The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqja20

The effects of belief on the spontaneous production of syllogistic conclusions

J. V. Oakhill [a] & P. N. Johnson-Laird [a]

[a] MRC Perceptual and Cognitive Performance Unit, Laboratory of Experimental Psychology, University of Sussex, Brighton, U.K.

[b] MRC Applied Psychology Unit, Cambridge, U.K.

Available online: 29 May 2007

To cite this article: J. V. Oakhill & P. N. Johnson-Laird (1985): The effects of belief on the spontaneous production of syllogistic conclusions, The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 37:4, 553-569

To link to this article: http://dx.doi.org/10.1080/14640748508400919

The Quarterly Journal of Experimental Psychology (1985) **37A**, 553–569

The Effects of Belief on the Spontaneous Production of Syllogistic Conclusions

J. V. Oakhill

MRC Perceptual and Cognitive Performance Unit, Laboratory of Experimental Psychology, University of Sussex, Brighton, U.K.

P. N. Johnson-Laird

MRC Applied Psychology Unit, Cambridge, U.K.

Two experiments examined the effects of subjects' beliefs on syllogistic inference. The first experiment showed that beliefs biased the spontaneous conclusions that subjects drew for themselves. These effects were more marked for indeterminate premises (which yield no non-trivial valid conclusions) than for determinate premises (which yield valid conclusions). There was also an effect of the nature of the beliefs: conclusions that were false by definition had a bigger effect on deductions than those that were false as a matter of fact. The second experiment replicated the finding for determinate syllogisms, using problems in moods in which the status of the valid conclusion could not be altered by conversion of the premises. Beliefs accordingly appear to affect the process of reasoning rather than the interpretation of premises.

GENERAL INTRODUCTION

Do beliefs and prejudices affect the way people reason? The question is an old one, but it has yet to receive a definitive answer. It has been claimed that subjects are biased by their beliefs when they assess

Requests for reprints should be sent to J.V. Oakhill, MRC Perceptual and Cognitive Performance Unit, Laboratory of Experimental Psychology, University of Sussex, Brighton BN1 9QG, Sussex, U.K.

This research was supported by a grant from the Social Science Research Council to the authors. We would like to thank Alan Garnham, Stephen Monsell, Paul Pollard and an anonymous reviewer for many helpful comments on earlier drafts of the manuscript.

whether stated conclusions follow from syllogistic premises (e.g. Wilkins, 1928; Janis and Frick, 1943; Morgan and Morton, 1944; Lefford, 1946; Henle and Michael, 1956; Thouless, 1959; Feather, 1964; Kaufman and Goldstein, 1967). All these studies, however, are open to criticism on methodological grounds, particularly those that used a multiple-choice procedure (see Wason and Johnson-Laird, 1972; Evans, 1982). Recently, Evans, Barston and Pollard (1983) have provided better support for the hypothesis that beliefs affect reasoning. They demonstrated substantial effects in a series of experiments, using carefully controlled materials, in which on each trial the subjects had to evaluate a single given conclusion. Their subjects were more likely to accept a conclusion that accorded with their prior beliefs. This tendency was stronger if the conclusion was invalid, an asymmetry suggesting that subjects tend to accept congenial conclusions uncritically but check their evaluations of uncongenial conclusions. Yet, even if there are genuine effects of bias in evaluating given conclusions, they may not reflect any distortion in the process of deductive reasoning: subjects may sometimes choose to accept a conclusion merely because they deem it to be true, or choose to reject a conclusion because they deem it to be false. Such judgements do not necessarily call for deductive thought and might arise simply because subjects misunderstand the nature of the task. The aim of our studies was accordingly to investigate whether beliefs influence the process of reasoning when subjects draw a conclusion for themselves rather than merely evaluate a conclusion presented by the experimenter.

There are two major ways in which beliefs could affect reasoning when subjects have to derive their own conclusions. First, a belief could distort the interpretation and representation of the premises. Second, it could directly influence the inferential machinery that derives and evaluates conclusions based on the representation of the premises. Since either or both of these effects could occur, we shall look more closely at each of them.

The Representation of Premises

Many theorists have argued that the fundamental deductive machinery of human thought is logically sound, and hence that bias can affect only the interpretation of premises, not the process of deriving conclusions from them (e.g. Falmagne, 1975; Revlin and Leirer, 1978). A long-standing hypothesis is, indeed, that subjects may illicitly convert premises. In particular, they may assume that "All A are B" is equivalent to "All B are A", and that "Some A are not B" is equivalent to "Some B are not A" (Chapman and Chapman, 1959). The error is readily apparent with concrete material: "Some women are not mothers" does

not imply its converse. Matte-Blanco (1965) suggested that schizophrenics make such conversions systematically; and Revlin and his colleagues have made the stronger claim that normal subjects routinely convert premises and then go on to reason logically from their converted forms (Revlis, 1975; Revlin, Leirer, Yopp and Yopp, 1980; though Revlin in a personal communication has expressed some reservations about this theory). Thus, subjects will automatically make errors unless an illicit conversion is blocked by the content of a premise, e.g. "All dogs are animals", or unless the converted premises yield the same conclusion as the originals. The "conversion" theory predicts that the only way in which beliefs can exert their effect is by influencing people's representation of the premises. Subjects will, for example, refrain from converting the premise "All priests are religious people" if they believe its converse to be false. Revlin and Leirer report findings that support their hypothesis: subjects made significantly fewer errors on those syllogisms where conversion had no effect on the validity of conclusions than on those where it altered which conclusions were valid, and, when conversion was blocked by the content of the premises, performance improved on just those problems where conversion of the premises would have led to the wrong answer. However, conversion does not appear to be as ubiquitous as Revlin would wish to claim. Newstead and Griggs (1983) showed that, when subjects were asked to evaluate the truth or falsity of converted statements with respect to their original form, only one-third of subjects showed evidence of a belief that *All* implied its converse. There were also very large individual differences in interpretation: only some of the subjects made conversion errors. We conclude that the notion that subjects routinely convert premises remains unproven, though it seems likely that illicit conversion does occur occasionally, and that it can be enhanced or diminished by the content of the premises. In short, beliefs sometimes influence the interpretation of premises.

Ironically, Revlin and Leirer also provide good evidence for the effects of emotional or prejudicial influences on the deductive process. Even when conversion did not alter the logically valid conclusion, they found that when logic and belief pointed to different conclusions, subjects were less accurate than when logic and belief pointed to the same conclusion. They also found that this variable interacted with type of belief: accuracy deteriorated more markedly when there was a conflict between logic and a conclusion that was true by definition as opposed to one that was true merely as a matter of fact. These results refute their theory, though the authors attempt to dismiss them. Moreover, the experiments have been criticized on methodological grounds (see Evans et al., 1983, who argue that Revlin and Leirer may have *overestimated* subjects' logical ability).

The Inferential Process

Several theorists have argued that deductive errors occur as a result of irrational influences on the inferential process (Nisbett and Ross, 1980; Evans, 1982; Evans et al., 1983). Few theorists, however, seem prepared to accept that beliefs could directly influence which particular steps are taken in order to reach a conclusion. If deduction depends on the existence of formal rules of inference in the mind, then beliefs would not lead to invalid conclusions unless they could somehow alter the form of such rules or else their application. Of course, mistakes could be made in using the rules, but it is difficult to see how beliefs could exert a systematic effect of this sort. It is possible that there are improper rules of inference in the mind too, and perhaps their use is triggered by certain beliefs. This kind of hypothesis has been put forward to account for the fallacious inferences allegedly made by schizophrenics (see e.g. Von Domarus, 1944). The trouble is that it predicts too great a degree of irrationality in ordinary individuals.

Let us suppose, however, that deductive reasoning depends, not on internalized rules of inference, but on the following procedure:

1. Imagine a state of affairs in which the premises are true, i.e. construct a mental model of them.
2. Formulate a conclusion that is true in the model (or all the models that have been constructed), and that relates terms not explicitly related in the premises.
   If there is no such conclusion, respond that there is no valid conclusion.
3. If there is such a conclusion:
   Try to imagine a model of the premises in which the conclusion is false.
   If there is no such model, respond with the conclusion, which is valid.
   If there is such a model, return to step 2.

This procedure of constructing mental models of the premises and searching for models that are counterexamples to putative conclusions has been described in detail elsewhere and has been corroborated by experiments with adults and children (see e.g. Johnson-Laird and Steedman 1978; Johnson-Laird, 1983; Johnson-Laird and Bara, 1984a; Johnson-Laird, Oakhill and Bull, in press). If it is true, then beliefs could play a direct role in the process of deduction. Subjects might be biased not to search for counterexamples to congenial conclusions, and biased to search hard for counterexamples to uncongenial conclusions.

Since, according to the theory, the process of search lies at the heart of deduction, beliefs would in this way directly affect deduction.

A similar idea has been proposed by Pollard (1982), who argues that the "availability heuristic" that people seem to employ in making probabilistic judgements (Tversky and Kahneman, 1973) also affects deductive reasoning. This heuristic leads subjects to accept conclusions that are readily available from general knowledge. In syllogistic reasoning, the available conclusions are those that are initially formulated on the basis of a model of the premises. If an initial and tentative conclusion accords with belief, then it may tend to be accepted without further ado; if it is contrary to belief, an assiduous search for a refutation may be made. This conjecture assumes that the content of a tentative conclusion may control the process of deduction by affecting the *search* procedure. It is possible, however, that when subjects realize that a putative conclusion is unbelievable, rather than conducting a search for alternative models or producing the unbelievable conclusion, they may instead modify the conclusion to one that they find more acceptable without further reference to models of the premises or else may respond that there is no valid conclusion.

Both potential effects of belief—on premise interpretation and on processing—may occur in deductive reasoning. Since previous studies have examined only the evaluation of given conclusions, our primary aim was to determine whether there were effects of belief bias when subjects draw conclusions for themselves. This procedure obviates the criticism that subjects may be making decisions solely on the basis of the congeniality of given conclusions. We also wanted to investigate two other factors: the nature of the belief and the determinacy of the premises. We manipulated the nature of belief along the same lines as Revlin and Leirer (1978) by constructing materials that yield conclusions that are false either by definition or as a matter of fact. We manipulated determinacy by comparing performance on premises that yielded valid conclusions with performance on premises that did not yield valid conclusions interrelating the end terms of the syllogism.

EXPERIMENT 1

Method

Subjects

Twenty-four subjects participated in the experiment. They were recruited from the student population of Sussex University and from a group of sixth-formers who were visiting our laboratory and were paid for their participation. The subjects had had no formal training in logic, and had not participated in any other experiment on syllogistic reasoning.

Materials

Two sorts of syllogism in the same mood were employed. One sort had a valid conclusion interrelating the end terms, and the other did not. The form of the syllogisms was as follows:

Some of the A are not B	Some of the A are not B
All of the C are B	All of the B are C
(∴ SOME OF THE A ARE NOT C)	(NO VALID CONCLUSION)

where we have stated the correct response in capital letters.

The materials were constructed in the following way: We began by devising a set of potential conclusions, half of which were *definitionally* false, and half of which were *empirically* false, i.e. false as a matter of fact (see Revlin and Leirer, 1978). Definitionally false statements are unbelievable because they violate a definition, e.g. "Some of the millionaires are not rich". Empirically false statements are merely contrary to general knowledge, and it is possible to think of circumstances in which they could be true, e.g. "Some of the athletes are not healthy". Each unbelievable statement was matched with a corresponding believable one, e.g. "Some of the healthy people are not athletes". The definitional/empirical distinction does not apply in the case of the control believable statements (i.e. there is little difference in believability between "Some of the women are not actresses" and "Some of the good swimmers are not lifeguards").

In order to assess the believability of the conclusions thus derived, the set of potential conclusions was rated by an independent group of 20 subjects on a seven-point scale, ranging from completely plausible to completely implausible. The subjects were asked merely to decide how plausible or implausible the statements were. Those statements that produced the most extreme and consistent ratings were selected for use in the study; they and their mean ratings are shown in the Appendix. The distinction between empirical and definitional falsity was borne out by the subjects' ratings: the definitionally false statements were rated as more strongly implausible.

For the valid form, two syllogisms were constructed with different contents: a content that led to a believable conclusion and one that led to an unbeliavable conclusion, both of which were valid. The premises were presented in the form "All of the A are B", since this form increases the acceptability of the premises by suggesting that the quantifier applies to a particular group of people or things rather than to every member of the class or an empty class (see Johnson-Laird and Bara, 1984b). Here are examples of the two types of premises:

Some of the healthy people are not vegetarians
All of the athletes are vegetarians
(∴ SOME OF THE HEALTHY PEOPLE ARE NOT ATHLETES:
a valid, believable conclusion)

Some of the athletes are not vegetarians
All of the healthy people are vegetarians
(∴ SOME OF THE ATHLETES ARE NOT HEALTHY PEOPLE:
a valid, unbelievable conclusion)

Similarly, the indeterminate syllogisms were presented with contents that made

the most frequently produced and erroneous conclusion, "Some A are not C" (see Johnson-Laird and Steedman, 1978; Johnson-Laird and Bara, 1984a) believable or unbelievable. For example, the syllogisms:

Some of the women are not beautiful
All of the beautiful people are actresses
(∴ some of the women are not actresses:
an invalid, believable conclusion)

Some of the actresses are not beautiful
All of the beautiful people are women
(∴ some of the actresses are not women:
an invalid, unbelievable conclusion)

In summary, there were eight different types of syllogism deriving from three variables: the determinacy of the premises, the believability of their potential conclusions, and the nature of the conclusions (definitional or empirical). In addition to these 8 syllogisms, which were all in the same mood, there were 3 fillers, which were each in a different mood, to distract subjects from the form of the syllogisms of interest and to ensure that each of the 5 possible forms of response was correct for at least one problem. The fillers were determinate syllogisms that yielded a conclusion of a different form to the 4 valid syllogisms in the experiment.

Design

Each subject received one example of each of the 8 types of problem, together with the filler items. The 11 problems were typed, one to a page, and stapled together in booklets. The different contents were rotated over problem types within strong and weak conclusions, producing four alternative sets of materials. The presentation order of the problems was randomized with the restriction that the filler items appeared in the same positions in each booklet: in 3rd, 5th and 9th places.

Procedure

The subjects were tested in small groups. The instructions were read to each group as a whole and a list of the possible forms of conclusion:

All of the _____ are _____.
Some of the _____ are _____.
None of the _____ is a _____.
Some of the _____ are not _____.
No valid conclusion.

was visible to the subjects throughout the experiment. The subjects were told that they should assume that the statements were true and that they should say what, if anything, followed from each pair of them solely on the basis of what could be deduced with absolute certainty. If they considered that there was no conclusion that followed, they should respond that there was no valid conclu-

sion; they were told that this was the correct answer in some cases. The subjects were given as long as they needed to complete the task.

Results and Discussion

The majority of erroneous conclusions (88%) were believable, if we exclude the mere repetition of premises and other responses that failed to interrelate the premises or included the middle term (9% of all responses). The only erroneous unbelievable conclusions drawn from determinate premises occurred when the correct response was, in fact, believable.

The percentages of correct responses are presented in Table I as a function of the determinacy of the premises, the believability of the potential conclusion, and its nature (true by definition or true as a matter of fact). The correct responses were submitted to an analysis of variance. There were more correct responses to determinate premises than to indeterminate premises, $F(1,20)=7.39$, $p<0.025$. This pattern may reflect what Revlis (1975) has termed a bias against accepting a "non-propositional" conclusion, i.e. a "No valid conclusion" response. There was no overall effect of believability, but this factor interacted with determinacy, $F(1,20)=4.51$, $p<0.05$: it had a smaller effect on determinate syllogisms than on indeterminate ones, which were more likely to yield the correct response of "No valid conclusion" when the obvious conclusion was unbelievable. However, all three factors of determinacy, plausibility and the nature of the belief interacted, $F(1,20)=4.84$, $p<0.05$. As Table I suggests, where a putative conclusion is false by definition, performance with determinate problems is poorer, but performance with indeterminate problems is enhanced. These effects are difficult to grasp at first, and we will spell out a rationale, supported by supplementary analyses.

With determinate premises, the subjects are quite competent in drawing valid conclusions (in agreement with the results reported by Johnson-Laird and Steedman, 1978). They are not affected by arriving at a conclusion that is empirically false, but they are disturbed when the conclusion violates a definition and are likely to respond "No valid conclusion". This result is supported by an analysis of subjects' errors; these data are shown in Table I. A test of the difference between two proportions (Hoel, 1971) revealed that the ratio of "No valid conclusion" responses to correct responses was significantly larger for definitionally false than for empirically false conclusions, $z=2.31$, $p<0.01$. With indeterminate premises, errors were generally higher than for determinate premises, with one exception: if the obvious, though invalid, conclusion was definitionally false, then subjects were more

Table I

The Percentages of Conclusions Produced in Experiment 1, as a Function of the Determinacy of the Syllogisms, the Believability of the Conclusion, and its Nature (Empirical or Definitional). The Correct Conclusion in Each Category is Presented First and in Capitals

	Unbelievable conclusions false by definition		Unbelievable conclusions empirically false	
	Believable control (Some of the women are not actresses)	Unbelievable (Some of the actresses are not women)	Believable control (Some of the healthy people are not athletes)	Unbelievable (Some of the athletes are not healthy)
Determinate premises				
"SOME A ARE NOT C" (correct)	58	38	50	58
"No valid conclusion" errors	13	46	17	13
Other errors[2]	29	16	33	29
Indeterminate premises				
"NO VALID CONCLUSION" (correct)	17	50	21	29
"Some A are not C" errors	46	17	29	8
Other errors[2]	37	33	50	63

[1]For indeterminate premises, believability refers to the rated believability of the most common *errors*, which are the same as the correct responses for determinate problems ("Some of the A are not C").

[2]The remaining responses in each category consist of conclusions in other moods, and conclusions that contain the middle term or fail to relate the end terms.

likely to respond correctly, "No valid conclusion". (Here, the difference in the corresponding pattern of errors was not reliable, principally because where a conclusion of the form "Some of the A are not B" was empirically false, there was a marked tendency for subjects to conclude "Some of the A are B".) In summary, subjects have a tendency to respond "No valid conclusion" whenever a putative conclusion is definitionally false, and therefore they do better in this condition with indeterminate premises but worse with determinate premises.

The data did not corroborate the conversion hypothesis. The conclusions that would have been produced by logical reasoning from the same premises in their converted form ("Some of the C are not A") accounted for only 2.6% of all responses. Thus, if errors do arise through misunderstanding the premises, the distortions must arise by some process other than conversion.

The present findings show that there are effects of belief bias when subjects *produce* conclusions of their own. Such effects occur with indeterminate premises and indeed are more marked for them than for determinate premises. Likewise, there are effects of the nature of the belief: definitionally false conclusions are less likely to be drawn than empirically false conclusions.

EXPERIMENT 2

We designed a second experiment to extend our research to syllogisms with "convertible" premises, i.e. those that can be converted without altering the logically valid conclusions. Such syllogisms were used by Evans et al. (1983) in their evaluation study, but they may not produce the same pattern of results in a production study because the syllogisms themselves are more difficult, and typically produce a far higher proportion of errors, than those used in our first experiment. We used the same two sorts of believability as in Experiment 1 but only determinate premises since there are no indeterminate premises in the same mood (i.e. pairs of premises in the "None ..., Some ..." mood always yield a valid conclusion precisely because both premises are convertible).

Method

Subjects

Sixteen subjects from the staff and student population of Sussex University were paid to participate in the experiment. None had done an experiment of this kind before, and none had any training in formal logic.

Materials

We constructed four syllogisms of the form:

None of the A is a B
Some of the C are B

which yield the valid conclusion "Some of the C are not A", and four syllogisms of the form:

Some of the A are B
None of the C is a B

which yield the valid conclusion: "Some of the A are not C".

The contents of the syllogisms were the same as those in Experiment 1, with two slight modifications as shown in the Appendix to produce more marked differences in rated plausibility.

Design

Each subject carried out 2 trials in the 4 experimental conditions deriving from two variables: the believability of the conclusions and the nature of the belief (definitional or empirical). The first trial was with one order of the premises and the second trial was with the other order. Thus, the overall number of different "moods" was the same as in Experiment 1. The contents of the syllogisms were rotated within the two levels of the empirical/definitional factor, thereby producing four alternative sets of materials.

We presented the subjects with two syllogisms of each type (believable/unbelievable, empirical/definitional), and one of each pair appeared in each half of the presentation booklet. Within these halves, the order of the problems was according to a balanced Latin square so that each problem was preceded and followed equally often by every other and appeared equally often in each position in the design as a whole. The four filler items consisted of those from Experiment 1, with the addition of one indeterminate filler, so that each of the 5 possible conclusions was correct for at least one problem. The syllogisms were typed, one to a sheet, and stapled together in a booklet, as for Experiment 1. The fillers appeared in 3rd, 5th, 8th and 10th positions in the booklets.

Procedure

The subjects were tested in small groups. Each subject was given a set of typed instructions that were the same as those in Experiment 1, and subjects worked through the booklet at their own pace.

Results

Of the erroneous conclusions interrelating the end terms, 89% were believable; and the percentage of unbelievable errors was only slightly higher when the correct conclusion was unbelievable (6.8%) than when it was believable (4.5%).

The percentages of correct responses are shown in Table II. An analysis of variance was carried out on the corresponding means. The pattern of results was exactly the same as for the determinate premises in Experiment 1. Although there was a marginally significant effect overall of the nature of belief—with empirical materials yielding a better performance than definitional ones, $F(1,12)=3.93$, $p=0.07$, this factor interacted with believability, $F(1,12)=11.30$, $p<0.025$. When the premises supported a believable conclusion, the subjects were more likely to deduce it if it was definitionally true than if it was merely empirically true; but when the premises supported an unbelievable conclusion the subjects were less likely to deduce it if it was definitionally false than if it was merely empirically false (the latter difference was highly reliable, $t(15)=4.04$, $p<0.005$). Indeed, as in Experiment 1, subjects tended to respond "No valid conclusion" whenever the premises yielded a valid conclusion that was definitionally false. The ratio of "No valid conclusion" responses to correct responses was significantly higher for those problems with definitionally false conclusions than for those with only empirically false conclusions, $z=3.70$, $p<0.001$. The order in which the premises were presented did not have an overall effect on performance and did not interact with any other factor.

GENERAL DISCUSSION

Both experiments have shown that beliefs can bias the conclusions that subjects draw for themselves from syllogistic premises. In particular, if premises lead validly to a conclusion, then subjects are more likely to respond with that conclusion if it is believable than if it is unbelievable, but only if the status of the conclusion rests on a matter of definition and not on a matter of fact. Where the premises do not support a valid conclusion, there is a corresponding effect: subjects are much more likely to respond "no valid conclusion" if the premises suggest a definitionally false conclusion. In short, subjects to not arrive at definitionally false conclusions—or if they do, they go on to reject them—with the result that they may be inhibited from making what is, in fact, a valid deduction, or, alternatively, encouraged to respond correctly that there is no valid conclusion.

The only other study known to us of the effects of beliefs on the spontaneous production of conclusions (J. Barston, personal communication) failed to detect the phenomenon. However, although this study used the same forms of premise as in our Experiment 2, the content of the problems yielded only *mildly* unbelievable conclusions. With the present type of design, it appears that definitionally false materials are needed in order to produce detectable effects. Revlin et al. (1980)

Table II

The Percentages of Conclusions Produced in Experiment 2, as a Function of the Believability of the Conclusion and its Nature (Empirical or Definitional)

	Unbelievable conclusions false by definition		Unbelievable conclusions empirically false	
	Believable control (Some of the women are not actresses)	Unbelievable (Some of the actresses are not women)	Believable control (Some of the healthy people are not athletes)	Unbelievable (Some of the athletes are not healthy)
Correct	53	19	44	66
"No valid conclusion"	13	34	19	6
Other errors	34	47	37	28

likewise only obtained effects of content on the evaluation of given conclusions when there was a conflict between a logical conclusion and a *definitionally* true one.

In the Introduction, we argued that beliefs could affect reasoning either by leading to distortions in the interpretation of premises or by influencing the processes of deduction and conclusion evaluation (or both). Our results provide no support for models of syllogistic reasoning that invoke "conversion" or "atmosphere" to explain errors. Our second experiment replicated the findings for determinate syllogisms with problems in moods in which conversion of the premises did not alter the valid conclusion. Hence, in the case of our materials, the effects of beliefs appear to arise in the process of making the inference and formulating and evaluating putative conclusions.

How could beliefs bias the process of inference? The answer depends, of course, on how people normally make inferences. If they rely solely on *formal* rules of inference, then, as we argued in the Introduction, it is difficult to see how beliefs could directly interact with such rules, which, by definition, make no reference to any specific content, and which, therefore, should be applied to materials of any degree of credibility. Once this notion of the uniform application of rules is abandoned, then of course the resulting theory of inference is no longer just a matter of mental logic. If, however, deductive reasoning depends, not on mental logic, but on the construction of mental models, then beliefs could exert their effects directly on the process of inference. The construction of a model of the premises, the formulation of a conclusion based on it, and the search for models refuting a putative conclusion, are all likely to be influenced by a subject's knowledge and experience. In particular, a conclusion that is credible is likely to be accepted without an assiduous search for counterexamples (i.e. alternative models of the premises), whereas a conclusion that is not credible is likely to lead to a suspension of the inferential process and the response that there is no valid conclusion. Such a mechanism yields precisely the pattern of results that we observed with the definitional materials. A similar pattern of results is reported by Revlin et al. (1980): they found that, when presented with syllogisms where logic and belief conflict, the predominant error response was "No valid conclusion".

The above account assumes that the content of a candidate conclusion may control the process of deduction by influencing the search process. It is possible, however, as we suggested in the Introduction, that the process of stating a conclusion in verbal form may engage certain beliefs, and these beliefs may affect the likelihood with which a particular conclusion is produced. A subject who comes up with an unbelievable conclusion, at whatever stage of the deductive process, may reject it in

favour of a "No valid conclusion" response or else produce a modified, more believable conclusion.

Our experiments do not enable us to distinguish between these two accounts of how subjects respond to unbelievable conclusions. If plausibility affects the search process, however, we can predict that it should have a considerable effect where several different models of the premises need to be considered in order to make an inference, and it should have no effect where there is only one possible model of the premises. If plausibility has an effect on the formulation and evaluation of conclusions, then its effects should be apparent even in one-model problems. Experiments are in progress to explore these questions.

REFERENCES

Chapman, L. J. and Chapman, J. P. (1959). Atmosphere effect re-examined. *Journal of Experimental Psychology*, **58**, 220–226.

Evans, J. St. B. T. (1982). *The psychology of deductive reasoning*. London: Routledge and Kegan Paul.

Evans, J. St. B. T., Barston, J. and Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition*, **11**, 295–306.

Falmagne, R. J. (1975). *Reasoning: Representation and process*. Hillsdale, N. J.: Lawrence Erlbaum Associates Inc.

Feather, N. T. (1964). Acceptance and rejection of arguments in relation to attitude strength, critical ability and intolerance of inconsistency. *Journal of Abnormal and Social Psychology*, **69**, 127–136.

Henle, M. and Michael, M. (1956). The influence of attitudes on syllogistic reasoning. *Journal of Social Psychology*, **44**, 115–127.

Hoel, P. G. (1971). *Elementary Statistics* (3rd. ed.). London: Wiley.

Janis, I. and Frick, P. (1943). The relationship between attitudes toward conclusions and errors in judging logical validity of syllogisms. *Journal of Experimental Psychology*, **33**, 73–77.

Johnson-Laird (1983). *Mental models*. Cambridge: Cambridge University Press.

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

Johnson-Laird, P. N. and Bara, B. (1984b). Logical expertise as a source of error: A reply to Boolos. *Cognition*, **17**, 183–184.

Johnson-Laird, P. N., Oakhill, J. V. and Bull, D. (in press). Children's syllogistic reasoning. *The Quarterly Journal of Experimental Psychology*.

Johnson-Laird, P. N. and Steedman, M. J. (1978). The psychology of syllogisms. *Cognitive Psychology*, **10**, 64–99.

Kaufman, H. and Goldstein, S. (1967). The effects of emotional value of conclusions upon distortion in syllogistic reasoning. *Psychometric Science*, **7**, 367–368.

Lefford, A. (1946). The influence of emotional subject matter on logical reasoning. *Journal of General Psychology*, **34**, 127–151.

Matte-Blanco, I. (1965). A study of schizophrenic thinking: Its expression in

terms of symbolic logic and its representation in terms of multi-dimensional space. *International Journal of Psychiatry*, **1**, 91–96.

Morgan, J. and Morton, J. (1944). The distortion of syllogistic reasoning produced by personal convictions. *Journal of Social Psychology*, **20**, 39–59.

Newstead, S. E. and Griggs, R. A. (1983). Drawing inferences from quantified statements: A study of the square of opposition. *Journal of Verbal Learning and Verbal Behavior*, **22**, 535–546.

Nisbett, R. E. and Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgement*. Englewood Cliffs, N.J.: Prentice-Hall.

Pollard, P. (1982). Human reasoning: Some possible effects of availability. *Cognition*, **12**, 65–96.

Revlin, R. and Leirer, V. (1978). The effect of personal biases on syllogistic reasoning: Rational decisions from personalised representations. In R. Revlin and R. E. Mayer (Eds.), *Human reasoning*. Washington, D.C.: Winston-Wiley.

Revlin, R., Leirer, V., Yopp, H. and Yopp, R. (1980). The belief–bias effect in formal reasoning: The influence of knowledge on logic. *Memory and Cognition*, **8**, 584–592.

Revlis, R. (1975). Two models of syllogistic reasoning: Feature selection and conversion. *Journal of Verbal Learning and Verbal Behavior*, **14**, 180–195.

Thouless, R. H. (1959). Effect of prejudice on reasoning. *British Journal of Psychology*, **50**, 289–293.

Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, **5**, 207–232.

Von Domarus, E. (1944). The specific laws of logic in schizophrenia. In J. S. Kasanin (Ed.), *Language and thought in schizophrenia*. Berkeley: University of California Press.

Wason, P. C. and Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. London: Batsford; Cambridge, Mass.: Harvard University Press.

Wilkins, M. (1928). The effect of changed material on ability to do formal syllogistic reasoning. *Archives of Psychology*, **16**, 83.

Revised manuscript received 14 May 1985

APPENDIX

The mean believability ratings of the conclusions in the two experiments (The ratings were derived from two groups of 10 subjects. Each subject rated an equal number of plausible and implausible statements, with different contents. The statements were rated on a 7-point scale, from 1 = highly implausible to 7 = highly plausible).

Materials for Experiments 1 and 2

	Empirically unbelievable (U)/believable (B) pairs	$\bar{X}$	*s.d.*
B	Some of the healthy people are not athletes	6.5	1.6
U	Some of the athletes are not healthy	3.5	2.0
B	Some of the good swimmers are not lifeguards	7.0	0
U	Some of the lifeguards are not good swimmers	2.4	1.2
B	Some of the highly-trained dogs are not police dogs	6.9	0.3
U	Some of the police dogs are not highly-trained	2.8	1.6
B	Some of the well-educated people are not professors	7.0	0
U	Some of the professors are not well-educated	4.2	2.0
	$\bar{X}$ believable	6.85	
	$\bar{X}$ unbelievable	3.23	
	Definitionally unbelievable (U)/believable (B) pairs		
B	Some of the rich people are not millionaires	6.1	1.9
U	Some of the millionaires are not rich	1.4	1.0
B	Some of the women are not actresses	7.0	0
U	Some of the actresses are not women	1.7	1.9
B	Some of the religious people are not priests*	7.0	0
U	Some of the priests are not religious*	3.1	2.2
B	Some of the scientists are not physicists*	5.8	2.5
U	Some of the physicists are not scientists*	1.4	0.7
	$\bar{X}$ believable	6.48	
	$\bar{X}$ unbelievable	1.9	
	Additional materials for Experiment 2		
B	Some of the doctors are not general practitioners	6.9	0.4
U	Some of the general practitioners are not doctors	1.1	0.4
B	Some of the married people are not husbands	6.0	2.0
U	Some of the husbands are not married	1.0	0
	Expt. 2 definitional materials:		
	$\bar{X}$ believable	6.5	
	$\bar{X}$ unbelievable	1.3	

*Materials replaced in Experiment 2.