67). Analogously, Rips (1994) states that "in situations where PSYCOP is expected to produce conclusions rather than to evaluate them, it can use only its forward rules to complete the task, since there is no conclusion-goal to trigger the backward rules" (p. 105). In the experiments conducted by Rips (1989), no evaluation task was used and participants had to produce their own conclusions within the margins of four or six response alternatives. In addition, Rips (1989) omitted the propositional backward inference rules from his meta-propositional ANDS model "since the protocol showed little evidence of subgoaling, ..." (p. 93). This means that, despite the fact that hypotheses form an integral part of the natural deduction approach advocated by Rips, generating hypotheses within the context of meta-propositional reasoning is not a direct specification of the theory.

The problem of generating hypotheses can be remedied quite easily, though other problems would arise (see Rips, 1994). Indeed, Rips (1983) suggests that eliminating the restriction of a subgoal for triggering the application of backward rules would augment the number of conclusions that can be produced. Placing rules in a "conclusion-production mode" is not yet a part of the PSYCOP program (see Rips, 1994).

Meta-Propositional Reasoning by Constructing and Manipulating Mental Models

According to the mental model theory (Johnson-Laird, 1983), reasoning is based on the construction and manipulation of mental models: representations of the possible state of affairs as described by structural relationships expressed in the premises. Reasoning by model is broken down into three processing stages (see e.g., Johnson-Laird & Byrne, 1991). The processing stages embody the principle of eliminating possible state of affairs when reasoners try to construct an integrative descriptive model of the premises. That is, on the basis of the meaning of the premises and other general knowledge triggered during the process of interpretation, reasoners first construct a set of mental models of the possible state of affairs. After this "comprehension" stage, reasoners derive the integrative model of the premises by conjoining the various models, while eliminating any inconsistent models. Reasoners will formulate a novel but putative conclusion on the basis of the integrative model that was constructed during this "description" stage. During the third "validation" stage a search is made for alternative models of the premises which falsify the putative conclusion. If there is no such model, then the conclusion is valid, that is, the denoted state of affairs is not just possibly, but is necessarily the case. If there is such falsifying model, then it is necessary to return to the second stage to determine whether there is any conclusion that holds over all the models constructed thus far. The theory's essential processing assumption is that the more models that have to be constructed the harder the inferential task will be. This principle gained support from studies of syllogistic reasoning (Johnson-Laird & Bara, 1984), spatial and temporal reasoning (Byrne & Johnson-Laird, 1989; Schaeken, Johnson-Laird, & d'Ydewalle, 1996a, 1996b; Vandierendonck & De Vooght, 1996), reasoning with multiple quantifiers (Johnson-Laird, Byrne, & Tabossi, 1991), and propositional reasoning (Johnson-Laird, Byrne, & Schaeken, 1992, 1994).

How did Johnson-Laird and Byrne (1990, 1991) extend the model for propositional reasoning to handle meta-propositional reasoning? I will present this model for meta-propositional reasoning, while pointing out some of the frailties in its original formulation. A first thing to notice is that Johnson-Laird and Byrne did not incorporate the distinction between implicit and explicit mental models of the assertions made by liars or truth-tellers. Johnson-Laird et al. (1992) proposed that reasoners initially represent as little information as possible explicitly, such that the load on working memory is reduced. The implicit information is not immediately available for processing, but becomes available when made explicit ("fleshing out"). For instance, A's disjunctive assertion in problem (1) "Person A is a liar or person B is a truth-teller, or both" would initially be represented by two partially explicit and one completely implicit model.

~A	
	B
...	

The first model denotes the state of affairs in which A is a liar ("~A"), while leaving implicit a representation of B in this model. The second model denotes the situation in which B is a truth-teller ("B"), while in this disjunct a representation of A is left implicit. The third completely implicit model denotes that other, as yet unspecified, situations are possible. The above set of implicit or explicit mental models is contrasted to the fully explicit representation:

~A	~B
A	B
~A	B

Johnson-Laird and Byrne (1990, 1991) discussed meta-propositional reasoning by model on the basis of the completely fleshed out set of mental models. Schroyens, Schaeken, and d'Ydewalle (1997) proposed to upgrade the model for meta-propositional reasoning by integrating the distinction between implicit and explicit mental models, as well as other recent developments of the mental model theory.

One will notice that there is a discrepancy between the mental models of the meta-propositional premise and the structural relations expressed by the first premise in problem (1) ("Person A asserts: Person A is a liar or person B is a truth-teller, or both"). The premise also contains the information that the disjunction is a proposition expressed by A. Schroyens et al. (1997) proposed that reasoners initially would also represent this information, which would be in accordance with the principle that mental models are homomorphic representations of the situations described by the premises (Johnson-Laird, 1983). That is, when constructing the initial mental model of the meta-propositional premise, the model of the assertion would be annotated by a token which represents the assertor who makes the assertion.

a: ~A
B

...

The annotated model "a" denotes explicitly that it is A who makes the assertion, but leaves implicit the status of A. Initially representing A's status implicitly, is consistent with the idea that reasoners initially represent as little information as possible (Johnson-Laird et al., 1992).

Johnson-Laird and Byrne (1990, 1991; see also Byrne, Handley, & Johnson-Laird, 1992, 1995; Byrne & Handley, 1992, 1997), and Rips (1989, 1990a) suppose that reasoners start solving problems like (1) by generating the hypothesis that the first assertor "A" is a truth-teller. This hypothesis is in accordance with one of the Gricean maxims of conversation: one's contribution to conversation should be truthful (Grice, 1975). Contrary to the assumption that reasoners initially already represent the fact that the first assertion is an assertion of a particular assertor, Johnson-Laird and Byrne (1990) assume that reasoners immediately annotate the mental models of the assertions with a model of the assertor as being a truth-teller when generating the "Gricean hypothesis". They propose that reasoners use the token "A", that is, the representation that A is a truth-teller to annotate the assertion, whereas I propose that "A" is a specification of A's truth-status. The mental model theory can provide an account of the thesis that a line of meta-propositional reasoning is initiated via the generation of the Gricean hypothesis "A" by making the assumption that initially, reasoners explicitly represent the link between the assertor and his or her assertion but leave the truth status of the assertor unspecified. Indeed, since mental models are "constructed on the basis of the truth conditions of the propositions expressed in the discourse" (Johnson-Laird, 1983, p. 407), the token "a" intended to represent that it is A who makes the assertion, actually implicitly represents that A is a truth-teller. It must be noted, however, that Johnson-Laird (1983) argued for the thesis that mental models represent the truth-conditions on the basis of the Gricean maxims.

Once people have generated a starting hypothesis, they follow up its consequences. It is assumed that given the hypothesis that someone is a truth-teller, the first inference would be that his or her assertion should be a true assertion (but see Schroyens et al., 1997). With respect to problem (1) in Figure 1 this means that the model for the annotated hypothesis that A is a truth-teller is integrated with the models which represent the assertion. As such, the elimination procedure in conjoining mental models is invoked when constructing a descriptive model of A's premise (both the assertor and the assertion). This procedure results in the "A-model", that is the model of A's premise in the case A is a truth-teller:

A B

The first model in which A was represented as a liar ("~A") is eliminated because it is inconsistent with "A". For the same reason, the third model is eliminated. However, the A-model is not an integrating model of the entire problem. The information in the second premise has yet to be conjoined with it. Since B is a truth-teller in the A-model, he or she should tell the truth. Hence, the second premise of problem (1) is represented as:

B: A

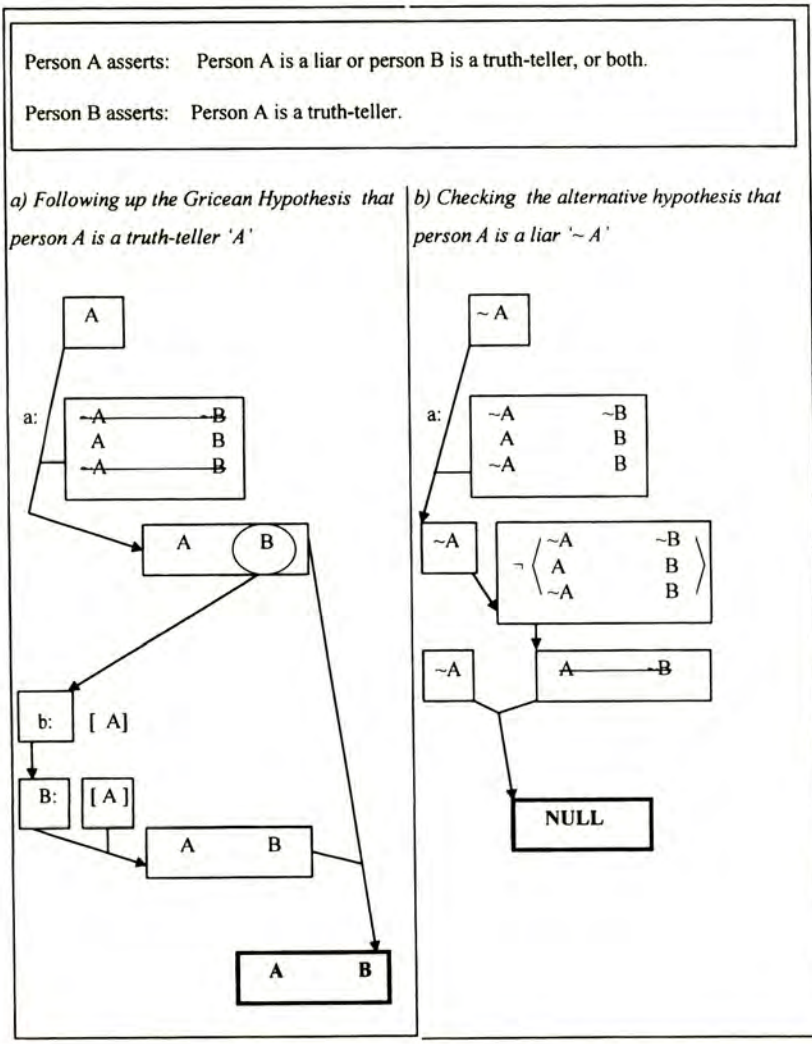

Figure 1. An example of meta-propositional reasoning by constructing and manipulating mental models in the frame of a full-chain procedure.

Conjoining this B-model, that is the model of B's premise in the case B is a truth-teller, with the A-model results in:

A B

This model yields the putative conclusion that A is a truth-teller and B is a truth-teller.

In order to validate this putative conclusion, reasoners search for alternative models which might falsify it. Given the starting hypothesis that A is a truth-teller, there is the obvious alternative possibility that A might be a liar. Following up this hypothesis requires constructing the mental models that represent the possible state of affairs when the disjunctive contingencies are uttered by a liar and are no longer true. Pursuing this negative hypothesis calls for constructing the complement of the set of models representing A's truthful assertion.

$$
\sim A \quad \neg \left\langle \begin{array}{cc} \sim A & \sim B \\ A & B \\ \sim A & B \end{array} \right\rangle
$$

The large brackets are not conventionally used. I use this notation to indicate the difficulty of negating a compound proposition: It requires that all models are represented explicitly and that the complement of the set of models is formed to represent that the assertion is not true. I use the negation symbol "5" to keep count of the difference between negation and falsity (see Rips, 1990a). Constructing the complement of the disjunctive assertion results in:

~A: A ~B

Integrating the annotated model of the lying assertor "~A" with the complementary mental models of the untrue assertion invokes the elimination procedure and results in the empty "Null" model. That is, nothing follows from the assumption that A is a liar. Hence, no alternative model of the premises of problem (1) could be constructed and the previous conclusion that A is a truth-teller and B is a truth-teller, could not be falsified and is a valid conclusion.

Johnson-Laird and Byrne (1990, 1991) labeled the above exhaustive procedure the "full-chain" strategy and noted that it is the mental model theory's analogue of the exhaustive proof procedure proposed by Rips (1989). However, Johnson-Laird and Byrne (1990) consider the full-chain strategy implausible because it would place too much processing load on working memory, regardless of whether it is based on the application of mental rules of inference or the manipulation and construction of mental models. Therefore they proposed several simpler strategies. Of those "less-then-full-chain" strategies proposed by Johnson-Laird and Byrne (1990, 1991) I will discuss only one: the so-called *hypothesize-and-match strategy*. I focus on this strategy because at present there seems to be a limited consensus about it. It was the only short-cut strategy to account for some of the variance in the data of Rips (1989; see Rips, 1990a), and recently some corroborating evidence has also been reported by Byrne and Handley (1997), and Schroyens et al. (1996).

The hypothesize-and-match strategy was stipulated to consist of the following procedure: "If the assumption that the first assertor is telling the truth leads to a contradiction, then attempt to match "~A" [A is a liar] with other assertions ..." (Johnson-Laird & Byrne, 1991, p. 159). Problem (2) is an example of a problem that can be solved by this hypothesize-and-match strategy, as well as by the full-chain strategy.

(2) Person A asserts: Person A is a truth-teller and person B is a truth-teller.
Person A asserts: Person A is liar.

As before, reasoners would initiate their line of reasoning by generating the Gricean hypothesis that the first assertor is a truth-teller ("A"; see Figure 2). Integrating hypothesis "A" which annotates the conjunctive assertion of A results in the A-model:

A B

Following up the consequences with respect to the second premise results in the B-model:

B: ~A

Conjoining the A-model and the B-model in order to construct an integrating descriptive model of the entire problem results in the Null-model. That is, nothing follows from the hypothesis that A is a truth-teller. Hence, A must be a liar:

~A

This definite assignment that A is a truth-teller matches the model of B's assertion:

b: ~A

That is, given that it is true that A is a liar and B indeed asserts that A is indeed a liar, B is a truth-teller because he or she would be telling the truth.

B: ~A

This means that without following up the consequences of "~A" with respect to A's assertion by forming its complement, reasoners can derive the valid conclusion for the problem: A is a liar and B is a truth-teller.

It should be noted, however, that the hypothesize-and-match strategy hinges upon a meta-propositional backward inference, and not particularly on matching. Byrne and Handley (1997), for instance, refer to the hypothesize-and-match strategy as the "simple backwards inference strategy". The matching mechanism can easily be seen as conjoining two mental models: the assignment of A's status and the assertion about A's status. Inferring the status of B on the basis of a comparison between, on the one hand, the previously inferred status of A and, on the other hand, the status of this A in B's assertion, makes up a meta-propositional backward inference: From the truth-value of an assertion (derived by matching), one infers the truth-status of the assertor. In the case of forward inference one infers the truth-value of an assertion (true/false) from the truth-status (truth-teller/liar) of the assertor. In the case of a full chain strategic line of inferences about problem (2), people would not apply the backwards matching inference, but would follow up the consequences of "~A" by a lengthier chain of forwards inferences. The fact that a single problem can be solved by means of different strategies illustrates Johnson-Laird and Byrne's supposition that reasoners' deductive machinery serves, as it were, as the subroutines for the strategies that they might employ in reasoning about the truth or falsity of assertions. In the following section I discuss the experiments that have been conducted to test the hypothesis that people use short-cut strategies in solving meta-propositional reasoning problems.

Evidence for Short-Cut Strategies in Meta-Propositional Reasoning

To test the plausibility of the hypothesize-and-match strategy (see above) usually a comparison is made between problems that cannot and those that can be solved by it. As a function of the status of the first assertor, these two types of problems can be referred to as "truth-teller problems" and "liar problems". For these problems the correct assignment of the first assertor is that she or he is a truth-teller or that she or he is a liar, respectively. Problem (1) is an example of a truth-teller problem (see Table 1 and Figure 1). Problem (2) is an example of a liar problem (see Figure 2). With respect to truth-teller problems, following up the Gricean hypothesis that the first assertor is a truth-teller by definition does not result in a contradiction. As such no definite assignment for A's status can be made and, thus, the matching-mechanism of the hypothesize-and-match strategy can not be applied (for other restrictions on the applicability of the hypothesize-and-match strategy, see Byrne, Handley, & Johnson-Laird, 1992). With respect to liar problems, the Gricean truth-teller hypothesis by definition results in a contradiction. The resulting assignment of the first assertor's status can then be matched with the second assertor's categorical assertion about the first assertor's status.

As noted before, Byrne and Handley (1997; Experiment 1) and Schroyens et al. (1996) found some corroborating evidence for the use of the hypothesize-and-match strategy. In general, the liar problems were solved better than the truth-teller problems. However, in both studies there is a possible confound in comparing the truth-teller and liar problems. This possible confound is present when the truth-teller problems contain a compound assertion with an inclusive disjunction (see Figure 1) and these problems are compared with liar problems in which the first assertor's assertion is a conjunction (see Figure 2). Representing a conjunctive assertion calls for constructing only a single mental model, whereas an inclusive disjunction would result in constructing three models. Hence, a difference in the percentages of correct responses on truth-teller and

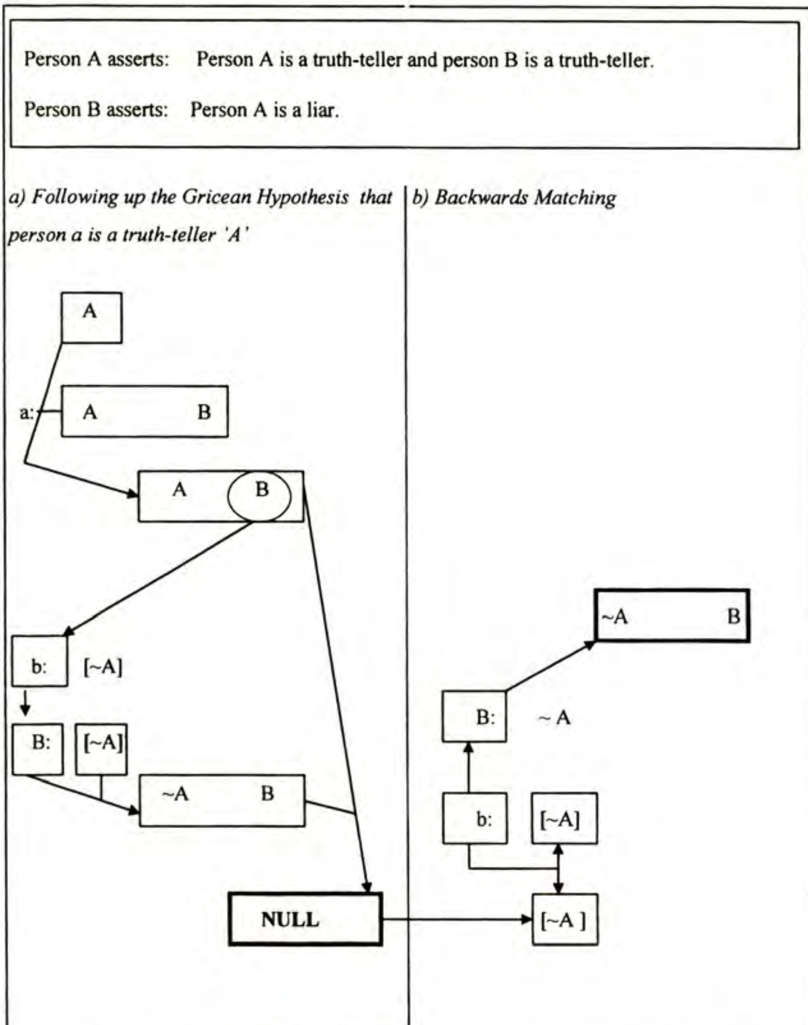

Figure 2. Example of solving a liar-problem (the first assertor is a liar "~A") by means of the hypothesize-and-match strategy.

liar problems cannot be accounted for unequivocally by the hypothesize-and-match strategy. The data of Experiment 1 in Byrne and Handley (1997) and especially a quick look at their Appendix A, seem to support the idea of a confound in the evidence interpreted as favouring the hypothesize-and-match strategy. On the one hand, a comparison of the disjunctive truth-teller problems with the conjunctive liar problems in their experiment reveals an overall difference of 21% in the percentages of correct solutions. On the other hand, when A's assertion was an exclusive disjunction in both the truth-teller and the liar problems, the overall difference on the percentages of correct solutions was only 7%.

It seems that in order to get a more stringent test of the hypothesize-and-match strategy one should control for the number of mental models that are constructed when representing the first assertor's compound assertion. Representing an exclusive disjunction ("either ..., or ..., but not both") or a bi-conditional ("If and only if ..., then ...") supposedly calls for constructing two mental models, though for the bi-conditional initially only one model will be represented explicitly. Byrne and Handley (1997, Experiment 2) reported results which were taken to indicate an effect of the applicability of the hypothesize-and-match strategy to the liar problems. This result was observed even though the number of models, that is, the connective in the compound assertion in these problems was controlled for: The first assertor's compound assertion was always a bi-conditional proposition. Byrne and Handley noted that the 18% difference between the truth-teller and the liar problems eliminated the possible confounds in the results they reported in Experiment 1. However, the 18% difference was not statistically significant at $\alpha = .05$ ($N = 11$).

Table 2 presents the results obtained in confronting nine groups of participants with problems in which the first assertor made either a bi-conditional assertion or an assertion with an exclusive disjunction (See the Appendix for a formal representation of the problems). In all problems the first assertor (A) made a compound assertion and the second assertor (B) made a categorical assertion about A ("_ab" problems; A's assertion comes before B's assertion). One-third of the problems in the first six experiments was presented without any ready-made hypothesis about A truth-status. These problems served as the control for the problems presented with various hypotheses about A's status. For instance, the hypothesis "Suppose that [A] is a truth-teller" ("Ha") was presented with half of the experimental problems: It is an appropriate hypothesis to the truth-teller problems and an inappropriate hypothesis to the liar problems. The inverse holds for the alternative hypothesis by which the participants were asked to make the supposition that A is a liar: It is appropriate to the liar problems and inappropriate to the truth-teller problems. In two of these experiments I used a more extended type of hypotheses: Suppose first that [A] is a truth-teller/liar, such that his/her assertion should be a true/false assertion. ("Ha+"). Other groups of participants received double hypotheses (Ha^2 and $Ha+^2$): Suppose first that [person a] is a truth-teller/liar (such that ...). Next suppose that [person a] is liar/truth-teller (such that ...). The data of the first six groups of participants are the results on the control problems, that is, the problems that were presented without a hypothesis. The final three experiments in Table 2 show the results of three groups of participants who served as the

Table 2

Mean Proportion and Standard Deviation of Correct Solutions Given by Nine Groups Presented With Problems in Which the First Assertor Made an Assertion With a Bi-Conditional or an Exclusive Disjunction, and the Second Assertor Made a Categorical Assertion

	First assertors' status			
	Truth-teller "A"	"~A" Liar	N	Z
Schroyens (1995)				
Ha_ab	.42 (.40)	.38 (.42)	25	- 0.431
Ha^2_ab	.42 (.40)	.42 (.37)	25	0.000
Ha+_ab	.33 (.42)	.43 (.36)	27	+0.931
$Ha+^2$_ab	.40 (.41)	.45 (.38)	21	+0.459
ab_Ha	.41 (.33)	.50 (.40)	29	+1.034
ab_Ha^2	.42 (.37)	.36 (.39)	26	- 0.770
Schroyens et al. (1997)	.33 (.29)	.39 (.30)	26	+1.955
Schroyens (1995)	.39 (.25)	.32 (.26)	17	- 1.137
Schroyens (in preparation)	.44 (.25)	.38 (.28)	28	- 2.452

Note. The number of problems per subject on which the mean proportions were calculated for the respective groups are: 2, 2, 2, 2, 2, 2, 4, 8, 8. See text and Appendix for the denotation of the different groups.

control group for those experiments in which the order of presentation was manipulated: truth-teller problems before or after the liar problems, as compared to the control group who did not receive these two types of problems in a blocked order. The final two groups of participants were confronted with a repeated presentation of the problems, that is, they had to solve each problem twice.

As can be seen in Table 2, there seems to be little evidence in favour of the hypothesize-and-match strategy. There was no overall reliable difference between the participants' performance on the problems, nor was there an overall reliable interaction as a function of the first assertor's status in these truth-teller or liar problems. Therefore the data were simply collapsed. There was no overall reliable difference between the truth-teller and the liar problems (40% vs. 41% on the liar problems). Meta-analyses resulted in a parallel pattern of results (see Rosenthal, 1994; for the applied procedures). A diffuse test of the Z-values associated with the p-levels in the respective groups was not statistically significant ($X^2(8) = 14.034$, $p = .08$), but nonetheless hinted at a heterogeneous pattern of results. Although the null-hypothesis of no differences between the truth-teller or liar problems could not be rejected, care should be taken when interpreting the overall pattern of results.

Discussion

In sum, the present section reviewed a series of experiment which controlled for the propositional connective that is used to construct the compound assertion of the first of the two assertors in the truth-teller or liar-problems. The analyses show that applicability of the hypothesize-and-match strategy originally proposed by the mental model theory did not result in better performance on the liar problems. At first sight, these results seem to imply a rejection of the mental model theory of meta-propositional reasoning, at least as concerns the proposed short-cut strategies.

However, Rips (1990a) argued that the strategies that were proposed by the mental model theory are independent of its general framework. As noted earlier, he found some evidence in favor of the hypothesize-and-match strategy. Indeed, simply adding the backwards meta-propositional inference rules:

If liar(x) and asserts(y, liar(x)) then truth-teller(y)

If liar(x) and asserts(y,truth-teller(x) then liar(y)

would allow for accommodation of the hypothesize-and-match strategy. Rips (1990a) states that some protocols show that reasoners would indeed not employ the systematic procedures, that is the full chain strategy they develop when confronted with meta-propositional reasoning problems. He therefore proposed to regard his model as a proposal about the way fully-fledged meta-propositional reasoners handle the problems at a steady-state of performance. This is more so since his theory makes no claims concerning the mechanisms responsible for the development of strategies; however, as argued by Rips, neither has the mental model theory. As a consequence, problem-specific strategies (like the hypothesize-and-match strategy) seem to be ad hoc amendments to, in the present case, the mental model theory. That is, the fact that Johnson-Laird and Byrne (1990) specified the strategy is not a sufficient condition for positive evidence in favor of this strategy being extrapolated to the theoretical framework which happens to have specified it first. However, the inverse also holds. The processing assumptions of the mental model theory are not falsified by a failure to corroborate the hypothesize-and-match strategy.

The Formal Structure of Meta-Propositional Reasoning Problems

Does the failure to find further evidence for short-cut strategies mean that reasoners engage in a exhaustive, full chain of meta-propositional inferences? It should be noted at the outset that the following discussion about the formal structure of meta-propositional reasoning problems is to a large extent independent of a rule or mental models based approach to reasoning. For this reason, reference is made to the "formal structure", as opposed to "proof structure" of the meta-propositional reasoning problems. Indeed, Rips's "Psychology of Proof" (1994) defends the thesis that reasoners will try to construct a mental proof of a problem's solution. However, once the concepts of a late or early, definite or indefinite assignment of the assertors' truth-status have been introduced, some restrictions are pointed out on the possible

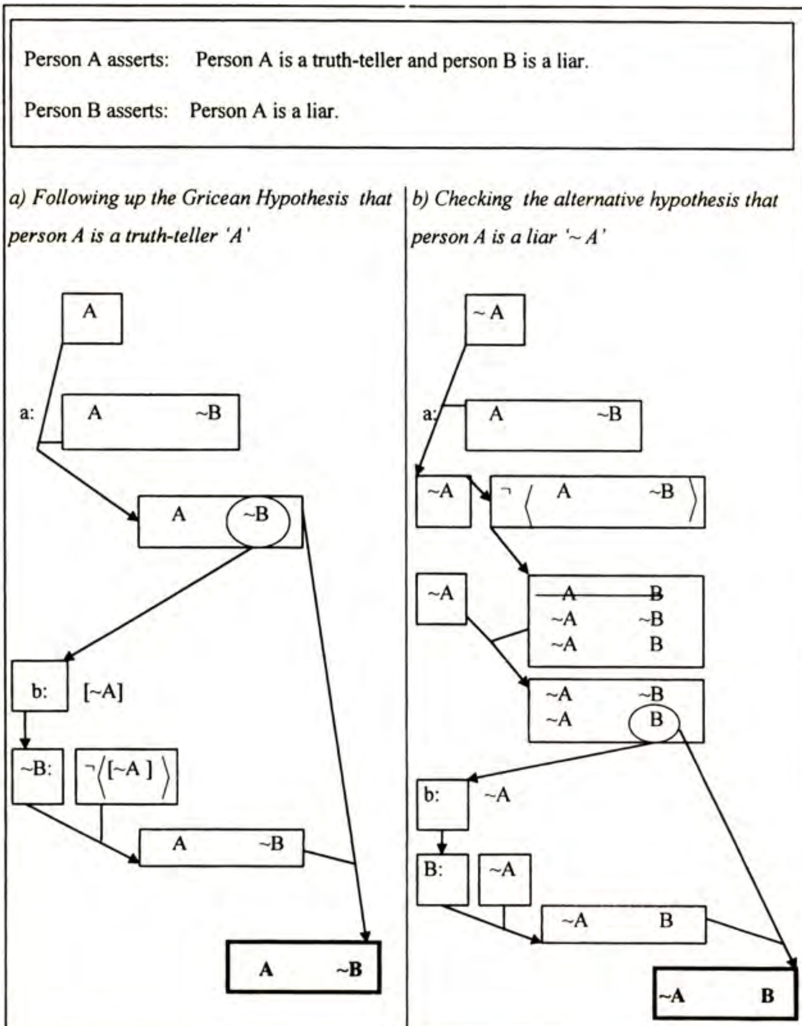

Figure 3. An example of a contigent meta-propositional reasoning problem in which the assignement of one assertor's status is contigent upon the assignment of the other person's status.

demarcation points at which reasoners might abandon a line of reasoning. That is, do people try to construct a mental proof or do they satisfice with an unvalidated, putative conclusion?

When reasoners start with the Gricean hypothesis that the first assertor is a truth-teller, the truth-teller problems yield a *late definite assignment* of the first assertor's truth-status. It is only after both hypotheses about A have been generated and tested that a mental proof (Rips, 1994) can be constructed of the first assertor's status in these problems. What is meant by a "definite assignment" can best be illustrated by a counter example. Problems can be constructed for which only a contingent solution exists. For the problem depicted in Figure 3, the hypothesis that A is a truth-teller results in an assignment of a truth-status to both assertors in the problems: If A is a truth-teller, then B is a liar. However, following up the alternative hypothesis that the first assertor is a liar also yields a consistent set of assignments: If A is liar, then B is a truth-teller. That is, an assignment of the second assertor's status is contingent upon an assignment of the first assertor's status and, hence, the correct solution to these problems is two-folded. In contrast with the late definite assignment that the first assertor is a truth-teller in the truth-teller problems, the Gricean hypothesis results in an *early definite assignment* of the first assertor's status in the liar problems. The Gricean hypothesis has contradictory consequences. As such, reasoners can infer (by some form of a Reductio Ad Absurdum argument) that the first assertor cannot be a truth-teller and, thus, given the knowledge about the context of persons who are either liars or truth-tellers, must be a liar.

The conceptual difference between an early and a late definite assignment for respectively the liar and the truth-teller problems does not imply that the liar problems would be solved better than the truth-teller problems (pace Byrne, Handley, & Johnson-Laird, 1992). Late definite assignment involves an assignment of *both the assertors'* truth-status: a complete solution for the truth-teller problems is inferred. In contrast, early definite assignment involves an assignment of only *one of the assertors'* truth-status: the problem has not yet been solved because only a partial solution for the liar problems is inferred. That is, when reasoners initiate a line of reasoning with an inappropriate hypothesis (e.g., the Gricean hypothesis is an inappropriate starting hypothesis with respect to the first assertor's liar status in the liar problems), they must construct a lengthier chain of inferences in order to derive a solution to the problem. However, if they first generate an appropriate starting hypothesis, then they need not follow up the alternative and inappropriate hypothesis. Even if they did, reasoners would already have a conclusion (which happened to be the correct one) on which to fall back. Hence, one might expect that entering the formal structure of the truth-teller problems by the Gricean hypothesis would result in better performance than tackling the formal structure of the liar problems by this hypothesis. That is, because reasoners can reach an early, though indefinite assignment of both the assertors in the truth-teller problems, these problems might be easier than the liar problems for which one can only make a late, but definite assignment of both the assertors.

Note that these principles would result in a pattern of results opposite to the one expected in applying the hypothesize-and-match strategy to the liar problems. However, when considering both the applicability of the hypothesize-and-match strategy and the formal structure of the problems, the relative performance on the liar and truth-teller problems can go either way. It is not just dependent on the liar problems being easier because the hypothesize-and-match strategy can be applied to them. The relative performance also depends on the type of strategy people use in tackling the truth-teller problems. If reasoners abandon a chain of reasoning prematurely, that is, they give a solution when they have one (regardless of whether this conclusion is proven or validated), then the truth-teller problems would be solved better than in the case where reasoners employed an exhaustive proof-procedure. As such liar problems need not be solved better than truth-teller problems.

A type of problem can be constructed which can be used to disentangle the applicability of the hypothesize-and-match strategy and the formal structure of the problems. For instance, Byrne and Handley (1997; see also Byrne et al., 1992) not only presented participants with the type of simple truth-teller or liar problems discussed thus far, but also constructed more complex problems. In these more complex problems both assertors make a compound assertion, as compared to the more simple problems in which the second assertor makes a categorical assertion about A. I will refer to these two types of problem according to the nature of the second assertor's assertion: *compound problems* (both assertions are compound assertions) and *categorical problems* (one person makes a categorical assertion). As with the distinction between the categorical truth-teller or liar problems, the compound problems are of two types depending on the first assertor's actual truth-status: problems in which the first assertor is a truth-teller, or problems in which the first assertor is a liar.

With respect to the compound truth-teller and liar problems, the results of Byrne and Handley (1997; Experiment 2) were in the expected direction (30% vs. 23%), but the effect was not reliable. In Table 3 the results are presented, obtained form twelve groups of participants given these two types of problems (Schroyens, 1995; Chapter 3). The results from the first six experiments were obtained from the same groups of participants discussed in the previous section (see Table 2). Besides the categorical control-problems (the "ab"-problems with first A's and next B's categorical assertion, presented without a hypothesis) these participants also received compound problems that served as controls for the other compound liar- and truth-teller problems presented with an accompanying hypothesis. The latter six groups of participants received the same premises, but presented in the reverse order (hence denoted as "ba"-

problems; see the Appendix). In the previous section the results on the categorical problems presented in these "ba-groups" were not included since the hypothesize-and-match strategy, especially the matching component of the strategy, is not applicable to them (see Byrne, Handley, & Johnson-Laird, 1992): the second assertor makes a compound assertion.

As expected because the problems shown in Table 3 served as the control problems in the series of experiments, there was neither an overall grouping effect, nor was there an overall interaction with the effect of the first assertor's status. Collapsing these data resulted in a reliable difference between the compound truth-teller and liar problems (37% vs. 26%; Wilcoxon $T = 3244.5$, $N = 144$, $Z = 3.941$, $p = .00008$). Meta-analytic analyses on the associated Z-value of the effect size p-values in the respective groups also showed no reliably diffuse pattern of effects. Combining the Z-values reveals a reliable effect of A's status in the compound truth-teller or liar problems ($Z = 3.717$, $p = .0001$). Indeed, all groups of participants showed an effect in the same direction.

Table 3

Mean Proportion (and standard deviations between brackets) of Correct Solutions in Twelve Groups of Participants Presented With Compound Problems in Which One Assertor Made a Bi-Conditional Assertion, Whereas the Other Assertor's Assertion was an Exclusive Disjunction

	First assertor=s status			
Groups	Truth-teller	Liar	N	Z
Ha_ab	.32 (.41)	.26 (.33)	25	+0.711
Ha²_ab	.32 (.38)	.28 (.36)	25	+0.384
Ha+_ab	.35 (.36)	.22 (.29)	27	+1.420
Ha+²_ab	.40 (.41)	.36 (.36)	21	+0.459
ab_Ha	.33 (.36)	.33 (.38)	29	+0.000
ab_Ha²	.48 (.37)	.34 (.35)	25	+1.350
Ha_ba	.37 (.43)	.17 (.24)	27	+2.167
Ha²_ba	.28 (.39)	.22 (.29)	23	+0.477
Hb_ba	.46 (.36)	.20 (.28)	28	+2.641
Hb²_ba	.28 (.33)	.24 (.33)	25	+0.561
Hb+_ba	.44 (.38)	.21 (.29)	26	+2.341
Hb+²_ba	.34 (.35)	.31 (.40)	16	+0.365

Note. The results are based on two observations per participant. See the Appendix for the notation that is used to denote the different groups of participants.

Though there was no overall grouping effect, a more detailed look at the data suggests that the type of hypothesis presented with the experimental problems might have exerted an influence on the performance on the control-problems represented in Table 3. When combining the groups presented with a double hypothesis in the experimental problems, the difference between the compound truth-teller and liar problems tends to be smaller than in the case where participants had to solve these control problems presented without a hypothesis in the context of other problems being presented with a single hypothesis (Mann-Whitney $U = 9780$, $n = 162$, $m = 165$, $Z = 1.707$, $p = .0879$).

Table 4

Differences Between the Proportion of Correct Solutions on the Compound Truth-Teller and Liar-Problems that Served as the Control Problems in the Groups of Participants Presented With Different Types of Single or Double Hypotheses to the Experimental Problems

Groups		Hypotheses		Z
		Single	Double	
Ha_ab	vs. Ha^2_ab	.06 (.41)	.04 (.53)	+ .053
Ha&_ab	vs. $Ha\&^2$_ab	.13 (.43)	.05 (.44)	+ .536
ab_Ha	vs. ab_Ha^2	.00 (.42)	.14 (.47)	-1.239
Ha_ba	vs. Ha^2_ba	.20 (.42)	.07 (.43)	+1.220
Hb_ba	vs. Hb^2_ba	.27 (.42)	.04 (.32)	+2.293
Hb&_ba	vs. $Hb\&^2$_ba	.23 (.41)	.03 (.34)	+1.798

Note. The results are based on two observations per participant. See the Appendix for the notation that is used to denote the different groups of participants.

This (apparent) interaction between the context of single or double hypotheses and the status of the first assertor in the problems can also be analyzed by the meta-analytic procedure of combining the Z-values of this interaction observed in comparing two groups with one another. Two specific groups can be matched since care was taken to maintain a conceptual link between the single and the double hypotheses that were presented to the different groups of participants (see the Appendix). The double hypotheses are matched with the single hypotheses because the first supposition in the double hypothesis is the same as the single hypothesis. For instance, in the *Ha-group* the participants received the experimental problems accompanied with a hypothesis of the form "Suppose that [A] is a truth-teller/liar", whereas the participants in the *Ha^2-group* received hypotheses such as "Suppose first that [A] is a truth-teller/liar. Suppose next that [A] is a liar/truth-teller". The Z-values of the differences between the compound truth-teller and liar-problems as a function of the groups presented with single or double hypotheses are represented in Table 4. Across the set of comparisons the interaction effect approached significance ($Z = 1.903$, $p = .057$, two-tailed).

This finding indicates that performance on the control problems was marginally affected by the context in which the problems were presented and thus questions the appropriateness of these problems as controls against which the effect of presenting a hypothesis can be assessed. Nonetheless, it is interesting to note that this finding provides some further evidence in favor of the importance of a problem's formal structure: entering the formal structure of the meta-propositional problems with an appropriate supposition might yield better performance than initiating a line of reasoning with an inappropriate supposition. When the participants in the double hypothesis groups are explicitly asked to follow up both the appropriate and the inappropriate hypothesis accompanying a problem, they seem to generalize this procedure to the control problems presented without such a double hypothesis. That is, the truth-teller problems that served as the control problems tended also to be solved through a procedure of generating both the appropriate and the inappropriate hypothesis. By assumption, the starting hypothesis would be the appropriate Gricean hypothesis (see above). However, the participants would be less inclined to stop after following up this hypothesis, which would result in the correct solution. They would also follow up the inappropriate hypothesis such that their line of reasoning would become a more extended one and they would be more likely to err — for instance, by deriving a second conclusion that is consistent with the inappropriate hypothesis (see Schroyens et al., 1996, 1997).

Discussion

How do these findings relate to the discussion between model-based or rule based reasoning? First, one should note that Rips' (1989) exhaustive meta-propositional inference procedure in fact delineates a deterministic proof procedure.

> Which step the computation takes next depends on the current contents of working memory and the conditions of the inference rules. No special "strategies" are necessary to determine stopping points, since inference rules (unlike mental models) supply this for free: problem-solving halts when the puzzle is solved or no more rules apply. (Rips, 1990a, p. 307)

Similarly, when Johnson-Laird and Byrne (1990, 1991) implement the model-theoretic reasoning mechanisms as subroutines within an algorithmic procedure that parallels the control mechanisms adhered to by Rips (1989), problem solving also halts when the puzzle is solved or no further alternative models of the problem can be constructed. However, finding a solution and interrupting the problem-solving procedures is used here in a restricted sense. In the case of an exhaustive proof-procedure one only halts making inferences when the solution, if any, is *proven and correct*. The full set of inference rules is applied. The same holds with respect to the full chain strategic manipulation of mental models: one stops if a previously inferred putative conclusion is validated by the consideration of all possible mental models.

Rips's (1990a) critique that constructing and manipulating mental models does not provide a means of delineating stopping points can be inverted. Trying to construct a mental proof seems to suggest that no end can be put to a line of reasoning besides the one that is marked by the proven conclusion or the empty pool of applicable inference rules. The mental model theory (Johnson-Laird & Byrne, 1991) also states that validating a putative conclusion is the standard. Indeed, it is the only stage in which "any essential deductive work is carried out: the first two stages are merely normal processes of comprehension and description." (p. 36). However, "it does not postulate that the search for counter examples is invariably complete - far from it, as such an impeccable performance would be incompatible with observed errors" (Johnson-Laird & Byrne, 1991, p. 39). For instance, with respect to the so-called "belief-bias", a series of experiments conducted by Newstead, Pollard, Evans, and Allen (1992; see also Evans, Newstead, Allen, & Pollard, 1994) provided corroborating evidence for the model theory's principle that believable or unbelievable conclusions respectively instigate or inhibit the search for counter-examples. As such, the present results seem to be more in line with the mental model theory.

A Perspective on Meta-Propositional Reasoning: Rules, or Mental Models, or Both?

There are two main findings from the studies I have reviewed. First, there is little evidence that reasoners use the hypothesize-and-match strategy (Johnson-Laird & Byrne, 1990) in solving meta-propositional reasoning problems. However, as argued on the basis of the ad hoc character of strategies, this failure to replicate Byrne and Handley (1997), and Schroyens et al. (1996) can hardly be taken as evidence against the mental model theory, and neither can the evidence in favour of this short-cut strategy be taken as evidence favouring the theory. In addition, a rule-theoretic approach can equally well incorporate or exclude this strategy from the procedures people develop when confronted with meta-propositional reasoning problems. As such, it seems that scrutinizing the strategies that reasoners use in meta-propositional reasoning does not provide a viable way of pitting a mental models approach and a rule-based approach.

Second, there is some evidence that the difficulty of meta-propositional reasoning is mediated by the way reasoners enter the formal structure of the meta-propositional reasoning problems. Generating a starting hypothesis by itself poses no major difficulty in solving meta-propositional reasoning problems, but has a precursor function to the difficulty of testing and revising hypotheses.

I argued that starting to solve problems with an inappropriate hypothesis precludes deriving a complete solution without making an exhaustive chain of inferences, whereas initiating a line of reasoning with an appropriate hypothesis results in the possibility of prematurely abandoning the line of reasoning. The data of a series of studies indicated that some reasoners do indeed show a "satisficing" tendency to stop reasoning prematurely at the point where they have inferred an as yet unvalidated solution for the problem (for arguments and evidence along the same lines see Schroyens et al., 1996). This tendency for seizing the possibility of abandoning a line of reasoning without making an exhaustive chain of inferences goes against a rule-theoretic approach advocating that people try to construct a mental proof of a problem's solution. Obviously, further research is needed to gain more insight in the strategies people use when solving problems. At present, however, theories of human deductive competence seem to fall short in specifying which strategies are developed, and the circumstances under which they are applied.

Nonetheless, given that reasoners apply a specific problem solving procedure, it remains debatable which theory can accommodate it best. At present the mental model theory seems to do the better job. First, it provides a sound answer to the question of how people start reasoning about the problems: Generating the Gricean starting hypothesis that a person is a truth-teller is accountable by the thesis that mental models represent the truth-conditions of the propositions expressed in discourse. Second, the mental model theory can reckon more easily for people's tendency to satisfice with a putative conclusion without an attempt to construct a mental proof of a problem's solution. Other findings such as the type of errors that people make in solving meta-propositional reasoning problems, which I have not discussed in detail (see Schroyens et al., 1996, 1997), also hint that the mental model theory provides a more complete account of how people traverse the route from the beginning to the end of a meta-propositional line of reasoning.

References

Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review*, *85*, 1-21.

Braine, M. D. S., Reiser, B. J., & Rumain, B. (1984). Some empirical justification for a theory of natural propositional logic. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 18, pp. 313-371). New York: Academic Press.

Byrne, R. M. J., & Handley, S. J. (1992). Reasoning strategies. *Irish Journal of Psychology*, *13*, 111-124.

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

Byrne, R. M. J., Handley, S. J., & Johnson-Laird, P. N. (1992). Advances in the psychology of reasoning: Meta-deduction. In M. T. Keane & K. J. Gilhooly (Eds.), *Advances in the psychology of thinking* (Vol. 1, pp. 127-145). London: Harvester Wheatsheaf.

Byrne, R. M. J., Handley, S. J., & Johnson-Laird, P. N. (1995). Reasoning from suppositions. *Quarterly Journal of Experimental Psychology, 48A*, 915-944.

Evans, J. St. B. T. (1989). *Bias in human reasoning: causes and consequences*. London: Erlbaum.

Evans, J. St. B. T. (1990). Reasoning with knights and knaves: A discussion of Rips. *Cognition, 36*, 85-90.

Evans, J. St. B. T., Newstead, S. E., Allen, J. L., & Pollard, P. (1994). Debiasing by instruction: The case of belief bias. *European Journal of Cognitive Psychology*, 6, 263-285.

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

Evans, J. St. B. T., & Over, D. E. (in press). Rationality in reasoning: the problem of deductive competence. *Current Psychology of Cognition.*

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3: Speech acts; pp. 41-58). New York: Academic Press.

Johnson-Laird, P. N. (1975). Models of deduction. In R. J. Falmagne (Ed.), *Reasoning: Representation and process in children and adults* (pp. 7-54). Hillsdale, NJ: Erlbaum.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge, MA: Cambridge University Press.

Johnson-Laird, P. N. (in press). An end to a contraversy? A Reply to Rips. *Mind and Language.*

Johnson-Laird, P. N., & Bara, B. G. (1984). Syllogistic inference. *Cognition, 16*, 1-62.

Johnson-Laird, P. N., & Byrne, R. M. J. (1990). Meta-logical problems: Knights, knaves and Rips. *Cognition, 36*, 69-84.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. London: Erlbaum.

Johnson-Laird, P. N., Byrne, R. M. J., & Tabossi, P. (1991). Reasoning by rule or model: The case of multiple quantification. *Psychological Review, 96*, 658-673.

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

Johnson-Laird, P. N., Byrne, R. M. J., & Schaeken, W. (1994). A defense of the model theory of propositional reasoning: A reply to Bonatti, and to O'Brien, Braine, and Yang. *Psychological Review, 101*, 734-739.

Osherson, D. (1975). Logic and models of logical thinking. In R. J. Falmagne (Ed.), *Reasoning: Representation and process in children and adults* (pp. 81-91). Hillsdale, NJ: Erlbaum.

Popper, K. R. (1963). *Conjectures and refutations: The growth of scientific knowledge.* London: Routledge and Kegan Paul.

Post, E. L. (1921). A general theory of elementary propositions. *The American Journal of Mathematics, 43*, 163-185.

Roberts, M. J. (1993). Human reasoning: Deduction rules or mental models, or both? *Quarterly Journal of Experimental Psychology, 46A*, 569-589.

Rosenthal, R. (1993). *Meta-analytic procedures for social research.* Londen: Sage

Publications.

Rips, L. J. (1983). Cognitive processes in propositional reasoning. *Psychological Review, 90*, 38-71.

Rips, L. J. (1988). Deduction. In R. J. Sternberg & E. E. Smith (Eds.), *The psychology of human thought* (pp. 116-152). Cambridge, England: Cambridge University Press.

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

Rips, L. J. (1990a). Paralogical reasoning: Evans, Johnson-Laird, and Byrne on liar and truth-teller puzzles. *Cognition, 36*, 291-314.

Rips, L. J. (1990b). Reasoning. *Annual Review of Psychology, 41*, 321-353.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking.* Cambridge, MA: MIT Press.

Schaeken, W., Johnson-Laird, P. N., & d'Ydewalle, G. (1996a). Mental models and temporal reasoning. *Cognition, 60*, 205-234.

Schaeken, W., Johnson-Laird, P. N., & d'Ydewalle (1996b). Tense, aspect and temporal reasoning. *Thinking and Reasoning, 2*, 309-327.

Schroyens, W. (1995). *Het redeneren over de waarheid of valsheid van beweringen.* Unpublished Licence thesis, University of Leuven, Leuven.

Schroyens, W., Schaeken, W., & d'Ydewalle, G. (1996). Meta-propositional reasoning with knight-knave problems: The importance of being hypothesized. *Psychologica Belgica, 36*, 145-169.

Schroyens, W., Schaeken, W., & d'Ydewalle, G. (1997). *Error and bias in meta-propositional reasoning: A case of the mental model theory.* Manuscript submitted for publication.

Smullyan, R. M. (1978). *What is the name of this book? The riddle of Dracula and other logical puzzles.* Englewoord Cliffs, NJ: Prentice Hall.

Vandierendonck, A., & De Vooght, G. (1996). A comparison of reasoning with time and space concepts. *Thinking and Reasoning, 2*, 249-272.

Appendix

The Appendix presents the 24 problems that were used in the Experiments of Schroyens (1995) and also discusses the design of the twelve groups of participants in these experiments. The correct solution with respect to the status of both assertors in the problems are underscored and placed in front of the assertion made by "person A" and "person B". "Person A" and "Person A" were represented by male and female proper names with a differing initial letter. "A" denotes that person A is a "truth-teller", like "B" denotes that person B is a truth-teller. "~A" represents that person A is a liar, like "~B" represents that person B is a truth-teller. The problems could contain either one or two compound assertions: respectively "categorical" and "compound" problems. A bi-conditional "If and only if ...., then ..." (<—>), or an exclusive disjunction "Either ..., or ...., but not both" (>—<), or both, was used as the connective in the assertions. Half of the problems were "A"-problems, that is, problems in which the correct assignment of person A's status is that she or he is truth-teller. In half of these A-problems, person B is a truth-teller "B". In the other six A-problems, person B is a liar (~B). Twelve other problems were "~A"-problems, that is, the problems in which person A is a "liar": In six of them person B is a truth-teller, in the other six ~A-problems, person B is a liar.

A specific short-hand denotes the different groups of participants which were all first-year psychology students. The groups in the experiments were: 1) Ha_ab, 2) Ha^{2}_ab, 3) Ha+_ab, 4) Ha+^{2}_ab, 5) ab_Ha, 6) ab_Ha^{2}, 7) Ha_ba, 8) Ha^{2}_ba, 9) Hb_ba, 10) Hb^{2}_ba, 11) Hb+_ba, Hb+^{2}_ba. Let's take the first group as point of reference. The participants received the problems presented with or without a ready made hypothesis about person A's status. The hypothesis presented with the problems ("Ha") was either an appropriate or an inappropriate hypothesis with respect to the actual status of person A. That is, each participant received eight problems presented with an appropriate hypothesis, eight without a hypothesis and eight problems with an appropriate hypothesis. The appropriate hypothesis for the A-problems is the hypothesis "Suppose that [person A] is a truth-teller". This hypothesis is the inappropriate hypothesis with respect to the ~A-problems. The inverse holds for the simple hypothesis "Suppose that [person A] is a liar". It's the appropriate hypothesis for the A-problems and the inappropriate hypothesis for the A-problems. The problems were presented with the hypothesis above the problem with first person A's assertion, and next person B's assertion (ab). In groups 7-12, the problems were presented with first person B's assertion and next person A's assertion (ba). In the groups 5 and 6, the problems were presented first, and below the problems the hypothesis (ab_Ha) was represented, whereas in the other groups the hypothesis was presented in front of the problems. In the groups 7-12 the given hypothesis was one with respect to the status of person B (Hb). The given hypothesis could also be an extended one. For instance ASuppose that [person A] is a truth-teller such that his or her assertion would be a true assertion (Ha&). All groups who received a single hypothesis (Ha, Hb,Ha+ or Hb+) were paralleled by groups who received a double hypothesis in the same format. For instance, in the Ha^{2}_ab-group, the appropriate double hypothesis with respect to the truth-teller problems would be "Suppose first that [person A] is a truth-teller. Next, suppose that [person a] is a liar."

Categorical First assertor's status				Compound First assertor's status			
Truth-teller "A"		Liar "~A"		Truth-teller "A"		Liar "~A"	
Second assertor's status: Truth-teller "B"							
A:	A <—> B	~A:	A <—> B	A:	A <—> B	~A:	A <—> B
B:	A	B:	~A	B:	~A >-< B	B:	A >-< B
A:	~A <—>~B	~A:	~A <—>~B	A:	~A <—>~B	~A:	~A <—>~B
B:	A	B:	~A	B:	~A >-< B	B:	A >-< B
A:	A >-<~B	~A:	A >-<~B	A:	A <—> ~B	~A:	A <—> ~B
B:	A	B:	~A	B:	A >-< ~B	B:	~A >-< ~B
Second assertor's status: Liar "~B"							
A:	A >-< B	~A:	A >-< B	A:	A <—> ~B	~A:	A <—> ~B
~B:	~A	~B:	A	~B:	~A >-< B	~B:	A >-<~ B
A:	A >-< B	~A:	A >-< B	A:	~A <—>~B	~A:	~A <—>~B
~B:	~A	~B:	A	~B:	A >-< ~B	~B:	~A >-< ~B
A:	A <—> B	~A:	A <—> B	A:	A <—>~B	~A:	A <—>~B
~B:	A	~B:	~A	~B:	A >-< ~B	~B:	~A >-< ~B

Received March 1997
Revision received December 1997
Accepted December 1997