

Only Reasoning

P. N. JOHNSON-LAIRD AND RUTH M. J. BYRNE

MRC Applied Psychology Unit, Cambridge CB2 2EF, England

Four experiments were carried out to investigate how people reason with "only" as a quantifier. An assertion such as "only artists are beekeepers" has the same truth conditions as "all beekeepers are artists," but we argue that it makes explicit both the relation between the two sets and the relation between their complements, i.e., an individual who is not an artist is not a beekeeper. Experiment 1 confirmed our expectation that this additional complexity would lead subjects to draw fewer logically correct conclusions from pairs of premises containing "only" than from equivalent pairs containing "all." We outline a putative representation of "only" in terms of a theory of reasoning based on mental models. Experiment 2 confirmed this theory's predictions about the most frequent errors and the relative difficulty of different sorts of inferences from a premise containing "only" and another premise in a mood based on "all," "some," "no," or "some—not." Experiment 3 corroborated the prediction that *modus tollens* would occur more often with an "only" premise than with an "all" premise, because of the former's explicit representation of the negative relation. Experiment 4 showed that the presence or absence of a definite article in the quantified noun phrase, e.g., "only the artists are beekeepers" had no marked effect on the interpretation of premises. © 1989 Academic Press, Inc.

The English word "only" is striking because it can be used as a sentential connective, as in

The game will be cancelled only if there's a frost
I'd leave now only there's no bus

as a quantifier

Only numerate people are mathematicians

or as a modifier of other quantifiers

Only some of the artists are beekeepers.

Its ubiquity has led one linguist to describe it as "a genuinely new and exciting quantifier" (Keenan, 1971).

Psychological studies of "only" have investigated its role as a sentential connective acting in concert with "if," and have established two principal phenomena. First, although the truth conditions of statements of the form *p only if q* and *if p then q* are

identical, there is a difference in emphasis, if not meaning. For example, it would be odd to paraphrase the assertion

If John stays sober then he keeps on his diet
as
John stays sober only if he keeps on his diet.

Both assertions would be false were John to stay sober but fail to keep to his diet, but the first assertion makes explicit that staying sober enables John to keep on his diet, whereas the second makes explicit that not keeping to his diet causes John not to stay sober. Wason and Johnson-Laird (1972, p. 74) found that subjects tended to adopt as their hypothesis *John stays sober* when they had to reason hypothetically from the first assertion but they tended to adopt *John keeps on his diet* when reasoning hypothetically from the second assertion (see also Rips & Marcus, 1977). This finding is consistent with the observation that a conditional is easier to understand when its antecedent refers to an event occurring prior to the event referred to by the consequent,

Reprint requests should be addressed to Dr. P. N. Johnson-Laird, MRC Applied Psychology Unit, 15 Chaucer Rd., Cambridge CB2 2EF, England.

whereas the opposite order is preferable for an "only if" assertion (cf. Evans & Newstead, 1977; Cheng & Holyoak, 1985).

A second, and related, phenomenon concerns patterns of inference. With a conditional premise, ordinary individuals are able to make a modus ponens inference (cf. Byrne, 1989)

If p then q

p

therefore, q

more readily than they are able to make a modus tollens inference

If p then q

not- q

therefore, not- p .

But, this difference disappears when the inferences are based on an "only if" premise (see Evans, 1977; Evans & Beck, 1981; Braine, 1978; Roberge, 1978).

How are these two phenomena to be explained? One view, put forward by Evans (1977) and Roberge (1978), is that a conditional, such as

If John stays sober then he keeps on his diet

asserts that the antecedent is sufficient for the consequent, whereas

John stays sober only if he keeps on his diet

asserts that the consequent is necessary for the antecedent. Evans and Beck (1981) add a further proposal: "the use of the word 'if' directs attention to the proposition which it modifies, irrespective of the presence of the logically critical word 'only'." Hence, there is a "directionality effect": forward inferences from the antecedent to the consequent are preferred with conditional assertions, but backward inferences from consequent to antecedent are preferred with "only if" assertions. Unfortunately, until we can explain why it is that "if" directs attention to the proposition that it modifies, and why it is that people ignore the force of "only," this proposal goes only

a little beyond a succinct description of the phenomena.

According to Braine (1978), a modus ponens inference with a conditional is easy because people are equipped with a formal rule of inference corresponding to it. A modus tollens inference, however, has to be made by deriving a *reductio ad absurdum* in the following chain of argument:

1. not- q [premise]
2. if p then q [premise]
3. hypothesis: p [hypothetical assumption]
4. q [modus ponens from 2 and 3]
5. q and not- q [conjunction of 4 and 1]
6. therefore not- p [3 led to the self-contradiction in 5].

Braine suggests that "only" functions like a double negation, and that "if" introduces a bias in direction: it carries us from information about the antecedent to information about the consequent. Hence, when "only" and "if" are combined in

p only if q

their effect, he says, is equivalent to

not- p if other than q

which is very similar to

if not- q then not- p .

The application of modus ponens to this assertion yields

not- q

if not- q then not- p

therefore not- p .

This inference is equivalent to modus tollens with p only if q . But, the inference is made in a single step without the need for a *reductio ad absurdum*. Thus, the theory explains the relative ease of modus tollens with "only if" premises.

There are two potential problems with this explanation. First, if modus tollens is easy with "only if" assertions because they

are treated as equivalent to "if not- q then not- p ," then modus ponens ought to become difficult because it would now depend on a *reductio ad absurdum*. In fact, modus ponens does not become particularly difficult (see Evans, 1977; Evans & Beck, 1981). Second, the theory merely provides a paraphrase of "only if"; it offers no account of how its meaning is mentally represented.

Our aims in the present paper are to report the results of some experiments using "only" as a quantifier, and, in the light of these findings, to propose a theory of the mental representation of the meaning of "only"—a theory that in principle makes sense of its use both as a quantifier and as a connective.

EXPERIMENT 1

Although it is not immediately obvious, as assertion such as "only criminals are psychopaths" is equivalent in meaning to "all psychopaths are criminals," i.e., statements of the form *only q 's are p 's* have the same truth conditions as statements of the form *all p 's are q 's*. However, the mental representation of an assertion containing "only" is likely to be more complex than the representation of the equivalent assertion containing "all." When you are told "only criminals are psychopaths," you immediately grasp that some criminals are psychopaths and that anyone who is not a criminal is not a psychopath. Keenan (1971) similarly argues that the first of these propositions is presupposed, and that the second is asserted by the "only" statement. (Since the presupposition can be denied, e.g., "but there aren't any criminals," some theorists argue that the proposition is merely an implicature, cf. Wilson, 1975. We shall not be concerned with the precise status of the proposition and will assume merely that in the normal interpretation of the sentence, people readily recover the information that some criminals are psychopaths.) There could also be criminals who are not psychopaths, although this possibil-

ity may not be immediately obvious. In contrast, when you are told "all psychopaths are criminals," you immediately grasp that fact, and perhaps that there may be criminals who are not psychopaths, but the negative relation that is so salient in the case of the "only" assertion is much less obvious. Hence, according to this hypothesis, the initial representation of "all" assertions is simpler than the initial representation of "only" assertions: the latter contain a negative component lacking from the former.

Our first experiment was designed to test this hypothesis. We predicted that subjects would find it easier, both in terms of latency and accuracy, to reason from "all" premises, such as

All of the bookkeepers are authors
All of the cyclists are bookkeepers

than to reason from the equivalent "only" premises, such as

Only the authors are bookkeepers
Only the bookkeepers are cyclists.

Method

Materials and design. Pairs of singly-quantified premises (i.e. "syllogisms") can be arranged in four distinct "figures," which depend on the arrangement of the terms in the two premises:

$A-B$	$B-A$	$A-B$	$B-A$
$B-C$	$C-B$	$C-B$	$B-C$
1	2	3	4

Thus, the previous example of an "only" syllogism is in the first of these figures. For each of the four figures, we constructed syllogisms containing the quantifier "only" in both premises, and syllogisms containing the quantifier "all" in both premises. Six of the resulting eight syllogisms have valid conclusions interrelating the end terms, and two do not, i.e., figure 4 for the "only" problem and figure 3 for the "all" problem do not yield valid conclusions. Sixteen filler items were also constructed consisting of

the four figures containing the quantifier "only" in one premise and a premise in one of the four orthodox moods (*A, I, E, O*). The lexical content of the problems was derived from triplets of nouns referring to hobbies and professions, which were selected so that subjects were unlikely to hold any strong *a priori* views about their relationships. The triplets were randomly assigned twice to the different types of syllogism in order to create two sets of experimental materials.

The subjects acted as their own controls and carried out the task for all of the problems, which were presented in a different random order to each of them. They were assigned at random to one of the two sets of materials.

Procedure. The subjects were tested individually. Their task was to construct their own conclusion, if possible, interrelating the people referred to in the end terms—a point that was explained by way of an example. Otherwise, they were to respond that there was no relation between these people. Each problem was printed on a separate page. The subjects were asked to read the sentences aloud and were allowed to peruse the page for as long as they liked. They were instructed to make their responses only when they were certain about them. Their responses were recorded on a cassette recorder, which ran uninterruptedly throughout the session. The latencies of the responses were measured from the time when the subject finished reading the sentences aloud to the point of commencing the oral response.

Subjects. Twenty-eight female subjects

from the subject pool of the MRC Applied Psychology Unit were paid £3 per hour to participate in the experiment, which lasted for about 20 min. The subjects' ages ranged from 19 to 59 years. We eliminated four subjects prior to the analysis of the data: two of them had failed to grasp the nature of the task, one had received tuition in logic, and one had received a series of trials that inadvertently omitted a problem.

Results and Discussion

The percentages of correct responses are shown in Table 1. The subjects performed reliably better than chance, but nevertheless rather poorly in comparison with other groups who have been tested on the "all" problems (cf. Johnson-Laird & Steedman, 1978). Since they were drawn from a wider sample than university students, these results may be more representative of the population at large. More than half of the subjects' responses retained the middle term in the conclusion, e.g.,

All of the artists are beekeepers and chemists

but we included such conclusions among the "correct" responses provided that they established a valid relation between the end terms. As we expected, there were significantly more correct responses to "all" problems than to "only" problems (Wilcoxon's $T = 12$, $N = 17$, $p < 0.005$, one-tailed). This result is corroborated by the latencies of the correct responses. After we eliminated those responses more than two standard deviations from the means, the mean latency to respond correctly to

TABLE 1
THE PERCENTAGES OF CORRECT RESPONSES FOR THE SYLLOGISMS OF EXPERIMENT 1

Mood	Figure				Means
	A-B B-C	B-A C-B	A-B C-B	B-A B-C	
Only-only	29	54	21	0	26
All-all	83	54	4	42	46

Note. Each percentage is based on the responses of 24 subjects to a single pair of premises. The "only" problem in figure 4 and the "all" problem in figure 3 have no valid conclusions.

"only" problems was 9.47 s, and the mean latency to respond correctly to "all" problems was 5.89 s. This difference was significant ($F(1,15) = 10.62, p < 0.01$). The subjects were evidently reluctant or unable to respond that there was no valid conclusion (see figure 3 for "all," and figure 4 for "only," in Table 1). This phenomenon has been reported in other studies (see, e.g., Rumin, Connell, & Braine, 1983). Their performance on problems that contained "only" in Fig. 1 was particularly poor (29% correct)—most of their errors on these problems were responses of "all *a*'s are *c*'s" (69%), which suggests both a reluctance to use the implicitly negative quantifier and a strong figural bias towards conclusions of the form *a*—*c* (as in studies of conventional syllogisms, e.g., Johnson-Laird & Bara, 1984).

The experiment confirmed that problems based on "only" tend to be more difficult than those based on "all," but there is clearly an effect of whether or not there is a valid conclusion that appears to override this basic phenomenon. We carried out a second experiment in which we investigated a wide variety of syllogisms containing an "only" premise in order to test the difficulty of various sorts of problems.

EXPERIMENT 2

Performance with orthodox syllogisms has been successfully accounted for by a theory of reasoning based on mental models (e.g., Johnson-Laird & Bara, 1984). The theory adopts a uniform representation that can be used for any sort of assertion, singly or multiply quantified (Johnson-Lain, Byrne, & Tabossi, 1989): finite sets of individuals are represented by finite sets of mental tokens, and further tokens represent properties and relations among individuals (see Johnson-Laird, 1983), and so we can specify the sort of model people should construct for premises that contain "only."

Given that there is no doubt about the existence of mathematicians, the assertion

All mathematicians are numerate

can be represented by a model with the following structure:

mathematician = numerate
 mathematician = numerate
 o numerate

The model represents the set of mathematicians, using an arbitrary number of tokens—two in this case—and the set of numerate individuals, again using an arbitrary number. To capture the content of the premise, each token representing a mathematician is tagged to indicate that the same individual is also numerate. There may, of course, be numerate individuals who are not mathematicians, and so one token representing such individuals has been included in the model. Since they may, or may not, exist in the domain of discourse, they are tagged (with an "o") to indicate that they are optional.

Our fundamental assumption about the meaning of "only" is that, unlike "all," it calls for a model that makes virtually all the information about the sets and their complements explicit:

$q = p$
 $q = p$
 . . .
 $\neg q = \neg p$
 $\neg q = \neg p$.

The three dots represent a region of uncertainty: could there be an instance of *q* that is not-*p*? In fact, as we mentioned earlier, the correct interpretation of "only" calls for just such a possibility, and so the complete model should be

$q = p$
 $q = p$
 $oq = \neg p$
 $\neg q = \neg p$
 $\neg q = \neg p$

Tags for representing an item as optional and for representing negation and other abstract notions, such as disjunction, are propositional-like in that they do not correspond directly to anything in the physical world. Although there are ways of avoiding

their use in models, e.g., by maintaining an independent linguistic representation of the premises (see Inder, 1987), there are reasons to believe that people do mentally represent abstract notions in models (see Johnson-Laird, 1983, Chap. 15, for some of the arguments). The use of tags to represent abstract concepts, such as negation, has also been advocated by Polk and Newell (1988) in their model-based theory of syllogisms (see also Erickson, 1974, 1978; Guyote & Sternberg, 1981). Some of our experimental evidence to be reported later also corroborates the representation of negative elements directly in models. Granted the use of tags, the procedures for constructing, manipulating, and evaluating models must be equipped with the appropriate semantics for them.

The theory assumes that the premises of a syllogism are integrated into a single model. For example, the premises

Some authors are bookkeepers
Some bookkeepers are cyclists

yield the initial model

author = bookkeeper = cyclist
author = bookkeeper = cyclist
o author *o* bookkeeper *o* cyclist

A procedure that interprets models to establish relations that are not asserted in the premises yields the conclusion

Some authors are cyclists.

The process of formulating conclusions from models should not be confused with the visual inspection of the diagram above (see the details of the program described by Johnson-Laird & Bara, 1984). In essence, it depends on establishing the nature of the link from one end item to a middle item, and from this middle item to an item at the other end of the model. This process has to be repeated for all relevant items, and the overall results amalgamated.

To test the validity of an inference, it is necessary to search for alternative models of the premises that falsify the conclusion. Since the models are finite, the search can

in principle be exhaustive. In practice, however, ordinary individuals lack any simple deterministic algorithm for searching for counterexamples, and we assume that they rely on a heuristic search procedure (cf. Newell & Simon, 1972).

In the case above, there is an alternative model that falsifies the conclusion

author = bookkeeper
author = bookkeeper
bookkeeper = cyclist
bookkeeper = cyclist
o author *o* bookkeeper *o* cyclist

and so the conclusion is not valid, and the two models taken together show that there is no valid conclusion interrelating the authors and cyclists.

The derivation of the models for each problem is a complex business, and we will not recapitulate the theory in detail here since it has been described elsewhere (see Johnson-Laird, 1983, Chap. 5; Johnson-Laird & Bara, 1984). Our principal concern is to which of the three categories a problem belongs. Consider the premise

All of the *a*'s are *b*'s

which supports a model of the form

a = *b*
a = *b*
ob
ob

Now, suppose that there is a second premise

All of the *b*'s are *c*'s.

In adding its information to the model, there is no choice about what to do: wherever there is a *b*, it must be tagged as a *c*. Hence, it is a one-model problem. But, suppose instead that the second premise is

Some of the *b*'s are *c*'s.

Now, there is a choice about which *b*'s to represent as *c*'s: the *b*'s that are *a*'s, or the *b*'s that are not *a*'s, or some mixture of the

two? Whenever there is such a choice, more than one model of the premises is possible. The precise number of alternative models depends on the particular procedures that are assumed to be used in constructing models: Johnson-Laird and Bara (1984) described two different sets of procedures that produced differing numbers of alternative models. What is common to both sets of procedures, however, is the set of one-model problems because there is no choice about how to add the information from the second premise. We have therefore made only one crucial distinction in the present experiment: does the inference call for one model yielding a valid conclusion, for multiple models yielding a valid conclusion, or for multiple models not yielding a valid conclusion?

Since the theory is already equipped with general procedures for comprehension and reasoning, it yields predictions just as soon as a semantics for a new quantifier is provided. Because "only" premises correspond in meaning to the converse of "all" premises, the status of syllogisms based on "only" premises is easy to derive. But, since *only a's are b's* has a more explicit initial representation than *all b's are a's*, we can predict that it will lead to greater difficulties in performance.

In the experiment, we examined the set of 36 syllogisms that can be constructed by combining one premise containing "only" with a premise based on either "only" or one of the four standard moods ("all," "some," "none," and "some—not"). These problems enabled us to assess whether performance was affected by the need to construct multiple models in order to make the right response for the right reason. The theory predicts that the task should be harder when more than one model has to be constructed. It also predicts that the typical errors will be conclusions that are based on only some of the possible models of the premises.

Method

Materials and design. Each problem con-

tained two premises drawn from a set of five possible moods ("only," "all," "some," "none," and "some—not"). The premises were combined so that the first or second premise (or both) contained "only," yielding problems in nine moods, which were assigned to each of the four figures. The resulting 36 problems had a lexical content deriving from triplets of nouns referring to professions and hobbies. The subjects acted as their own controls and were tested with the complete set of 36 problems presented in a different random order to each of them. Half the subjects received one set of lexical materials, and the other half received the other set. Of the 36 problems, 17 were one-model with a valid conclusion, 8 were multiple-model with a valid conclusion, and 11 were multiple-model with no valid conclusion.

Procedure. The subjects were tested individually. They were given the instructions used in Experiment 1 except that they were told to write their conclusions on a separate sheet of paper for each syllogism. The latencies of the responses were not recorded in this experiment.

Subjects. Twenty-two female subjects from the MRC Applied Psychology Unit subject pool were paid £3 per hour to participate in the experiment, which lasted for about half an hour. Their ages ranged from 23 to 71 years. None had received tuition in logic.

Results and Discussion

The detailed results for each of the 36 problems are presented in Tables 4 to 7 in the Appendix. The percentages of correct conclusions were reliably affected by whether or not it was necessary to construct more than one model. There were 55% correct conclusions for one-model problem with valid conclusions, 15% correct for multiple-model problems with valid conclusions, and only 3% correct for multiple-model problems with no valid conclusions. The theory does not make any prediction about the difference between the multiple-model problems with and without

valid conclusions, which in any case would be confounded by the qualitative difference in response between them. The difference between one-model and multiple-model problems overall is highly significant: every single subject showed the predicted effect ($p = 0.5^{22}$). Likewise, every single subject performed more accurately with one-model problems than with multiple-model problems with valid conclusions ($p = 0.5^{22}$), and more accurately with one-model problems than with multiple-model problems with no valid conclusions ($p = 0.5^{22}$).

The overall level of accuracy was again poor in comparison to other groups that have been tested with conventional syllogistic problems. In general, as in the previous experiment, the subjects seemed to be reluctant to respond that there was no valid conclusion. Hence, performance on the problems that did not support valid conclusions interrelating the end terms was very poor. There is also considerable variety in the conclusions that subjects draw—a phenomenon invariably observed in studies in which subjects are asked to frame their own conclusions. According to the theory, the ultimate source of such variability is the lack of a simple deterministic procedure for syllogistic inference. Reasoners are free to construct their model starting with the first or the second premise; they are free to interpret models starting with either set of end terms; they are free to search, or not to search, for alternative models of the premises; and, lacking a deterministic algorithm for search, the process itself may be systematic or haphazard. What the theory does predict, however, is the relative difficulty of different problems, and the nature of the most frequent sorts of errors; i.e., they will be based on only a subset of the possible models of the premises.

The order of the terms in the conclusions was influenced by the figure of the problem. There were 85% conclusions in a forward direction ($A-C$) for the first figure and 58% conclusions in a backward direction ($C-A$) for the second figure (18 out of

the 22 subjects showed this effect with one tie, Sign test, $p < 0.001$). This “figural effect” has been observed in previous experiments where subjects were free to frame conclusions in their own words (see Johnson-Laird & Steedman, 1978; Johnson-Laird & Bara, 1984; Johnson-Laird, Oakhill, & Bull, 1986).

The mental model theory predicts that errors occur as a result of a failure to examine all possible models of a pair of premises. The majority of the errors (56%) could indeed be accounted for in this way. A further 23% could be explained in terms of two types of error previously identified by Johnson-Laird and Bara (1984): 13% as a result of omitting optional tokens, and 10% as a result of responding “Some of the A are C ,” where the predicted response was “Some of the A are not C .” The latter error has been called a Gricean response in earlier accounts since it presumably derives from the pragmatic conventions governing language (see Grice, 1975), and, in particular, from the fact that the actual conclusion is an implicature of the predicted conclusion. Another 8% of errors were conclusions that included modal verbs, such as “Some of the A may be C ,” which, though sensible, lie outside the scope of the current theory of syllogisms, and a further 6% of errors were conclusions that omitted one or other end term—presumably as a result of memory lapses. Only 6% of the subjects’ responses were wholly inexplicable.

One final observation casts doubt on the generality of the “atmosphere” effect, i.e., the alleged tendency to draw conclusions that match the mood of the premises (Woodworth & Sells, 1935; Revlis, 1975). Where the two premises both contained “only,” there were merely 16% of conclusions containing this same quantifier; where one of the premises contained “only,” just 2% of the conclusions contained it.

EXPERIMENT 3

The aim of this experiment was to exam-

ine the relative difficulty of modus ponens and modus tollens depending on whether the inference was based on an "all" premise or an "only" premise. We asked subjects what followed from eight sorts of premise pairs. Four pairs contained an "only" premise, and four contained an "all" premise:

1. Modus ponens:
 Only the beekeepers are artists.
 Lisa is an artist.
 and
 All the artists are beekeepers.
 Lisa is an artist.
2. Modus tollens:
 Only the beekeepers are artists.
 Lisa is not a beekeeper.
 and
 All the artists are beekeepers.
 Lisa is not a beekeeper.
3. Denying the antecedent:
 Only the beekeepers are artists.
 Lisa is not an artist.
 and
 All the artists are beekeepers.
 Lisa is not an artist.
4. Affirming the consequent:
 Only the beekeepers are artists.
 Lisa is a beekeeper.
 and
 All the artists are beekeepers.
 Lisa is a beekeeper.

From the models based on the semantics of "only," we can predict that the difference in difficulty between modus ponens and modus tollens should be significantly reduced when subjects reason from "only" premises in comparison with "all" premises. The same prediction follows, of course, from the previous studies of "only" as a connective (see the Introduction). We suggest that the previous results arise because the initial model of "only" makes explicit information about the sets and their complements. We can also predict that the tendency to deny the antecedent and affirm the consequent should be greater for "only" premises than for "all" pre-

mises, because reasoners should be more likely to omit or to drop the optional element from the more complex model needed for "only" premises. These predictions are contrary to what one would expect if subjects represent "only" premises solely by the negative relation: with "all" premises modus ponens should be easier than modus tollens, but the difference in difficulty should switch round with "only" premises. Likewise, there is no reason to suppose that denial of the antecedent and affirmation of the consequent should be more likely to occur with "only" premises than with "all" premises.

Method

Materials and design. There were eight sorts of problems: modus ponens, modus tollens, denial of antecedent, and affirmation of consequent, based on a first premise containing either "only" or "all." Each subject carried out two instances of each sort of problem, making a total of 16 inferences in all.

The lexical content of the problems referred to hobbies and professions, and the specific individuals referred to in the second premises were identified by proper names: Half of the names were female and half were male, and the hobbies and professions were neutral with respect to gender. The problems were randomly assigned twice to 16 different sets of lexical materials, and the subjects were assigned at random to one of the two sets of materials. The problems were presented in a different random order to each subject. The subjects acted as their own controls and constructed a conclusion in their own words to each problem.

Procedure. The subjects were tested individually. They were given similar instructions to those used in the previous experiments, and the task was explained by way of an example. The subjects had to construct a conclusion, if possible, relating the person referred to in the second premise to the hobbies or professions referred to in the

first premise. Otherwise, they were to respond that there was not enough information in the premises to draw a definite conclusion. Each problem was printed on a separate card. The subjects read the sentences aloud and wrote their responses on separate pages.

Subjects. Twenty-six subjects (18 female and 8 male) from the Applied Psychology Unit subject panel were paid £3 per hour to participate in the experiment, which lasted for about half an hour. Two subjects were replaced during the experiment because they had failed to grasp the nature of the task. The resulting 26 subjects were between 21 and 63 years of age, and none had received tuition in logic.

Results and Discussion

Table 2 shows the percentages of the four sorts of inference as a function of whether there was an "all" or "only" premise. The results corroborated our predictions. The difference in difficulty between *modus ponens* and *modus tollens* was reliably reduced in the case of "only" premises in comparison with "all" premises, and this interaction was significant (Wilcoxon's $T = 10$, $N = 11$, $p < 0.025$). Moreover, the interaction was not produced by a decline in *modus ponens* from one sort of premise to the other: there was no reliable difference in the percentages for "all" and "only" premises (Wilcoxon's $T = 3$, $N = 5$, $p > 0.05$). There was, however, a significant improvement in *modus tollens* with "only" premises in comparison with "all" premises (Wilcoxon's $T = 3$, $N = 7$, $p < 0.05$).

The subjects also made more fallacies from "only" premises (78%) than from

"all" premises (36%), and this predicted difference was reliable (Wilcoxon's $T = 2.5$, $N = 19$, $p < 0.001$). The difference was reliable both for the denial of the antecedent (Wilcoxon's $T = 4$, $N = 18$, $p < 0.001$), and for the affirmation of the consequent (Wilcoxon's $T = 2$, $N = 12$, $p < 0.005$).

The results corroborated our predictions. The models that subjects construct of "only" premises do appear to contain information about the members of each set and their complements. That is, the models do represent negative information of the sort that we propose to capture by the use of tags. In other words, if subjects are constructing models, then these models contain abstract propositional-like tokens. There is, however, an alternative hypothesis about the interpretation of assertions of the form "Only the artists are beekeepers." The presence of the definite article in the quantified noun phrase may lead ordinary individuals to make an interpretation equivalent to: all and only the artists are beekeepers. Previous linguistic analyses of the assertion lacking the definite article, e.g., "only artists are beekeepers," have suggested that it implies that at least some of the artists are beekeepers (see Keenan, 1971). The presence of the definite article may imply that the artists as a whole are beekeepers (A. J. Marcel, personal communication). We have no clear intuitions about this possibility. We believe that it is perfectly feasible to assert, for example, "only the Republicans are monetarists" without being committed to the view that all the Republicans are monetarists. But, other cases may well differ. We have been unable to find any discussion of such sen-

TABLE 2
THE PERCENTAGES OF THE FOUR SORTS OF INFERENCE MADE FROM "ALL" PREMISES AND "ONLY"
PREMISES IN EXPERIMENT 3

	Modus ponens	Modus tollens	Affirmation of the consequent	Denial of the antecedent
All	96	73	38	33
Only	90	86	75	85

tences in the literature, and linguists whom we have consulted have also expressed uncertainty (G. Gazdar, personal communication). The difference between the two interpretations is not critical for the model-based theory, which in either case makes the same predictions. All that is at stake is whether or not in the example above there could be Republicans who are not monetarists, and hence whether the denial of the antecedent and the affirmation of the consequent are valid or invalid inferences. Nevertheless, we have carried out a simple experiment in order to collect the intuitions of ordinary individuals about this aspect of the meaning of "only."

EXPERIMENT 4

The purpose of this experiment was to gather a set of judgements about whether or not assertions containing "only" are judged to imply assertions containing "all," and vice versa. We examined eight immediate inferences from assertions with one quantifier to assertions with the other quantifier. Four of these inferences concerned quantifiers with no accompanying definite article:

1. *Only* artists are beekeepers.
Therefore, *all* artists are beekeepers?
2. *All* artists are beekeepers.
Therefore, *only* artists are beekeepers?
3. *Only* artists are beekeepers.
Therefore, *all* beekeepers are artists?
4. *All* artists are beekeepers.
Therefore, *only* beekeepers are artists?

The other four inferences had the same forms but the definite article was included in the quantified noun phrases of both premisses and conclusions.

"Only artists are beekeepers" is equivalent to "all beekeepers are artists," and so the inferences from one to the other should be judged as valid. When the order of the two terms is held constant, however, as in the pair "only artists are beekeepers" and

"all artists are beekeepers," the two sentences are not equivalent, and so the inferences from one to the other should be judged as invalid. In the case of the assertions containing the definite article in the quantified noun phrases, we expected the same pattern of results. Clearly if "only the artists are beekeepers" were taken to mean that all and only the artists are beekeepers, then all four patterns of inference should be judged as valid.

Method

Design and materials. The subjects acted as their own controls and judged all eight inferences. Half the subjects judged first the four inferences based on quantifiers accompanied by definite articles, and then the four inferences based on quantifiers with no accompanying definite articles; and half the subjects made the judgments, in the opposite order. The order of the inferences within each half of the experiment was randomized for each subject. There were eight different sets of lexical materials so that no subject encountered the same lexical items in more than one inference: the materials were assigned to the different inferences in eight different ways by rotating them over the inferences. Each set of materials consisted of a pair of terms that referred respectively to a profession and to a hobby.

Procedure. The materials were made up into booklets and the subjects were tested in a single group. Each problem was printed on a separate page, and an example of such a page is as follows;

Only the electricians are rowers
Does it follow that:
All the rowers are electricians.
Yes. . . . No. . . . Can't tell. . . .

In order to avoid questions of factual truth and falsity for the assertions without the definite articles, such as "all rowers are electricians," the printed instructions told the subjects to imagine that the sentences referred to a community of people. The subjects' task was to decide whether the

second sentence was implied by the first, and to indicate their response by ticking one of the three options shown on each page. The subjects worked through the booklets at their own pace.

Subjects. Nine female subjects from the Applied Psychology Unit subject pool were paid £3.60 per hour to participate in the experiment, which lasted for about 15 min. Their ages ranged from 22 to 57 years. None of the subjects had received tuition in logic.

Results and Discussion

Table 3 presents the percentages of judgements of validity of the eight sorts of inference: the balance of the responses were judgements of invalidity with the exception of a single "can't tell" response. Evidently, the subjects were not treating *only the a's are b's* as meaning that *all and only the a's are b's*, since such in interpretation calls for all four patterns of inference to be judged as valid. Indeed, as the table shows, the presence or absence of the definite article had no striking effect on performance: none of the four inferences differed reliably (in each case, Wilcoxon's $T = 2$, $N = 3$, $p > 0.05$).

One unexpected finding was that although the majority of subjects correctly evaluated three of the four inferences, one inference conspicuously misled them, namely, the argument from a premise of the form:

All A are B

to a conclusion of the form:

Only B are A

regardless of whether or not a definite article occurred in the quantified noun phrases. These inferences were judged as valid on only 17% of occasions, whereas the inferences in the opposite order were judged as valid on 83% of occasions (Sign test, $N = 8$, $p < 0.005$).

The experiment suggests that the presence or absence of the definite article within "only" and "all" noun phrases has no major effect on the judgement of immediate inferences. The only surprise in the results was the failure of the subjects to evaluate the inferences from *all a's are b's* to *only b's are a's* as valid. A possible, though *post hoc*, explanation of this phenomenon can be derived from the account based on mental models. An assertion of the form "only b's are a's" calls for the following initial model:

$$\begin{aligned} b &= a \\ b &= a \\ ob &= \neg a \\ \neg b &= \neg a \\ \neg b &= \neg a \end{aligned}$$

in which the information that all a's are b's is available. However, the model for all a's are b's is of the sort

$$\begin{aligned} a &= b \\ a &= b \\ ob \end{aligned}$$

which does not make explicit that entities that are not b's are not a's. This information is, of course, necessary if one is to draw the conclusion that only b's are a's. In

TABLE 3
THE PERCENTAGES OF JUDGEMENTS OF VALIDITY FOR THE EIGHT IMMEDIATE INFERENCES
IN EXPERIMENT 4

	Only A are B All B are A	All A are B Only B are A	Only A are B All A are B	All A are B Only A are B
Definite article	78	25	22	0
No definite article	89	11	22	0

Note. Inferences taken as a function of whether or not the quantified noun phrases in premise and conclusion contained definite articles.

short, the phenomenon is compatible with the asymmetry between the two models: the model for "only" makes explicit the information for "all," but not vice versa.

There is one remaining loose end as Martin Braine (personal communication) has pointed out. In Experiment 3, the subjects tended to affirm the consequent and to deny the antecedent with the "only" premises. We predicted this phenomenon on the grounds that people would tend to omit the optional item in their initial representations of these assertions. Why, then, do they correctly reject the inferences from "only *a*'s are *b*'s" to "all *a*'s are *b*'s", and vice versa? One possible explanation is that when the order of the two terms is the same, the subjects merely consider whether or not the two quantifiers are synonymous. Since "all" plainly does not mean the same as "only," they conclude that the inference

is invalid. This strategy was not available to the subjects in Experiment 3, who were presented with one "only" assertion and one categorical assertion and were asked to state what followed from them.

GENERAL DISCUSSION

Our results show that people are able to reason with premises of the form *only the p's are q's*, though, as Experiment 1 established, they generally find such premises harder to cope with than the logically equivalent premises *all the q's are p's*. We predicted this difference on the grounds that the semantics for "only" calls for an explicit representation of more information than the semantics of "all": a model of the "all" statement represents each *q* as a *p*; whereas a model of the "only" statement represents in addition each not-*p* as not-*q*. The more complete representation will ac-

TABLE 4
FIGURE: A-B, B-C

Premise		First									
Second	Only		All		Some		No		Some—not		
Only	ONLY A-C	5	All A-C	12	Some A-C	16	NO A-C	14	SOME A-NOT C	3	
	ALL C-A	2	All C-A	3	Some C-A	3	NO C-A	1	Some A-C	10	
	SOME A-C	2	NVC	1	NVC	0	?Some A-C	2			
	?All A-C	8									
	One-model		Multiple-model	Multiple-model	Multiple-model	One-model		Multiple-model			
All	SOME A-C	3									
	SOME C-A	1									
	?All A-C	9									
	?All C-A	4									
	?Only A-C	2									
Some	One-model										
	SOME A-C	14									
	SOME C-A	3									
	?All A-C	2									
No	One-model										
	No A-C	16									
	SOME A-NOT C	0									
Some—not	Multiple-model										
	SOME A-NOT C	10									
	Some A-C	7									
	Multiple-model										

Note. Conclusions are drawn from the syllogisms of Experiment 2.

cordingly preoccupy working memory to a greater extent, and so will lead to a reduction in inferential accuracy. However, the more complete representation is advantageous for simple inferences. It enhances performance with modus tollens and so the difference in difficulty between it and modus ponens is smaller for "only" premises than for "all" premises. But, the denial of the antecedent and the affirmation of the consequent are more likely to occur with "only" premises than with "all" premises. Both these predictions were confirmed in Experiment 3.

Some theorists have argued that although people use mental models in reasoning,

these models are direct physical representations of the world and do not contain tokens corresponding to abstract notions, such as negation (see, e.g., Inder, 1987). On this account, negation occurs only in the linguistic representation of the premises, and not in models, which can represent only affirmative content. Our results, however, suggest that this claim is wrong: a model of a premise containing "only" does contain an explicit representation of negative elements, and so the modus tollens inference is easy (see Table 2). Where both approaches concur, however, is in the need for a semantics that ensures that the verbal representation, or the mental tag, is appro-

TABLE 5
FIGURE: *B-A, C-B*

Premise		First								
Second		Only	All	Some	No	Some—not				
Only	ONLY C-A	3	SOME C-A	1	SOME C-A	10	No C-A	7	Some A-not C	2
	ALL A-C	7	SOME A-C	7	SOME A-C	3	No A-C	5	SOME C-NOT A	10
	SOME C-A	3	?All C-A	7	?All A-C	3	SOME C-NOT A	0	Nvc	2
	SOME A-C	3	?All A-C	3			Nvc	3	Some A-C	3
							Some C-A	2	Some C-A	2
							Some A-C	2		
	One-model		One-model	One-model	Multiple-model	Multiple-model				
	All C-A	7								
	All A-C	4								
	NVC	0								
All	?Some C-A	2								
	?Some A-C	3								
	Multiple-model									
	Some C-A	13								
	Some A-C	2								
	NVC	1								
Some	?No C-A	2								
	Multiple-model									
	No C-A	10								
	No A-C	3								
	?Some C-A	4								
No	One-model									
	SOME C-NOT A	4								
	Nvc	2								
	Some C-A	5								
	Some A-C	4								
Some—not	?No C-A	2								
	Multiple-model									

Note. Conclusions are drawn from the syllogisms of Experiment 2.

propriately interpreted with respect to states of the world.

The model-based theory of reasoning, once it is equipped with the semantics for "only," immediately yields a set of predictions about the difficulty of syllogistic reasoning with the quantifier. In Experiment 2, we confirmed that syllogisms requiring only one model to be constructed are reliably easier than those requiring more than one model. Likewise, as we predicted, the systematic errors that occur with "only" largely correspond to conclusions that would follow from just *some* of the possible models of the premises. The salience of both negative and affirmative elements in the meaning of "only" perhaps explains subjects' reluctance to draw conclusions containing "only"—a reluctance that is

damaging to the alleged "atmosphere" effect. Where both premises contain "only," the conclusion can be expressed in either an affirmative way using "all" or in an implicitly negative way using "only." Given the greater difficulty of "only," it is not surprising that subjects prefer to express their conclusions affirmatively.

Our semantic analysis of "only" applies to its uses both as a sentential connective (in conjunction with "if") and as a simple quantifier. It can also be used to modify other quantifiers, where again it appears to have the same essential meaning. Thus, the following assertion:

Only *some* of the *p*'s are *q*'s
can be paraphrased as some of the *p*'s are *q*'s and anything that is not among this set

TABLE 6
FIGURE: A-B, C-B

Premise	First										
Second	Only		All		Some		No		Some—not		
Only	SOME A-C	4	ALL A-C	7	SOME A-C	9	No A-C	13	Some A-not C	3	
	?All A-C	10	ONLY C-A	1	SOME C-A	3	No C-A	3	NVC	0	
	?Nvc	2	SOME A-C	2	?All A-C	2	SOME C-NOT A	0	Some A-C	9	
			SOME C-A	2	?Some A-not C	2			Some C-A	3	
			?All C-A	3							
	One-model		One-model		One-model		Multiple-model		Multiple-model		
	All	ALL C-A	7								
		SOME A-C	2								
		SOME C-A	3								
		?All A-C	6								
One-model											
Some	SOME C-A	8									
	SOME A-C	8									
	One-model										
No	No C-A	7									
	No A-C	5									
	SOME A-NOT C	0									
	Nvc	4									
	Multiple-model										
Some— not	Some C-not A	2									
	NVC	3									
	Some C-A	4									
	Some A-C	5									
	?No C-A	2									
	Multiple-model										

Note. Conclusions are drawn from the syllogisms of Experiment 2.

TABLE 7
FIGURE: *B-A, B-C*

Premise	First									
	Second	Only	All	Some	No	Some—not				
Only	All <i>A-C</i>	10	ALL <i>C-A</i>	3	Some <i>C-A</i>	6	NO <i>C-A</i>	2	Some <i>C-not A</i>	7
	Only <i>A-C</i>	2	ONLY <i>A-C</i>	2	Some <i>A-C</i>	9	NO <i>A-C</i>	14	NVC	1
	NVC	0	SOME <i>A-C</i>	1	NVC	0	?Some <i>C-A</i>	2	Some <i>A-C</i>	6
	?Some <i>A-C</i>	6	?All <i>A-C</i>	11	?All <i>A-C</i>	2			Some <i>C-A</i>	3
	Multiple-model		One-model		Multiple-model		One-model		Multiple-model	
All	ALL <i>A-C</i>	10								
	SOME <i>C-A</i>	4								
	?All <i>C-A</i>	5								
	One-model									
Some	Some <i>A-C</i>	17								
	NVC	0								
	Multiple-model									
No	NO <i>A-C</i>	15								
	NO <i>C-A</i>	1								
	One-model									
Some—not	Some <i>A-not C</i>	11								
	NVC	0								
	Some <i>A-C</i>	7								
	Multiple-model									

Note. Conclusions are drawn from the syllogisms of Experiment 2.

of *p*'s is not a *q*. As many linguists have noted (see, e.g., Keenan, 1971), "only" modifies different constituents in a sentence depending on its structural role within the sentence (though contrastive stress may affect the interpretation). Consider, for example, the following cases:

Only John telephoned Mary today. (I.e., no-one else telephoned her.)

John only telephoned Mary today. (I.e., he did not interact with her in any other way.)

John telephoned only Mary today. (I.e., he did not telephone anyone else.)

John telephoned Mary only today. (I.e., he did not telephone her on any other day.)

We believe that all of these uses can be captured in the uniform semantics. One use of "only," however, appears to be rather different. The last example above can also be paraphrased as:

He telephoned her as recently as today.

This meaning seems to be idiomatic, and it is restricted to temporal expressions that can be construed as relatively recent in relation to the reference time of the utterance. Thus, the assertion "John telephoned Mary only here," means merely that he did not telephone her anywhere else; it lacks an interpretation analagous to the "as recently as" case.

Although a theory based on formal rules might be contrived to explain reasoning with "only," there is as yet no such theory, and so our results present a challenge to proponents of formal rules. Indeed, there is no such theory for conventional syllogisms that accounts satisfactorily for the relative difficulty of different inferences and the pattern of systematic errors. It may be the case, of course, that reasoning depends on both formal rules and mental models. Unfortunately, this possibility is very difficult to falsify, and so it may be prudent for psychologists to continue to pursue the alter-

native theories independently. The advantage of the model-based theory is that it already has general procedures for constructing models and for reasoning by searching for alternative models that refute putative conclusions. There is no need to alter these procedures, which carry over directly to novel quantifiers. To extend the theory to encompass a novel quantifier, it is necessary only to frame an account of the contribution made by the quantifier to the truth conditions of assertions. This account of the meaning of the quantifier will be needed in any case to explain how people grasp these truth conditions. Hence, once reasoners have acquired the meaning of "only," they can make inferences using models of the premises. Strictly speaking, they have no need to acquire specific formal rules of inference for the term.

APPENDIX

Tables 4 to 7 present the 36 pairs of premises used in Experiment 2 together with the subjects, responses. Only those responses made by two or more subjects are included, except in the case of correct responses which are reported regardless of the number of subjects making them. Each of the four tables refers to one figure and each cell in a table corresponds to a pair of premises. The correct valid conclusions are shown in capitals; conclusions depending on Gricean implicatures are italicized. Responses of "no valid conclusion" are indicated by "Nvc," and a "?" indicates a response not predicted by the theory—including those responses that could be accounted for by the omission of optional elements. The details may be found in Johnson-Laird and Bara (1984, pp. 52–59).

ACKNOWLEDGMENTS

We are indebted to Mark Keane, Patrizia Tabossi, Martin Braine, Marcel Just, and an anonymous reviewer for their insightful criticisms of an earlier draft of this paper.

REFERENCES

- BRAINE, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review*, 85, 1–21.
- BYRNE, R. M. J. (1989). Suppressing valid inferences with conditionals. *Cognition*, in press.
- CHENG, P. N., & HOLYOAK, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17, 391–416.
- ERICKSON, J. R. (1974). A set analysis theory of behaviour in formal syllogistic reasoning tasks. In R. Solso (Ed.), *Loyola symposium on cognition*, Vol. 2. Hillsdale, NJ: Erlbaum.
- ERICKSON, J. R. (1978). Research on syllogistic reasoning. In R. Revlin and R. E. Mayer (Eds.), *Human reasoning*. Washington, DC: Winston.
- EVANS, J. ST. B. T. (1977). Linguistic factors in reasoning. *Quarterly Journal of Experimental Psychology*, 29, 297–306.
- EVANS, J. ST. B. T., & BECK, M. A. (1981). Directionality and temporal factors in conditional reasoning. *Current Psychological Research*, 1, 111–120.
- EVANS, J. ST. B. T., & NEWSTEAD, S. E. (1977). Language and reasoning: A study of temporal factors. *Cognition*, 8, 265–283.
- GRICE, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics*, Vol. 3, *Speech acts*. New York: Academic Press.
- GUYOTE, M. J., & STERNBERG, R. J. (1981). A transitive-chain theory of syllogistic reasoning. *Cognitive Psychology*, 13, 461–524.
- INDER, R. (1987). Computer simulation of syllogism solving using restricted mental models. Unpublished Ph.D thesis, Edinburgh University.
- JOHNSON-LAIRD, P. N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge: Cambridge University Press.
- JOHNSON-LAIRD, P. N., & BARA, B. G. (1984). Syllogistic inference. *Cognition*, 16, 1–61.
- JOHNSON-LAIRD, P. N., & TABOSSI, P. BYRNE, R. M. J. (1989). Reasoning by model: The case of multiple quantification. *Psychological Review*, in press.
- JOHNSON-LAIRD, P. N., OAKHILL, J. V., & BULL, D. (1986). Children's syllogistic reasoning. *Quarterly Journal of Experimental Psychology A*, 38, 35–58.
- JOHNSON-LAIRD, P. N., & STEEDMAN, M. (1978). The psychology of syllogisms. *Cognitive Psychology*, 10, 64–99.
- KEENAN, E. (1971). Quantifier structures in English. *Foundations of Language*, 7, 255–284.
- NEWELL, A., & SIMON, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- POLK, T., & NEWELL, A. (1988). Modeling human syllogistic reasoning in Soar. *Proceedings of the*

- Tenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.
- REVLIS, R. (1975). Two models of syllogistic reasoning: Feature selection and conversion. *Journal of Verbal Learning and Verbal Behaviour*, 14, 180–195.
- RIPS, L. J., & MARCUS, S. L. (1977). Suppositions and the analysis of conditional sentences. In M. A. Just and P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, NJ: Erlbaum.
- ROBERGE, J. J. (1978). Linguistic and psychometric factors in propositional reasoning. *Quarterly Journal of Experimental Psychology*, 30, 705–716.
- RUMAIN, B., CONNELL, J. & BRAINE, M.D.S. (1983). Conversational comprehension processes are responsible for reasoning fallacies in children as well as adults: IF is not the biconditional. *Developmental Psychology*, 19, 471–481.
- WASON, P. C., & JOHNSON-LAIRD, P. N. (1972). *Psychology of reasoning: Structure and content*. London: Batsford; Cambridge, MA: Harvard University Press.
- WILSON, D. (1975). *Presuppositions and non-truth-conditional semantics*. New York: Academic Press.
- WOODWORTH, R. S., & SELLS, S. B. (1935). An atmosphere effect in formal syllogistic reasoning. *Journal of Experimental Psychology*, 18, 451–460.

(Received June 21, 1988)

(Revision received November 1988)