

HUMAN REASONING: LOGICAL
AND NONLOGICAL EXPLANATIONS.

BY

PAUL POLLARD

Department of Psychology, Plymouth Polytechnic

Thesis submitted to the Council for National

Academic Awards

In partial fulfilment of the requirements for the

Degree of Doctor of Philosophy

November, 1979

PLYMOUTH POLYTECHNIC LIBRARY	
	Accu no
	Class no
	Cover no

STACK

PLYMOUTH POLYTECHNIC LIBRARY	
Accn. No.	5500034
Class. No.	T 153.43 Pol
Contl. No.	X700113885

ABSTRACT

HUMAN REASONING: LOGICAL AND NONLOGICAL EXPLANATIONS. BY PAUL POLLARD

This thesis is concerned with subjects' responses to psychological 'reasoning' tasks. Theoretical interpretations of responses to such tasks can be broadly categorised into logical explanations, that assume that subjects perform an analysis on the structure of the problem and nonlogical explanations, that explain responses as determined by some, logically irrelevant, feature of the problem. In part 1, data from relevant reasoning studies are reviewed with particular reference to identifying separate effects of logical and nonlogical factors. It is concluded that both factors appear to play a role in determining responses.

In part 2, seven experiments are presented in five chapters. The first chapter reports experiments which offered subjects the opportunity of affirming or denying a set of valid or invalid inferences from a conditional rule. The relative influence of logical and nonlogical factors was considered and the results related to comparable findings discussed in part 1. The remaining chapters report experiments on the effect of perceived truth status of a conditional rule on 'selection task' responses. Truth status was manipulated both experimentally, by use of probability learning tasks, and by use of relevant thematic material. Logical and non-logical interpretations of results were tested in later experiments and an association theory developed to explain the main effects.

In the general discussion, the association theory is discussed as a special case of the 'availability' theory of Tversky and Kahneman (1973). This general theory is applied to various effects of truth status and to other nonlogical responses. It is postulated that such responses may be acquired because they have some utility in 'real life' and this hypothesis is extended to other types of preferred responses. Finally, the distinction between 'logical' and 'nonlogical' factors is reassessed and it is concluded that the distinction may be artificial, as practically all preferred responses may be a function of simple cues that the subject has learned to attend to.

ABSTRACT

HUMAN REASONING: LOGICAL AND NONLOGICAL EXPLANATIONS. BY PAUL POLLARD

This thesis is concerned with subjects' responses to psychological 'reasoning' tasks. Theoretical interpretations of responses to such tasks can be broadly categorised into logical explanations, that assume that subjects perform an analysis on the structure of the problem and nonlogical explanations, that explain responses as determined by some, logically irrelevant, feature of the problem. In part 1, data from relevant reasoning studies are reviewed with particular reference to identifying separate effects of logical and nonlogical factors. It is concluded that both factors appear to play a role in determining responses.

In part 2, seven experiments are presented in five chapters. The first chapter reports experiments which offered subjects the opportunity of affirming or denying a set of valid or invalid inferences from a conditional rule. The relative influence of logical and nonlogical factors was considered and the results related to comparable findings discussed in part 1. The remaining chapters report experiments on the effect of perceived truth status of a conditional rule on 'selection task' responses. Truth status was manipulated both experimentally, by use of probability learning tasks, and by use of relevant thematic material. Logical and non-logical interpretations of results were tested in later experiments and an association theory developed to explain the main effects.

In the general discussion, the association theory is discussed as a special case of the 'availability' theory of Tversky and Kahneman (1973). This general theory is applied to various effects of truth status and to other nonlogical responses. It is postulated that such responses may be acquired because they have some utility in 'real life' and this hypothesis is extended to other types of preferred responses. Finally, the distinction between 'logical' and 'nonlogical' factors is reassessed and it is concluded that the distinction may be artificial, as practically all preferred responses may be a function of simple cues that the subject has learned to attend to.

ACKNOWLEDGEMENTS

I should like to acknowledge the very great amount of attention and help I have received from my supervisor, Dr. J. St. B.T. Evans, both as regards the design and conduct of the experiments reported herein and the production of the final draft of the thesis. In addition, the many interesting and rewarding theoretical discussions I have had with Dr. Evans throughout the last three years have been important contributors to my own theoretical development.

I should also like to thank Plymouth Polytechnic for financial support, enabling this thesis to be completed, and Dr. S.E. Newstead for his supporting role in the supervision.

DECLARATIONS

- 1) While registered for this degree, I have not been a registered candidate for another award of the CNAA or of a University.

- 2) The contents of Chapter 5 form the basis of a paper to be published in the British Journal of Psychology entitled "The effect of prior beliefs in reasoning: An associational interpretation".

- 3) A course of advanced study has been completed, in partial fulfilment of the requirements for the degree, consisting of guided reading in the area of deductive reasoning (supervised by Dr Evans), attendance at an advanced course on language and thought (BA Psychology special option, Plymouth Polytechnic) and attendance at relevant conferences.

CONTENTS

<u>INTRODUCTION</u>	6
<u>PART 1 - HISTORICAL REVIEW</u>	8
CHAPTER 1: CATEGORICAL SYLLOGISMS	12
CHAPTER 2: PROPOSITIONAL REASONING	59
CHAPTER 3: THE WASON SELECTION TASK	105
CHAPTER 4: THEORETICAL APPROACHES TO THE STUDY OF REASONING	138
<u>PART 2 - EXPERIMENTAL WORK</u>	150
CHAPTER 5: CONDITIONAL INFERENCE EXPERIMENTS	151
CHAPTER 6: TRUTH STATUS EFFECTS ON SELECTION TASK BEHAVIOUR	203
CHAPTER 7: THE EFFECTS OF EXPERIMENTALLY MANIPULATED TRUTH STATUS ON SELECTION TASK BEHAVIOUR	218
CHAPTER 8: THE EFFECT OF EXPERIMENTALLY MANIPULATED TRUTH STATUS ON SELECTION TASK RESPONSES ACROSS ALL FOUR FORMS OF A CONDITIONAL RULE	245
CHAPTER 9: AN EXPERIMENT ON THE EFFECT OF IMPLICIT TRUTH STATUS, USING THEMATIC RULES	274
<u>PART 3 - GENERAL DISCUSSION</u>	286
CHAPTER 10: TRUTH STATUS EFFECTS	288
CHAPTER 11: ABSTRACT REASONING TASKS: SOME POSSIBLE EFFECTS OF PRIOR EXPERIENCE	317
CHAPTER 12: A REEVALUATION OF THE POSITION OF EVANS (1972a, 1977b)	339

INTRODUCTION

This thesis is concerned with subjects' responses to 'logical problems'. That is, problems whose structure may be analysed in terms of formal logic. On the basis of such analysis, the subjects' response may be categorised (with certain exceptions) as either valid or invalid. As will be seen, subjects frequently fail to make the logically valid response and thus, from the standpoint of logic, subjects may be said to make errors.

Evans (1972a) has analysed explanations of these errors into three main types, as follows:

- 1) Logical explanations assume that the subject has misinterpreted the problem in some way and that he has 'reasoned' in accordance with formal logic on the basis of this misinterpretation. The response is thus seen as the 'right' answer to the 'wrong' question.
- 2) Illogical explanations assume that the subject has employed an invalid 'strategy'. Such 'strategies' are often held to be valid in a 'natural', real life, logic. The subject is thus seen as 'reasoning' logically in accordance with an alternative system. A slightly different form of 'illogical' explanation is that the subject's (invalid) response is due to a characteristic error that occurs during (valid) analysis of the structure of the problem.
- 3) Nonlogical explanations assume that the response is a resultant of a logically irrelevant influence, unrelated to the structure of the problem. The subject is thus not seen as performing any logical analysis on the problem

It will be noted that the first two alternatives both assume that the subject attends to, and performs an analysis upon, the logical structure of the problem, whereas the third alternative does not. The first two alternatives may thus be 'collapsed', producing a basic dichotomy between 'logical' and 'nonlogical' explanations (in fact, Evans also adopts a similar dichotomy). This dichotomy may be extended to logically valid responses as, clearly, the subject may make a valid response because he has correctly analysed the logical structure of the problem, or simply because some nonlogical influence fortuitously leads to the logically correct answer.

Reference will frequently be made to logical and nonlogical factors. As defined here, a logical (or nonlogical) factor is that mediator of responses that is posited by a logical (or non-logical) explanation. The relative merits of logical and nonlogical explanations, in general, may thus be investigated by evaluating the evidence for the effect of various specific logical and nonlogical factors. Such an investigation is carried out in the historical review. As will be seen, although in certain cases there is reasonably clear-cut evidence for the influence of both logical and nonlogical factors, in other cases there are competing logical and nonlogical explanations of the same data. This 'competition' is a theme that continues into the experimental section and is a central topic of chapter 5. In the general discussion, however, logical and nonlogical effects are considered separately and a theoretical position is developed that suggests that these factors may have a similar source and that the distinction between them is far from clear-cut.

PART 1

HISTORICAL REVIEW

This review deals with experiments which have presented subjects with problems that can be analysed in terms of formal logic. Such tasks are typically referred to as 'reasoning tasks'. In this context, 'formal logic' refers to systems of logic that may be traced back to classical philosophers, notably Aristotle. Before considering these problems in more detail, a brief explanation of the subject matter of formal logic is in order.

Formal logic is concerned with propositions. A proposition has been defined as "anything which can be said to be true or false" (Cohen & Nagel, 1934) and, to be stated, must be contained within a sentence. Sentences and propositions should not be confused, however, as different sentences may contain the same proposition (for instance, 'David is smaller than Goliath' and 'Goliath is larger than David') and the same sentence may be used to convey different propositions (for instance, the sentence 'my party should govern the country', uttered by persons having different political affiliations).

There are three basic 'laws' of logic that apply to propositions: 'if a proposition is true, it is true' (the principle of identity), 'a proposition cannot be both true and false' (the principle of noncontradiction) and 'all propositions are either true or false' (the principle of excluded middle). However, it need not be known whether a proposition is true or false. The truth or falsity of a proposition may be the subject of disagreement or may be only ascertainable in the future and it may never be possible to ascertain the truth or falsity of certain propositions. In fact the concept of absolute truth or falsity is inapplicable to many propositions. For instance, psychology and other statistically based branches of science view truth as a probabilistic, rather than absolute concept. However, whether a proposition is true or false is not a

question with which the logician concerns himself. The logician is solely concerned with sets of propositions structured in the form of logical arguments, specifically, with whether such arguments are valid or invalid.

Copi (1972) defines a logical argument as "any group of propositions of which one is claimed to follow from the others, which are regarded as providing grounds for the truth of that one". The proposition claimed to follow from the others is termed the conclusion and the other propositions are termed premisses. Such arguments are valid if, and only if, it is impossible for the premisses to be true and the conclusion false. The logical validity of an argument is thus not necessarily dependent on the actual truth or falsity of its components. In the extreme cases, it is possible for a valid argument to be entirely composed of false propositions, and for an invalid argument to be entirely composed of true propositions. Of particular importance is that the logical validity of an argument is not dependent on the truth or falsity of its conclusion.

The experiments reported in this review have investigated the extent to which human reasoning responses are in accordance with those responses indicated by logical validity. In classical terminology, chapter 1 concerns categorical syllogisms whereas much of chapter 2 is concerned with hypothetical and disjunctive syllogisms. However, classical writers concentrated almost exclusively on categorical syllogisms and other forms have provided the basis for the 'propositional calculus' developed by more recent logicians. Accordingly, the terms 'hypothetical syllogism' and 'disjunctive syllogism' are not used in the text and categorical syllogisms are referred to simply as 'syllogisms'.

Although none of the experimental work reported in later chapters is directly concerned with syllogisms, research on syllogisms is included in this review. There are two reasons for this. Firstly the initiation of psychological research on syllogisms predates research in other areas of reasoning and thus provides an historical perspective. Secondly, it will be seen that the issues raised by research on syllogisms are comparable to those raised by research on the propositional calculus and that any theory of reasoning must attempt to explain data from both of these areas.

CHAPTER 1
CATEGORICAL SYLLOGISMS

SECTION 1.1	The formal logic of Syllogisms: Mood	13
SECTION 1.2	Experiments on the Effect of 'Mood'	16
SECTION 1.3	Thematic Content: The Effect of Truth Status	31
SECTION 1.4	The Formal Logic of Syllogisms: Figure	43
SECTION 1.5	Experiments on the Effect of 'Figure'	45
SECTION 1.6	Conclusions	54

SECTION 1.1 THE FORMAL LOGIC OF SYLLOGISMS: MOOD

Syllogisms involve propositions that affirm or deny that one class is included in another. The following are examples of such propositions:

All cats are animals

All cats are good pets

These propositions link the class of 'cats' with the class of 'animals' and the class of 'things that are good pets'. They are both affirmative and 'universal' (that is, they apply to the whole class of cats) and are termed 'A' propositions. The negative of universal proposition, 'all cats are not good pets' (traditionally expressed as 'no cat is a good pet') is termed an 'E' proposition. It is also possible to construct ('particular') propositions that do not necessarily apply to the whole class; for instance: 'some cats are good pets' ('I' type propositions) and 'some cats are not good pets' ('O' type propositions). Using A and B to stand for any two classes, the four types of proposition are listed below:

A	All A are B	(universal, affirmative)
E	No A are B	(universal, negative)
I	Some A are B	(particular, affirmative)
O	Some A are not B	(particular, negative)

The formal meaning of 'some' is 'at least some' and

thus the proposition 'some A are B' allows the possibility that 'all A are B'. For this reason, an I proposition does not imply an O proposition and vice versa. It is also worth noting that, although 'no A are B' implies that 'no B are A' and 'some A are B' implies 'some B are A', 'all A are B' does not imply that 'all B are A' and 'some A are not B' does not imply that 'some B are not A'. Put simply, it is valid to 'convert' E and I propositions but invalid to convert A and O propositions.

A syllogism is a deductive argument containing three propositions of the type discussed above. There are two premisses, having only three terms between them (i.e. there is a common middle term) and a conclusion having the two terms other than the middle term. For instance, the following is a (valid) syllogism:

All artists are perfectionists
All artists are creative (people)
∴ Some creative people are perfectionists

The premisses have three terms ('artists', 'perfectionists' and 'creative people'); one of these ('artists') is a middle term and the other two are identical to the terms of the conclusion. The premiss containing the predicate (second term) of the conclusion is referred to as the major premiss and the premiss containing the subject (first term) of the conclusion is referred to as the minor premiss.

Each of the three component propositions may be any one of the four types (A, E, I or O) and there are thus 64 distinct types of syllogisms. These are referred to as moods. A simple way to give the mood of a syllogism is to express it in the form

of three letters, for instance, AIE, which defines the mood of the major premiss, the minor premiss and the conclusion, respectively. Similarly, the mood of a premiss pair may be expressed in the form of two letters.

The conclusion of a syllogism may be either (logically) valid or invalid and there are a variety of formal rules for determining whether a given conclusion is valid. However, the general rule is that a conclusion is logically valid if, and only if, the conclusion must be true if both of the premisses are true.

SECTION 1.2 EXPERIMENTS ON THE EFFECT OF 'MOOD'

Most psychological experiments on syllogisms have asked subjects to evaluate abstract syllogisms. That is, letters are used as the terms of the constituent propositions to reduce extraneous influences on the task. All experiments reported in this and the following sections have utilised abstract syllogisms unless otherwise stated.

The first experiments of this type were performed by Woodworth & Sells (1935) and Sells (1936). Subjects were asked to describe given conclusions as either 'absolutely true', 'probably true' (these being scored as agreement), 'indeterminate' or 'absolutely false' (these latter being scored as disagreement). Clearly the logically correct response would be to agree with valid, and disagree with invalid, syllogisms.

The authors looked at the percentage of agreement with the conclusions of invalid syllogisms. Some syllogisms were correctly evaluated (i.e. disagreed with) by almost all the subjects. However, other syllogisms yielded up to 80% incorrect agreement. The type of conclusion incorrectly agreed with varied greatly between syllogisms but was relatively stable, across subjects, on individual syllogisms. Thus, for each syllogism, there appeared to be a specific conclusion that led subjects into error.

This pattern was explained in terms of an 'atmosphere effect': the drawing of conclusions on the basis of the 'global impression' of the premisses. This theory holds that subjects will agree with a conclusion if it accords with the 'atmosphere' of the premisses. Affirmative (A or I) premisses produce an 'affirmative atmosphere', negative (E or O) premisses produce a 'negative atmos-

phere', universal (A or E) premisses produce a 'universal atmosphere' and particular (I or O) premisses produce a 'particular atmosphere'. Additionally, particular and negative atmospheres were held to be dominant so that, in combination, a universal and a particular premiss tend to produce a particular atmosphere and an affirmative and negative premiss tend to produce a negative atmosphere.

This theory was held to account for almost all the data obtained. However, Sells noted that particular conclusions were often agreed with when both premisses were universal. Accordingly, a further principle, termed 'caution' was added. This was defined as a tendency to prefer 'weak' to 'strong' conclusions.

These explanations view subjects' reasoning responses as being determined by factors other than logical validity. The correct rejection of some of the invalid syllogisms is seen as primarily fortuitous: subjects rejecting the conclusion because it did not accord with the atmosphere of the premisses not because it was invalid. The failure of subjects to reject invalid conclusions that did accord with atmosphere clearly lends strong support to this viewpoint. This agreement with a conclusion on the basis of its 'atmosphere' may be referred to as a response bias. More generally, for the purposes of this thesis, a response bias is defined as 'a tendency to adopt, or not adopt, a particular response alternative dependent upon its having, or not having, some logically irrelevant feature'.

This view of subjects as being essentially incapable of reasoning logically on syllogistic tasks remained unchallenged for over two decades, until Chapman & Chapman (1959) put forward a fundamentally different explanation. They pointed out that, if subjects were self consistent, the acceptance of a universal conclusion would automatically lead them to accept a particular conclusion

(to the same pair of premisses) presented in a different part of the task. It is, however, consistent to accept a particular conclusion without accepting the universal and thus acceptance of (invalid) particular conclusions would be expected to be the most frequent type of error. Observed preference for particular conclusions was thus argued to be a function of the Woodworth & Sells methodology.

Chapman & Chapman presented subjects with pairs of premisses and offered them five alternative conclusions:- one conclusion in each of the four moods and a 'none of these' alternative. Subjects were asked to choose one alternative only and thus could not affirm both a universal and a particular conclusion to the same pair of premisses.

This design has been used by almost all subsequent experiments and will henceforth be referred to as the standard design. Subjects are effectively being asked to construct one of four possible syllogisms. Premiss pairs to which at least one of the given conclusions can be added to form a valid syllogism are referred to as determinate. If no valid syllogism is possible (i.e. if the correct answer is 'none of these'), the premiss pair is indeterminate. (A convention has arisen in the literature whereby premiss pairs alone are referred to as syllogisms, this convention will not be used, however, as it can lead to confusion.)

Chapman & Chapman presented subjects with 42 indeterminate premiss pairs. The principle of caution was not confirmed. When both premisses were universals, selection of particular conclusions was very infrequent. The predictions of atmosphere theory were predominantly confirmed. The dominance of particular atmosphere, however, was not always evident when one or more premiss was negative.

In fact, for IE premiss pairs, the preferred conclusion was a universal negative. The authors would thus seem to be correct in arguing that the previously observed bias towards particular conclusions was a function of the design.

The results of this study were interpreted as invalidating the atmosphere theory, the authors remarking that "since the atmosphere predictions are not substantiated, we must look for other principles of explanation". They offer two main suggestions. Firstly they proposed that subjects interpret the premisses to mean that the converse is true (an interpretation that is logically invalid for A and O propositions) and, secondly, that subjects' responses are mediated by 'probabilistic inference'.

For all premiss pairs that include an A premiss, the type of (invalid) conclusion preferred is in accordance with both atmosphere theory and the conversion hypothesis. However, the conversion hypothesis cannot explain the preferred conclusions for the remaining premiss pairs. To explain these, Chapman & Chapman suggested that subjects believe that things that share a common property are related and that things that lack a common property are not related ('probabilistic inference'). For instance, in the case of an IO premiss pair in the second figure ('some A are B', 'some C are not B'), the subject reasons that some A and some C do not share the common quality of B and, therefore, that some C are not A. To support their ideas, Chapman & Chapman referred to the way in which subjects reason in real life situations. To support the conversion hypothesis, they suggested that their subjects' prior experience of A propositions was primarily obtained from mathematics courses in which 'are' means 'are equal to' and, in support of 'probabilistic inference', they argued that subjects tend to accept probable conclusions rather than reason

in an entirely deductive way.

There are various criticisms that may be levelled at 'probabilistic inference'. The authors remark that they can only offer 'intuitive evidence' for their explanations, but it is far from 'intuitively evident' how the explanation fits some of the data. Unfortunately, the authors give few examples, but not even all the examples chosen are self evident. For instance, for two E premisses ('no A are B', 'no C are B') the conclusion 'no C are A' is held to be a result of probabilistic inference as the middle term, B is not shared. However, this can hardly be an example of subjects' (real life) tendency to accept probable conclusions, as the lack of quality B in both A and C increases (albeit marginally) the probability of their being related. In fact, they both share the common quality of non-B.

The unclarity of the exposition of probabilistic inference has been noted by most authors in the area. The 'conversion hypothesis, however, has proved very popular. It should be noted that it presents an entirely different view (to atmosphere theory) of subjects' reasoning abilities. Subjects are viewed as capable of valid reasoning within the framework of an alternative 'real-life' logic in which A and O propositions are defined as convertible.

Many workers in the field have endeavoured to determine whether characteristic errors are best explained by the conversion or atmosphere hypothesis.

Begg & Denny (1969) considered the question to be empirical rather than theoretical and focussed upon the discrepancies between the Chapman & Chapman data and those of Sells. In essence, there are two main differences:

1) Chapman & Chapman found little evidence of a dominance of particular atmosphere when the argument contained an E premiss.

2) Chapman & Chapman found no evidence for the 'caution' hypothesis (that there is some bias towards particular conclusions, even when both premises are universal).

Begg & Denny reported results that agreed with the Sells data in the case of (1) above and with the Chapman & Chapman data in the case of (2). Unfortunately, Begg & Denny failed to note the Chapman & Chapman argument that the apparent preference for particular conclusions is a function of allowing subjects to affirm more than one conclusion to the same premiss pair. Begg & Denny allowed subjects to affirm more than one conclusion and thus the differences between their data and those of Chapman & Chapman were wholly consistent with the latter's arguments.

Similarly, that Begg & Denny did not observe an effect of 'caution' does not constitute a failure to replicate the Sells data. Begg & Denny asked subjects to choose one or more conclusions to a premiss pair, whereas Sells presented subjects with separate sylogisms. It may well be that particular conclusions to universal premiss pairs are only affirmed when there is no alternative.

Both Simpson & Johnson (1966) and Dickstein (1975) attempted to specifically train one group of subjects against atmosphere, and one group of subjects against conversion, errors. However these results lend little clarity to the issue as Simpson & Johnson reported an effect of 'anti-atmosphere training', whereas Dickstein reported no effect of 'anti-atmosphere training' but did report an effect of 'anti-conversion training'

There are, in fact, good grounds for the argument that 'training' experiments cannot resolve the issue. Firstly, the fact that 'anti-atmosphere' or 'anti-conversion' instructions reduced errors could only suggest that those factors usually mediate typical responses. Secondly, the effect of instructions will be dependent upon the exact nature of the wording. For instance, Dickstein (1975) appears to have given a more strongly worded warning against 'atmosphere', than against 'conversion' errors. Thirdly, in the Simpson & Johnson experiment, there was some indication that a warning of any kind produced a tendency to avoid all four propositional conclusions and underline 'none of these'. The control group, who received no warning, made more mistakes on the indeterminate premiss pairs but made less mistakes on the determinate 'filler' items. Thus a warning may reduce errors on indeterminate premiss pairs simply by creating a response bias against propositional conclusions.

Ceraso & Provitera (1971) pointed out that propositions may be 'modified' to make their meaning clear. For instance, the following proposition makes one specific meaning of an A premiss clear:

M1 'All A are B, but some B are not A'

Under the conversion hypothesis, subjects are held to interpret the usual A form ('all A are B') to mean the following:

M2 'All A are B and all B are A'

Ceraso & Provitera presented subjects with a set of premiss pairs in traditional form (T) and a set of 'modified' premiss pairs (M) in which A premisses were sometimes given in the form of M1 and

sometimes in the form of M2. They argued that, if subjects usually 'illicitly convert' premisses, then data from 'traditional' premiss pairs containing A premisses should be similar to data from 'modified' premiss pairs containing modifications of the M2 form.

Ceraso & Provitera reported that the above result was indeed obtained and thus concluded that "the subject treats the T syllogism" (sic)" as if it were the corresponding M syllogism, which we take as direct evidence for premiss-misinterpretation". Their study thus appears to provide good evidence for the conversion hypothesis.

However, there are two very good reasons for considering their conclusion to be invalid. Firstly, in certain cases where Ceraso & Provitera report that T data is similar to that from modified premiss pairs using the M2 modification of A premisses, they did not use a modified version of the same premiss pair using the M1 modification of A premisses. Thus, in these cases, there was no 'control'. Similarly, they use only one modification of I premisses and argue that data from premiss pairs including these is similar to data from 'traditional' premiss pairs. Again, there was no 'control' to establish that other modifications would not also produce data similar to the T data.

Secondly, in those cases where traditional premiss pairs containing A premisses are compared with modified premiss pairs using both the M1 and M2 modifications, Ceraso & Provitera report that the T data was similar to that from modified premiss pairs using the M2 modification, but neglect to mention that the T data were equally similar to those from modified premiss pairs using the M1 modification. Clearly, that the T data are similar to the M2 data is no evidence for conversion if they are also similar to the M1 data. In only two cases (their comparison 3 and 7 and comparison 4 and 8) were differences observed between M1 and M2 modifications of the same premiss

pair. In one of these, the M2 modification did produce greater similarity to the T data and Ceraso & Provitera point this out in their discussion. However, in the other case, the M1 modification produced greater similarity to the T data.

Thus, in most cases where the M1 and M2 modifications were properly compared, there was no difference between them and, in the remaining two cases, one difference was in favour of the conversion hypothesis and one difference was against it.

Revlis (1975a) used a standard design to present subjects with a wide range of premiss pairs, half of them determinate, in an attempt empirically to compare the atmosphere and conversion hypotheses. He developed two information processing models based on the two theories and concluded that results yielded more support for the atmosphere theory. The atmosphere model accounted for 88.9% of responses on indeterminate, and 85.2% on the determinate, premiss pairs, whereas the conversion model accounted for only 6.8% and 8.4% of these responses respectively.

However, Revlis interpreted conversion as reformulation of a premiss as its converse. This interpretation leads to some peculiar predictions for certain premiss pairs. For instance, consider the following indeterminate AA premiss pair:

- (1) All A are B
All C are B

Revlis's conversion model assumes that the subject reformulates this as the following premiss pair:

- (2a) All B are A
All B are C

From the above premiss pair, the conclusion 'some C are A' is valid and thus the Revlis conversion model predicts an I conclusion. However, this is very different from Chapman & Chapman's theory of conversion. According to the latter, premiss pair (1) is reformulated as:

- (2b) A is equivalent to B
C is equivalent to B

From (2b) it is deducible that A is equivalent to C, and hence that 'all A are C'. This conclusion accords with atmosphere theory and is typically observed. Thus much of the data which contradicted the Revlis conversion model did not contradict the Chapman & Chapman hypothesis. The Revlis model presumes that premisses are converted and a consistent conclusion looked for. If no conclusion is found, the subject is held to 'unconvert' the premisses (i.e. return them to their original form) and again search for a consistent conclusion. Clearly, this model does not fit the data, although it produces the same predictions as the Chapman & Chapman hypothesis for most premiss pairs.

Revlis (1975b) presents an information processing model of syllogistic reasoning that is essentially the same as the conversion model of Revlis (1975a) with the atmosphere model added. The Revlis (1975a) conversion model had three sequential stages: search for a conclusion to the converted premisses, then (if none found) a search for a conclusion to the original premisses and, finally, (if still no conclusion is found) a guess. Revlis (1975b) replaces the guessing stage with a decision stage based on atmosphere.

The Revlis (1975b) model thus incorporates the idea that the subject uses only the converse of premisses on the first pass. It was demonstrated above (premiss pairs (1), (2a) and (2b)) that this model produces predictions that differ both from those made by the Chapman & Chapman hypothesis and from the observed data. There is no question that the analysis given here misinterprets Revlis, as Revlis (1975b) actually shows premiss pairs (1) and (2a) as an example of his theory (page 108). This is a somewhat peculiar thing for Revlis to do as any reader with some knowledge of the literature would know that the prediction is not empirically confirmed.

More peculiar, however, is the fact that Revlis (1975b) uses the data of Revlis (1975a) to support his model. These data show the conversion model to be hopelessly inadequate. How, then, can these data be used to support what is (up to pass 3) the same conversion model? Revlis (1975b) achieves this by confining his analysis to error rates but he does not say what type of errors. Where his model predicts erroneous responses he quotes high observed error rates as supporting his model and thus tacitly implies that these errors are in accordance with his predictions. (Revlis, 1975a, gives error rates and percentage of decisions correctly predicted; the poor predictive accuracy of the model being evident from the latter but not from the former.)

It is also worthy of note that reference to Revlis (1975a) reveals that his classification of premiss pairs as 'sames' (those that lead to the same conclusion whether or not converted) and 'differents' (those that lead to a different conclusion when converted) is somewhat idiosyncratic but conveniently enables certain other predictions of his model to be validated. However, as the Revlis conversion model does not fit data from syllogistic reasoning tasks, further discussion is not warranted.

Dickstein (1978a) argues in favour of the Chapman & Chapman conversion hypothesis but also extends this position to explain data from those premiss pairs that cannot be accounted for entirely by the theory of conversion. This explanation is in three parts as follows:

- 1) EE, EO, OE and OO premiss pairs all indicate that the middle term is not related to either term of the conclusion. The subject assumes that, if the terms of the conclusion are not related through the middle term, then they are not related. This leads to acceptance of E or O conclusions.

- 2) II premiss pairs lead to invalid I conclusions because the subject assumes that the 'some' referred to in each premiss is the same 'some'. Thus, given 'some A are B', 'some B are C', the subject assumes that some of the B that are A are the same B as are C. This explanation extends to IO and OI premiss pairs although, for these, a conversion operation is involved, leading to an invalid O conclusion.

- 3) IE premiss pairs are explained similarly to the first category with the addition of a conversion operation.

His analysis of data obtained from Dickstein (1978b) provides some support for the cohesiveness of these categories. Percentage correct responses were 73.9, 72.7, 72.7 and 77.3 for the four premiss pairs in category 1 as compared with 50.0, 55.7 and 58.0 for the three premiss pairs in category 2. Dickstein argues that the category 2 error is more pervasive since "it corresponds to general language usage in which the same term in two successive sentences almost always refers to the same referent".

Percentage correct responses to IE premiss pairs was 37.5. Dickstein argues that this is the most difficult as two errors (assumption of nonrelation due to nonrelation with middle term and conversion) are involved. However, this is somewhat spurious as one could equally well argue that OI and IO premiss pairs involve two errors. Given that the conversion step is not counted as an additional error on these two, there is no apparent reason (other than that it would not fit the data) that IE premiss pairs should not have been included in category 1. IE premiss pairs appear to pose problems for all theories, even for atmosphere theory as there are two (E and O) errors that are frequently observed.

The median correlation (of valid responses) between different premiss pairs was 0.648, for those in category 1, and 0.768, for those in category 2. However, the median correlation between premiss pairs in category 1 and premiss pairs in category 2 was only 0.166. The total correct responses on category 1 premiss pairs yielded a correlation of only 0.157 with the total correct responses on category 2 premiss pairs. Thus a subject who was incorrect on one premiss pair was more likely to have also been incorrect on another premiss pair within the same category but not more likely (than other subjects) to have also been incorrect on premiss pairs in the other category. This appears to lend strong support to the hypothesis that a different explanation is required for each category. This does not, however, necessarily imply that Dickstein's explanations are correct, although it should be noted that neither probabilistic inference nor atmosphere theory would predict different error frequencies.

Dickstein's work provides an elaborate and complex explanation of syllogistic reasoning that depends on, at least, the following three factors, other than logical validity:

- 1) Conversion of A and O premisses.
- 2) The assumption that, if both terms of the conclusion are not related to the middle term, then they are not related to each other.
- 3) The assumption that the 'some' in one premiss refers to the same 'some' as does the other premiss.

The complexity is increased by Dickstein's further suggestion that acceptance of 'probable conclusions', and a bias against the 'none of the above' alternative (first proposed by Revlis, 1975a), may also play some role in determining responses.

Of the work considered in this section, Dickstein's position represents the most viable alternative to atmosphere theory, as it explains at least most of the typical atmosphere responses. Although certain of his arguments are similar to those of Chapman & Chapman, regarding both conversion and probabilistic inference, he explains these in more detail. However, although Dickstein's arguments are plausible, there is no real evidence that subjects do behave in the manner that he suggests. The variety of the explanations are an inherent weakness of his position and an explanation in terms of atmosphere must be seen as being far more parsimonious.

However, before concluding this section, it should be noted that atmosphere theory cannot explain all the data. The following two findings indicate that the logical validity of a conclusion plays at least some role in mediating responses:

- 1) Several workers (e.g. Ceraso & Provitera, 1971, Roberge, 1970, Revlis 1975a) have reported that atmosphere effects are most apparent on those determinate premiss pairs whose logically valid conclusion accords with atmosphere.

2) Certain indeterminate premiss pairs (in particular, those referred to as category 1 and category 2 by Dickstein, 1978a) yield high frequencies of valid conclusions. That is, subjects indicate that 'none of the above' (conclusions) follows from the premisses. Such conclusions, which will henceforth be referred to as nonpropositional conclusions, clearly cannot accord with atmosphere.

SECTION 1.3 THEMATIC CONTENT: THE EFFECT OF TRUTH STATUS

Janis & Frick (1943) point out that it is a "widely-held belief that people are likely to be satisfied with unsound arguments if they accept the conclusion to which the arguments lead, and conversely, that they are likely to be unduly critical of sound arguments if they reject the conclusion". This "widely-held belief" has generated much research in the syllogistic literature, although unfortunately, design errors appear to have occurred in some of the earlier studies.

Janis & Frick presented subjects with (whole) syllogisms to evaluate as 'valid' or 'invalid', and subsequently presented them with an 'attitude test' comprised of the conclusions to the syllogisms. On the attitude test, subjects were asked whether they agreed or disagreed with each statement. Analysis of the errors in evaluating the syllogisms showed that significantly more errors occurred when subjects had to accept a conclusion they disagreed with, and least errors occurred when subjects had to reject a conclusion they disagreed with. These results support the "widely-held belief" quoted above in that they suggest that the 'truth status' of a conclusion mediates evaluation of its logical validity. However, Janis & Frick appear to have used Chi-squared tests to analyse data comprised of more than one response per subject and the thematic syllogisms used were, in some cases, very obscure.

Lefford (1946) presented subjects with 20 syllogisms of a "socially controversial nature, which may tend to excite an emotional reaction in the subject" and 20 syllogisms of "a neutral nature and not intended to excite any particular emotional reaction". Subjects were asked to rate the conclusions as valid or invalid and then to go back through the test indicating whether they believed the

conclusions to be 'true' or 'false'. Subjects made more errors on 'emotional' syllogisms and there was a relationship between judged validity and agreement for both types of content. Lefford argued that, whereas attitudes affected 'emotional syllogisms', previous knowledge affected 'neutral syllogisms'. In fact, both may be regarded as 'prior belief', as there is little operational difference between belief acquired from fact or prejudice. However, the data were not fully reported, the analyses were atypical and subjects may well have been influenced by their prior validity judgements when indicating their beliefs. Subjects should at least have been given a separate sheet on which to indicate belief so that they could not see their original evaluations of validity.

One of the most widely quoted studies on belief effects, and possibly the most badly designed, is that by Morgan & Morton (1944). In their main experiment, they presented subjects with 15 thematic premiss pairs and 15 abstract premiss pairs of the same structure. Comparisons showed differences in responses for 14 of the pairs. They claimed that these differences could be explained in terms of a bias towards conclusions in which the subjects believed and concluded from their results that, although responses to abstract syllogisms are primarily determined by atmosphere, responses to thematic syllogisms are primarily determined by personal convictions. They stated that "a person is likely to accept a conclusion which expresses his convictions with little regard for the correctness of the inferences involved".

None of the 'biases' observed in the thematic materials were predicted and thus their post hoc explanations of why subjects preferred particular conclusions must be rigorously scrutinised. Such investigation reveals that few of their explanations are self-evident,

even when considered in their historical context.

Henle & Michael (1956) presented subjects with ten of the pairs used by Morgan & Morton, followed by an attitude questionnaire comprised of all four possible conclusions to each of the ten thematic premiss pairs. Some wording of the thematic problems was changed to make them of current interest. Although results were reported as "agreeing well" with those of Morgan & Morton, preferred conclusions on the reasoning task did not usually match preferred statements on the attitude questionnaire. Further, Henle & Michael point out that the very fact that their data conformed to that of Morgan & Morton's presents problems for the latter's position, as the changes made would be expected to produce different attitudes. For instance, on one problem, Henle & Michael changed the wording from 'India' to 'Russia' and point out that "it would be rash to assume that the attitude of our subjects toward the people of Russia was the same as the attitude of the subjects of the previous investigation to the people of India".

Henle & Michael performed a further experiment using ten thematic premiss pairs concerning Russia or Communism, matched with abstract premiss pairs having the same structure. Subjects were subsequently asked to indicate their attitude toward Russia. 'Anti-Russian' and 'neutral' groups were compared and significant differences found on only one problem, relating to atheism. 'Anti-Russian' subjects were less likely to evaluate 'no atheists are found in Russia' as valid and more likely to evaluate 'some Russians are atheists' as valid.

Given the shortcomings in earlier studies, and the data of Henle & Michael, it would appear that the effect of truth status is weaker than was originally thought. However, there are a variety of

experiments that do provide reasonable support for the effect, although the size of the effect, relative to that of atmosphere, is open to question.

For instance, Gordon (1953) used premiss pairs whose conclusions involved statements about Russia, embedded in a selection of 'neutral' items. Subjects were subsequently presented with an attitude questionnaire comprised of the conclusions to the Russian problems. 11 of the 28 subjects were identified as making non-atmosphere errors consistently favourable or unfavourable towards Russia and, in ten of these cases, the bias was the same as that indicated by the attitude questionnaire. However, Gordon reports that most errors were in accordance with atmosphere, although there were more non-atmosphere errors on the 'Russian' problems.

The initial study reported by Morgan & Morton involved abstract problems, matched with thematic problems which all used the same three terms ('insults', as subject, 'unpleasant', as predicate, and 'troubles' as the middle term). This experiment essentially showed that the nature of the differences between response profiles from the abstract and thematic content varied according to the structure of the problem. However, reanalysis of their data, to produce a comparison across all problems¹, reveals that the thematic group showed a strong bias to reject the statement 'no insults are unpleasant'. Analysis by sign test reveals a highly significant indication of a difference between abstract and thematic problems in the case of 'E' responses ($p < 0.001$) but no significant difference in the case of A, I, O or 'none' responses. There is some, nonsignificant, indication that subjects made more logically valid responses on thematic problems, but this cannot be advanced as an explanation of the differences in E responses.

The above result makes certain of the Frase (1966a, 1968a) findings somewhat surprising. He used conclusions that were measured as to compatibility, in terms of subjects' responses to a semantic differential test. Thus a conclusion such as 'no insults are unpleasant' would presumably have high incompatibility. However, Frase reports very small, but significant, effects, such that compatibility facilitated correct responses and argues that, as three-quarters of his syllogisms were invalid, this must imply that compatible conclusions were more often (correctly) rejected. Thus, according to Frase, invalid incompatible conclusions, such as 'no insults are unpleasant', are more likely to be accepted than invalid compatible conclusions. The Frase results are, in fact, apparently contrary to all the other reported effects of truth status. However, in the absence of support for his position, the results may be overlooked for two reasons:

- 1) The effects are marginal and not entirely consistent across the two studies, and it is debatable whether 'compatibility' is in all cases correlated with truth status. It is possible to pair two terms, both having a high positive or negative affective 'tone', and produce a meaningless statement.
- 2) Frase does not separately analyse valid and invalid syllogisms. It is thus merely an assumption that invalid 'compatible' conclusions are more often rejected than invalid 'incompatible' conclusions.

Thouless (1959) did fail to find any significant indication of an effect of belief in a group of student subjects, although he obtained such an effect on a group of adult subjects. He reported that some subjects had apparently guessed the intention of the task. However, Thouless's method was to ask subjects whether they agreed

or disagreed with the conclusions before asking them to indicate whether they believed them to be validly derived. This may well 'alert' test-sophisticated subjects.

In fact, the procedure of asking subjects to evaluate given conclusions, both on the basis of logical validity and truth status, within a short period of time, may cause distortion of results, whatever the order in which the subject is asked to perform these two tasks. Kaufman & Goldstein (1967) avoided this difficulty by using syllogisms with conclusions that a different sample had rated for agreement or disagreement. They selected conclusions that showed 80%, or above, agreement or disagreement (referred to as 'positive affect', PA, and 'negative affect', NA, respectively) and formulated a selection of items that were "presumably neutral" (N).

They observed strong effects of truth status, although these effects were complex and difficult to interpret. Essentially, they observed three important findings:

- 1) Although subjects rejected less valid, and accepted more invalid, conclusions to PA than NA arguments (in accordance with the effect of truth status), a significant effect was only observed for acceptance of invalid arguments. This result may well explain why Henle & Michael only observed an effect on one of the ten syllogisms used in their second experiment. Six of these syllogisms were valid and it appears, from the Kaufman & Goldstein data, that a strong effect of truth status is only observed on acceptance of invalid arguments.
- 2) There was also little effect of truth status on evaluation of syllogisms having universal conclusions. Thus the only large difference observed between responses to PA and NA

sylogisms was on the acceptance of invalid particular conclusions (68% to 36%). However, it should be noted that they only studied affirmative conclusions and that the initial Morgan & Morton (1943) experiment suggests that a strong effect may be obtained on negative universals.

- 3) More errors were observed on neutral items than on either type of affective item. This was primarily due to a far greater tendency to accept invalid universal conclusions to 'neutral' items (48% for N, as against 15% for PA and 8% for NA). In fact, across all syllogisms, subjects accepted less particular than universal conclusions to 'neutral' items, but accepted more particular than universal conclusions to 'affective' items. (This interaction was significant).

These last results could be related to the Frase (1966b) finding that there is a relationship between quantification and belief, such that subjects show a greater tendency to indicate belief in quantified than in nonquantified statements. Thus, both PA and NA 'affective' items may have been more 'believable' when quantified, whereas 'neutral' items, not relating to belief, would possibly not be expected to be affected in this way by quantification. However, it is unclear whether the data can be explained in these terms as, unfortunately, Kaufman & Goldstein do not give examples of the type of content used and give little indication of the nature of the 'neutral' items.

Another well-designed experiment, that uncovered new information as to the effect of truth status, was that by Feather (1964). He presented subjects with syllogisms concerning religion (half with favourable, and half with unfavourable conclusions) randomly embedded in a selection of 'neutral syllogisms'. A religious attitude test was

administered one week later, together with a test of 'intolerance of ambiguity' developed by Budner (1962). Separate analyses were carried out for subjects classified as 'proreligious' and (a much smaller number of) subjects classified as 'anti-religious'.

For the proreligious group, significantly more (reasoning) errors were in the proreligious direction. Further, the extent of this bias was significantly correlated with both the strength of proreligious attitude and intolerance of ambiguity, and inversely correlated with the number of (logically) correct evaluations on the 'neutral' syllogisms.

The errors of the anti-religious group, however, showed no evidence of bias. This was possibly due to the finding that anti-religious subjects had a significantly lower intolerance of ambiguity, and made significantly less errors on the neutral syllogisms, than proreligious subjects.

In general, the results of Feather (1964) clearly demonstrate that truth status is a source of bias, but suggest that susceptibility to this bias is a function of strength of attitude and of certain personality factors. However, the more able a subject is correctly to evaluate ('neutral') syllogisms, the less susceptible he is to the bias.

Revlin & Leirer (1978) have attacked the view that reasoning responses may be influenced by personal bias. They suggest that typical results concerning 'belief' may be explained in terms of 'conversion blocking'. Their theory is best explained with reference to the following (invented) syllogism which is invalid:

All honest people are good

All religious people are good

∴ All honest people are religious

If truth status affected evaluation of this syllogism, then proreligious subjects would be expected to accept the conclusion more often than non-religious subjects. Revlin & Leirer would explain such a response, not as due to the subjects' belief that religious people are honest, but due to his conversion of the premiss 'all religious people are good'. The point is that a proreligious subject may have a greater 'set equivalent' perception of 'good' people and 'religious' people. To the nonreligious subject, it will be more readily apparent that not all 'good' people are religious, and thus his conversion will be blocked. That is, according to Revlin & Leirer, his experience will indicate that the conversion is invalid and he will thus be less likely to accept the invalid conclusion.²

This theory was tested in their first experiment by use of 'neutral' premiss pairs, that did not relate to 'real life' knowledge, but which led to conclusions that did relate to 'real life' knowledge. According to their theory, there thus should be no effect of truth status of conclusion, as there should be no differential (premiss) conversion. For instance, in one such problem, the two premisses expressed a relationship between an unnamed committee (the middle term) and 'women' and 'U.S. senators'. Subjects would thus have 'real life' knowledge relating to the relationship between the terms of the conclusion ('women' and 'U.S. senators'), but no knowledge about the relationship between either of these terms and the middle term. In some cases the logically valid conclusion agreed with the 'true' (real life) conclusion, and in others it did not.

Although observing an effect of truth status, the authors report that it was very marginal. However, this can hardly be taken as contradicting earlier studies for a variety of reasons. Firstly, the Kaufman & Goldstein results suggest that a strong effect is only

observed on invalid syllogisms and stronger effects thus may³ be expected on indeterminate premiss pairs, which Revlin & Leirer did not use. Secondly, on those arguments for which Revlin & Leirer claim that the valid answer conflicts with the 'true' answer, this is not entirely the case. For instance, although it is true that 'no Arabian sheiks are U.S. senators', the 'conflicting' logically valid answer 'some Arabian sheiks are not U.S. senators' is also true. Thus subjects are not necessarily accepting valid conclusions that they do not believe. Thirdly, it appears that Revlin & Leirer did not control for the effect of atmosphere. As the logically valid conclusions accorded with atmosphere, and as Gordon (1953) has specifically reported truth status to be weaker than atmosphere bias, it is not at all surprising that the effect of truth status was "quite limited". Finally, it is worth noting that reported 'accuracy' (Table 3) was 83% when 'belief' accorded with logic, and 67% when it conflicted with logic. Given the above points. the size of the effect is reasonably compatible with that observed in other studies.

However, it should be noted that Revlin & Leirer employed an excellent control for differential 'conversion blocking' by using premisses that did not relate to 'real life' knowledge. In fact, conversion itself was controlled for, as only E and I premisses were used. Conversion of E or I premisses does not affect the logic of the argument. Thus on the basis of the above argument, that the size of the truth status effect observed is compatible with other studies, their experiment appears to indicate that the 'conversion blocking' explanation is wholly unfounded.

In their second experiment, Revlin & Leirer used premiss pairs that were claimed to have "real world truth values (of a sort)", and conclusions that did not. Subjects were also presented with a

questionnaire that investigated their knowledge of the relationships between various terms used in the premisses. For instance, for the premiss "all blacks in Neuberg are welfare recipients", subjects were asked on the questionnaire to assess the percentage of blacks who are welfare recipients and the percentage of welfare recipients who are black. Revlin & Leirer argued that, the higher these percentages are thought to be, the more likely it is that the premiss, "all blacks in Neuberg are welfare recipients", will be converted.

On the basis of the questionnaire responses, the authors predicted conversion errors on the syllogistic reasoning task and reported a reasonable degree of confirmation of these predictions. The data were not fully reported, but it appears that subjects made more errors consistent with the conversion hypothesis, if they reported high 'set equivalent' percentages on the questionnaire. Unfortunately, the authors present this result as a demonstration that truth status bias is, "at least to a considerable degree", a resultant of differential premiss encoding", an interpretation that was seen to be unlikely on the basis of the results of their first experiment.

Clearly, if certain responses on abstract problems are a resultant of premiss conversion, then it may well be that such conversion may be blocked on thematic problems using premisses such as 'all rabbits are animals', as subjects know that not all animals are rabbits. Thus this may well play some role in mediating observed differences between data from thematic and abstract problems. For instance, Morgan & Morton (1944) report a variety of differences, which Henle & Michael (1956) have pointed out cannot all be due to the effect of truth status of conclusion. It is quite possible that some of these differences were due to an effect of the 'real life' truth status of the premisses.

However, it would be very difficult to explain how 'conversion blocking' could explain all the results of 'truth status' experiments, particularly those that have compared one thematic problem with another. In fact, there is direct evidence from Revlin & Leirer's first experiment that 'conversion blocking' cannot explain all the data and their study appears to indicate that there may be separate effects of truth status on both premisses and conclusion.

SECTION 1.4 THE FORMAL LOGIC OF SYLLOGISMS: FIGURE

There are four possible ways in which the subject (S) and predicate (P) terms of the conclusion and the middle term (M) may be arranged in the two premisses. These are referred to as figures of the syllogism and are shown in Table 1.1.

TABLE 1.1

The Four Possible Arrangements (Figures) of the Subject (S),
Predicate (P) and Middle Term (M).

	<u>FIGURE 1</u>	<u>FIGURE 2</u>	<u>FIGURE 3</u>	<u>FIGURE 4</u>
<u>MAJOR PREMISS</u>	M-P	P-M	M-P	P-M
<u>MINOR PREMISS</u>	S-M	S-M	M-S	M-S
<u>CONCLUSION</u>	S-P	S-P	S-P	S-P

It is traditional to write the major premiss first, but the premiss pairs are logically interchangeable. Thus the following argument is also in the first figure:

$$\begin{array}{c}
 S-M \\
 \underline{M-P} \\
 \therefore S-P
 \end{array}$$

Due to this interchangeability of premisses, it should be noted that four further figures cannot be obtained by reversing the conclusions. If the conclusion of a figure 2 syllogism is reversed it simply becomes a (different) figure 2 syllogism with interchanged

premisses. The same applies to figure 3. If the conclusion of a figure 1 syllogism is reversed then it becomes a figure 4 syllogism with interchanged premisses. Similarly, figure 4 syllogisms transpose into figure 1.

This reveals an important relationship between figure 1 and figure 4: they are identical as far as premiss pairs are concerned. Premiss pairs alone can define a syllogism as figure 2 or as figure 3, but can only define a syllogism as either figure 1 or figure 4. In this latter case, the figure of the syllogism is defined by the direction of its conclusion. For instance, consider the following premiss pair:

All A are B

No B are C

This may become either a figure 1 or a figure 4 syllogism dependent upon the conclusion. Consider first the (valid) conclusion:

No A are C

This produces a figure 1 syllogism (with reordered premisses). Alternatively, consider the (equally valid) conclusion:

No C are A

This produces a figure 4 syllogism. There are thus four possible structures for syllogisms but only three possible structures for premiss pairs. In fact, Aristotle recognised only three syllogisms, making no distinction between figure 1 and figure 4.

It will be noted that the 'standard' design of presenting subjects with four alternative propositional conclusions, all in a given direction, defines the figure of the syllogism for figures 1 and 4. In this context, then, premiss pairs may be said to be in any one of the four figures. However, if no direction of conclusion is specified, then no distinction may be made between figure 1 and figure 4 and premiss pairs can be said to be in one of only three figures.

SECTION 1.5 EXPERIMENTS ON THE EFFECT OF 'FIGURE'

Frase (1968b) was the first worker to consider specifically the effect of figure on syllogistic reasoning. His idea was that subjects solve syllogisms by setting up associations between the terms and then assessing the strength of association in the conclusion. The building of associations is comparable, he argued, to the processes involved in paired associate learning.

Frase was primarily concerned with those premiss pairs that may become figure 1 or figure 4 syllogisms dependent on direction of conclusion. Henceforth, these will be referred to as figure 1/4 premiss pairs. Given that premiss order is logically irrelevant, these premiss pairs are all in the following form:

A-B (minor premiss for figure 1, major premiss for figure 4)

B-C (major premiss for figure 1, minor premiss for figure 4)

Frase suggested that this may lead to an associated chain (A-B-C) and that, if the given direction of conclusion specifies a figure 1 syllogism (A-C), it is in accordance with this chain, but, if the given direction of conclusion specifies a figure 4 syllogism (C-A), the subject must evaluate a 'backwards' chain. He reported two experiments, presenting subjects with whole syllogisms, which together yield reasonable evidence that subjects make less errors when offered an A-C conclusion to figure 1/4 premiss pairs, than when offered a C-A conclusion. Roberge (1971a) also reports data showing that most errors are observed on figure 1/4 premiss pairs for which the given conclusions specify a figure 4 syllogism.

Pezzoli & Frase (1968) worked with figure 2 and figure 3 syllogisms. The Frase position is that figure 2 is comparable to a stimulus equivalence paradigm in paired associate learning, in that the same (middle) term is paired to subject and predicate of conclusion. Figure 3 is comparable to a response equivalent paradigm, in that two different terms (subject and predicate) are paired to one stimulus term (the middle term). Pezzoli & Frase argued that, as Figure 3 requires different responses to be associated to the same stimulus, it can produce interference and hence more reasoning errors, whereas figure 2 syllogisms should facilitate reasoning. Further, they argued that the stronger the association between the premiss terms, the stronger will be the interference effect on figure 3 syllogisms and thus the more apparent these differences should become.

In one condition of their experiment, realistic terms were used such that, although the terms of the conclusion were not associated, there was a high association between each of these terms and the middle term. This condition led to considerably more figure 3 than figure 2 errors. The finding supports their theory that associations will increase interference on figure 3 syllogisms and they concluded that "the present study has shown that mediated associations play a significant role in deductive reasoning". However, differences between figure 2 and figure 3 need to be investigated in a wider context as this study used only five syllogisms, all having I conclusions.

It should be noted that Pezzoli & Frase argued that differences between figure 2 and figure 3 should be increased by the presence of associations but that Pezzoli & Frase, and Roberge (1971a), found no difference between figure 2 and figure 3 when abstract content was used. Thus, although the Pezzoli & Frase study shows that

the presence of association may affect reasoning behaviour, the results do not necessarily imply that abstract syllogisms are solved by an associational process.

However, although the position regarding figure 2 and figure 3 is unclear, it does appear that figure 1 yields less errors than figure 4. The best explanation of this appears to be Frase's theory that subjects prefer 'forwards' conclusions.

Dickstein (1978b) argues that, if subjects do prefer 'forwards' conclusions, then figure 4 will lead them to conclusions that are the reverse of those given and he suggests that subjects convert these to match them to the given alternatives. Clearly, if the conclusion is valid in both directions, this conversion will yield the correct answer. Where the conclusions differ, more errors would be expected on figure 4 than figure 1.

To test this possibility it is preferable primarily to consider only those (nine) premiss pairs that are consistent in the above respect across all four figures. Dickstein reports two experiments that employed a standard design to present subjects with all premiss pairs in all four figures. Although there was a highly significant figure effect overall, specific support was obtained for Dickstein's theory. Two main findings are of primary importance:

- 1) A significant figure effect (such that figure 4 yielded most, and figure 1 least errors) was observed for EI premiss pairs. There is a valid conclusion for each of these but no valid reverse conclusion. Most (invalid) 'no conclusion' responses were observed on figure 4. A similar significant figure effect (although figure 4 did not produce the most errors) was observed for IE premiss pairs. There is no valid conclusion for each of these but there is a valid reverse conclusion. Most (valid) 'no conclusion' responses were observed on figure 1.
- 2) No figure effect was observed on EE, EO, OE, OO, II, IO and OI premiss pairs which, in all four figures, yield no valid conclusion in either direction.

These results lend strong support to the view that figure 1 is least prone to error resulting from invalid conversion of reversed conclusions. Dickstein reports further corroboration of this viewpoint based on certain premiss pairs in some figures only. He points out that conversion of conclusions is an extension of Chapman & Chapman's theories. Essentially, Dickstein argues that the subject is "assuming a symmetrical relation between the classes represented by the terms". This assumption of symmetry thus applies to both premisses and conclusion. Dickstein's suggestion that subjects may convert conclusions is, of course, additional to those explanations given for Dickstein (1978a) (see section 1.2).

The definitive experiment on figure was performed by Johnson-Laird & Steedman (1978). Other workers have noted that figure 4 produced more errors than figure 1 and have presumed that this was due to a preference for 'forward processing'. Johnson-Laird

& Steedman presented subjects with premiss pairs only and allowed them to write their own conclusions. Subjects constructed more 'forwards' than 'backwards' conclusions for figure 1/4 premiss pairs, except in the case of one of the two premiss pairs that allow only backwards conclusions. In cases where conclusions may be stated in either order, this bias was overwhelming and valid deductions compatible with the bias were made on over 80% of occasions, whereas valid deductions incompatible with the bias were made on only 20% of occasions.

The authors report no bias for figure 3 or indeterminate figure 2 premiss pairs. However, for determinate figure 2 premiss pairs which allow only one valid conclusion, subjects made significantly less errors when the subject of the conclusion appeared in the second premiss (67.5%, as opposed to 30%, correct). There was also an effect of premiss order on figure 1/4 premiss pairs such that order (a) produced less errors than order (b):

(a) A-B
B-C

(b) B-C
A-B

This represents further confirmation for the idea of 'forward chain processing' as (a) clearly facilitates this. Further, the above difference only applied to determinate premiss pairs. Facilitation of the construction of a chain can only facilitate valid responses when a valid conclusion can be derived.

The authors present a theory of syllogistic reasoning processes (a preliminary account of which is given by Johnson-Laird, 1975) that is based, in part, on this preference for forward processing. Essentially, subjects are held to formulate a representation of the premisses by imagining one class and 'tagging' it with (or without) the attribute represented by the other class as follows:

<u>All A are B</u>	<u>Some A are B</u>	<u>Some A are not B</u>	<u>No A are B</u>
a a	a (a)	a a (a)	a a
↓ ↓	↓	⊥ ⊥ ↓	⊥ ⊥
b b (b)	b (b)	b b b	b b

↓ denotes a connection and ⊥ denotes that two terms are known not to be connected. An absence of either symbol denotes an indeterminate connection. Brackets indicate that there may be a further instance of the class. Thus, for the I premiss, the representation denotes that there may be some As that are not Bs and (or) some Bs that are not As.

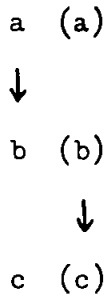
When premisses are combined, the next premiss is added on to the first representation. Two sorts of error are held to arise at this stage, as the theory postulates a bias towards forming connections. Examples of these are shown below:

- (1) Some A are C, Some B are C (2) No A are B, No B are C

a (a)	a a
↓	⊥ ⊥
b (b)	b b
↓	⊥ ⊥
c (c)	c c

The above will lead to an invalid I conclusion at (1), due to the middle term, B, having been assumed to be common to both premisses. Although (2) is a correct representation, it may lead to error due to the subject noting that there is a ⊥ symbol in all paths between A and C. Once these invalid conclusions are derived,

the theory holds that the subject puts them to a sequence of tests that manipulate the representation. Successful subjects, for instance, will alter (1) to:



There is now no line linking A and C (although the brackets show that a link is possible) and the premiss pair will be correctly judged to be indeterminate. (1) and (2) above are examples of Dickstein's category 2 and category 1 premiss pairs respectively. However, it should be noted that the Johnson-Laird & Steedman data do not show a differential error rate between category 1 and category 2 premiss pairs, error rates being low in both cases.

The most appealing aspect of the Johnson-Laird & Steedman model is its utilisation of forward processing. Consider the representation of the following AA premiss pairs:

Figure 1/4

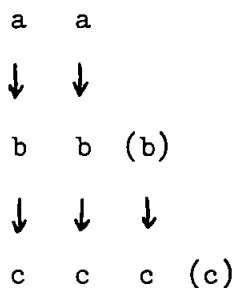


Figure 2

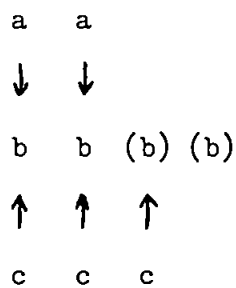
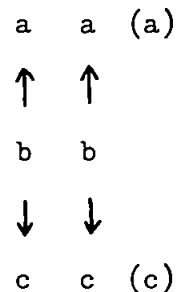


Figure 3



Connections between A and C are more difficult to derive from figures 2 and 3 as the A-B and B-C connections are in opposite directions. The authors point out that this is consistent with their finding that there were less errors on indeterminate figure 2 and 3 premiss pairs. The readily available connection in figure 1/4 increases the possibility of making an invalid propositional conclusion on indeterminate premiss pairs. More important, however, is the direction of the connection in figure 1/4. It is not surprising that A-C conclusions are constructed, and that error rates are high, in cases where the only valid conclusion goes from C to A, against the direction of both arrows. The forward processing findings are thus central to the theory of how the premisses are represented.

Although Johnson-Laird & Steedman's model is presented as an alternative to atmosphere theory, it is of interest to note that the preferred propositional errors were in accordance with atmosphere in almost all cases. Of more interest are the effects of atmosphere on determinate premiss pairs. There are two such effects as follows:⁴

- 1) All determinate figure 1/4 premiss pairs that yield 'forwards' conclusions yield conclusions in accordance with atmosphere, and frequencies of valid conclusions were very high on these problems (81% and 88% one week later). However, determinate figure 1/4 premiss pairs that yield only 'backwards' conclusions do not always yield conclusions in accordance with atmosphere. In the case of these premiss pairs, the percentage of logically valid conclusions was higher when the conclusion accorded with atmosphere, than when it did not (20% to 3%, and 33% to 8% one week later).

2) For figure 3 premiss pairs, the percentage of logically valid solutions was far greater when the valid conclusion accorded with atmosphere than when it did not (63% to 23% and 82% to 45%, when the same subjects were tested one week later).⁵

SECTION 1.6 CONCLUSIONS

Much of the research on syllogisms has focussed on the question of whether atmosphere bias is, or is not, a satisfactory explanation of the data. It should be noted that the data of Chapman & Chapman (1959) were not inconsistent with the atmosphere predictions, except in one respect: the presence of an E premiss reducing the dominance of particular atmosphere. However, their (alternative) conversion hypothesis has proved very popular, although none of the experiments discussed in section 1.2 has produced results indicating that it is a better explanation.

A subsidiary finding of Revlis (1975a) suggests that both conversion and atmosphere explanations may be applicable. He reported that subjects made more errors on indeterminate premiss pairs that become determinate when premisses are converted than on indeterminate premiss pairs that cannot be made determinate by conversion. Thus, more atmosphere responses are observed when the response accords with the conversion hypothesis. As error rates increase when both atmosphere and conversion may be said to be operating, this suggests that both factors may influence the data.

Similar to the above is the finding that atmosphere responses are most frequent when they accord with a logically valid conclusion (to a determinate premiss pair), which clearly suggests that logical validity is a further factor that may be said to mediate responses. Further support for this is provided by the fact that valid nonpropositional conclusions are frequently observed on certain premiss pairs. However, although an increase in atmosphere responses, when they are valid, attests to an effect of logical validity on responses, it does not indicate that all such responses are determined by validity. In fact, subjects frequently fail to draw the valid

conclusion when it does not accord with atmosphere, which suggests that atmosphere is, at least, an equally important determinant.

It is implicit in the conversion hypothesis that the subject makes a logically valid inference, given the misinterpretation of one or more premisses as implying their converse. Just two distinct factors (logical validity and atmosphere) may thus be identified as mediating responses, one relating to the structure of the problem and the other a response bias. These may be referred to as 'logical' and 'nonlogical' factors respectively, after Evans (1972a). It will be noted from the above that response frequencies are highest when both factors favour the response.

A further 'nonlogical' factor is the effect of the truth status of an argument's conclusion. Although some of the earlier work on this is open to criticism, certain of the later studies, in particular those of Kaufman & Goldstein (1967) and Feather (1964), appear to be well designed. The results of the Revlin & Leirer experiments suggest that use of thematic content may affect both the logical and nonlogical factors that influence responses on abstract tasks. A new nonlogical factor (truth status of conclusion) may come into operation and, if the subject attends to the structure of the problem, his responses may well be influenced by 'real life' knowledge or belief that relates to the premisses.

The idea that responses to syllogistic tasks are partly mediated by nonlogical factors has been recently challenged by Dickstein (1978a, 1978b) and also by Johnson-Laird & Steedman (1978). Dickstein (1978a) discusses a variety of apparently unrelated factors that may mediate responses but, unlike Johnson-Laird & Steedman, he does not present these within the context of a cohesive theory. In terms of

parsimony, the multiplicity of these factors is an inherent weakness of his position and he fails to demonstrate that they provide a better explanation than that provided by the more parsimonious atmosphere theory.

The Johnson-Laird & Steedman (1978) model explains the data in a similar way to Dickstein. For instance, the model accounts for possible errors resulting from conversion or a preference for 'forwards' processing and for errors arising from the use of particular or negative premisses. However, the authors bring a cohesiveness to these explanations by attempting to show that they are all a function of the way in which the subject represents the premisses. The foundation of this model is centred upon the observed preference for 'forwards' conclusions, which appears to have been well validated. However, there is no direct evidence that subjects attempt to solve syllogistic reasoning problems in the way that Johnson-Laird & Steedman suggest.

They attack 'atmosphere' theory on the grounds that it cannot explain preferred direction of conclusion. However, this is an attack upon the extreme position that atmosphere is the only determinant of responses. Atmosphere theory only makes predictions about mood but other factors may influence other features of the conclusion.

On indeterminate premiss pairs, the preferred propositional errors observed by Johnson-Laird & Steedman were generally in accordance with atmosphere and thus their explanations of these can only be seen as an alternative to atmosphere theory. In the case of determinate premiss pairs, however, it is difficult to envisage the model being able to provide an alternative explanation

to that of atmosphere. Thus, although it is possible that further development of the Johnson-Laird & Steedman model may provide a viable alternative point of view, there is no good reason at present to abandon the view that responses are determined by a combination of logical and nonlogical factors. However, the Johnson-Laird & Steedman data do suggest that certain designs may increase the relative effect of logical factors.

FOOTNOTES

1. This comparison may be derived from Table 1 of Morgan & Morton. However, the reader should note that it is apparent, from their discussion and, particularly, from Table 2 and Table 3, that the designations 'S' and 'C' in Table 1 are in the wrong order in all cases except that of problem 1.
2. It should be noted that Revlin & Leirer make no attempt to show, for instance, by content analysis of various studies, that conversion blocking can explain truth status effects on individual problems used by other experimenters.
3. It is, of course, accepted that determinate premiss pairs are not strictly comparable with valid syllogisms, as the former offer subjects one valid and three invalid conclusions
4. Based on analysis of the data given in the Appendix of the paper. Only two such analyses of the effect of atmosphere on determinate premiss pairs can be made, as all determinate figure 2 premiss pairs yield valid conclusions in accordance with atmosphere.
5. Six of the nine determinate figure 3 premiss pairs yield a conclusion in accordance with atmosphere.

CHAPTER 2
PROPOSITIONAL REASONING

SECTION 2.1	Logical Constants within the Propositional Calculus	60
SECTION 2.2	Truth Table Experiments	68
SECTION 2.3	Conditional Inference Tasks	77
2.3.1	Abstract Conditional Inference Tasks	77
2.3.2	Thematic Conditional Inference Tasks	83
SECTION 2.4	Disjunctive Inference Tasks	90
2.4.1	Differential Error Rates	90
2.4.2	The Denial of a Negative: An Illusory Problem?	92
2.4.3	Thematic Disjunctive Inference Tasks	98
SECTION 2.5	Conclusions	100

SECTION 2.1 LOGICAL CONSTANTS WITHIN THE PROPOSITIONAL CALCULUS

The propositional calculus is concerned with relationships between propositions. A proposition has been defined as "anything which can be said to be true or false" (Cohen & Nagel, 1934). For any two propositions, (P and Q), there are thus four possible combinations of truth or falsity as shown below:

<u>P</u>	<u>Q</u>
T	T
T	F
F	T
F	F

These may be referred to as the TT, TF, FT and FF truth table cases.

There are four basic logical constants used to connect propositions and these may be defined in terms of their truth tables. That is, tables that define their truth or falsity in each of the four truth table cases. The definition of the four basic connectives is shown in Table 2.1.

The \cdot relation, known as conjunction, is relatively simple and expresses the assertion that both P and Q are true.

The \supset relation is true unless P is true and Q false (the TF case) and is referred to as expressing material implication.

The \vee relation is true unless both P and Q are false (the FF case) and may be referred to as inclusive disjunction. In common parlance, exclusive disjunction is often used, which does not have a special symbol. This relation is false not only when both P and Q are false, but also when both P and Q are true. Thus exclusive disjunction demands that one component be true and that the other

component be false.

The \equiv relation is referred to as expressing material equivalence and is true if either both or neither component is true. It is logically equivalent to the assertion: $(P \supset Q) \cdot (Q \supset P)$.

Table 2.1

Truth Table Definitions of Four Logical Connectives Used to Connect Two Propositions P and Q

<u>TRUTH TABLE CASE</u>		<u>FORMS OF CONNECTION</u>			
<u>P</u>	<u>Q</u>	<u>P · Q</u>	<u>P ⊃ Q</u>	<u>P ∨ Q</u>	<u>P ≡ Q</u>
T	T	T	T	T	T
T	F	F	F	T	F
F	T	F	T	T	F
F	F	F	T	F	T

The constituent propositions (P and Q), of the above relations, may refer to classes of events that occur on more than one occasion. In such cases, one falsifying case is sufficient to prove the relation false (however many verifying cases have been observed). Thus, for instance, $P \supset Q$ is only true if there are no TF instances.

The falsity of a proposition, P, may be referred to as \bar{P} which, in set terminology, may be interpreted as 'anything that is not P'. Similarly, the falsity of Q may be expressed as \bar{Q} . Only expressions linking P and Q have so far been considered, but \bar{P} and \bar{Q} may also be involved in a combination. There are thus four possible implication relationships, and four possible (inclusive) disjunctive relationships, each with its own truth table. A convenient way to refer to these is

as AA, AN, NA, or NN (implication or disjunctive) rules, which denotes the affirmation or negation of each component. The truth tables for all these rules are shown in Table 2.2, which is drawn in such a way as to show the connection between implication and inclusive disjunction. For each material implication rule in the first column, there is an inclusive disjunction rule in the last column that has the same truth table.

Table 2.2

The Truth Tables for all Four Possible \supset and \vee Relations

MATERIAL IMPLICATION RULES	TRUTH TABLES				INCLUSIVE DISJUNCTION RULES
	P Q	P Q	P Q	P Q	
	<u>T T</u>	<u>T F</u>	<u>F T</u>	<u>F F</u>	
$P \supset Q$ (AA)	T	F	T	T	$\bar{P} \vee Q$ (NA)
$P \supset \bar{Q}$ (AN)	F	T	T	T	$\bar{P} \vee \bar{Q}$ (NN)
$\bar{P} \supset Q$ (NA)	T	T	T	F	$P \vee Q$ (AA)
$\bar{P} \supset \bar{Q}$ (NN)	T	T	F	T	$P \vee \bar{Q}$ (AN)

Similar tables may be constructed for material equivalence and exclusive disjunction rules, which are also logically connected.

It is important to note that the middle columns of Table 2.2 refer to the truth or falsity of P and Q. If, however, the cases are made to apply to the actual propositions contained in the rule, so that, for instance, the TT case of $P \supset \bar{Q}$ is $P\bar{Q}$, then all four rules will have the same standard truth table (i.e. that given for the AA rule). This latter notation has certain advantages and will be adopted henceforth.

The \supset and \vee constants are fundamental to the propositional calculus and could be argued to be fundamental to the construction of logical arguments in real life. However, if subjects use of these constants is to be investigated, a way must be found to express them verbally.

The best translation of the inclusive disjunction relation is probably 'either P or Q'. The phrase 'or both' possibly needs to be added to make the meaning wholly clear and to differentiate inclusive from exclusive disjunction, which may be expressed as 'either P or Q but not both' (corresponding to an \equiv relation). Strawson (1952), although pointing out that 'or' is not an exact translation of \vee in all cases, concludes that 'or' is the best available translation.

However, the translation of \supset is more difficult. A word used by logicians is 'entails', where 'P entails Q' is defined as 'P and not Q is inconsistent' (Strawson). This word would appear to be the solution to the problem of translating \supset . Unfortunately, 'entails' is an infrequently used word in ordinary discourse and cannot be used in a 'natural' way to express implication, except in a few cases. The closest, frequently used, natural equivalent of $P \supset Q$ that embodies the concept of entailment, is 'if P then Q'. Such statements are known as conditionals, the first component (P) being referred to as the antecedent and the second component (Q) as the consequent.

There are, however, various other possibilities, for instance: 'never P without Q'; 'whenever P, then Q'; 'Q if P'; the universal, 'all P are Q'; or even the disjunctive, 'either not P or Q'. All these forms may be used to express implication and are thus logically equivalent but it would be naive to assume that they are necessarily psychologically equivalent.

Further, Strawson points out that the truth of $P \supset Q$ may be established in cases which would not constitute verification of a conditional. The falsity of P is sufficient for the truth of $P \supset Q$, whereas conditionals make statements about Q , contingent upon the occurrence of P , and are not usually considered verified when P fails to occur. Kneale & Kneale (1962) point out that persons are not generally considered to have kept (conditional) promises, or obeyed conditional orders, when the conditions for those promises or orders have not been fulfilled.

Strawson also points out that counterfactual conditionals do not correspond to $P \supset Q$ statements. Counterfactuals all have antecedents that are known to be false and thus, if they were the same as implication, would all be necessarily true. For instance, the counterfactual statements, 'if Hitler had had the atomic bomb, he would have won the war' and 'if Hitler had had the atomic bomb, he would (still) have lost the war', would both be necessarily true, as Hitler was not in possession of such a weapon. Strawson argues that such statements cannot be taken as always true in real life, otherwise there would be no point in making them. Lewis (1973) argues that counterfactual conditionals cannot be described in terms of implication, if so, as logically valid arguments may be constructed that have true premisses but false conclusions.

However, it is not the case that conditionals have only recently been considered (and rejected) as possible translations of implication. Kneale & Kneale report a history of debate on conditionals in their own right originating in antiquity. Much of this debate concerns whether a conditional implies implication. Kneale & Kneale conclude that a conditional is best described by assigning it no truth value when the antecedent is false. This is referred to as defective

implication and yields the following truth table for the statement 'if P then Q'.

<u>P</u>	<u>Q</u>	<u> </u>
T	T	T
T	F	F
F	T	?
F	F	?

Similar tables can be constructed for the three other possible conditional rules. It can be seen that the first two entries are as for $P \supset Q$. Kneale & Kneale explain that the ? (they use ..) does not mean that the conditional is something other than either true or false, but simply that it is inapplicable. This also appears to handle the problem of counterfactuals as they become neither true nor false. It is clearly impossible to actually verify or falsify a counterfactual and thus the FT or FF case represented by the real state of affairs is inapplicable to evaluating it.

This defective truth table can also be applied to other forms that use 'if', such as 'Q if P' and to the universal. These forms, then, cannot be taken as an exact translation of $P \supset Q$, although there are certain features in common. Strawson concludes that the best expression of implication is 'not both P and not Q'. However, this is somewhat too simplistic for psychological purposes. The expression of $P \supset \bar{Q}$ would be 'not both P and Q'. This does not appear to convey any impression of entailment. A reasonable expression of implication may be disjunction but 'if P then Q' appears to be psychologically different to 'either not P or Q' and the latter form, again, does not seem to convey much impression of

entailment. In any case, the disjunction cannot be used to express implication if it is desired to investigate both disjunction and implication.

Thus the conditional (or similar forms) appears to be the best choice for the psychologist who wishes to study entailment. The conditional has a long history of association with implication and thus is perhaps a better choice than other, similar, forms. It should be remembered, however, that this does not represent a study of $P \supset Q$.

There are inferences that may be drawn from the conditional 'if P then Q' that parallel inferences which may be drawn from $P \supset Q$. Two valid inferences are modus ponens (MP) and modus tollens (MT). The MP inference infers 'Q', given 'P', and the MT inference infers 'not P', given 'not Q'. It can be seen that these inferences depend upon the fact that 'P' and 'not Q' cannot occur together if the conditional is true.

Two invalid inferences are the affirmation of the consequent (AC) and the denial of the antecedent (DA). The AC inference infers 'P', given 'Q', and the DA inference infers 'not Q', given 'not P'. These inferences are invalid because a conditional (if in accord with implication) only states that Q must be true if P is true; it does not state that Q cannot also be true when P is false. These inferences would be valid if the rule expressed equivalence (they are valid for $P \equiv Q$). Clearly, interpretation of 'if P then Q' as also meaning 'if Q then P' is a form on conversion and, for this reason, the conversion hypothesis is frequently stated, in the propositional literature as the hypothesis that subjects interpret a conditional rule as implying material equivalence. In fact, conditionals aresometimes used in real life to express an equivalence relationship.

As, with the addition of negatives, there are four possible conditional rules, it is not always the case that, for instance, MP infers 'Q' from 'P'. Given the (NN) conditional 'if not P then not Q', MP infers 'not Q' from 'not P' and MT infers P from Q. A convenient standard notation is to refer to the true or false case of the antecedent (TA or FA) and to the true or false case of the consequent (TC or FC). Thus, TA and TC are P and Q for the AA rule, but 'not P' and 'not Q' for the NN rule. This allows the following standard expressions of the four inferences, which apply to all rules:

MP (valid)	TA \therefore TC
MT (valid)	FC \therefore FA
AC (invalid)	TC \therefore TA
DA (invalid)	FA \therefore FC

Inferences may also be drawn from disjunctive rules. Given that one component is false, it is valid to infer that the other is true. However, given that one component is true, it is not valid to infer that the other is false (unless exclusive disjunction is specified). There are thus two valid, and two invalid inferences that are the equivalent to those that may be inferred from the corresponding conditional.

Finally, a note on the difference between 'not P' and ' \bar{P} ' is in order. These will be used to stand for explicit and implicit negation respectively. Thus, given the proposition (P) 'the number is a 7', the conclusion (or denial) 'the number is not 7' will be expressed as 'not P'. However, if there were a choice of conclusion between '7' and '6', this would be expressed as 'P' and ' \bar{P} ' (because '6' is 'something which is not 7').

SECTION 2.2 TRUTH TABLE EXPERIMENTS

The idea that subjects may interpret conditionals in terms of defective implication was first introduced into the psychological literature by Wason (1966). It was Johnson-Laird & Tagart (1969), however, who first attempted to establish which type of truth table was psychologically appropriate for the various ways in which implication may be expressed. They studied the four alternative forms shown below:

- (1) If P then Q
- (2) There is never P without there being Q
- (3) There isn't P if there isn't Q
- (4) Either there isn't P or there is Q

It can be seen that (1), (2) and (4) are all possible translations of $P \supset Q$, (1) being the conditional, and (4) the disjunctive form. Although the conditional at (3) could be taken as a translation of $P \supset Q$, it is best described as a translation of $\bar{Q} \supset \bar{P}$ (which has the same truth table as $P \supset Q$). Thus it differs in two ways from the conditional at (1); firstly, it contains negatives, and secondly the order of terms is reversed as the antecedent is the second term.

The rules expressed relationships between letters and numbers that appeared on the left and right hand sides of cards having a line drawn down the centre (e.g. 'if there is an A on the left, then there is a 7 on the right'). Each subject had to classify a particular set of cards as to whether each card proved the rule true, proved the rule false or was irrelevant to the rule. The sets included instances of TT, TF, FT and FF cases (e.g. for the above

rule, A7, A6, B7 and C6, respectively) so that subjects evaluated each truth table case. The P and Q terms were falsified in three different ways; either by a different letter (or number), a geometric shape, or a blank. There were no reported response differences between these three methods.

Data were reported in terms of derived truth tables for each subject. However, they were not fully reported as a 'miscellaneous' category was used which included all truth tables that occurred no more than twice throughout the experiment. On some rules this category accounted for a large proportion of the data.

As expected, most subjects (19 out of 24) were classified as interpreting rule (1) as defective implication. There was only one (material) implication, and four miscellaneous, classifications. The same tendency, although weaker, was observed on rule (2). 14 subjects were classified as interpreting it as defective implication and only three were classified as giving an implication interpretation. Rule (3), however produced a wide range of responses. Only five subjects were classified as interpreting it as defective implication (of $\bar{Q} \supset \bar{P}$) and no other classification occurred more than twice. This result reveals a drawback of the method used. Each subject evaluated 16 stimuli for each rule and thus repeated measures were used. The subject's truth table was derived (presumably) from his most frequent evaluations. If all the data had been analysed, it is possible that a greater consistency across subjects may have become apparent.

Two important points emerge from these results. Firstly, (AA) conditional statements are most frequently interpreted as defective implication. It is of interest to note that there was no apparent tendency to interpret the conditional as implying

equivalence (i.e. they were not converted). Secondly, the interpretation of $P \supset Q$ is dependent upon its linguistic expression. Interpretations were less consistent for the nonconditional rule (2). The reversed, negated conditional was apparently too complex for most subjects. The root of this difficulty is unclear, however, as it may be the reversal or the presence of negatives (or both). It is unfortunate that the study did not investigate rules of the form 'there is a Q if there is a P' (reversed) and 'if there is not a P, then there is not a Q' (negated).

The disjunctive rule (4) was evaluated very differently from the conditional rule (1). Responses were more varied and classification times were (on average) more than twice as long. However, there were less 'irrelevant' responses and eight of the subjects were classified as interpreting it as implication. Four subjects responded as though the disjunctive was true when the first (negated) part was true and false otherwise, thus appearing to ignore the second half of the rule and no other truth table was ascribed to more than two subjects. Thus one very clear finding of this study is that, although $P \supset Q$ is logically equivalent to $\bar{P} \vee Q$, the AA conditional is not psychologically equivalent to the NA disjunctive.

The investigation of conditional rules was extended by Evans (1972b) to include all four conditional rules formed by systematic negation of antecedent and consequent. Evans points out that this procedure entails that "overall the effect of instances matching (affirming) or mismatching (negating) values named in the rules should cancel out". This is demonstrated in Table 2.3.

It can be seen from Table 2.3 that each of the four matching cases (i.e. PQ , $P\bar{Q}$, $\bar{P}Q$ and $\bar{P}\bar{Q}$) appears exactly once as each of the four truth table cases.

Table 2.3

Table Demonstrating How Use of Four Rules Controls for Matching

		<u>TT</u>	<u>TF</u>	<u>FT</u>	<u>FF</u>
<u>CONDITIONAL</u>		<u>CASE</u>	<u>CASE</u>	<u>CASE</u>	<u>CASE</u>
IF P THEN Q	(AA)	PQ	$\bar{P}\bar{Q}$	$\bar{P}Q$	$P\bar{Q}$
IF P THEN NOT Q	(AN)	$\bar{P}\bar{Q}$	PQ	$\bar{P}Q$	$P\bar{Q}$
IF NOT P THEN Q	(NA)	$\bar{P}Q$	$\bar{P}\bar{Q}$	PQ	$P\bar{Q}$
IF NOT P THEN NOT Q	(NN)	$\bar{P}\bar{Q}$	$\bar{P}Q$	$P\bar{Q}$	PQ

The subjects were not offered an 'irrelevant' category but were presented with an array of P, \bar{P} , Q and \bar{Q} instances and asked to construct true and false instances of given (abstract) conditional rules. They were asked to give further instances until they said there were no more. Irrelevance was thus simply denoted by a failure to construct a particular case as either verifying or falsifying. Table 2.4 shows the frequency of selection of the four truth table cases summed across the four rules.

Table 2.4

Frequency of Construction of the Four Truth Table Cases Summed Across the Four Rules (data obtained from Evans, 1972b, Table 3)

N=96 (24 subjects x 4 rules)

<u>TRUTH TABLE CASE</u>	<u>VERIFICATION</u>	<u>FALSIFICATION</u>	<u>NOT CONSTRUCTED</u>
TT	95	0	1
TF	3	77	16
FT	13	33	50
FF	32	22	42

As can be seen, there was a strong bias to construct TT to verify and TF to falsify. Treatment of FT and FF was more variable but across the four rules it was more common to ignore these two cases than to construct them as either true or false instances.

An interesting finding emerges if this data is analysed, not by truth table case, but by actual values selected, summed across the four rules. This is done in Table 2.5, which shows Evans' use of negated conditionals to be an effective technique, as it disambiguates these values and truth table case.

Table 2.5

Frequency of Selections Summed Across the Four Rules (Data from Evans, 1972b, Table 3) N=96 (24 subjects x 4 rules)

	<u>VERIFICATION</u>	<u>FALSIFICATION</u>	<u>NOT CONSTRUCTED</u>
PQ	33	50	13
$P\bar{Q}$	39	32	25
$\bar{P}Q$	38	26	32
$\bar{P}\bar{Q}$	33	24	39

It can be seen that falsification, but not verification, classifications show a trend down the table. In particular, PQ was constructed far more frequently than either $P\bar{Q}$, $\bar{P}Q$ or $\bar{P}\bar{Q}$. This effect is even more noticeable when initial constructions are looked at. PQ was constructed as a falsifying instance 46 times as opposed to 21 ($P\bar{Q}$), 13 ($\bar{P}Q$) and 13 ($\bar{P}\bar{Q}$). Evans refers to this tendency to prefer to construct those values mentioned in the rule as 'matching bias'.

Evans (1975) reports data that replicated the above study using an evaluation task, in which subjects were presented with instances of truth table cases and asked whether they were true or false cases of (or irrelevant to) the rule. He extended the investigation to rules of the form 'P only if Q' (referred to as OI conditionals), which have the same truth table as the usual 'if...then' (IT) conditional but a somewhat different form. The evaluations of the four truth table cases, summed across the four rules, are shown in Table 2.6.

Table 2.6

Evaluations of the Four Truth Table Cases Summed across the Four Rules

(Data from Evans, 1975, Table 1) N=96 (24 subjects x 4 rules)

	<u>IF...THEN</u>			<u>ONLY IF</u>		
	TRUE	FALSE	IRRELEVANT	TRUE	FALSE	IRRELEVANT
TT	85	5	6	79	2	15
TF	9	78	9	11	56	29
FT	18	28	50	12	55	29
FF	29	11	55	42	15	38

As can be seen by comparison with Table 2.4, data for the IT rule are comparable to those obtained from the construction task. Striking differences may be observed, however, between the IT and OI data. Firstly, the tendency to evaluate FF as a verifying instance is more noticeable on the OI rule, 'true' being the modal response to FF. More noticeable is that there is no difference on the OI rules between responses to TF and FT, which both show a strong bias

towards 'false' evaluations.

The similarity between TF and FT responses, together with the increased bias towards evaluating FF as a 'true' case, suggest that subjects may interpret the OI rule as implying equivalence. However, it is also noticeable that the amount of 'irrelevant' responses summed across both TF and FT is approximately the same for the two rules (59 and 58). It could thus be that, on the OI rule, some subjects evaluated FT as false and TF as irrelevant. This would suggest a conversion to 'if TC then TA' treated as defective implication.

The effect of matching on the IT rule was somewhat different from that observed in the data of Evans (1972b). It was very noticeable, in that data, that matching only affected falsification classifications and that its primary effect was to produce a high construction frequency for the double matching case. On the Evans (1975) data, matching appeared to have some effect on 'true' as well as 'false' evaluations, and its primary effect appeared to be to increase 'irrelevant' evaluations of the double mismatching case. The Evans (1972b) suggestion that matching affects 'irrelevant' responses appears to be the best explanation of the evaluation task data. Matching bias had an equally strong effect on responses to the OI rule. This is of particular interest, as it shows that the bias extends to a different formulation of the rule.

Evans & Newstead (1977) also investigated subjects' evaluations of IT and OI rules. A strong matching effect on 'irrelevant' responses was observed and, again, the effect was most noticeable on $\bar{P}\bar{Q}$. The mean number of 'irrelevant' evaluations (per 16 subjects) was 7.88 for $\bar{P}\bar{Q}$ as opposed to 3.75 ($\bar{P}Q$), 2.75 ($P\bar{Q}$) and 1.95 (PQ). It seems possible that the effect of matching is qualitatively different between construction and evaluation paradigms,

primarily leading to the construction of double matches, and the evaluation of double mismatches as irrelevant.

Evans & Newstead investigated the apparent conversion of OI rules noted in the data of Evans (1975) but concluded that such conversion is only a minor tendency. If OI rules are converted then FT becomes TF, and vice versa. The OI data were compared with the IT data to see whether a better match was obtained with, or without, the assumption of conversion of OI rules and a better fit was obtained without this assumption.

Truth table experiments have indicated that the interpretation of the \supset relation is highly dependent on the way it is expressed. For any particular expression, responses are very varied and often no consistent interpretation, across subjects, can be determined. Conditionals appear to be often interpreted as implying defective implication, although data from the FT and FF cases suggest that they may also be occasionally interpreted as implying equivalence. There appears to be some tendency on OI conditionals to also convert them to 'if TC then TA' and interpret the converted form as implying defective implication. Some subjects on conditional tasks, and many subjects on disjunctive tasks, yield inconsistent or unclassifiable truth tables.

Smedslund (1970) has pointed out that there is a 'circular relation between understanding and logic' as it is only possible to find out whether the subject has reasoned logically if his interpretation is assumed, but his interpretation may only be inferred by assuming that he has reasoned logically. Although Smedslund was referring to theories of reasoning in general, the point is well made with reference to truth table experiments. The only way in which truth tables may be inferred is on the assumption that subjects reason correctly on the basis of their interpretation. The existence of matching response bias is one indication that this is

an invalid assumption.

The ability of truth table tasks to determine interpretation is thus debatable. They have served, however, to demonstrate that different response profiles are obtained from different forms of (logically equivalent) rules. Clearly, different responses must be expected to tasks using disjunctive, conditional or OI conditional rules.

SECTION 2.3 CONDITIONAL INFERENCE TASKS

SECTION 2.3.1 ABSTRACT CONDITIONAL INFERENCE TASKS

In general, inference tasks present subjects with a major premiss (a conditional or disjunctive rule) and a minor premiss (the affirmation or denial of one component of the rule), and require them to draw (or evaluate) a conclusion about the affirmation or denial of the other component.

As was described in section 2.1, the following four inferences can be drawn from a conditional rule:

MP:	TA \therefore TC
MT:	FC \therefore FA
AC:	TC \therefore TA
DA:	FA \therefore FC

Whatever the interpretation, MP and MT are valid as they depend on the restriction that TA and FC may not both be true. However, AC and DA are invalid unless the rule is interpreted as implying equivalence. Some work in this field has been directed at inferring underlying truth tables and thus determining the interpretation of the conditional: if subjects make a large quantity of AC and DA inferences, they are presumed to have interpreted the conditional as implying equivalence and, if they make few such inferences, they are presumed to have interpreted the conditional as implying implication. This approach has two drawbacks. Firstly many subjects behave inconsistently and are not easily classified,

and secondly it suffers from the Smedslund (1970) objection of circularity. It is preferable to consider actual frequencies of the AC and DA inferences, rather than classifications of subjects on the basis of their presumed interpretation.

Taplin & Staudenmayer (1973) presented subjects with sets of three statements, in which the antecedent and consequent of conditionals referred to letters of the alphabet, such as the following:

If there is a Z, then there is an H	(major premiss)
There is a Z	(minor premiss)
There is an H	(conclusion)

Subjects were asked to mark each conclusion true or false and all four inferences were evaluated as 'true' in most cases. In their second experiment, Taplin & Staudenmayer offered subjects a choice of three possible conclusions: 'always true', 'sometimes but not always true' and 'never true'. (For half the subjects the word 'false' was used in place of 'true' but this had no effect on responses.) There was a marked reduction in the frequency of 'always true' responses to AC and DA conclusions (no specific data are given), and subjects took advantage of the 'sometimes but not always' categories.

Taplin & Staudenmayer suggest that one reason for the high frequency of acceptance of AC and DA conclusions in the first experiment is that subjects misinterpreted the response categories. However, it is the authors themselves who appear to have done this. They state that, "in logic", conclusions that follow from premisses are true and that "any other conclusion is false", and appear to

confuse 'true' and 'false' with 'valid' and 'invalid'. In fact, given two true premisses, valid conclusions are true but invalid conclusions may be either true or false. Forced to make a choice, subjects in the first experiment presumably believed that AC and DA conclusions were more likely to be true than false. Overall, the results of this study thus appear to suggest that most subjects believe that AC and DA conclusions (from true premisses) are probably, but not necessarily, true.

Evans (1972c) noted that, for the AA rule, some inferences involve negative conclusions and others involve affirmative conclusions, and that this difference may well affect relative frequencies of acceptance of these inferences. Accordingly, he used all four conditional rules and thus the extent to which each inference produced affirmative or negative conclusions was balanced across the four rules. For instance, MT produces a negative conclusion on AA and AN rules, but an affirmative conclusion on NA and NN rules. Evans studies only MT and AC and thus all conclusions involved the antecedent of the rule. For each problem, subjects were presented with three alternative conclusions: 'P', 'not P' and 'indeterminate'. Each subject had the opportunity of constructing each inference twice. The percentage frequency with which each inference was made, for each of the four rules, is shown in Table 2.7. (Percentages are based on two responses from each of 16 subjects).

Significantly more MT inferences, and less AC inferences, were made on rules having affirmative antecedents and thus, for both inferences, subjects affirmed significantly more negative conclusions. It is thus possible to interpret these data as suggesting a response bias that produces a preference for negative conclusions. However, there are two alternative possibilities worth considering:

- 1) NA and NN rules may be confusing and thus yield more errors
If AC is considered invalid, then confusion would thus yield more AC and less MT responses.
- 2) NA and NN rules may be more frequently interpreted as implying equivalence. This would explain the greater frequency of AC responses. (It should be noted, however, that this explanation does not account for differential frequencies of MT responses.)

Table 2.7

The Percentage Frequency of MT and AC inferences in Exp. 1, Evans, 1972c

(Data from Table 2, Evans, 1972c)

	MT	AC
IF P THEN Q *	91	32
IF P THEN NOT Q	75	35
IF NOT P THEN Q	38	61
IF NOT P THEN NOT Q	41	55

*The data for the AA rule are slightly suspect as the subjects were given pretests with feedback for both inferences on this rule.

Evans (1977a) studied all four inferences, on all four rules, using an evaluation task. Subjects were asked to indicate ('yes' or 'no') whether the conclusion followed from the given premisses. All subjects affirmed the MP inference on all four rules, but significant differences between the rules were observed for the other three inferences. The Evans (1972c) results were replicated in that significantly more MT inferences, and less AC inferences, were affirmed on rules having affirmative antecedents. Furthermore,

more DA inferences were affirmed on rules having affirmative consequents.

This latter result is of particular importance because, again, it suggests a bias towards negative conclusions. Both of the above alternative explanations would predict that more DA responses would be observed on rules having negative antecedents. Thus, as the observed DA data do not accord with these alternative explanations, the best available explanation is that subjects are, in fact, biased towards negative conclusions. This will be referred to henceforth as 'conclusion bias'. However, although it appears that a nonlogical factor influences responses, there is also clear evidence of the influence of a logical factor, as all subjects affirmed MP on all four rules.

Evans (1977a) also investigated the OI forms of conditional rule and found no significant evidence of the bias. However, subjects did affirm more negative than affirmative conclusions on all four rules (which result is, in itself, close to significance) and a rank correlation, performed across the cells of the Evans data (rule form by inference), yielded a significant correlation of 0.76 between the two rules. This suggests that the data from the two rule forms 'move' in the same way and there is thus good reason to believe that the bias did have a (weaker) effect on the OI responses. The data from the OI rules suggested that some subjects were converting them to 'if TC then TA' (in accordance with the findings of Evans, 1975, and Evans & Newstead, 1977).

Roberge (1971b) ran a comparatively large sample (110 subjects), presenting them with all four inferences to be evaluated on all four conditional rules. Subjects were asked to respond 'yes', 'no' or 'maybe'. The problem with this study is that Roberge

does not differentiate between inference type and thus some of his findings are impossible to interpret. For instance, he reports that there is no difference in error rates between subjects' evaluation of affirmative and negative conclusions. However, this does not necessarily mean that one type of conclusion was not affirmed more often than another, as agreement with a conclusion may or may not be an error, depending on the inference type. If subjects had accepted all negative conclusions and rejected all affirmative conclusions, then the error rates would have been identical.

Roberge analysed the data in terms of mean errors per rule, on the assumption of an implication interpretation (i.e. AC and DA were scored as errors). Least errors were made on AN rules and most errors were made on NA rules, with the error rate on AA rules about the same as that on NN rules. This overall pattern does coincide with the data of Evans (1977a) but no more precise comparisons can be drawn, as Roberge does not give separate data for each inference.

SECTION 2.3.2 THEMATIC CONDITIONAL INFERENCE TASKS

Taplin (1971) used thematic rules to investigate all four main inferences plus the invalid inference: TA .'. FC. Subjects evaluated each inference 12 times and thus 60 problems (all based on affirmative conditional rules) were used. As in the Evans (1977a) study, subjects were asked whether a given conclusion necessarily followed from the premisses. The invalid inference, TA .'. FC, was rejected on 94% of occasions and MP was affirmed on 91% of occasions. However, the error rates are surprisingly high, particularly as Evans (1977a) reported 100% correct MP responses on an abstract task complicated by the presence of negatives. MT was affirmed on only 63% of occasions which again reveals a higher error rate than that reported for the AA rule by Evans (1972b and 1977a). As Taplin's subjects were presented with 60 problems, the error rate could possibly be explained by boredom (or other factors) resulting from the length of the task. However, Taplin & Staudenmayer (1973) presented subjects with 96 (abstract) problems but still obtained frequencies of 99% (MP) and 87% (MT) in experiment 1.

The position is unclear as the Taplin & Staudenmayer methodology may have increased the frequency of all inferences in experiment 1 and, unfortunately they do not report the raw data from experiment 2. However, as they stand, these data suggest that use of thematic rules appears to have a small inhibitory effect on the affirmation of valid (MP and MT) inferences. Further support for this view is provided by the data of Roberge (1978) who compared use of thematic and abstract content on a within subject design. Although Roberge reports no analysis of the effect of content on conditional rules only, his data do show higher frequencies of MT affirmation on

abstract than on thematic problems (70% to 61% on the AA rule and 67% to 58% across all four rules).

Taplin's data show that AC was affirmed on 57%, and DA on 61% of occasions. However, when thematic rules are used, the frequencies of these inferences would be expected to be related to the content of the rules. For instance, subjects might be expected to make more AC and DA inferences on the rule 'if the number is even then it is divisible by two' than on the rule 'if it is a rabbit then it is an animal'.

Unfortunately, although it is possible to construct statements that clearly imply equivalence or implication, many thematic statements, although not strictly implying equivalence, may suggest that the AC and DA inferences are very likely. For instance, Geis & Zwicky (1971) have suggested that conditional promises and threats, 'invite' the (DA) inference that, if the conditional behaviour is not performed, then the promised reward, or threatened punishment, will not be forthcoming. Fillenbaum (1975, 1976) has provided experimental support for this view. Most subjects conclude that, for instance, the 'real life' utterance of the statement 'if you mow the lawn I will give you five dollars' implies that no money will be paid if the lawn is not mown.

It is not necessarily the case, however, that such 'invited inferences' denote a consistent equivalence interpretation. As Kneale & Kneale (1962) point out, promises are not usually considered verified if the condition on which the promise was made is not fulfilled. Thus a person told 'if you mow the lawn I will give you five dollars' may accept the inference that if he doesn't mow the lawn he will not receive five dollars; but, if he does not mow the lawn (and does not receive five dollars) he may not accept this as

verification that he would have received five dollars if he had mown the lawn. That is, his 'inference task' responses may appear to indicate an equivalence interpretation, whereas his 'truth table' responses may appear to indicate a defective implication interpretation. The argument against assumption of interpretation on the basis of responses thus holds at least as well for thematic rules as for abstract rules. This point should be borne in mind during the following discussion of the effect of content on frequencies of AC and DA affirmation.

Taplin reports that the 60 conditional rules used "contained a mixture of instances where the two propositions, P and Q, were causally related and where the connection between the propositions was either arbitrary or ambiguous". However, the relationship expressed by the conditional was reported to have no effect on responses, each inference being affirmed with approximately equal frequency on each problem. The data of Staudenmayer (1975), however, although not fully reported, appear to suggest that more AC and DA inferences are drawn from abstract rules of the form 'P causes Q' than from abstract rules of the form 'if P then Q'. (It is not clear whether the same result applied to thematic rules).

A useful idea introduced by Staudenmayer is that acceptance of AC and DA may be related to the perceived necessity of the antecedent. TA may be defined as necessary if there are no alternatives to TA that lead to TC. It appears from Staudenmayer's data that statements such as 'if the switch is turned on then the light will go on' led to more AC and DA inferences than statements such as 'if I turn on the switch then the light will go on'. Staudenmayer argues that the antecedent of the second statement is not necessary as it is possible for somebody else to turn the light on.

A more sophisticated form of the above distinction would be to view necessity as a dimension. That is, the more alternatives to TA that lead to TC, the less necessary TA may be said to be. Bucci (1978), using affirmative universal statements, compared performance on statements drawn from either end of this dimension. 'Broad predicate' items (e.g. 'all football players are strong', 'every cat has whiskers') were such that many alternatives to TA led to TC. For instance, strength is not peculiar to football players and various species have whiskers. 'Narrow predicate' items (e.g. 'every dog can bark', 'all oak trees have acorns') were such that there were either no alternatives or only a few alternatives. There are only a few species, other than dogs, that can bark and there are no trees, other than oaks, that have acorns. Adult subjects affirmed significantly more AC and DA inferences on 'narrow predicate' items than on 'broad predicate' items.

The problem with the Staudenmayer and Bucci findings is that they do not necessarily reveal anything about human reasoning. The necessity of TA is clearly directly related to the likelihood of the AC and DA inferences. If a problem is constructed such that the subject's experience leads him to assume that TA is not necessary for TC, then that same experience will also tell him that the AC and DA inferences are not necessarily true. For instance, for the conditional 'all football players are strong', it is clear that the inference 'all strong people are football players' is not true. Thus no reasoning is required as subjects may directly evaluate the truth of the conclusion. Bucci himself believed that subjects were assessing the plausibility, rather than validity, of conclusions to thematic problems.

Rips & Marcus (1977) suggest a way in which prior beliefs may have a more subtle effect on responses. Their main experimental work follows up a truth table experiment by Legrenzi (1970). Legrenzi showed subjects a machine in which a ball could roll down one of only two channels. When the ball reached the end of the channel, a red or green light came on, and Legrenzi used conditionals such 'if the ball rolls right, the red light comes on'. A high number of FT - false and FF - true responses were observed and he argued that the binary condition had led to a higher than usual tendency to assume that the conditional implied equivalence.

Rips & Marcus extended the design to include two further types of material: cards with letters and numbers on them and a set of pictures of fish that varied in colour and markings. Their data were not fully reported but, even when all three problems were made binary, they appear to have obtained more FT - false and FF - true evaluations, and more AC and DA inference affirmation, on problems involving the machine context. As this difference could not have been due to the binary nature of the task, it must have been due to a factor more specific to the machine context itself. Rips & Marcus argue that it is the relationship that the subject presupposes between the antecedent and consequent that defines the interpretation of a conditional rule.

In a further experiment, subjects were told that each of three antecedent values was associated with one (only) of three consequent values (the authors refer to this as correlation). Two thirds of subjects' responses were consistent with an equivalence interpretation and there was no difference between the three contexts. Rips & Marcus claim that this finding explains earlier results as it

appears to demonstrate that the one-to-one relationship is usually more often assumed for the machine context. They conclude that conditionals are more likely to be interpreted as implying equivalence if "subjects believe a priori that a simple correlation exists between the dimensions".

This is somewhat vacuous, however, as to tell subjects that the dimensions are correlated on a one-to-one basis is also to tell them that the conditional rules imply equivalence. Similarly, belief in such a correlation is belief that an equivalence relation exists. Thus Rips & Marcus appear to be arguing that subjects are more likely to interpret a rule as implying equivalence if they believe a priori that an equivalence relation exists! The crucial question, which Rips & Marcus fail to answer, is why subjects appear to have a greater tendency to assume an equivalence relation in the machine context.

One possible explanation is that subjects assumed a causal relation between the ball and the light. Rips & Marcus believed that subjects did assume a causal relation, but argued that this did not affect interpretation. Their conclusion is based on the fact that the machine context did not yield more responses consistent with an equivalence interpretation when subjects were told that the antecedent and consequent values were not correlated. However, there are two reasons for rejecting this argument. Firstly, Rips & Marcus assume that the instructions about non-correlation did not affect subjects' belief in a causal connection. This is clearly debatable. They overlook the fact that although correlation does not imply causality, causality does imply correlation. Secondly, when analysing subjects' responses to 'uncorrelated' problems using the machine context, they ignore over half of their data. Thus the finding that the machine context yields more responses consistent

with an equivalence interpretation (when subjects are not told that the dimensions are, or are not, correlated) may be due to perceived causality. The Staudenmayer (1975) results suggest that, when a causal relation is made explicit, more AC and DA inferences are affirmed on abstract rules. Although Taplin (1971) reports no difference in response profiles between rules expressing a causal relation and rules expressing an arbitrary relation, it is possible that statements that Taplin viewed as having causally related propositions were not perceived in this way by the subjects. The example quoted is "if food is constantly supplied to them, then the very fiercest creatures live peaceably together". It is plausible that such statements do not relate to the subjects' experience and are thus treated as arbitrary, but that subjects' experience (for instance, of electronic games) does relate to balls (apparently) turning lights on.

The quantitative model proposed by Rips & Marcus is of some interest but they give no proper explanation of how the fit to the data is achieved. This fit is surprisingly good, which is perhaps, at least partly, a resultant of fitting the model to only part of the data. An interesting suggestion contained within the model (stage 1) is that subjects first look for a match between the minor premiss and the antecedent of the major premiss. If a match is found, subjects will accept the (MP) inference if the consequent matches the conclusion and reject the inference if it does not. This suggestion explains the lower error rate generally observed on these inferences. However, the model cannot explain most data from other experiments as it gives no account of the circumstances under which the error parameters will vary, is only applied to AA conditionals and takes no account of conclusion bias.

SECTION 2.4 DISJUNCTIVE INFERENCE TASKS

SECTION 2.4.1 DIFFERENTIAL ERROR RATES

Roberge (1974) compared performance on all four conditional rules with all four disjunctive rules. He studied frequencies of affirmation of those inferences that are invalid, given a material implication interpretation of the rule. Over all four rules and inference types, the frequency of acceptance of invalid inferences was approximately the same for both types of expression. It would thus apparently be an oversimplification to suggest that one form of rule is intrinsically more difficult than the other. However, there was a significant interaction between the form (i.e. AA, AN, NA or NN) and type (i.e. conditional or disjunctive) of rule. Significantly more inferences were affirmed on NA, than on any other form of conditional rules; and significantly less inferences were affirmed on AA, than on any other form of disjunctive rules. (In this and later Roberge studies, subjects could respond 'yes', 'no' or 'maybe').

These findings reveal a striking difference between the NA conditional and the logically equivalent AA disjunction and, in general, lend further support to the Johnson-Laird & Tagart (1969) finding that logical equivalence does not necessarily lead to psychological equivalence. Unfortunately, Roberge (1974) pooled data from more than one inference and thus further analysis of the type of error involved is not possible.

Roberge (1976a) studied one valid disjunctive inference only: the affirmation of the second component of a disjunctive rule given the denial of the first component, and compared performance on

all four inclusive rules (e.g. 'Either A or B or both') with performance on all four exclusive rules (e.g. 'Either A or B but not both'). For inclusive disjunction, significantly less errors were again observed on AA rules. For exclusive disjunction, there were far more errors on AN and NA rules than on both AA and NN rules. The average error frequency on NN rules was only slightly greater than on AA rules. Thus the presence of negatives per se does not always produce a high error frequency in disjunctive reasoning. Significantly less errors were observed on exclusive rules and this result was entirely due to the reduced error frequency on AA and, particularly, NN rules. Roberge suggests that the 'but not both' instruction "appears to reduce the confusion" of "homogeneous" major premisses. However, these findings may be due to a tendency to ignore negatives and treat all disjunctive rules as if they were AA. This hypothesis provides a simple explanation of the very low error frequencies observed on the AA rule, and it is particularly well equipped to deal with the reduction of NN errors on exclusive rules. If an NN exclusive rule is treated as if it were an AA exclusive rule, this will not result in error as the two rules are logically equivalent.

SECTION 2.4.2 THE DENIAL OF A NEGATIVE: AN ILLUSORY PROBLEM?

The two premiss pairs shown below both yield the valid conclusion "John is intelligent".

<u>Problem A</u>	<u>Problem B</u>
1) Either John is intelligent or he is rich	1) Either John is intelligent or he is <u>not</u> rich
2) John is <u>not</u> rich	2) John is rich

Johnson-Laird & Tridgell (1972) predicted that the valid conclusion would more frequently be derived from problem A than from problem B. Subjects were presented with premiss pairs such as those shown above and asked "to determine what followed from them in virtue of logic alone". Results confirmed the prediction, subjects making more errors on problems such as B above. Mean response times were also higher on these problems. These results were interpreted as showing "that it is easy to grasp that a negative denies an affirmative, but exceedingly difficult to grasp that an affirmative denies a negative".

Johnson-Laird & Tridgell argue that "the proper function of affirmatives is to make assertions, and of negatives to make denials" (a similar argument has been advanced by Wason, 1965). Thus, in problem A above, the negative is fulfilling its proper function and causes little difficulty. In fact, Johnson-Laird & Tridgell found that explicit negation of an affirmative appeared to cause less difficulty than an implicit negation such as 'John is poor'. In problem B, however, a negative component of the disjunctive rule is denied by an affirmative and this is held to cause problems,

because the affirmative and negative are not fulfilling their natural functions.

This hypothesis has been utilised by Roberge (1976a, 1976b, 1978) to explain several aspects of his data and there is thus apparently considerable empirical support for the view that denial of a negative is a determinant of error. Such error may potentially occur in the following two ways:

- 1) Subjects may have difficulty in inferring a conclusion from an affirmative minor premiss that denies a negative component.
- 2) Subjects may have difficulty in evaluating (or inferring) an affirmative conclusion that denies a negative component.

These two possibilities will be considered in order. Roberge (1976a) studied arguments involving the affirmation of the second component, given denial of the first component, of the rule, and Roberge (1978) studied arguments involving the affirmation of the first component, given denial of the second. In both cases he reports that, in accordance with the data of Johnson-Laird & Tridgell, more errors were made when an (affirmative) minor premiss denied a negative component.

However, this result may simply be due to the very low error frequency observed on AA rules, as the reported finding does not hold up for comparisons not involving the AA rule. There was little difference in error frequencies between (inclusive) AN and NN rules in the experiment of Roberge (1976a), or between NA and NN rules in the experiment of Roberge (1978). Both of these comparisons compare an argument in which the minor premiss denies an affirmative component with an argument in which the minor premiss denies a negative component. Further, for these comparisons, the

reported difference is reversed on exclusive rules. For instance, in the experiment of Roberge (1978) error rates were 58% for NA, and 33% for NN, (exclusive) rules. Thus there is no real evidence in the data of Roberge (1976a, 1978) that more errors occur when a negative component is denied by the minor premiss. This argument clearly also applies to the results of Johnson-Laird & Tridgell, as these authors only used AA and AN disjunctive rules. In fact, an explanation of the results of Johnson-Laird & Tridgell, as due to the low error rate typically observed on AA disjunctive rules, has been proposed by Evans (1972a).

Roberge (1976b) used exclusive disjunction rules to study two types of valid inference:

- 1) Principle 1 arguments involved the denial of one component, given the affirmation of the other.
- 2) Principle 2 arguments involved the affirmation of one component, given the denial of the other. (The same form of inference as that studied by Roberge, 1976a, 1978).

Roberge reports that more errors were observed on principle 2 than on principle 1 arguments and argues that, as principle 2 arguments involve minor premisses that deny a component, this represents further evidence that more errors are observed when the minor premiss denies a negative. However, the greater error frequency on principle 2 arguments was at least as noticeable on the AA rule as on the other rules. The differential error frequency between principle 1 and principle 2 arguments thus cannot be explained as due to difficulties arising when the minor premiss denies (only) a negative component of the rule, and some other explanation must be sought.¹

Thus, on the basis of the above, there is no evidence

that the denial of a negative in the minor premiss is a determinant of error. The argument that difficulty arises from inferences whose conclusions deny a negative component is proposed by Roberge (1976b). He reports two findings in support of this point of view:

- 1) For principle 1 arguments, most errors occurred due to a failure to accept affirmative conclusions that validly denied a negative. That is, most errors were made on affirmative conclusions when the valid response was 'yes'.
- 2) For principle 2 arguments, the most prevalent error was failure to deny affirmative conclusions that invalidly denied a negative. That is, most errors were made on affirmative conclusions when the valid response was 'no'.

However, these two findings are again based on a comparison across the four rules and may be the resultant of low error rates on the AA rule. It is possible that no valid analysis can properly be made but, as Roberge used arguments that affirmed both first and second components of the rule, both the NA and AN rules yield both valid and invalid affirmative and negative conclusions, for each principle of argument. A comparison for these rules is thus of some interest and is shown in Table 2.8. (Data from AN and NA rules is 'pooled' as there were only negligible differences between these.)

The two crucial frequencies are denoted by an asterisk. As may be seen from Table 2.8, for principle 2 arguments, it is not the case that more errors occurred on arguments demanding a 'no' evaluation of affirmative conclusions (that denied negative components). Error rates were the same for both affirmative and negative conclusions

for which the valid response was 'no'.

Table 2.8

Average Percentage Error Frequencies on AN and NA (Exclusive) Disjunctive Rules, Compared by Polarity of Given Conclusion and Valid Response.

(Data obtained from Roberge, 1976b, Table 2)

<u>POLARITY OF CONCLUSION</u>	<u>PRINCIPLE</u>	<u>VALID RESPONSE</u>	
		<u>YES</u>	<u>NO</u>
AFFIRMATIVE	PRINCIPLE 1	68*	47
	PRINCIPLE 2	60	60*
NEGATIVE	PRINCIPLE 1	51	41
	PRINCIPLE 2	54	60

Thus, for principle 2, there is no evidence that conclusions that deny a negative are a determinant of error on NA and AN rules, and no way to test this possibility on AA and NN rules.

For principle 1, the highest error rate (68%) in table 2.8 was observed on evaluation of (valid) conclusions that denied a negative component. Subjects made more errors on affirmative conclusions that denied a negative than on those that did not and this clearly could be taken as (the only) evidence that denial of negatives causes error. However, the following two observations suggest that this may be a premature interpretation:

- a) Invalid conclusions yielded less errors on principle 1 arguments in the case of both affirmative and negative conclusions.

b) Considering valid conclusions separately (i.e. the 'yes' column of Table 2.8), there is clearly a possibility that future experimentation would fail to produce any significant difference in error rates between principle 1 and principle 2 arguments.

It is thus possible that the high error rate, on acceptance of principle 1 conclusions that deny negatives, is a resultant of other factors. In conclusion, then, the following points may be made:

- 1) There is no evidence that affirmative minor premisses that deny negatives are a determinant of error.
- 2) There is no evidence that presentation of affirmative conclusions that deny negative components leads to error on principle 1 arguments.
- 3) There is some indication that presentation of affirmative conclusions that deny negative components may lead to error on principle 2 arguments, although this result may possibly be a resultant of other factors.
- 4) Given that the only support is that mentioned at (3) above, it is a reasonable conclusion that, at least at present, the hypothesis that denial of negatives is a determinant of error is not sufficiently supported by the data to be accepted.

SECTION 2.4.3 THEMATIC DISJUNCTIVE INFERENCE TASKS

Roberge (1977) compared thematic AA rules such as those used by Johnson-Laird & Tridgell with abstract AA rules. The rules specified inclusive disjunction (by the addition of 'or both') and subjects were asked to evaluate conclusions to both valid principle 2 and invalid principle 1 arguments. The latter are invalid for inclusive disjunction as nothing may be inferred from the affirmation of one component. As usual for AA rules, error rates were fairly low (about 20%) and there was no difference between the abstract and thematic problems.

However, Roberge also included 'contradictory' thematic rules in which the components were semantically incompatible (e.g. 'either John is intelligent or he is stupid'). For the valid principle 2 inferences, significantly fewer errors were observed on the 'contradictory' rules, than on other thematic or abstract rules. For instance, given 'John is not intelligent', subjects made few errors in affirming 'John is stupid'. However, on principle 1 inferences more errors were observed on the 'contradictory' rules, subjects tending to respond in accordance with an interpretation of exclusive disjunction. These responses can hardly be regarded as errors, however, as subjects must have found it somewhat strange to be presented with problems such as the following:

"Suppose you know that:

Either the car is old or it is new (or both)

The car is old

Then would this be true:

The car is not new"

55% of subjects agreed with such conclusions and 50% denied conclusions such as 'the car is new' to the same premiss pair, whereas the logically valid ('maybe') response was made on only 38% and 44% of occasions, respectively. Given the question 'would this be true', it is perhaps surprising that there were so many 'maybe' responses. Although a somewhat extreme example, these results clearly demonstrate that the content of thematic rules is of crucial importance. The major premiss indicated inclusion syntactically but exclusion semantically and both indications appear to have affected subjects' responses. However, many subjects may simply have evaluated the conclusion on the basis of the minor premiss alone and thus it is perhaps debatable whether these were truly disjunctive reasoning problems.

SECTION 2.5 CONCLUSIONS

The findings of Evans (1972b, 1975) that 'matching bias' affects responses on truth table tasks has two important implications. Firstly, it demonstrates the effect of a 'nonlogical' bias on a propositional task, that parallels the finding of non-logical bias on syllogistic tasks, discussed in Chapter 1. Secondly, the existence of a response bias complicates the investigation of psychological truth tables. As Evans points out, "it is possible, of course, that some difference in interpretation between the rules may be reflected in the truth tables elicited". However, no explanation based solely on interpretation could adequately cope with the different frequencies of TF construction observed and there is thus good reason to believe that matching bias has, at least, some effect on responses. It is almost impossible to determine, however, to what extent differences between the rules are a function of matching and to what extent, if any, they are a function of interpretation.

However, although specific interpretations of individual rules cannot be inferred, certain information may be derived from an analysis across all four rules. Such an analysis controls for matching bias and reveals that there is a very strong tendency (validly) to evaluate the TT case as verifying and the TF case as falsifying. In addition, there appears to be a residual tendency to evaluate FT as falsifying and FF as verifying. This latter tendency is more noticeable on OI rules.

Thus truth table responses appear to be mediated by a combination of logical validity (possibly sometimes determined by 'conversion') and a nonlogical bias, in the same way as was argued to be the case in syllogistic reasoning. The same position applies

to abstract conditional inference tasks, which appear to be mediated by logical factors, at least in the case of the MP inference, but also by a nonlogical bias in favour of negative conclusions.

Both matching and conclusion bias may be compared to atmosphere bias in some respects. Revlis (1975a) argued that atmosphere bias was similar to matching bias, in that both produced conclusions that matched 'features' of the premisses. Conclusion bias may also be likened to the dominance of negative atmosphere and thus it would appear that nonlogical factors may be producing consistent effects across paradigms.

When considering the effect of thematic content on conditional inferences, it is best to consider separately the effects on those inferences that are, assuming implication, valid and invalid. There is some suggestion in the data that thematic content may have some inhibitory effect on the drawing of valid (MP and MT) inferences although no studies have reported significant differences between abstract and thematic content for these inferences and thus, clearly, the effect of thematic content awaits further investigation. However, it is evident that thematic content has no general facilitatory effect on the drawing of valid inferences.

The investigation of the effect of thematic content on AC and DA inferences is very difficult as, the more the major premiss encourages conversion, the more 'true' the AC conclusion will appear. The validity of the inference is thus confused with the truth status of the conclusion. As Bucci (1978) suggests, it is likely that subjects evaluate thematic inferences such as 'all strong people are football players' on the basis of plausibility rather than validity.

The well validated finding of disjunctive inference

tasks is that a very low percentage of errors is observed on AA rules and a high percentage of errors are observed on all rules containing negatives, with the exception of exclusive NN rules. There is also a noticeable difference between conditional and disjunctive inference tasks in that, on the former, subjects who fail to accept MT, AC, or DA usually evaluate them as 'indeterminate', whereas, on disjunctive tasks, subjects frequently deny valid inferences and affirm invalid inferences.

There appears to be a tendency for subjects to treat all problems as if they are reasoning with AA rules, an interpretation that is consistent with the data of Wason & Johnson-Laird (1969a) as well as with the data of the Roberge experiments. In 'real life', conversation, negated disjunctions, particularly those of the NA and AN forms are rarely, if ever, used and subjects are presumably depending on the rules of inference they have acquired for the AA form.

The use of thematic content by Johnson-Laird & Tridgell (1972) and Roberge (1977) appears to have produced response profiles comparable to those obtained from abstract content, except in the case of Roberge's 'contradictory' content. For both conditional and disjunctive thematic tasks, the crucial factor appears to be whether a clear relationship between the terms is expressed. For instance in the 'machine context', studied by Rips & Marcus (1977), the balls' movement appears to be the reason for the light and, for the 'contradictory' items of Roberge (1977), there is a clearly (semantically) expressed disjunctive relationship between terms such as 'intelligent' and 'stupid'. If such a relationship is not perceived, it appears likely that rules with thematic terms will be responded to as though the terms were abstract.

However, although use of thematic terms may affect responses, there is no reason to believe that such an effect is due to the terms having clarified the interpretation of the rule. Many subjects fail to draw, or agree with, the MT inference, which is valid on any interpretation of a conditional; and Roberge's work yielded high error rates on disjunctive rules, for which the interpretation was made clear to the subject by the use of 'or both' or 'but not both'. There is thus no reason to believe that 'interpretation' can be further clarified. The semantic effect of thematic materials appears to be, not to clarify the logical structure of the problem, but to clarify the truth status of the conclusion. The results of Roberge (1977) demonstrate that, when truth status accords with validity, then thematic content may appear to facilitate logical behaviour but that, when truth status conflicts with validity, far more 'errors' will be observed. Irrespective of consideration of validity, subjects will conclude (for instance) that it is not the case that 'all strong people are football players', that it is very likely that balls in machines will turn lights on and that old cars are not new cars.

FOOTNOTES

1. One possibility is that the instruction 'but not both' stresses the invalidity of the conjunction and that such statements are usually made in 'real life' to demonstrate that one of the components cannot be the case. Denial of a component (principle 1) would thus be a more 'natural' conclusion.

CHAPTER 3

THE WASON SELECTION TASK

SECTION 3.1	Introduction to the Task	106
SECTION 3.2	Early Work and the 'Therapy' Experiments	108
SECTION 3.3	Information Processing Models	113
SECTION 3.4	Matching Bias: An Explanation of Behaviour	118
SECTION 3.5	The Dual-Process Hypothesis: An Explanation of Introspection	121
SECTION 3.6	Use of Thematic Content	128
SECTION 3.7	Conclusions	136

SECTION 3.1 INTRODUCTION TO THE TASK

The paradigm developed by Wason (1966) is variously referred to as 'the Wason selection task', 'the selection task' and 'the four card problem'. It will be referred to here as 'the selection task'. The simplest way to explain the task is to give an example. Suppose a subject is shown a pack of cards all having a letter on one side and a number on the other. The experimenter then takes four such cards and lays them on the table in front of the subject to produce an array such as that shown below:



The subject is then shown a typewritten conditional rule that specifies a relationship between a particular letter and a particular number. For instance, the rule might be 'if the letter is an A, then the number is a 7'. The subject is told that the rule is only supposed to apply to the four cards on the table and is asked which cards it would be necessary to turn over in order to find out whether the rule is true or false. The four cards are always chosen so as to present subjects with one instance each of TA, FA, TC and FC and thus the task essentially involves evaluation of these four cases.

The only combination that falsifies a conditional rule is TA and FC, and thus the logically correct response is to select those cards that may potentially yield this combination (i.e. TA and FC). For instance, for the above example, the rule is false

if the A does not have a 7 on the back or if the 3 has an A on the back, otherwise it is true. It is irrelevant what is on the back of the D and the 7.

If a conditional rule is true, then the two valid inferences, MP and MT, must hold. The solution to the selection task can thus be seen as testing these two inferences. Selection of TA tests MP (which infers that TC is on the other side) and selection of FC tests MT (which infers that FA is on the other side). However, if the rule is interpreted as implying equivalence, then the AC and DA inferences (which may be tested by selecting TC and FA, respectively) are also valid and thus the logically correct solution is to select all four cards.

All experiments reported hereafter have used AA conditional rules only, unless otherwise stated. Typical results from such studies are that most subjects select TA but that many also select (invalidly) TC, and that FA and FC are both infrequently selected. Thus, in failing to select FC, subjects behave as though they do not possess the MT inference. Wason & Johnson-Laird (1972) report the following frequencies of initial selections from four experiments:

TA and TC	59 subjects
TA only	42 subjects
TA, TC and FC	9 subjects
TA and FC	5 subjects
All other combinations	<u>13</u> subjects

128

SECTION 3.2 EARLY WORK AND THE 'THERAPY' EXPERIMENTS

Wason (1968a) suggested that conditionals are usually only used in real life if they are true, and thus that subjects perform the task with the assumption that the rule is true. If the subject has a defective implication interpretation of the conditional, then the only true case will be the joint occurrence of TA and TC, and thus subjects' responses may be aimed at revealing this verifying case. To direct the subjects' attention towards the possibility of falsity, Wason showed subjects four cards prior to a selection task and asked them to pick out the one card that falsified, and the one card that verified the rule. However, although subjects were able to do this, the procedure had no effect (as compared with a control group) on subsequent responses.

Wason & Johnson-Laird (1970) pointed out that one possible explanation of failure to select FC, and invalid selection of TC, is that the referents 'one side of the card' (which refers to the antecedent) and 'the other side of the card' (which refers to the consequent) may cause confusion. Specifically subjects may confuse 'the other side of the card' with the side that is face down. They used conditionals whose antecedent and consequent referred to the centres and borders of the cards and presented subjects with the four alternatives by masking relevant parts. However, although the stimuli were all on the same side of the card, subjects performed no differently to a control group who were presented with the usual form of the selection task.

Wason (1969a) adopted two modifications to the task. Firstly, binary stimuli (red or blue circles or triangles) were used.

Secondly, he suggested that the conditional form "has the possible undesirable connotation that there is a temporal, or even a causal, relation between antecedent and consequent". He avoided any such implied causality by use of rules such as the following:

"Every card which has a red triangle on one side has a blue circle on the other"

Typical results were obtained and thus presumably neither the binary nature of the task, nor the linguistic form (which is a type of universal), had any significant effect. Wason introduced three stages of 'therapy', subsequent to initial selections aimed at directing the subject's attention towards falsification of the rule, and thereby raising the frequency of FC selections and reducing the frequency of TC selections. The three stages of the procedure were as follows:

- 1) Firstly, the subjects (who were not allowed actually to turn over their initial selections) were asked what might be on the other side of TA. If necessary, they were told that FC may be on the other side and thus that the statement possibly might be false. Subjects were given the opportunity of revising their selections after this, and later, stages.
- 2) If a 'correct' revised selection was not obtained, subjects were prompted to conclude that, if a TA instance were on the other side of the FC card, then this would again falsify the statement (and to realise that it was possible for this to occur).
- 3) If the FC card was still not selected, the experimenter eventually turned this card over, revealing a TA instance on the other side. The subject was then asked

whether he still believed the statement to be true and, after the FC card had been turned back over, was again invited to revise his previous selections.

After these 'therapies', some subjects still failed to make the combination of selections required by the experimenter (i.e. TA and FC). These subjects were explicitly informed that their selections were wrong and once again invited to change them. Of the 32 subjects, only two initially selected FC and this was progressively raised to 10 (after stage 1), 16 (after stage 2) and 30 (after stage 3). After the final decision stage, 28 subjects selected FC. Selection of TC was hardly affected, although it was selected slightly more often in the middle of the experiment, selection frequencies being 15 initially, 20, 20 and 21 after the three 'therapy' stages, and 14 after the final decision stage. The procedure was presumed to facilitate 'insight', that is, some cognitive state in which the subject reasoned in accordance with formal logic. However this interpretation may well be unwarranted. The evident bias of the experimenter towards falsification in general, and the FC card in particular, may be sufficient to explain changes in subjects' responses. A verbal reinforcement explanation would account for the fact that only FC selections were affected whereas, if the subject is assumed to have acquired some general state of insight, it is difficult to explain why TC selections were unaffected.

Wason & Golding (1974) used statements about cards that were divided into two vertical sections. The top section either did, or did not, contain a letter and the bottom section either did or did not, contain a number. Rules referred (nonspecifically) to 'numbers' and 'letters', and thus FA and FC were represented by the absence of a stimulus rather than by a different stimulus to that

mentioned in the rule. The following formulations of the relationship were used:

- 1) Whenever there is a number below the line there is a letter above the line.
- 2) There is a letter above the line whenever there is a number below the line.
- 3) There is a letter above each number.
- 4) A letter is above each number.
- 5) Above each number is a letter.

None of these alternatives appeared to produce responses significantly different from those usually obtained or to each other. There was a slight suggestion in the data that statements in which the consequent appears first (2,3 and 4) may be slightly easier. This experiment also used 'therapy' but it was found that, after responses had been changed by 'therapy' on one task, responses on a subsequent task were no different to those usually obtained. Clearly, FC on a subsequent task, although a different card, is conceptually the same. A verbal reinforcement explanation of 'therapy' refers only to the selection of a specific card. An insight explanation, however, presumes that the subject gains a conceptual understanding of the logical relevance of the FC card and such an understanding would be expected to transfer to a subsequent task.

However, the results of Smalley (1974) do suggest that some understanding of the problem can be induced. Subjects were group tested and the four cards (red and blue circles and triangles) were placed on a chalk rail at the front of the room. The task booklet explained the selection task and asked subjects to give their solution and then asked them to give (written) reasons for their selections. Subjects were subsequently given the opportunity

of revising these selections. They were then asked to consider each card at a time. Subjects were told (in the booklet) what the two possible figures on the other side of the card could be and asked in each instance whether the statement would be true or false. Following consideration of each card, they were again asked for their selections (and reasons for change if they had changed). In the final phase, the experimenter turned each card over, one at a time. When they were turned over, TA revealed a TC instance on the back but FC revealed a TA (falsifying) instance. After each card had been turned over, subjects were asked what conclusion they would draw about the truth or falsity of the statement and (after the cards were turned back over) were again asked to indicate their present choice of cards together with reasons for change.

This procedure had little effect on TC selections but produced a large increase in the frequency of FC selection (although not so marked as that observed by Wason, 1969a) After the final stage, 61% of the subjects selected FC. These results do suggest that the 'therapy' induces the subject to realise the relevance of searching for falsifying cases. Thus, although verbal encouragement from the experimenter may be necessary to induce very high frequencies of FC selection, many subjects, given sufficient 'cues', are apparently able to learn 'for themselves' to select FC.

SECTION 3.3 INFORMATION PROCESSING MODELS

Johnson-Laird & Wason (1970) proposed an information processing model to explain subjects' responses on the selection task. This basically assumes three cognitive states through which subjects progress during therapy. The three states are defined as follows:

- 1) 'No insight' - subjects focus only on those cards mentioned in the rule. Subjects who interpret the rule as implying implication focus only upon the TA card whereas subjects who interpret the rule as implying its converse focus on both TA and TC. These cards are selected as they may verify the rule.
- 2) 'Partial insight' - the subject realises two things: that all cards should be considered and that cards should also be selected if they may falsify the rule. Subjects thus select TA, TC and FC.
- 3) 'Complete insight' - subjects select only those cards (TA and FC) that may falsify the rule.

Subjects, although usually initially in a 'no insight' state, may start in any of the three states but, if they are not in a state of 'complete insight', they will progress towards this state during therapy. Goodwin & Wason (1972) asked their subjects to give reasons for their selections and claimed that these protocols represented supportive evidence for the model. Subjects who chose TA and TC only, explained TA selection in terms of verification, whereas subjects who included FC in their selections tended to explain both FC and TA selections in terms of falsification.

However, several criticisms have been levelled at the model. For instance, it does not make it clear why a subject who

interprets the rule as implying implication will select TC, because it may verify, when he is in a 'partial insight' state but fail to do so when he is in a 'no insight' state. Also, although state 1 presumes that some subjects interpret the rule as implying its converse, the second two states apparently presume that no subjects interpret the rule in this way (as, if the rule implies its converse, all four cards are potential falsifiers). Smalley (1974) accordingly proposed an alternative model that, although retaining the three insight stages, attempted to take account of various possible interpretations of the problem. The main differences were as follows:

- a) Subjects who interpret the rule as implying implication are held to select both TA and TC in state 1 unless they do not appreciate the reversibility of the stimulus cards (i.e. that TA, with TC or FC, on the back is the same as TC or FC, with TA on the back). If they do not appreciate this reversibility, they will select TA only in all three states.
- b) Subjects who interpret the rule as implying its converse are held to select all four cards in states 2 and 3. However, if they do not appreciate the reversibility of the cards, they will select only TA and TC in all three states.

Smalley appears to argue that perception of reversibility, unlike 'insight', cannot be induced by 'therapy'. However, this is somewhat peculiar, as it implies that subjects who select TA only (implication interpretation with no perception of reversibility) will not change their selections during 'therapy', whereas the results of Wason (1969a), which Smalley refers to, clearly indicate that they do. Smalley reports that the protocols gathered in his experiment (described earlier) lend support to his model. That is,

subjects tended to reveal interpretations of the rule and states of insight (i.e. bias towards verification or falsification) consistent with his interpretation of their selections.

A fundamental criticism of both the Johnson-Laird & Wason and Smalley models concerns the use of the term 'insight'. 'Partial insight' is defined as a set towards both verification and falsification which, in the Johnson-Laird & Wason model is itself defined uniquely as the selection of TA, TC and FC. Similarly, this model defines 'complete insight' as selection of TA and FC. As Evans (1977b) has pointed out, the state of 'insight' is deduced from the behaviour and thus simply describes the selections. The label is circular, as subjects are held to have selected TA and FC because they are in a state of 'complete insight', but the only 'evidence' for their being in a state of 'complete insight' is their selection of TA and FC. Similarly, in the Smalley model, both state of insight and interpretation are inferred from, and held to be determinants of the selections. It should be noted that supportive protocol evidence does not break this circularity as, in this case, 'insight' is inferred from, and held to be the determinant of, protocol evidence of a falsification bias. Thus, if introspective data were acceptable, protocol data may be taken as (non-circular) evidence of a verification or falsification bias, but neither selections nor protocols may be taken as evidence of 'insight'. The contention that subjects are in a state of 'insight' when they give the logically correct answer is thus non-falsifiable and, in terms of modern philosophy of science (e.g. Popper, 1959), must be held to be outside the province of experimental psychology.

Brée (1973) proposed a model that replaced the three insight states with three strategies that refer to neither insight

nor a verification or falsification bias. The three strategies (as stated by Brée & Coppens, 1976), are shown below:

- A Selection of "an object if the visible symbol requires the presence of a particular hidden symbol" (selection of TA only, or TA and TC if the converse is assumed to hold).
- B Selection of "any object which could have a symbol (hidden or visible) requiring the presence of any other particular symbol" (selection of TA, TC and FC, or all four cards if the converse is assumed).
- C Selection "as in A. Then select any object which could have a hidden symbol that would necessitate the presence of a symbol other than the symbol already visible" (selection of TA and FC or all four cards if the converse is assumed).

This model assumes that subjects do not interpret the conditional as implying the MT inference (i.e. FC is not taken to 'require' FA). Moshman (1978) points out that the model may be taken as making the MT inference "a matter of strategy rather than of interpretation".

Brée & Coppens (1976, 1978) presented subjects with both a truth table task and a selection task. Subjects' interpretation of the conditional was inferred from the truth table data and there was significant evidence of a relationship between inferred equivalence interpretations and 'TA and TC' selections on the selection task. All 11 subjects who were judged to have interpreted the conditional as implying its converse selected TA and TC. These subjects were thus all assumed to be adopting strategy A. However,

the nine subjects judged to have interpreted the rule as implying implication gave a variety of selection task responses (although only two selected TA and TC). This implies some interaction not accounted for by the model, between interpretation and strategy adopted.

The Brée model does not appear to be particularly useful as the definition of the three strategies is no less circular than the definition of the three insight states in the Johnson-Laird & Wason model. Also, as Brée & Coppens point out, it is no better at explaining the data and, unlike the Johnson-Laird & Wason model is incapable of predicting any selections in those cases where the subjects' interpretation cannot be inferred.

SECTION 3.4 MATCHING BIAS: AN EXPLANATION OF BEHAVIOUR

The studies discussed so far have shown that the most frequent initial selection is TA and TC and most workers, other than Bree, have explained this as a verification bias, an interpretation that is supported by subjects' own explanations. However, all these studies used only the AA conditional and thus TA and TC are confused with P and Q, the values mentioned in the rule. There was thus a possibility that selections could be explained in terms of the matching bias observed, on a truth table task, by Evans (1972b). The crucial experiment was performed by Evans & Lynch (1973). They presented subjects with four selection tasks, using all four forms of conditional rule. The data showed four significant effects, all of which would be expected if subjects were displaying a tendency to match their selections to the values named in the rule:

- 1) More TA cases were selected on rules having affirmative antecedents.
- 2) More FA cases were selected on rules having negative antecedents.
- 3) More TC cases were selected on rules having affirmative consequents.
- 4) More FC cases were selected on rules having negative consequents.

Over all four rules, subjects selected far more TA than FA, and thus there is a clear tendency to select TA, independent of matching bias. This is possibly related to subjects' ability to derive the MP inference. There was no evidence of a verification bias, in fact, subjects were found to select significantly more FC than TC, across all four rules. However, this latter result,

significant at the 5% level, should be treated with caution. Selection of FC on rules having affirmative conclusions was unusually high - one third of the subjects selected FC on the AA rule. Manktelow & Evans (1979) report two replications of the Evans & Lynch study and, although replicating the matching effects, their data show no overall difference between frequencies of TC and FC selections. They comment: "taking the Evans & Lynch results together with the present data, it seems that the idea of a verification (or falsification) tendency in the selection task should be used with extreme caution".

Explanation of selection task responses as a function of matching has been attacked on various grounds by several workers in the field. Bracewell (1974), for instance, claimed that results can be wholly explained by the way in which the subject interprets the rule. Taking written protocols as subjects progressed through a series of selection tasks, he argued that these written explanations could account for all the data. He criticised matching bias theory on the grounds that protocols clearly demonstrated that subjects were reasoning with a particular interpretation and that they were attempting to verify or falsify. However, Evans (1974) replied that it was a non-proven assumption that introspection mirrors the causes of behaviour although, on the selection task, it frequently correlates with behaviour. Evans also argued that the matching bias explanation cannot easily be overlooked as it is the only explanation of results that was originally observed on a different task and subsequently predicted and observed on the selection task.

Van Duyne (1973) published what is perhaps the most comprehensive attack upon the matching bias theory. He argues that the propositional calculus should not be used as a baseline for assessing logical performance and that a theory is required that

refers to the logic used in 'natural language'. He suggests that such a theory may well explain the data, and thus that search for such a theory would prove more fruitful than a search for theories such as matching bias. Unfortunately, he neglects to mention how a 'natural logic' theory should be searched for, or why the need for such an (unknown) theory automatically invalidates the (known) effects of matching.

Van Duyne argues that matching bias is limited primarily to performance on the AA conditional, as the 'P and Q' combination is rarely chosen on disjunctive tasks and many other forms yield this combination only infrequently. To support this criticism, Van Duyne supplies data from selection tasks using the following three rules:

'Either P or Q'

'Not P and not Q'

'If not P then Q'

The logically correct answer is ' \bar{P} and \bar{Q} ' in all three cases and these rules yield a very low incidence of selection of the 'P and Q' combination. The results of Wason & Johnson-Laird (1969a) suggest that matching bias is not found on disjunctive rules but, in the case of the conjunction and the conditional rule, Van Duyne appears to have missed the point. The matching bias theory merely predicts a tendency to select P and to select Q, not that the P and Q combination will be most frequently observed on all rules. This tendency is significantly noticeable across all four conditional rules, but the data from one rule alone can demonstrate neither the presence, nor the absence of, the bias. Further, Evans (1975) points out that matching bias is found on other tasks, as it was initially observed on a truth table task, and that it extends to the OI form of the conditional.

INTROSPECTION

Perhaps the most persistent objection to the matching bias hypothesis was that subjects' introspective protocols suggest that their responses are due to such factors as interpretation or verification bias. As has been mentioned, such protocols have been used to support theories of reasoning by a variety of workers (e.g. Bracewell, 1974, Goodwin & Wason, 1972, Smalley, 1974) although Evans (1974) pointed out that such protocols do not necessarily mirror the causes of behaviour. Van Duyne (1973) went so far as to argue that if an experiment wishes to demonstrate matching bias, then it should be supported by protocol data. This reliance on introspection (particularly the absurd suggestion that a response bias should be introspectable) is somewhat peculiar as, in most other areas of psychology, introspection has not been accorded such an important place in research for many years. In fact, protocol data may be explained by the results of Wason & Evans (1975).

Wason & Evans presented 24 subjects with two rules to evaluate, one of the (AA) form, 'if P then Q', and one of the (AN) form, 'if P then not Q', and asked them to give written reasons for their selections. There were 15 (correct) 'P and Q' selections on the AN rule and 12 (incorrect) 'P and Q' selections (and no correct selections) on the AA rule. Responses were thus in accordance with matching bias predictions. However, subjects tended to explain selection of 'P and Q' as aimed at verifying the AA rule but as aimed at falsifying the AN rule. These explanations are consistent with the selections, but it is difficult to believe that subjects attempt to verify one rule and falsify the other. The

subject appears to give the most plausible post hoc explanation of his selections. It is clear from this result that introspective data do not reveal the causes of behaviour and thus cannot be used to support theories about such causes. It is not surprising that protocols have appeared to support such theories as, on the AA rule, the most plausible explanation (verification bias) was adopted by both experimenters and subjects.

Wason & Evans propose a 'dual process' theory of reasoning which posits:

- a) Processes underlying behaviour, that are not usually available to introspection.
- b) Introspections, which reflect a tendency to construct a justification of behaviour - that is, to rationalise the response that has been made.

This explanation of introspections as constituting rationalisations appears to be a useful synthesis that is consistent with all the previous data. Further, it is not an explanation that need be restricted to the field of human reasoning. Nisbett & Wilson (1977), who discuss data from a wide range of social and cognitive psychological experiments, also conclude that subjects' introspections do not reveal underlying causes of behaviour and that such introspections are essentially post hoc rationalisations.

It is of some historical interest to note that Wason & Johnson-Laird (1969b) came close to producing a similar theory to that of Wason & Evans, some years before, as the following quote shows:

"It is as if two parallel thought processes were going on: a non-verbal unconscious decision process, connected with the selection or rejection of cards, which controls (and distorts) a more superficial, conscious

verbal process connected with the evaluation of the cards with respect to the truth or falsity of the rule. Thus, what the subject says is a compromise between a need to have a card fully revealed and a need to maintain a semblance of rationality".

Clearly, there are differences between this position and the 'dual-process' hypothesis. For instance, the two processes are seen as parallel. However, there is a similarity in that a distinction is made between an unconscious decision process (selection behaviour) and verbal report of a verification or falsification bias. Also, the quote suggests that the reasons given are 'controlled' by the selections made. This is exactly the point that was demonstrated by Wason & Evans.

Wason & Evans point out that the strong form of their 'dual-process' hypothesis would be to argue that responses determine conscious thought and suggest that a comparison could be drawn between this strong form of the hypothesis and the James-Lange theory of emotion. However, they also point out that a subject's rationalisations may be "wholly appropriate when the problem lies within his competence or experience". It should be noted, however, that, although it is quite possible that a subject's introspections will yield the causes of his behaviour on certain occasions, the crucial point is that the observer can never be sure when such an occasion will occur. Thus introspections can never be used to infer the process underlying behaviour.

Evans & Wason (1976) specifically investigated the theory that introspections are rationalisations of selections. Wason (1969b) had previously presented subjects with the logically correct answer and asked them to explain it. All 20 subjects were able to

do this. Evans & Wason note that "at the time Wason supposed that the subjects had been prevented from imposing their own erroneous structure on the problem" but suggest that, in the light of the Wason & Evans (1975) results, a more parsimonious explanation is that subjects had been simply "rationalising someone else's solution". They presented four groups of subjects each with a different solution to explain. The four solutions used were those most commonly observed: 'TA', 'TA and TC' and 'TA, TC and FC', and the correct solution 'TA and FC'. Most subjects in each group were willing to agree with, and explain, the given solution.

Subjects were asked to indicate their confidence in their explanations on a four point scale and most subjects indicated 'confident agreement' (defined as 'highly confident' or 'confident' ratings). The authors note that results could be explained in terms of the Asch (1956) work on conformity but argue that this high rate of confident agreement suggests that the results cannot be merely due to compliance. 75% of subjects confidently agreed with selection or rejection of TC or FC. Subjects were thus equally confident in agreeing with selection or rejection of these cards which suggests that presentation of a solution involving selection of FC did not facilitate any genuine understanding of the problem. In fact, confidence ratings for TA selection and FA rejection tended to be lower in the groups given the 'TA, TC and FC' or 'TA and FC' solutions than in the groups given the, more typically observed, 'TA' and 'TA and TC' solutions. Evans & Wason conclude that these results indicate that "the presentation of the correct solution does not induce enlightenment" and that they support the rationalisation hypothesis.

In discussing this tendency to rationalise responses,

the authors invoke the ideas of Festinger (1957) in suggesting that subjects' motive is to appear, both to themselves and to others (for instance the experimenter), as consistent in their behaviour. Thus, knowing that they have made a particular response, they construct reasons to show why this response was correct. Earlier studies (e.g. Wason & Golding, 1974) show that subjects occasionally produce highly inconsistent arguments in order to avoid admitting that their initial selection was wrong.

A similar explanation may be applied to the results of Smalley (1974) whose 'therapy' procedure produced some increase in FC selection even though no feedback from the experimenter was received. Some subjects may have failed to answer 'therapy' questions in a manner consistent with their previous selections, and may then have made a (revised) selection consistent with their answers to the 'therapy' questions. One substrate of the effect of 'therapy', then, may be the number of occasions on which the subject is required to rationalise. If a break in the rationalisation chain is induced, producing a valid response, then the subject will continue to respond consistently (and thus validly) from that point. Wason (1964) presented subjects with a task in which they were repeatedly given the opportunity of drawing an AC inference. Subjects who made AC tended to consistently continue to do so. However, when presented with the opportunity of making a valid inference, whose conclusion was inconsistent with a previous AC conclusion, subjects subsequently tended to avoid making AC inferences. This result also could be taken as an example of a consistent sequence of fallacious reasoning encountering an induced break in consistency and, from that point, transposing into a consistent sequence of valid reasoning.

Results from selection task studies using all four

conditional rules suggest that there is an overall tendency to select TA and reject FA, and a matching tendency, mediating selection of each card, which may well be the only determinant of TC and FC selection. This position views selection of each card as independently determined. Earlier workers, however, has assumed that card selections were influenced by each other and that particular combinations of cards had psychological relevance (e.g. the models of Johnson-Laird & Wason, 1970, Brée, 1973, and Smalley, 1974) Evans (1977b) tested this assumption of association between card selections.

Two problems are encountered in this respect. Firstly the incidence of TA selection is usually so high, and the incidence of FA selection so low, that it is not possible to answer questions about associations between these and other cards. Secondly, many studies report the data in terms of combinations of responses, with often a fairly large number of (atypical) combinations being reported simply as 'others' or 'miscellaneous'. Evans bypassed the first problem by looking, specifically, for association between selection of TC and FC. This association is probably the most crucial as it is the frequency of selection of these cards that is the most variable. He overcame the second problem by utilising some of his own data and by obtaining full data from other experiments by correspondence.

Evans analysed the data from a variety of experiments and found no case of a significant association between TC and FC selections. Two results were fairly close to significance at the 5% level but this is perhaps only to be expected, as 17 separate blocks of data were analysed. Of particular interest was the lack of association during the successive stages of therapy in the experiment of Wason (1969a). The theory that therapy produces an

('insight') state of greater understanding would have to predict that understanding of the relevance of FC is associated with the understanding of the non-relevance of TC. Evans points out that one cannot prove a null hypothesis but that, in the light of the analyses carried out, it is reasonable to treat selections as independent. Such independence clearly lends further support to the view that selections, particularly consequent selections, are not mediated by any overall understanding (or misunderstanding) of the problem.

SECTION 3.6 USE OF THEMATIC CONTENT

Johnson-Laird, Legrenzi & Legrenzi (1972) presented selection tasks using four envelopes in place of cards. In the abstract condition, the rule referred to letters and numbers on the front and back of these envelopes. In the thematic condition, rules were of the form 'if a letter is sealed then it has a 5d stamp on it' and subjects were presented with the back of a sealed (TA) and unsealed (FA) envelope, and the front of envelopes stamped with 5d (TC) and 4d (FC) stamps. Each subject performed a selection task under each of these two conditions and under two similar conditions that used the OI form of the conditional rule.

The authors report a very high frequency of logically correct (TA and FC) solutions on thematic problems. For the standard conditional, 87% of subjects were correct in the thematic condition as against only 8% in the abstract condition. For the OI rule, the comparison was 75% to 21%. Apart from the very clear effect of thematic material, two subsidiary results are apparent in the data. Firstly, there was no significant difference in performance between tasks using the standard, and OI, forms of the conditional and, secondly, the facilitation of correct responding observed on the thematic problems did not transfer to the abstract problems. This lack of transfer suggests that the thematic condition did not induce an appreciation of the logical structure of the problem.

It should be noted that, although at the present date the question of whether or not an envelope is sealed is irrelevant to the value of the stamp required, near the time of the study unsealed envelopes could be posted at a cheaper rate. Subjects were asked to imagine that they were Post Office workers, sorting letters, and to discover whether or not the rule had been violated.

Given the form of the instruction, and the fact that a postal sorter would be specifically looking for incorrectly stamped letters, it is not surprising that almost all subjects played the role and 'checked' the 4d stamp. In fact, many subjects may well have done this to their own letters in 'real life'. It is thus debatable whether this was truly a 'reasoning' task and the results are perhaps best explained in terms of the falsifying case being readily available on the basis of the subjects' experience.

A similar explanation may be applied to the results of Legrenzi (1971). He presented subjects with four cards, divided into two sections. The cards depicted each of the four possible combinations of permuting a triangle or circle on the left, with a triangle or circle on the right. Three of these were labelled underneath with a plus, and one with a minus, and subjects were told that the plusses and minus showed whether the cards did, or did not conform to a rule, which the subjects were asked to deduce and write down. Subjects were then presented with a selection task using a rule expressed in the same way as the subject's written formulation of the deduced rule. A 'matched' control group were also given a selection task using these same formulations of the rule. Significantly more of the experimental group gave the correct solution. Legrenzi repeated the experiment, asking the subjects to deduce the rule and then complete the conditional sentence 'if there is....on the left, then there is....on the right'. 10 of the 11 subjects did this correctly and all ten selected 'TA and FC', on the subsequent selection task, as against only two in the control group.

Subjects were initially exposed to four cards, three of which conformed to the rule to be evaluated and one of which did not. It is thus likely that, when presented with the selection task,

they would expect there to be, again, three conforming, and one nonconforming, instances. Also, in deducing the rule in the first part of the experiment, the 'minus' card would have been most salient. The selection task results could thus be explained in terms of the subject having been led to expect, and focus upon, the falsifying case. As for the Johnson-Laird, Legrenzi & Legrenzi (1972) study, these results may thus be seen as a function of the subjects' experience.

Wason & Shapiro (1971) used thematic rules whose antecedent referred to the name of a town and whose consequent referred to a means of travel (e.g. 'every time I go to Leeds I travel by car'). Subjects were told that the rule referred to journeys the experimenter had made on four separate days. The four cards (with 'Manchester', 'Leeds', 'car' and 'train' printed on them) each had a different day of the week in smaller type at the top. Results showed a significant effect of use of such material, the authors reporting that 10 of the 16 subjects gave the logically correct solution, as against only 2 out of 16 in a control group who were presented with abstract material.. (Unfortunately, no further analysis of responses was given).

The use of all four possible stimuli in the rule was "systematically rotated between the subjects to control for any possible preconceptions about the relation between destinations and modes of transport". Thus it cannot be argued, in this case, that the subjects' experience would lead him to select the falsifying case. Wason & Johnson-Laird (1972) suggest that "the 'story' provides a framework into which the subjects can project themselves by an act of imagination" thus allowing "the conditional nature of the rules to be grasped".

Van Duyne (1974) used 'student record cards' and rules expressing a relationship between degree topic and place of study. Comparable results to those of Wason & Shapiro were observed, the correct solution being obtained from 50% of subjects using a conditional, and 58% of subjects using a universal rule. Again, the rules used appear to represent a believable 'story' as (student) subjects would be aware that different institutions often offer different courses. However, Van Duyne reports that use of thematic materials had no effect (as compared with a group given abstract materials) when the rules were expressed in the disjunctive form 'either not P or Q', or in the form 'not both P and not Q' (argued by Strawson, 1952, to be the most unambiguous expression of implication).

Bracewell & Hidi (1974) investigated whether the effect of thematic materials is due primarily to the fact that the terms used have some everyday meaning or to the fact that there is a meaningful relationship between the terms used (e.g. the concept of travelling). They used four types of universal rule, permuting thematic and abstract terms (names of towns and transport or letters and numbers) with a thematic or abstract relation ('travelling' or reference to stimuli being on one, or the other, side of a card). Thus one condition (A) essentially replicated the Wason & Shapiro study, whereas another condition (D) was a standard abstract task.

The authors report that a significant effect of relation, but not of terms, was observed. However, this result may well have been a function of their method of analysis. Visual inspection of their data reveals that response profiles from conditions B and C differed little from those obtained from the standard abstract task (condition D), whereas data from conditions A and D were noticeably different. 75% of subjects in condition A

made the correct solution and thus the data were compatible with those of Wason & Shapiro. A disturbing feature of the results however, is the very low incidence of TC selection. The authors state in their procedure section that "subjects were told that the conditional was not reversible" but they do not mention how this was done. The form of such an instruction is crucial and it would appear from the data (in particular, the data from the standard abstract condition) that the instructions may have inadvertently cued subjects to ignore the TC card. However, it is unlikely, given the results of Evans (1977b), that this can explain why 11 out of 12 subjects in condition A selected FC.

The best explanation of these results, then, is to conclude, with certain reservations, that they represent a replication of the results of Wason & Shapiro (1971), and that they suggest that both thematic terms and relation are necessary to produce a noticeable effect. It may be that the statement has to appear wholly non-arbitrary in order to make the story 'believable'. The authors report a study in preparation using rules such as 'every time I think of Ottawa I remember car', which apparently had no effect. This type of statement may again be less believable as, although it is common to use means of transport when travelling to places, it is not common to remember means of transport when thinking of places. A similar argument may be applied to the results of Lunzer, Harrison & Davey (1972) who appear to have found no effect of thematic materials except when an unusual presentation (of TC and FC cases only) was used. Their rules also expressed a seemingly arbitrary relationship (between colour of lorries and whether or not they carried coal).

However, Bracewell & Hidi also used reversed forms of rule such as "I travel by car every time I go to Ottawa". For these

rule forms, the thematic content apparently had no effect on responses, as there was no difference between the conditions. This is a very surprising result that suggests that the effect of thematic materials is far more specific than even the Van Duyne (1974) study indicated. The 'story' context generated by 'I travel by car every time I go to Ottawa' appears to be the same as that generated by 'every time I go to Ottawa I travel by car'. Bracewell & Hidi suggest that the former rule may be more often interpreted as implying its converse, but this is unlikely and the suggestion cannot explain their data, as there is no evidence of differential TC selection.

Gilhooly & Falconer (1974) report a study identical to that of Bracewell & Hidi, except that conditional rules were used and English towns were referred to. 'Reversed' rules were not used. They report that it is thematic terms, and not relations, that produce an effect. However, their effect appears to be very marginal. Only 36% of subjects selected FC in the two conditions using thematic terms, as against 26% in the two conditions using abstract terms (data from Evans, 1977b). If the thematic terms and relation condition is compared with the abstract condition, significantly more correct solutions were observed (11 out of 50 to 3 out of 50), and thus the results do provide a further replication of the Wason & Shapiro (1971) study. However, the reported effect is considerably weaker than that observed by Wason & Shapiro and these results, together with the 'reversed' rule results of Bracewell & Hidi, indicate that the effect of thematic materials is less stable or pervasive than was originally believed.

It appears from the above results that, for rules that create a 'believable' story, an effect of thematic content (although variable) is observed and that no effect is observed when more

arbitrary thematic rules are used. However, the results of Manktelow & Evans (1979) challenge both of these interpretations.

They report data from four experiments using thematic rules expressing a relationship between what the experimenter eats and drinks at particular meals (e.g. 'if I eat haddock then I drink gin'). Two of these experiments utilised all four forms (AA, AN, NA and NN) of conditional rule. In no case did response profiles on thematic rules differ from those on abstract rules. This finding may be interpreted as compatible with previous results as, for instance, there is no more relationship between drinking gin and eating haddock than there is between remembering car and thinking of Ottawa. However, Manktelow (1978) reported that differences were observed between these thematic and abstract rules on a number of truth table tasks. All four forms of conditional rules were used and subjects classified significantly more double mismatching ($\bar{P}\bar{Q}$) cases as 'irrelevant' on thematic, than on abstract rules. This result held for both standard and OI conditional rules and also for disjunctive rules, although in the latter case it is questionable whether the result should be described as a function of matching. Thus, although this thematic content had no effect on responses to the selection task, it cannot be considered to be simply equivalent to abstract content.

Experiment 5 of Manktelow & Evans (1979) is very difficult to interpret. The design and the thematic rules used were as in the Wason & Shapiro (1971) study but they found no effect of content. Only two out of 16 subjects selected FC in the thematic condition. Although Manktelow & Evans accept that the Wason & Shapiro study "contains few obvious defects", they point out that, in the Gilhooly & Falconer (1974) study, the thematic (terms and

relation) condition "shows a generally low level of correct responding - only 22%". This is correct but the fact remains that, as was demonstrated above, appropriate analysis of the data of Gilhooly & Falconer reveals that they provide a significant replication of the Wason & Shapiro result. Similarly, the Bracewell & Hidi (1974) replication is criticised by Manktelow & Evans on the grounds that "only the relation and order main effects and relation - order interaction are significant", but this hardly accounts for the fact that, in the directly comparable thematic condition, 75% of subjects gave the correct solution and 92% selected FC. In their concluding paragraph, Manktelow & Evans summarise their criticisms of various studies but fail to mention both Gilhooly & Falconer and Bracewell & Hidi and, in fact, at no point in their discussion do they present an explanation, either of why these two studies replicated that of Wason & Shapiro, or of why their own study failed to do so.

One difference between the Manktelow & Evans experiment (5) and that of Wason & Shapiro is that in the former the experimenter laid four cards on the table under four headings giving the day of the week, whereas in the latter experiment, each card was actually labelled with a day of the week. The (latter) procedure of Wason & Shapiro may have helped convince the subject that the cards did represent an actual record of the experimenter's journeys on particular days. The point is that thematic content may only have an effect on selection task responses if it leads the subject to believe that he is dealing with a 'real', rather than an artificial, problem.

SECTION 3.7 CONCLUSIONS

The crucial determinants of responses to abstract selection tasks appear to be matching bias and the validity of TA, and invalidity of FA, selections. Responses can thus once again be described in terms of a combination of logical and nonlogical factors. It should be noted, however, that the results of Manktelow & Evans (1979) suggest that consequent selections are determined by non-logical factors only. The 'rationalisation' hypothesis of Wason & Evans (1975) and Evans & Wason (1976) adequately accounts for the fact that subjects' introspections do not reveal the existence of such nonlogical tendencies. In line with Wason & Johnson-Laird (1972), it has been argued that use of thematic content on the selection task will only affect responses when it creates a 'believable' story'. For instance, in the case of the 'towns and transport' content, it is possible that an effect is obtained only if the subject is led to believe that the experimenter is making a real, rather than a hypothetical claim about journeys he has actually made.

When thematic content does have an effect on selection task responses, it appears to be to 'cue' the falsifying (FC) card. However, this effect is not necessarily due to the content having aided appreciation of the logical structure of the problem. In fact, that it is not due to an appreciation of logical structure, is suggested by its failure to transfer to a subsequent abstract task (Johnson-Laird, Legrenzi & Legrenzi, 1972).

An alternative explanation, which accords with similar arguments developed in chapters 1 and 2, is that the effect is a function of the subjects' experience. The way in which the subjects' experience would have 'cued' the falsifying case in the experiments

of Legrenzi (1971) and Johnson-Laird, Legrenzi & Legrenzi (1972) has been discussed in chapter 3.6. However, it is not so readily apparent how the subjects' experience cues the falsifying case when the 'towns and transport' content is used. One possible explanation is that, if the subject believes that the rule is a 'real' claim about the experimenter's actual journeys, his responses may be influenced by the context within which such claims are made in 'real life'. If a person makes a statement about his own behaviour and requests someone else to find out whether this statement is 'true or false', the request usually implies: "this is what I claim and I challenge you to prove me wrong".

CHAPTER 4

THEORETICAL APPROACHES TO THE STUDY OF REASONING

SECTION 4.1	'Logical' Explanations	139
SECTION 4.2	The Need for 'Nonlogical' Explanations and the Theories of Evans (1972a,1977b)	146

SECTION 4.1 'LOGICAL' EXPLANATIONS

Nineteenth century logicians clearly believed that the structure of formal logic was a model of the thought processes involved in human reasoning (Henle, 1962). A similar view is implicit in Piagetian theory which holds that adults are capable of 'formal operations', that is, that they are capable of logical reasoning with abstract stimuli. Data from reasoning tasks, however, show this position to be untenable as subjects frequently make logically erroneous responses. Thus, although subjects' responses may be partly determined by the influence of logic, the observed data disconfirm the hypothesis that their behaviour is wholly logical.

The only way to maintain the position that subjects' reasoning is wholly logical is to argue that the subject misinterprets the premisses and that his apparent errors are thus the right answer to the wrong question. The main proponent of this viewpoint has been Henle (1962) who reported a 'qualitative' analysis of subjects' responses to syllogisms involving thematic content relating to 'real life' issues. She argued that errors could be explained in four main ways:

- 1) Restatement of a premiss or conclusion so that its meaning is changed (i.e. the premiss is misinterpreted).
- 2) Omission of one or both premisses leading to direct evaluation of only one premiss and the conclusion, or of the conclusion only.

- 3) The addition of a premiss which is inferred from one of the other premisses.
- 4) A general "failure to accept the logical task".

Subjects pay attention to the content rather than the structure.

There are several criticisms that may be levelled at this argument. Firstly, she places great reliance on subjects' explanations of their responses. The Wason & Evans (1975) results suggest that this reliance is misplaced. Evans (1976) refers to rationalisations creating the 'illusion' of the subject's awareness of his thought processes and this 'illusion' is likely to be even stronger when the content of the problems relates to issues that both the subject and the experimenter have certain preconceptions about. More importantly, her analysis presents no evidence for her theory that subjects are always capable of logical reasoning. She appears to be arguing that subjects would have reasoned logically if they had not ignored premisses, made invalid assumptions and misinterpretations, and generally showed themselves incapable of attending to the logical structure of the problem. Further, most of these explanations cannot be applied to data from abstract tasks in which, for instance, the given premisses are unlikely to imply additional premisses. Admittedly, the existence of response biases suggests that subjects do directly evaluate conclusions and 'fail to accept the logical task', but it would be ridiculous to take this as evidence in favour of logical competence. The only explanation that can feasibly be applied to abstract tasks is premiss-misinterpretation.

A possible misinterpretation, that has figured prominently in the literature, is the interpretation of universals and conditionals as implying their converse (conversion). This was

first proposed by Chapman & Chapman (1959) as a partial explanation of errors in syllogistic reasoning, but has since been used in explaining data from both syllogisms and the selection task (e.g. Bracewell, 1974, Brée, 1973, Dickstein, 1975, Johnson-Laird & Wason, 1970, Smalley, 1974). On conditional inference tasks it has often been assumed that AC and DA inferences are uniquely determined by conversion (e.g. Rips & Marcus, 1977, Staudenmayer, 1975, Taplin, 1971). Many workers in the field have argued that conditionals cause confusion because, when used in 'real life' discourse, their interpretation (as implication or equivalence) is dependent on the context (e.g. Bracewell & Hidi, 1974, Brée, 1974, Bucci, 1978, Fillenbaum, 1975, 1976, Geis & Zwicky, 1971, Legrenzi, 1970, Rips & Marcus, 1977, Staudenmayer, 1975). A somewhat different view was proposed by Tsai (1977) who presented subjects with groups of statements such as "either Pat (-) Bob or Pat (-) Fred" and found that (-) was usually interpreted as being symmetrical. He argued that conversion (as evidenced by AC affirmation) is mediated by a preference for symmetrical relationships.

The prevalent view is thus that, on abstract tasks devoid of contextual 'cues', the conditional will be interpreted, sometimes as implying implication and sometimes as implying equivalence. This theory has been attacked by Braine (1978), who refers to it as the 'chameleon' theory. He presents four specific criticisms as follows:

- 1) The theory fails to explain why more errors are made on MT than on MP.
- 2) The Johnson-Laird & Tridgell (1972) results, suggesting that subjects usually interpret the conditional as implying defective implication, cause problems for the

theory as it cannot then explain why, on inference tasks, most subjects appear to interpret the conditional as implying either (material) implication or equivalence.

- 3) None of the three possible interpretations (material or defective implication, or equivalence) resolves the problem of counterfactuals.
- 4) Different formulations of the conditional have the same truth table and thus should be interpreted in the same way. However, for instance, the IT and OI forms of the conditional rule are psychologically distinct, as certain content can be used on one form but not the other (e.g. 'if one pulls out the knob, the television goes on' cannot sensibly be transposed into 'one pulls out the knob only if the television goes on').

Points 1 and 2 are only valid if it is claimed that 'chameleon' theory explains all the data. However, no theorists have claimed this and it is a separate problem as to why subjects fail to make MT, or make more AC and DA responses on inference tasks than FT - false and FF - true responses on truth table tasks. It should be noted, however, that the apparent defective implication interpretation observed on truth table tasks is not wholly inconsistent with data from other tasks. There is very little evidence that subjects ever interpret a conditional as implying material implication.

Point 3 is somewhat unfair as 'chameleon' theory is not aimed at resolving the problem of counterfactuals. Braine notes, however, that "the defective truth table provides no meaning at all for such sentences". Thus, from a logical positivist viewpoint (e.g. Ayer, 1936), the defective truth table does cope adequately with

counterfactuals

To support point 4, Braine (1978) reports data from Braine (1976), comparing inference frequencies on IT and OI conditional rules. He reports that, for the OI rules, less errors, and shorter latencies, were observed on MT than on MP and argues that this is due to OI rules being processed from consequent to antecedent. However, this finding is inconsistent with the data of Evans (1977a) although he did find that more MT, and less MP inferences were affirmed on OI, than on IT rules. He agrees with Braine that there is a tendency to process OI rules from consequent to antecedent, as they are often used when the consequent precedes the antecedent in time. However, it is not clear why Braine considers this view to be inconsistent with 'chameleon' theory. In fact, point 4 appears to be based on a misunderstanding of 'chameleon' theory as it is specifically because 'real life' usages vary that interpretation is held to be variable. If different forms of rules are used in different ways in 'real life', then these rule forms would be expected to be interpreted differently on abstract tasks.

Thus, Braine's arguments in no way invalidate the theory that subjects sometimes, but not always, convert conditional rules. There is reasonable evidence that conversion does play some role in determining responses. For instance, on syllogistic tasks, more errors are observed on problems involving an A premiss and, on truth table tasks, although the most prevalent response to FT and FF is 'irrelevant', there is a residual tendency to classify FT as false rather than true and FF as true rather than false. However, the existence of matching and conclusion bias shows that it cannot be the case that these truth table classifications, or affirmation of AC and DA inferences, are uniquely determined by conversion. Also, many

observed errors (such as errors on certain syllogistic premiss pairs, failure to make the MT inference and failure to select the FC card on the selection task) cannot be explained in terms of conversion and it is thus impossible to explain the data solely in terms of premiss misinterpretation.

However, there are other forms of 'logical' explanation. such as the hypothesis that errors are due to the application of inappropriate reasoning strategies. For instance, Chapman & Chapman (1959) proposed that subjects reason by 'probabilistic inference', accepting not only necessary, but also probable conclusions. It is possible that, in 'real life' reasoning, subjects are often concerned with how likely a conclusion is, given a set of premisses, and thus, although the subject is not viewed as reasoning in accordance with formal logic, he is seen as reasoning in a consistent, 'logical' manner.¹ Braine (1978) argues that 'real life' reasoning conforms to a pattern of 'natural' logic and presents a set of inference rules (the basis of which is attributed to Gentzen, 1964) upon which this 'natural' logic is based. However, he presents no evidence for his system and makes no effort to show how it may explain errors on reasoning tasks. Although it is certainly possible that subjects may consider plausible inferences to be valid in 'real life' reasoning, it is unlikely that any set of 'real life' inferences could fully explain all the data.

Reference has been made to 'logical' theories concerning premiss misinterpretation and the acceptance of formally invalid inferences that are considered valid in 'real life'. A third type of 'logical' approach is to focus upon specific aspects of the problem that cause difficulty and thus lead to characteristic errors. For instance, it has been proposed that denial of a negative causes

problems on certain disjunctive tasks (e.g. Johnson-Laird & Tridgell, 1972, Roberge, 1976a, 1976b). There is also some evidence from latency data that negatives per se cause problems on reasoning tasks (e.g. Evans, 1977, Evans & Newstead, 1977, Lippman, 1972), although the results of Johnson-Laird & Tridgell (1972) show that this is not necessarily always the case. One finding that any general theory of reasoning must take into account is the preference for 'forward processing' observed on syllogistic tasks. Subjects find it very difficult to reason 'backwards' through a chain of premisses (e.g. Dickstein, 1978b, Frase, 1968b, Johnson-Laird & Steedman, 1978) and this difficulty also appears to apply to conditional problems. Subjects affirm more MP, than MT, inferences and select more TA, than FC, cards on the selection task. Thus more valid responses are observed on problems concerning the antecedent than on problems concerning the consequent, of a conditional rule.

SECTION 4.2 THE NEED FOR 'NONLOGICAL' EXPLANATIONS AND THE
THEORIES OF EVANS (1972a, 1977b).

All the 'logical' explanations, considered in the last section, presume that the subject attends to the logical structure of the problem. Logically correct responses are accepted as being due to the subjects' appreciation of their logical validity and logically incorrect responses are explained in one, or more, of the following three ways:

- 1) As due to a valid inference drawn from a misinterpretation of a premiss (e.g. the conversion theory of Chapman & Chapman, 1959).
- 2) As due to the application of an invalid strategy that is valid in 'real life' logic (e.g. Chapman & Chapman's 'probabilistic inference' theory).
- 3) As due to a characteristic error that occurs during analysis of the structure of the problem (e.g. the sources of error proposed in the Johnson-Laird & Steedman, 1978, model).

However, many responses are not susceptible to a 'logical' explanation. At present, it would be impossible to present a 'logical' explanation of matching bias and very difficult to do so for conclusion bias. Thus, as has been argued throughout this review, it is apparent that responses to problems involving conditionals are mediated by both 'logical' and 'nonlogical' factors.

In the case of syllogistic reasoning, although there are alternative information processing theories, data can also be explained in terms of a combination of logical and nonlogical factors. At the end of chapter 1, it was concluded that the latter

explanation is preferable, and consideration of data from tasks using conditional rules lends further support to this conclusion for two reasons. Firstly, in the case of the selection task, information processing approaches have had little success. An explanation of syllogistic data in terms of a combination of 'logical' and 'nonlogical' factors is thus preferable, as it generalises across paradigms. Secondly, there are noticeable similarities between the nonlogical biases. Both 'atmosphere' and 'matching' are feature matching biases, and the dominance of 'negative atmosphere' may be related to the preference for negative conclusions observed on conditional inference tasks.

At least in the area of conditional reasoning tasks, the view that subjects' responses are mediated by a combination of 'logical' and 'nonlogical' factors is by no means novel. This theory was first proposed by Evans (1972a) who argued that responses on reasoning tasks are mediated by a combination of interpretational factors (tendencies to respond resulting from the subjects' comprehension of the rule) and operational (nonlogical) factors. Evans (1977b) reformulated this theory in mathematical terms, proposing that responses are determined by a weighted addition of interpretational (I) and response bias (R) tendencies.

Formally, the probability of a response (P(R)) is given by:

$$P(r) = \alpha I + (1 - \alpha) R$$

$$\text{where } 0 \leq \alpha \leq 1$$

$$0 \leq I \leq 1$$

$$\text{and } 0 \leq R \leq 1$$

The model assumes reasoning responses to be probabilistic within subjects. Evans points out that it is "not, of course, suggested that individual differences do not affect reasoning behaviour. If, however, behaviour is probabilistic at the level of individual subjects, we might regard this as the principal source of observed variation within an homogeneous sample such as undergraduate students". Evans compares this approach to the 'classical' learning theory of Hull (1943) and points out that, although the idea of stochastic behaviour is unusual in the study of reasoning, the "assumption is common in the study of learning".

Evans fitted this model to the selection task data of Evans & Lynch (1973), and with the adoption of certain assumptions, an impressively good fit was obtained. Some of the assumptions adopted in fitting the model were derived from the data to which it is fitted and thus the extent of the support which this fit lends to the general model is debatable (as Evans admits). However, the Evans theory in general does provide the best published explanation of reasoning task data in that it focusses upon a combination of logical and nonlogical factors.

FOOTNOTES

1. A clear distinction can be made between the ('logical') hypothesis that subjects may accept plausible or 'probabilistic' inferences and the ('nonlogical') hypothesis that subjects may accept conclusions that they know to be true or likely to be true. In the former case, the subject is held to be attending to the logical structure of the problem and inferring a conclusion that is likely, given the premisses. In the latter case, the subject is held to be evaluating the conclusion directly, without reference to the premisses.

PART 2

EXPERIMENTAL WORK

CHAPTER 5

CONDITIONAL INFERENCE EXPERIMENTS

EXPERIMENT 1	156
EXPERIMENT 2	169
APPENDICES	192

All the reasoning task paradigms reviewed have produced evidence for the existence of nonlogical response biases. However, the 'conclusion bias' observed on conditional inference tasks (Evans, 1972b, 1977a) has been little investigated as most studies in the area have used only AA rules. This is unfortunate as little attempt has been made to assess the influence of logical factors when all four forms of conditional rule (formed by systematic negation of the components) are used. Clearly, the extent of a nonlogical bias can only be assessed relative to the effect of such logical factors. For instance, it has been found that matching bias is probably the major determinant of consequent selections on the selection task but that logical validity is an equally strong determinant of antecedent selections (Evans, 1977b, Evans & Lynch, 1973, Manktelow & Evans, 1978).

It would thus be of interest to investigate further the relative effect of logical factors on an inference task using all four forms of conditional rule. The effect of such factors should be reflected in the relative frequencies of the different inferences and in relationships between them. In particular, if some subjects affirm MT because they 'understand' that it is logically valid, then a subject who affirms MT on one rule form should be more likely to do so on another.

The choice of inference task used in this experiment was influenced by the notable dearth of research on the contrapositive, converse and inverse inferences, which are logically related to the MT, AC and DA inferences, respectively. These inferences may be referred to as conditional inferences. For instance, on the AA rule, the MT inference infers 'not P', given 'not Q', whereas the contra-

Table 5.1

Form of Two Valid and Two Invalid Inferences, and the Related Conditional Inferences, from the
Conditional Rule 'If P then Q'

<u>INFERENCE</u>		<u>RELATED CONDITIONAL INFERENCE</u>				<u>VALIDITY</u>
<u>DESIGNATION</u>		<u>GIVEN</u>	<u>INFER</u>	<u>DESIGNATION</u>	<u>FORM</u>	(assuming implication)
MODUS PONENS	(MP)	P	Q	CONDITIONAL	If P then Q	VALID
MODUS TOLLENS	(MT)	Not Q	Not P	CONTRAPOSITIVE	If not Q then not P	VALID
AFFIRMATION OF THE CONSEQUENT	(AC)	Q	P	CONVERSE	If Q then P	INVALID
DENIAL OF THE ANTECEDENT	(DA)	Not P	Not Q	INVERSE	If not P then not Q	INVALID

positive infers 'not P' if 'not Q', and thus may be expressed as a conditional statement. The full form of the conditional inferences related to MP, MT, AC and DA is shown, for the AA rule form, in Table 5.1.

It will be noted from Table 5.1 that the (conditional) inference related to MP is simply the conditional statement from which the inference is made. For the other three inferences, the equivalent conditional inference is different from the conditional statement from which the inference is made.

Clearly, there is also a conditional inference related to each inference on the AN, NA and NN forms of conditional rule. For purposes of clarity, Table 5.2 provides a reference table for the relationships, concerning MT, AC and DA, on all four forms of conditional rule. It will be seen from the table that the contrapositive, converse and inverse are all conditional statements and may thus be referred to as AA, AN, NA, NN, which defines their antecedent and consequent polarity. It should be noted, however, that, in the case of the contrapositive and converse, Q is the antecedent value and P is the consequent value. In the text of chapter 5, these statements are referred to as AA, AN, NA or NN conclusions to distinguish them from the rule forms from which they are derived.

It would be predicted that subjects' behaviour concerning the contrapositive, converse and inverse would be comparable with behaviour concerning the MT, AC and DA inferences, respectively, but this has not, as yet, been tested. In such case, conclusion bias would appear as a tendency to affirm conditionals with negative consequents.

Experiment 1 was thus designed to investigate whether such results would be obtained and to further investigate the relative effects of both nonlogical and logical factors on inference affirmation.

Table 5.2

Relationship between the Contrapositive, Converse and Inverse and
MT, AC and DA, respectively on the Four Forms of Conditional Rule

<u>Rule Form</u>	<u>MT</u>		<u>Contrapositive</u>	
	<u>Given</u>	<u>Infer</u>		
If P then Q (AA)	'not Q'	'not P'	If not Q then not P	(NN)
If P then not Q (AN)	'Q'	'not P'	If Q then not P	(AN)
If not P then Q (NA)	'not Q'	'P'	If not Q then P	(NA)
If not P then not Q(NN)	'Q'	'P'	If Q then P	(AA)

	<u>AC</u>		<u>Converse</u>	
	<u>Given</u>	<u>Infer</u>		
If P then Q (AA)	'Q'	'P'	If Q then P	(AA)
If P then not Q (AN)	'not Q'	'P'	If not Q then P	(NA)
If not P then Q (NA)	'Q'	'not P'	If Q then not P	(AN)
If not P then not Q(NN)	'not Q'	'not P'	If not Q then not P	(NN)

	<u>DA</u>		<u>Inverse</u>	
	<u>Given</u>	<u>Infer</u>		
If P then Q (AA)	'not P'	'not Q'	If not P then not Q	(NN)
If P then not Q (AN)	'not P'	'Q'	If not P then Q	(NA)
If not P then Q (NA)	'P'	'not Q'	If P then not Q	(AN)
If not P then not Q(NN)	'P'	'Q'	If P then Q	(AA)

EXPERIMENT 1

M E T H O D

Design

A sixteen item reasoning task was presented in booklet form. Each question was presented on a separate page and involved a conditional rule of the form 'If P then Q'. The components were systematically negated to yield four types of rule. Each rule was presented four times.

Each of the sixteen rules (four instances of each of the four types of rule) was printed at the top of a page and labelled 'KEY STATEMENT'. Below this was a list of seven alternative conditionals which the subject was requested to underline if he 'felt' that they 'followed from' the key statement. Eight distinct forms of a conditional can be constructed and these are shown (symbolically) below:

- 1) If P then Q
- 2) If P then not Q
- 3) If not P then Q
- 4) If not P then not Q
- 5) If Q then P
- 6) If Q then not P
- 7) If not Q then P
- 8) If not Q then not P

Each of the key statements was of one of the forms 1 - 4, and the seven alternative conditionals used were the remaining seven members of the list. Thus each question gave the subject the opportunity to affirm a (valid) contrapositive conclusion and six invalid conclusions, including the converse and inverse. Additionally,

they were offered the alternative of underlining the statement 'none of the above' (always appearing last on the list) if they 'felt' that none of the alternatives followed from the key statement.

The order in which the seven alternative conditionals were presented was randomised for each of the sixteen questions. The order in which the questions were presented was randomised for each subject with the constraint that there should be four blocks of four questions in which one of each type of rule appeared and that no two adjacent questions should use the same type of rule.

Subjects

Forty undergraduates at Plymouth Polytechnic acted on a paid volunteer basis. They were tested in groups of from two to fifteen.

Materials

Conditionals involving capital letters and numbers were used. If the antecedent referred to a letter, then the consequent referred to a number and vice versa. Thus the 'key statement' and the seven alternative conditionals were all sentences such as 'if the letter is a C then the number is a 7'. Whether or not the antecedent referred to the letter was randomly determined for each question. The letters and numbers used were randomly determined with the constraint that no letter should be used on more than one question. Each problem was presented on an A4 page.

Procedure

Each subject was first shown a few white index cards with a vertical line drawn down the centre and a capital letter on

the left hand side and a number on the right. The letters used on these cards did not appear in any of the questions. Instructions were presented on the first page of the task booklet. A copy of this instruction page is shown in Appendix A.

RESULTS

Each rule form was presented four times to 40 subjects. The percentage frequency (out of 160) of affirmation of the three main alternatives is shown in Table 5.3 (the other four alternatives were invalidly affirmed on only 1.7% of occasions). The following two points about Table 5.3 should be noted:

- 1) The column headings refer to the form of the conclusion affirmed and not to the form of the rule from which the conclusion is drawn (i.e. the 'key statement').
- 2) Reference to the form of conclusion as simply AA, AN, NA or NN allows a standardised comparison between contrapositive and converse conclusions, which argue from Q to P, and inverse conclusions, which argue from P to Q.

As it was predicted that the experiment would yield similar data to that obtained from studies of MT, AC and DA, the data are compared in the table with those of Evans (1977a). In each cell, each inference is compared with the related conditional inference, in accordance with the relationships shown in Table 5.2.

Two factors indicate that the two sets of data shown in Table 5.3 are comparable:

- 1) The average cell frequencies are about the same - 51% for the data from this experiment and 49% from Evans (1977a).
- 2) If the two sets of data are ranked from 1 to 12 across the cells, a rank correlation of 0.886 is obtained ($p < 0.001$).

Table 5.3

Percentage Frequencies of Affirmation of each of the Four Forms of Contrapositive, Converse and Inverse Conclusions in Experiment 1 (160 data points per cell) compared with Percentage Frequencies of Affirmation of MT, AC and DA obtained by Evans (1977a) (16 data points per cell)

<u>RESPONSE</u>	<u>FORM OF CONCLUSION AFFIRMED</u>			
	<u>AA</u>	<u>AN</u>	<u>NA</u>	<u>NN</u>
<u>CONTRAPOSITIVE</u>	44	72	34	59
<u>MT (EVANS 1977a)</u>	25	56	12	75
<u>CONVERSE</u>	66	72	37	64
<u>AC (EVANS 1977a)</u>	75	81	31	81
<u>INVERSE</u>	37	47	30	54
<u>DA (EVANS 1977a)</u>	19	50	12	69

In general, it is reasonable to conclude that contrapositive, converse and inverse frequencies are similar to MT, AC and DA frequencies, although it would be premature to conclude that they are psychologically equivalent. However, for purposes of brevity and clarity, the contrapositive, converse and inverse will henceforth be referred to as MT, AC and DA, respectively.

A three-way analysis of variance was performed on the data from MT, AC and DA, conclusions, analysed by:

- 1) Conclusion type (MT, AC or DA).
- 2) Polarity of the consequent of the conclusion.
- 3) Polarity of the antecedent of the conclusion.

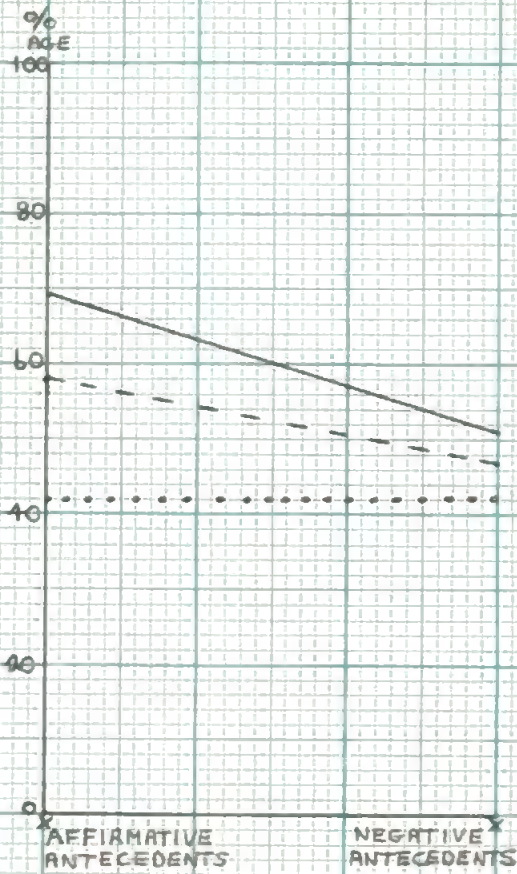
Results of this analysis were as follows:

- 1) There was a significant main effect of conclusion type ($p < 0.01$). Across all rules, subjects affirmed AC on 60% of occasions, MT on 52% of occasions and DA on 42% of occasions. Paired comparisons showed that subjects affirmed significantly less DA than MT ($p < 0.05$), or AC ($p < 0.01$), conclusions. There was no significant difference between frequencies of AC and MT affirmation.
- 2) There was a significant main effect of consequent polarity, such that subjects affirmed more conclusions having negative consequents ($p < 0.01$). This result is consistent with the conclusion bias observed by Evans (1972c, 1977a).
- 3) There was a significant main effect of antecedent polarity, such that subjects affirmed more conclusions having affirmative antecedents. However, there was a significant interaction between antecedent polarity and conclusion type ($p < 0.05$). This interaction is shown in Figure 1. The source of the interaction appeared to be that antecedent polarity had no effect on DA affirmation, as the figure clearly shows.

Further analysis of MT responses was carried out to test for consistent logical behaviour across the four forms of rule. As MT is a valid response on all possible interpretations of a conditional rule, it should be equally valid on all four rules. Thus if

FIGURE 1

ANALYSIS OF THE EFFECT OF ANTECEDENT POLARITY ON THE FREQUENCY OF AFFIRMATION OF THE THREE CONCLUSION TYPES



KEY

- AC CONCLUSIONS
- - - - MT CONCLUSIONS
- DA CONCLUSIONS

MT affirmation is a function of 'logical' behaviour, then MT responses should be associated across the four rules (i.e. those subjects affirming a high frequency of MT on one rule should do so on another rule). This was tested (by median tests) and found to be the case. However, MT responses were also found to be associated with AC and DA responses on other rules and, similarly, there was a strong association between all three inferences on each particular rule (the relevant analyses are tabulated in Appendix B). What these results appear to reveal is that subjects have differential tendencies to affirm conclusions.

If logical consistency across rule forms is to be investigated, this differential tendency to affirm conclusions must be controlled for. One method of accomplishing this is to express, for each subject, on each rule, frequencies of affirmation as a proportion of all affirmations by that subject on that rule. This expresses the extent to which each conclusion is affirmed relative to other conclusions. If subjects affirm MT for logical reasons, then subjects who affirm relatively (to other conclusions) more MT conclusions on one rule should affirm relatively more MT conclusions on another rule.

The data were transformed in the manner described above (the convention $0/0=0$ being adopted) and median tests on MT affirmation across rules were carried out. The results of these tests are shown in Table 5.4 .

The data of Table 5.4 reveal no evidence of a consistent positive relationship between MT affirmation on one rule and MT affirmation on other rules.

Table 5.4

Comparison (by Median Tests) of MT Conclusions on one Rule with MT Conclusions on Other Rules
(Data from Experiment 1, Expressed as a Proportion of all Affirmations on that Rule.)

RULE FORM COMPARISON	BOTH ABOVE MEDIAN	<u>FIRST</u> ABOVE MEDIAN	<u>FIRST</u> MEDIAN OR BELOW	BOTH MEDIAN OR BELOW	Z ¹
		<u>SECOND</u> MEDIAN OR BELOW	<u>SECOND</u> ABOVE MEDIAN		
AA v AN	7	6	11	16	+ 0.44
AA v NA	6	7	13	14	0.00
AA v NN	8	5	12	15	+ 0.68
AN v NA	3	15	16	6	- 3.21
AN v NN	6	12	14	8	- 1.59
NA v .NN	10	9	10	11	0.00
<hr/>					
Average Z					- 0.61

DISCUSSION

Subjects affirmed significantly more conclusions with negative consequents and this finding is a parallel of the conclusion bias results of Evans (1972b and 1977a). Evans (1977a) also observed higher frequencies of AC responses when the minor premiss was affirmative. This latter finding was extended in the present study, as subjects affirmed more AC, and MT, conclusions having affirmative antecedents.

There are thus apparently two 'nonlogical' effects of polarity mediating responses: a preference for conclusions having negative consequents and a preference for conclusions having affirmative antecedents. The bases of these biases are not at present known. However, on the basis of the results of this experiment, it appeared possible that they reflect a heuristic that maximises the subjects' chances of accepting statements that are unlikely to be proved wrong. The larger the class of event to which a statement refers, and the more specific its predictions, the more potentially falsifiable the statement. Thus, in most non-binary situations, a preference for affirmative antecedents and negative consequents would produce less potentially falsifiable statements, that refer to a small class of events (the specific antecedent value) and make nonspecific predictions (that anything but the specific consequent value will be the case). For instance, a statement such as 'if P then not Q' can only be falsified by the concurrence of P and Q, whereas the statement 'if not P then Q' can be falsified in many ways. In conditions of uncertainty about the content, subjects may have learned that the former type of statement is 'safer' to adopt.

There is little evidence that logical validity played a role in mediating responses. The strong association between all conclusion types, across all forms of major premiss, indicates that subjects differ in their tendency to affirm a given conclusion, but this difference is unlikely to have a 'logical' substrate. The highest, and most consistent, associations appear to be between MT and DA (see tables in Appendix B) and it would be very difficult to construct an 'interpretational' explanation of this finding.

The analysis of consistency in MT responses across the four rules, shown in Table 5.4, controlled for this differential tendency to affirm conclusions. There is clearly no evidence of a consistent positive relationship between MT affirmation on one rule and another, and thus no evidence that MT affirmation results from an appreciation of its logical validity.

One aspect of the results of experiment 1 that might be considered evidence of subjects' logical ability is the very low frequency of (erroneous) affirmation of the other (four) invalid conclusions offered, which were affirmed on only 1.7% of occasions. These alternatives have a logical relationship with MP, MT AC and DA and may be referred to as their opposites. This relationship is shown for the AA rule in Table 5.5.

The relationship shown in Table 5.5 can be easily extended to inferences expressed in the form of conditional statements. For instance, the contrapositive opposite is 'if not Q then P'.

Table 5.5

The Relationship Between the MP, MT, AC and DA Inferences and their
Opposites on the Conditional Rule: 'If P then Q'

<u>INFERENCE</u>	<u>FORM</u>	<u>OPPOSITE INFERENCE</u>	<u>FORM</u>
MP	Infer Q, given P	MP (O)	Infer 'not Q', given P
MT	Infer 'not P', given 'not Q'	MT (O)	Infer P, given 'not Q'
AC	Infer P given Q	AC (O)	Infer 'not P', given Q
DA	Infer 'not Q', given 'not P'	DA (O)	Infer Q, given 'not P'

As may be seen from Table 5.5, if MP, MT, AC or DA are considered true then their opposites should be considered false and, if they are considered indeterminate, then their opposites should also be considered indeterminate. In fact, these 'opposites' have been used by various workers as direct measures of the four inferences (e.g. Evans, 1972c, Roberge, 1971b, Taplin, 1971, Taplin & Staudenmayer, 1973). This use has rested on an assumption that subjects reason consistently. However, inferences that are logically equivalent are not necessarily psychologically equivalent. For instance, although logically related AC affirmation cannot be used as a direct measure of DA affirmation.

It would thus be of interest to investigate these 'opposites' in more detail, as the design of experiment 1 did not allow subjects to evaluate them as either 'false' or 'indeterminate'.

Accordingly, the experiment was repeated, asking subjects to respond 'true' (a tick), 'false' (a cross) or 'neither' (a question mark) to each conclusion. Apart from providing a replication of the results of experiment 1, this design allows for the following additional tests:

- 1) If subjects understand the logical relationship between a conclusion and its opposite, then, if they evaluate a conclusion as 'true', they should evaluate its opposite as 'false', and, if they evaluate a conclusion as 'indeterminate', they should evaluate its opposite as 'indeterminate'. Thus an association would be expected between 'true' responses on MT, AC and DA and 'false' responses on MT (0), AC (0) and DA (0) respectively.
- 2) As MT (0) is invalid on any interpretation of the major premiss, then, if MT (0) denial is a function of 'logical' behaviour, such denials should be associated across the four rules. The influence of logical validity can thus be investigated by testing (with suitable control for differential tendencies to affirm or deny) for consistency across the four rules in the case of both MT and MT (0).
- 3) If the effects of polarity of antecedent and consequent of conclusion, observed in experiment 1, are the resultant of a preference for less falsifiable statements (as suggested above), then subjects would be expected to respond 'false', more often to statements that are more falsifiable (than others). The broader the referent, and the more specific the prediction, the

more falsifiable the statement and thus more
denial of 'opposite' conclusions having negative
antecedents, and affirmative consequents, would be
expected.

EXPERIMENT 2

METHOD

This experiment used identical test booklets to those used in experiment 1 and again used 40 undergraduates at Plymouth Polytechnic as subjects on a paid volunteer basis.

The only difference was that, on the instruction page of the booklet, subjects were requested to put a tick by any statement that they thought necessarily true if the key statement was true, a cross by any statement that they thought necessarily false and a question mark by a statement if they thought it to be neither necessarily true nor false. A copy of the instruction page is shown in Appendix A.

RESULTS

The percentage frequencies of responses on all seven conclusions is shown in Table 5.6. Again, it should be noted that, as in Table 5.3, the column headings refer to the form of conclusion evaluated and not to the form of the rule from which the conclusion is drawn.

Results of ANOVA on 'True' Responses

As in experiment 1, a three-way analysis of variance was performed on affirmation ('true' responses) of MT, AC and DA conclusions.

There was a significant main effect of conclusion type ($p < 0.01$). Across all rules, subjects affirmed AC on 81% of occasions, MT on 62% of occasions and DA on 56% of occasions. Paired

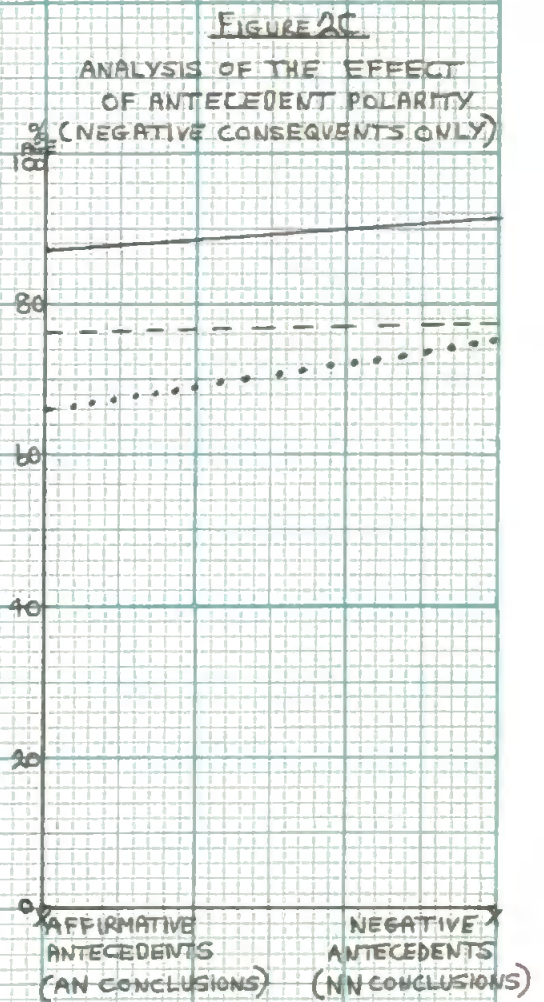
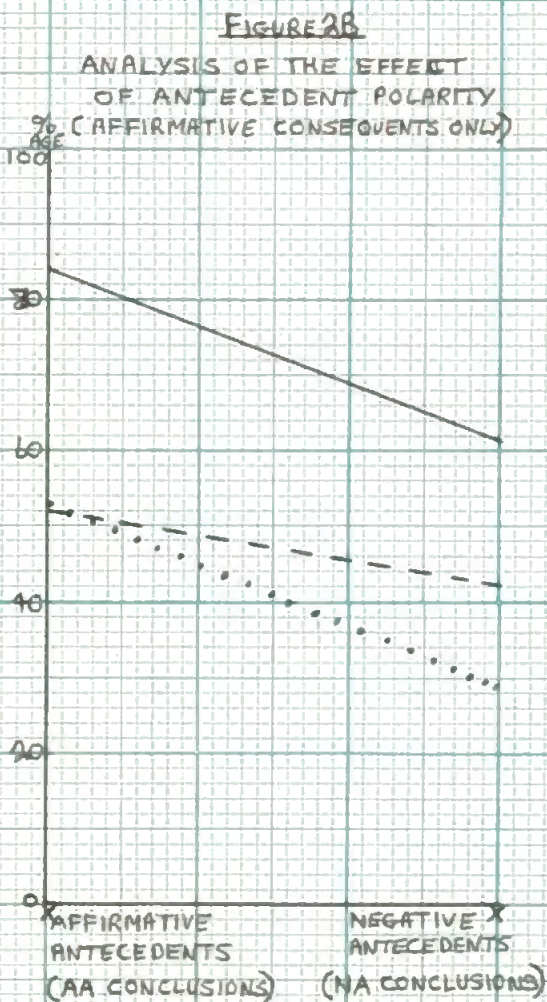
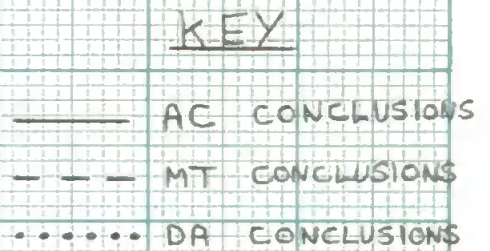
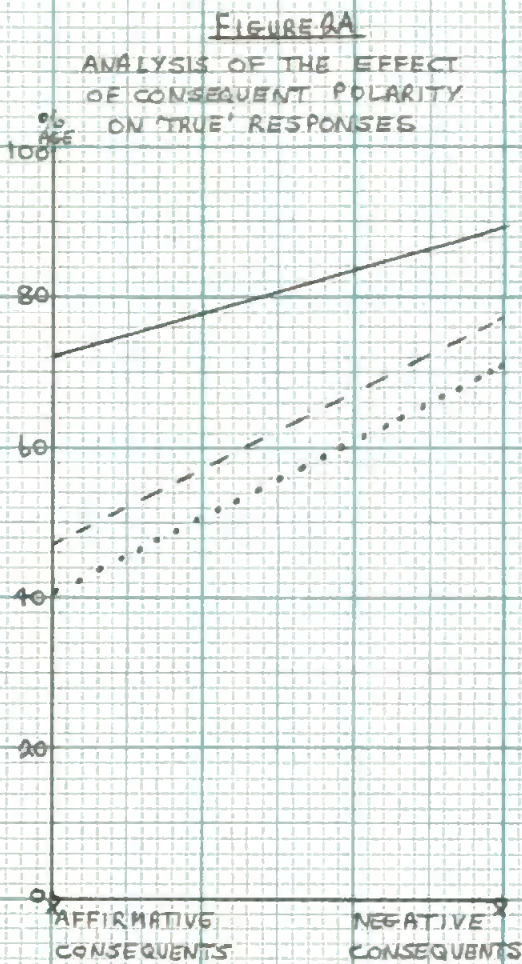
comparisons showed that subjects affirmed significantly more AC than MT ($p < 0.01$), or DA ($p < 0.01$), conclusions. There was no significant difference between frequencies of MT and DA affirmation. However, the latter comparison was significant in experiment 1 and, on the basis of the two studies, it is reasonable to conclude that subjects tend to affirm more AC than MT, and more MT than DA, conclusions.

Table 5.6

Percentage Frequencies of T, F and ? Evaluations of all Seven Conclusions in Experiment 2, Analysed by Form of Conclusion Evaluated

<u>INFERENCE</u>	<u>FORM OF CONCLUSION EVALUATED</u>											
	<u>AA</u>			<u>AN</u>			<u>NA</u>			<u>NN</u>		
	<u>T</u>	<u>F</u>	<u>?</u>	<u>T</u>	<u>F</u>	<u>?</u>	<u>T</u>	<u>F</u>	<u>?</u>	<u>T</u>	<u>F</u>	<u>?</u>
MT	52	14	34	76	11	13	42	14	44	77	4	19
AC	84	2	13	87	2	11	61	6	33	91	3	6
DA	53	13	34	66	11	22	29	15	56	75	6	19
MP (0)	2	96	2	2	96	1	0	89	11	4	79	17
MT (0)	1	93	6	5	72	23	2	86	12	11	54	35
AC (0)	2	90	7	4	82	14	3	72	24	6	52	42
DA (0)	4	84	12	5	69	26	2	76	22	4	44	52

As in experiment 1, there was a significant main effect of consequent polarity, such that subjects affirmed significantly more conclusions having negative consequents ($p < 0.01$). However, there was a significant interaction between consequent polarity and conclusion type ($p < 0.05$). This interaction is



shown in figure 2A. The significant interaction is possibly due to the apparently weaker effect of consequent polarity on AC conclusions. However, the effect of consequent polarity was significant on all three conclusion types.

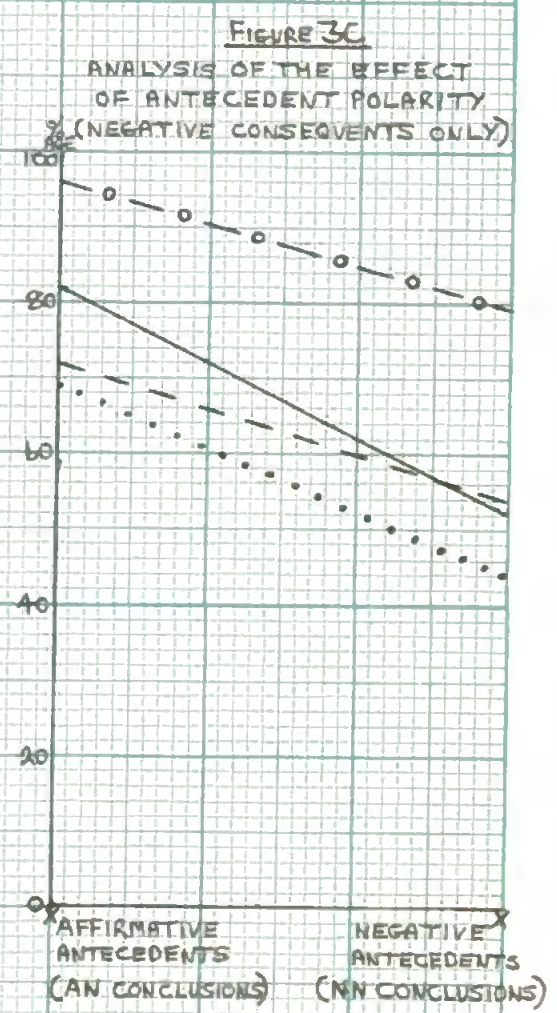
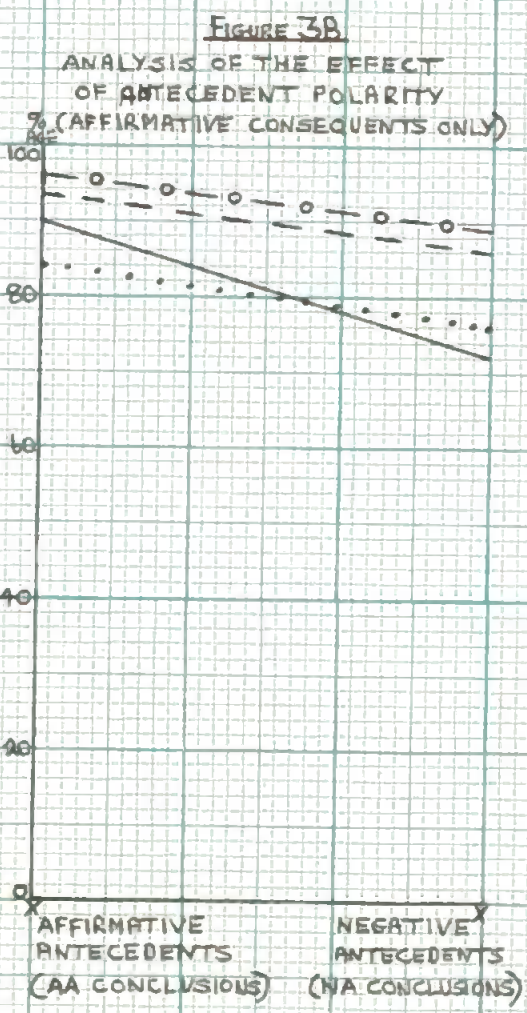
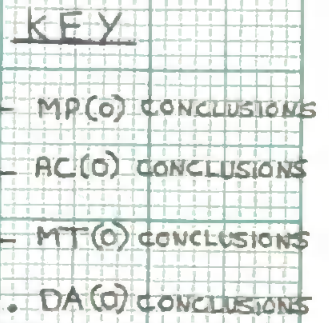
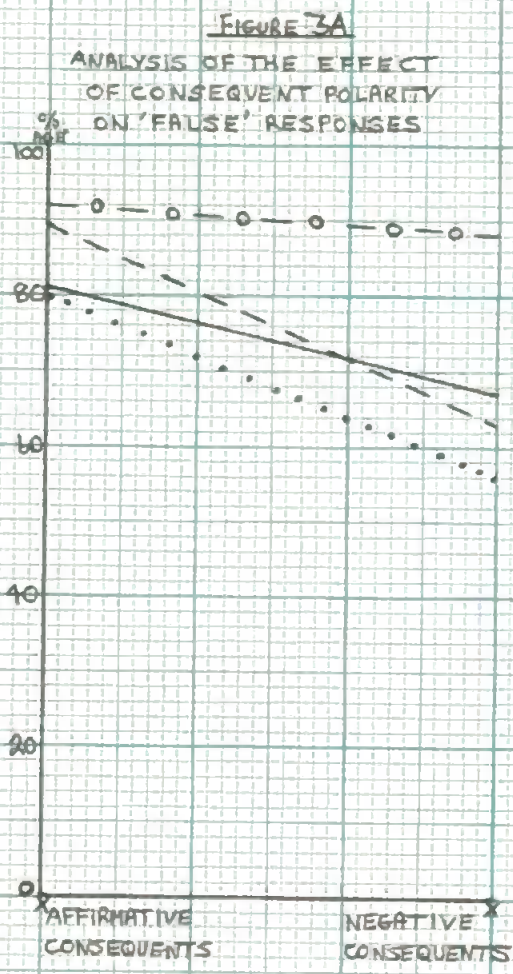
As in experiment 1, there was a significant main effect of antecedent polarity, such that subjects affirmed significantly more conclusions having affirmative antecedents ($p < 0.01$). However, there was a significant interaction between antecedent, and consequent, polarity. This interaction is readily apparent from figures 2B and 2C. There was no significant effect of antecedent polarity on conclusions having negative consequents (figure 2C). Reanalysis of the results of experiment 1 showed that a similar result obtained, although the interaction between antecedent and consequent polarity was not significant.

Results of ANOVA on 'False' Responses

A three-way analysis of variance was performed on denial ('false' responses) of MP (0), MT (0), AC (0) and DA (0) conclusions.

There was a significant main effect of conclusion type ($p < 0.01$). Across all rules, subjects denied MP (0) on 90% of occasions, MT (0) on 76% of occasions, AC (0) on 74% of occasions and DA (0) on 68% of occasions. Paired comparisons showed that subjects denied significantly more MP (0), than any other type of conclusion. No other comparisons were significant.

There was a significant main effect of consequent polarity, such that subjects denied significantly more conclusions having affirmative consequents. This accords with the hypothesis that subjects would prefer to deny conclusions that make specific,



affirmative, predictions. However, as in the case of 'true' responses, there was a significant interaction between consequent polarity and conclusion type ($p < 0.01$). This interaction is shown in figure 3A. The interaction appears to be due to two factors. Firstly that the effect of consequent polarity was not significant on MP (0) conclusions and, secondly, that the effect appears to be weaker (although significant) on AC (0) conclusions.

There was a significant main effect of antecedent polarity, such that subjects denied more conclusions having affirmative antecedents ($p < 0.01$). This finding is not in accordance with the hypothesis that subjects would prefer to deny (more falsifiable) statements that have 'broad' (i.e. negative) referents. As in the case of 'true' responses, there was a significant interaction between the effects of antecedent, and consequent, polarity ($p < 0.05$). This interaction may be inferred from figures 3B and 3C. The effect of antecedent polarity is clearly stronger for conclusions having negative consequents. Thus, as for 'true' responses, antecedent polarity has a greater effect when the consequent polarity does not favour the response. However, for 'false' responses, separate analyses showed that the effect of antecedent polarity was significant on conclusions having negative and conclusions having affirmative, consequents.

It is worth noting that the effect of antecedent polarity appeared to be stronger on 'false', than on 'true', responses ($F = 48.255$ compared with $F = 9.416$).

Chi-Squared Analyses

Associations between evaluation of a conclusion as 'true' and its 'opposite' as 'false', were compared by median tests on each

of the four rule forms and a further analysis was conducted across all four forms of rule. As in experiment 1, subjects' response to a given conclusion was expressed in terms of a percentage of total 'true' (or 'false' in the case of opposite conclusions) responses, by that subject, on that rule. (the $\frac{0}{0} = 0$ convention was not necessary in this experiment). The results of the median tests are shown in Table 5.7.

A relationship between MT and MT (0) and DA and DA (0) conclusions is clearly apparent from Table 5.7. For these two comparisons, there is a positive association on each of the four rules and, although in most cases these associations do not reach significance, analysis across all four rules yields a significant association in both cases ($p < 0.03$, two-tailed). However, no such relationship exists between AC and AC (0).

As data are expressed as a proportion of total conclusions affirmed (or denied), data from each conclusion type are not independent within affirmation or denial, and thus statistical comparisons cannot be made between (MT, AC and DA) conclusions, or between 'opposite' conclusions. It is interesting to note, however, that only five subjects were classified as above median (across all four rules) on both DA and AC or on both MT and AC, whereas 15 subjects were classified as above median on both MT and DA. This latter result would yield a significant positive association if tested by chi-squared, although such a test would be technically invalid. However, the violation of the chi-squared assumption of independence should, on these data, favour negative associations and increase the likelihood of a Type 2 error when testing positive associations.²

Table 5.7

Comparison of 'True' Responses to Conclusions with 'False' Responses to their Opposites, on each of the Four Rules in Experiment 2. (Data for each Conclusion expressed as a proportion of all

Affirmations or Denials on that Rule.

COMPARISON	RULE	Affirmations or Denials on that Rule.				Z
		BOTH ABOVE MEDIAN	1st ABOVE MEDIAN 2nd MEDIAN OR BELOW	1st MEDIAN OR BELOW 2nd ABOVE MEDIAN	BOTH MEDIAN OR BELOW	
MT v MT (0)	IF P THEN Q	5	3	8	24	+ 1.60
	IF P THEN NOT Q	13	6	7	14	+ 1.90
	IF NOT P THEN Q	13	7	6	14	+ 1.90
	<u>IF NOT P THEN NOT Q</u>	11	7	8	14	+ 1.24
	ACROSS ALL RULES	14	6	6	14	+ 2.21
AC v AC (0)	IF P THEN Q	7	9	12	12	- 0.06
	IF P THEN NOT Q	8	6	12	14	+ 0.33
	IF NOT P THEN Q	10	8	9	13	+ 0.60
	<u>IF NOT P THEN NOT Q</u>	6	13	12	9	- 1.30
	ACROSS ALL RULES	10	10	10	10	0.00
DA v DA (0)	IF P THEN Q	12	8	7	13	+ 1.27
	IF P THEN NOT Q	13	7	6	14	+ 1.90
	IF NOT P THEN Q	10	10	8	12	+ 0.32
	<u>IF NOT P THEN NOT Q</u>	14	6	6	14	+ 2.21
	ACROSS ALL RULES	14	6	6	14	+ 2.21

There is thus quite good indication that, even when differential tendencies to affirm statements are controlled for, there is a positive association between affirmation of MT and DA. This interpretation is compatible with the suggestion made in experiment 1 that, when the raw (untransformed) frequencies are considered, the associations between MT and DA appear, in general, to be somewhat stronger, and more consistent, than those between MT and AC or those between DA and AC.

Finally, the consistency of MT affirmation, and MT (0) denial, across rules may be considered. The association between MT affirmation across rules was tested as in experiment 1, and the results are shown in Table 5.8. As the same null hypothesis was tested in both experiments, the table shows the Z values obtained from both experiments, together with an estimate of the combined Z value calculated by the Stouffer method.³

Across the two studies, there is no indication of a consistent positive association between MT affirmation across all rules. There is one significant positive association, between MT affirmation on NA and NN rules, ($p < 0.02$, two-tailed), but there is also one significant negative association, between MT affirmation on AN and NA rules, ($p < 0.03$, two-tailed). It should be noted, however, that there are a sufficient number of high Z values to indicate that MT affirmations are not independent of MT affirmations on other rules.

As MT (0) is equally invalid on all possible interpretations of a conditional rule, similar tests for logical consistency may also be performed on this inference. The results of such tests are shown in Table 5.9. As for MT affirmation, although there is one significant association (between MT (0) denial on NA and NN rules, $p < 0.03$, two-tailed), there is no indication of a consistent positive association between MT (0) denial across the four rules

Table 5.8

Comparison (By Median Tests) of Affirmation of MT on one Rule with Affirmation of MT on Other Rules (Data from Experiment 2, expressed as a proportion of all affirmations on that rule) Together with Z Values Obtained in Experiment 1 and Combined Estimate of Z Values.

<u>RULE FORM</u>	<u>BOTH</u>	<u>1st ABOVE</u>	<u>1st MEDIAN</u>	<u>BOTH</u>	<u>Z VALUE</u>	<u>Z VALUE</u>	<u>COMBINED</u>
<u>COMPARISON</u>	<u>ABOVE</u>	<u>MEDIAN</u>	<u>OR BELOW</u>	<u>MEDIAN</u>		<u>OBTAINED IN</u>	<u>ESTIMATE</u>
	<u>MEDIAN</u>	<u>2nd MEDIAN</u>	<u>2nd ABOVE</u>	<u>OR BELOW</u>		<u>EXPERIMENT 1</u>	<u>OF Z</u>
		<u>OR BELOW</u>	<u>MEDIAN</u>				
AA v AN	7	1	12	20	+ 2.14	+ 0.44	+ 1.82
AA v NA	3	5	17	15	- 0.40	0.00	- 0.28
AA v NN	5	3	13	19	+ 0.72	+ 0.68	+ 0.99
AN v NA	10	9	10	11	0.00	- 3.21	- 2.27
AN v NN	10	9	8	13	+ 0.60	- 1.59	- 0.70
<u>NA v NN</u>	15	5	3	17	<u>+ 3.50</u>	<u>0.00</u>	<u>+ 2.47</u>
Average Z					+ 1.09	- 0.61	+ 0.34

Table 5.9

Comparison (By Median Tests) of Denial of MT (0) on one Rule with Denial of MT (0) on Other Rules
Data from Experiment 2, Expressed as a Proportion of all Denials on that Rule.

<u>RULE FORM</u> <u>COMPARISON</u>	<u>BOTH ABOVE</u> <u>MEDIAN</u>	<u>1st ABOVE</u> <u>MEDIAN</u> <u>2nd MEDIAN</u> <u>OR BELOW</u>	<u>1st MEDIAN</u> <u>OR BELOW</u> <u>2nd ABOVE</u> <u>MEDIAN</u>	<u>BOTH MEDIAN</u> <u>OR BELOW</u>	<u>Z</u>
AA v AN	7	6	13	14	0.00
AA v NA	7	6	12	15	+ 0.22
AA v NN	6	7	13	14	0.00
AN v NA	7	13	12	8	- 1.27
AN v NN	11	9	8	12	+ 0.63
<u>NA v NN</u>	13	6	6	15	<u>+ 2.20</u>
Average Z					+ 0.30

DISCUSSION

Main Effects of Polarity

In both experiments 1 and 2, subjects affirmed more MT, AC and DA conclusions having negative consequents and more such conclusions having affirmative antecedents. The effect of consequent polarity was significant on all three types of conclusion, in both experiments. The effect of antecedent polarity was significant only for AC and MT in experiment 1, but was significant for DA in experiment 2. It is thus reasonable to conclude that both polarity of antecedent and consequent affects the affirmation of all three conclusions. In both experiments, however, there was no significant effect of antecedent when the consequent was negative.

In experiment 2, subjects denied more 'opposite' conclusions having affirmative consequents and more such conclusions having affirmative antecedents. These effects of polarity were significant on all four conclusions except MP (0), for which the effect of consequent polarity was not significant.

In the discussion of experiment 1, the hypothesis was developed that the effects of polarity on affirmation are the result of a preference for less falsifiable statements. That is, statements that have narrow (affirmative) referents and that make nonspecific (negative) predictions. This hypothesis can explain why subjects affirm more conclusions having affirmative antecedents and more conclusions having negative consequents. On the basis of this hypothesis, it was presumed that subjects would display a preference for denial of more falsifiable statements. That is, statements that have broad (negative) referents and that

make specific (affirmative) predictions. Accordingly, it was predicted that subjects would deny more conclusions having negative antecedents, and more conclusions having affirmative consequents. The first of these predictions was disconfirmed, as subjects denied significantly more conclusions having affirmative antecedents.

However, somewhat unfortunately, it was realised in retrospect that this (disconfirmed) prediction did not follow from the hypothesis. It is true that statements with broad referents are more falsifiable, than statements with narrow referents, but the denial of statements with broad referents is also more falsifiable than the denial of statements with narrow referents. In effect denial of a statement merely reverses the polarity of the consequent. For instance, to state that 'it is not the case that, if not P then Q', is equivalent (presuming defective implication) to stating that 'if not P then not Q'.

If subjects are attempting to make decisions that are least likely to be wrong, they thus should tend to avoid both affirmation and denial of statements with negative antecedents (i.e. with broad referents), and the data from experiment 2 are consistent with this 'caution' hypothesis. In summary, this effect of 'caution' may be seen as operating in the following two ways:

- 1) Subjects tend to avoid making any prediction about statements with broad (negative) referents. In most non-binary situations, such statements provide more opportunity for the prediction to be tested (and thus more opportunity for the prediction to be disconfirmed).
- 2) In most non-binary situations, negative predictions are less specific than affirmative predictions. Subjects may thus be seen as affirming more, and denying

less, statements making nonspecific, than specific, predictions. Nonspecific predictions, when tested, are more likely to be confirmed.

However, there are alternative 'logical' explanations that must be considered.

For each conclusion, its opposite will have the same polarity of antecedent and the opposite polarity of consequent. For instance, on the AA rule, the MT conclusion has a negative antecedent and a negative consequent ('if not Q then not P'), whereas the MT (0) conclusion has a negative antecedent and an affirmative consequent ('if not Q then P'). Thus, given the direction of the polarity effects, if they are in favour of a conclusion, they will also be in favour of its opposite.

It could thus quite reasonably be objected that effects of antecedent and consequent polarity are not nonlogical biases, but simply reflect the fact that there is a logical relationship between affirming a conclusion and denying its opposite. The significant associations between affirmation of MT and denial of MT (0), and between affirmation of DA and denial of DA (0), support the idea that subjects did perceive the logical relationship between conclusions and their opposites.

However, although apparent nonlogical effects of polarity on 'false' responses may be explained in terms of a perceived logical relationship between conclusions and their opposites, this cannot explain polarity effects on 'true' responses,⁴ Further, in chapter 2 it has been pointed out that no ('logical') interpretational explanation of the effects of polarity on 'true' responses is possible, for two main reasons:

- 1) Differences in MT affirmation cannot be explained.
- 2) The effects of polarity of conclusion apply to different rule forms in the case of AC than they do in the case of DA, but it is impossible to construct an interpretational theory in which the validity of AC does not 'move' with the validity of DA.

Thus, as there is good reason to view the consistent effects of polarity of conclusion on 'true' responses as nonlogical, there are two possible interpretations of the effects of polarity on both 'true' and 'false' responses, as follows:

NONLOGICAL	As stated earlier
(SEMI) LOGICAL	That nonlogical biases affect 'true' responses but not 'false' responses, which are determined by an appreciation of the relationship between a conclusion and its opposite.

The second interpretation lacks parsimony and appears to assume that subjects make 'false' responses after making 'true' responses. It is thus argued that, to be acceptable, a ('logical') explanation of the effects of polarity on 'false' responses, as solely due to the subjects' appreciation of the logical relationship between a conclusion and its opposite, would need to also present a viable 'logical' explanation of the effects of polarity on 'true' responses.

As it appears that no such explanation is possible, it is concluded that 'false' responses are not simply a 'mirror image' of 'true' responses, determined by the logical relationship between them, but that nonlogical effects of polarity affect both 'true' and 'false' responses. In fact, the effect of antecedent

polarity appears stronger on 'false' responses.

Implications of the Results of the Chi-Squared Analyses

If a subject affirms relatively more MT (than other) conclusions, or denies relatively more MT (0) conclusions, due to an appreciation of their logical validity (or invalidity), then this behaviour should be consistent across the four rules. That is, if certain subjects understand the logical necessity of MT, then these subjects should consistently affirm relatively more MT, and deny relatively more MT (0), than other subjects. However, controlling for differential tendencies to affirm, the tests of consistency in MT affirmation across the four rules strongly suggest that MT affirmation is not mediated by an appreciation of its logical validity.

It is noticeable, however, that there is a significant positive association between MT affirmation on the NA and NN rules, and a near significant positive association between MT affirmation on the AA and AN rules ($p < 0.07$). In both cases, consequent polarity has the same effect on both rules. However, there is a significant negative association between MT affirmation on NA and AN rules. Polarity of both antecedent and consequent favours MT affirmation on the AN rule and is against MT affirmation on the NA rule.

Thus, although providing no evidence of logical consistency, the data provide some indication of consistency in susceptibility to the effects of polarity. No 'logical' explanation could cope with a negative association, as how could a subject, who is more likely (than others) to appreciate the validity of MT on one rule, be less likely to appreciate its validity on another rule?

However, if a subject was more susceptible to the effects of polarity than other subjects, he would be expected to be more likely (than other subjects) to affirm MT conclusions on the AN rule, when both effects of polarity favour affirmation, and to be less likely to affirm MT conclusions on the NA rule, when both effects of polarity act against affirmation. Also, as the effect of consequent polarity is stronger than that of antecedent polarity, his responses would be expected to be positively associated on conclusions having the same polarity of consequent.

The idea that subjects are consistent in their susceptibility to the effects of polarity thus accounts for the main (positive and negative) associations observed between MT affirmation on different rules. The tests of association between MT (O) denial on different rules yield somewhat weaker support for this interpretation (for instance, given the nonsignificant negative association between MT(O) denial on AN and NA rules) but again yield no evidence of logical consistency.

If subjects are consistent in their susceptibility to the effects of polarity, then this is clearly an alternative explanation of the associations between MT affirmation and MT (O) denial, and between DA affirmation and DA (O) denial, which may well not have been due to an appreciation of the logical relationship between a conclusion and its opposite. For instance, subjects who are more likely to affirm MT when its consequent is negative, may be more likely to deny MT (O) when its consequent is affirmative, simply because they are more susceptible to the effects of polarity.

Admittedly, if subjects are consistent in their susceptibility to the effects of polarity, it is difficult to explain why AC affirmation was not associated with AC (O) denial.

However, this problem equally applies to the argument that subjects perceive a logical relationship between conclusions and their opposites. A possible explanation of the lack of association between AC affirmation and AC (O) denial will be considered later.

Preferences for Affirmation or Denial of the Different Conclusions

Although the effects of polarity appear to be non-logical factors, there is evidence of the apparent effect of logical factors on the data. Subjects showed a strong tendency to affirm MT, AC and DA and to deny MT (O), AC (O), and DA (O) and, if they did not make these evaluations, they usually evaluated these conclusions as 'indeterminate'. Subjects thus appear to be fairly capable of determining that MT, AC and DA are not 'false' and that MP (O), MT (O), AC (O) and DA (O) are not 'true'. However, the amount of appreciation of logical structure required for this is debatable.

There was some divergence from this pattern. For instance, MT, which was evaluated as indeterminate on 27.3% of occasions, was denied on 10.8% of occasions.

Main Effects of Inference

In both experiments, subjects affirmed more AC than MT, and more MT than DA, conclusions. It would possibly be expected that a main effect of conclusion type would be evidence for the effect of logical factors. However, that subjects affirm far more AC than MT and DA, is almost impossible to interpret in this way.

As has been mentioned, the highest AC frequencies are not yielded by the same rules as yield the highest DA frequencies. Given this, and the much higher frequency of AC, than DA, affirmations,

it would be very difficult to attribute AC affirmation to a perception of the conditional rule as implying equivalence. In any case, as AC can never be more valid than MT, the higher frequency of AC, than MT, affirmation does not appear to be explicable in terms of a 'logical' theory.

One possible explanation, of the strong tendency to affirm AC, is that it represents an exact restatement of the terms of the major premiss (in a different order). This explanation is of particular use, as apart from explaining the higher frequencies of AC affirmations, it is able to explain certain other anomalies in the data. Firstly, it would explain why subjects affirmed more AC, than MT, conclusions but did not deny more AC (O), than MT (O) conclusions. AC affirmation, but not AC (O) denial, is mediated by one extra favourable factor (i.e. that it restates the terms of the major premiss) than is the affirmation or denial of any other conclusion type (except MP). Thus, although subjects affirm more AC than MT, there is no reason why they should deny more AC (O) than MT (O). Secondly, as there is no comparable factor affecting AC (O) denial, it would explain the lack of association between AC affirmation and AC (O) denial.

This hypothesis would also explain an apparent discrepancy between data from inference and truth table tasks. Subjects appear to interpret the conditional as implying equivalence more often on inference tasks, than on truth table tasks. However, if AC affirmation is explained as due to its restating the terms of the major premiss, then this is seen not to be the case. The apparent discrepancy is due to an unwarranted comparison of AC affirmation with evaluations of the FT and FF truth table cases. However, the most relevant truth table case to AC affirmation may

well be the TT case. Most subjects appreciate that the TT case, which restates the two components of the rule, is a 'true' case of the rule and this appreciation may well be one factor mediating AC affirmation

Conclusions

The high frequency of MP (0) denial suggests that its logical invalidity strongly influenced responses. This accords with the results of other studies that have reported very low error frequencies for MP affirmation. (e.g. Evans, 1977a, Taplin, 1971). Affirmation of DA was less frequent than affirmation of AC and MT in both experiments (although not significantly less frequent than affirmation of MT, in experiment 2) and it is thus reasonable to assume that the logical status of this conclusion may have influenced responses. Data concerning DA (0) were consistent with this assumption, as denial of DA (0) was less frequent than denial of MT (0) and AC (0), although these comparisons were not significant.

That logical validity may influence responses to MP and DA may be compared with the finding that logical factors play an important role in mediating TA and FA selections on the Wason selection task. Thus, in both cases, logical factors appear to influence behaviour concerning the antecedent of the rule. However, the ('inhibitory') effect of logic appears to have less influence on DA affirmation (and much less influence, if any, on DA (0) denial) than it does on FA selection.

It can also be argued that behaviour on AC and MT conclusions, and their opposites (which argue from the consequent of the conditional rule), is comparable with behaviour concerning consequent selections on the selection task, in that 'logical'

factors play little role in mediating responses. This discussion will be concluded by listing the three main reasons for this point of view:

- 1) There is no evidence of consistent MT affirmation, or MT (0) denial, across the four rules and thus no evidence that affirmation or denial of these conclusions is mediated by an appreciation of the logical validity of MT. There is, however, some indication of consistency in susceptibility to the effects of polarity.
- 2) That there was no significant difference between denial of MT (0) and AC (0) strongly suggests that logical factors had little influence on denial of these conclusions. That subjects affirmed significantly more AC than MT conclusions, does attest to a distinction between these conclusions, but this distinction cannot be due to the influence of logical factors, as no interpretation of the rule can make AC more valid than MT. The difference can be accounted for by the explanation given earlier.
- 3) Consideration of Table 5.6 shows that frequencies of MT and AC (and DA) affirmation are more similar when compared by form of conclusion (as is done in the table), than when compared by the form of the rule from which the conclusion is derived. This point is particularly true of the 'opposite' conclusions. For instance, percentage denial of conclusions of the AA form was 93% for MT (0) and 90% for AC (0), whereas percentage denial of conclusions of the NN form was 54% for MT (0)

and 52% for AC (0). Comparison by major premiss would yield percentage frequencies of MT (0) and AC (0) denial, respectively, of 54% and 90% on the NA rule and 93% and 52% on the AN rule.

The above comparisons clearly indicate that it is the form of conclusion that primarily determines the response. Affirmation of MT and AC, and denial of MT (0) and AC (0) are primarily determined by the nonlogical effects of polarity on responses to a conclusion, irrespective of the type of argument by which the conclusion is supposedly derived. Similarly, on the selection task, consequent selections are primarily determined by matching bias, irrespective of the logical status of the card selected.

FOOTNOTES

1. Z values are derived from chi-squared tests. All chi-squared tests carried out in this, and later, experiments have employed Yates' correction in accordance with standard textbooks (e.g. Siegal, 1956). However, it should be noted that the results of Camilli & Hopkins (1978) indicate that, in the words of the authors, "in all instances the Yates' correction decreases the accuracy of the probability statements". They conclude that use of Yates' correction results "in an unnecessary loss of power" and thus the chi-squared tests used in this chapter should be regarded as conservative.

2. When a subject's MT, AC and DA affirmations are expressed as a proportion of his total affirmations, if one conclusion gains a high proportion, then other conclusions should tend to have a low proportion. Thus, if responses were random, subjects affirming an above median (relative) frequency of DA would be less likely (than other subjects) to also affirm an above median (relative) frequency of MT. The nonindependence thus acts against positive associations.

3. The Stouffer method is attributed to Stouffer et al (1949) and is given by Rosenthal (1978). The formula is:

$$\text{Est. } (Z) = \frac{\sum z}{\sqrt{N}}$$

As Z is a unit normal variate under the null hypothesis, the standard error of the average of a sample of Zs is thus $\frac{1}{\sqrt{N}}$ Hence:

$$\text{Est } (Z) = \frac{\sum z}{N} \text{ divided by } \frac{1}{\sqrt{N}} = \frac{\sum z}{\sqrt{N}}$$

4. Strictly, this explanation can account for the effects of polarity on either 'false' or 'true' responses (i.e. by explaining one set of responses on the grounds that it 'moves' with the other), but the crucial point is that it cannot account for the effects of polarity on both 'true' and 'false' responses.

APPENDICES TO CHAPTER 5

APPENDIX A	Copies of the Instruction Pages on the Task Booklets Used in Experiments 1 and 2	193
APPENDIX B	Tables Showing the Associations Within and Between Inferences in Experiment 1 (Untransformed Data)	196
APPENDIX C	Tables for ANOVAS Performed in Experiments 1 and 2	199

APPENDIX A

COPIES OF THE INSTRUCTION PAGES ON THE TASK BOOKLETS USED
IN EXPERIMENTS 1 AND 2

Instruction Page Heading Test Booklets in Experiment 1

You have been shown a set of cards all having a letter on the left hand side and a number on the right hand side.

You will be given a series of rules to consider, each of which apply to these kinds of cards and define which letters may be paired with which numbers (or vice versa).

For instance, here are some examples of the type of rules you may be given:

EG (1) If the letter is not a V then the number is a 6

EG (2) If the letter is not an S then the number is not a 9

EG (3) If the letter is a C then the number is a 5

EG (4) If the number is a 2 then the letter is not an M

You will find on each page a KEY STATEMENT which is a rule of this kind. Under this are seven other statements.

Your task is to underline any of the other statements that you believe to follow logically from the KEY STATEMENT.

If you do not believe that any of the listed statements necessarily follow from the KEY STATEMENT, you should underline the eighth alternative ('None of the above').

Thus you may underline

EITHER one or more of alternatives 1 to 7

OR alternative 8

Instruction Page Heading Test Booklet in Experiment 2

You have been shown a set of cards all having a letter on the left hand side and a number on the right hand side.

You will be given a series of rules to consider, each of which apply to these kinds of cards and define which letters may be paired with which numbers (or vice versa).

For instance, here are some examples of the types of rules you may be given:

- EG (1) If the letter is not a V then the number is a 6
- EG (2) If the letter is not an S then the number is not a 9
- EG (3) If the letter is a C then the number is a 5
- EG (4) If the number is a 2 then the letter is not an M

You will find on each page a KEY STATEMENT which is a rule of this kind. Under this there are seven other statements.

Your task is to assume that the KEY STATEMENT IS TRUE and on the basis of this assumption to:

- a) Put a tick by any of the other statements that are therefore also necessarily TRUE
- b) Put a cross by any of the other statements that are therefore necessarily FALSE
- c) Put a question mark by any statement that is neither necessarily TRUE or FALSE

APPENDIX B

TABLES SHOWING THE ASSOCIATIONS WITHIN AND BETWEEN
INFERENCES IN EXPERIMENT 1 (UNTRANSFORMED DATA)

Table 5.B.1.

Comparison by Median Tests of MT on one rule with MT, AC and DA on
other rules. (Experiment 1)

<u>INFERENCE</u> <u>COMPARISON</u>	<u>MAJOR</u> <u>PREMISS</u> <u>COMPARISON</u>	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>Z</u> *
MT v MT	AA v AN	11	9	6	14	1.28
	AA v NA	12	8	2	18	2.98
	AA v NN	9	11	5	15	0.99
	AN v NA	10	7	4	19	2.38
	AN v NN	9	8	5	18	1.71
	<u>NA v NN</u>	8	6	6	20	<u>1.81</u>
	Average Z					1.86
MT v AC	AA v AN	13	7	4	16	2.56
	AA v NA	17	3	2	18	4.43
	AA v NN	13	7	5	15	2.22
	AN v NA	8	9	11	12	0.00
	AN v NN	10	7	8	15	1.19
	<u>NA v NN</u>	10	4	8	18	<u>2.13</u>
	Average Z					2.09
MT v DA	AA v AN	13	7	7	13	1.58
	AA v NA	11	9	4	16	1.96
	AA v NN	13	7	2	18	3.27
	AN v NA	12	5	3	20	3.39
	AN v NN	10	7	5	18	2.06
	<u>NA v NN</u>	12	2	3	23	<u>4.28</u>
	Average Z					2.76

* All associations are positive

For key to columns see Table 5.B.2

Table 5.B.2.

Median Tests comparing the three inferences across the four types
of Major Premiss (Experiment 1)

<u>COMPARISON</u>		<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>Z</u> *
<u>INFERENCES</u>	<u>FORM OF MAJOR PREMISS</u>					
MT v AC	AA	14	6	3	17	3.20
	AN	10	7	7	16	1.47
	NA	11	3	8	18	2.56
	<u>NN</u>	8	6	10	16	<u>0.80</u>
	Average Z					2.01
MT v DA	AA	15	5	3	17	3.50
	AN	13	4	7	16	2.56
	NA	11	3	4	22	3.59
	<u>NN</u>	11	3	4	22	<u>3.59</u>
	Average Z					3.31
AC v DA	AA	15	2	3	20	4.40
	AN	13	4	7	16	2.56
	NA	10	9	5	16	1.55
	<u>NN</u>	9	9	6	16	<u>1.15</u>
	Average Z					2.41

* All associations are positive

Key to Columns

- A - Both above median
- B - 1st above median; 2nd median or below
- C - 1st median or below; 2nd above median
- D - Both median or below

APPENDIX C

TABLES FOR ANOVAS PERFORMED
IN EXPERIMENTS 1 AND 2

TABLE OF ANOVA RESULTS. EXPERIMENT 1

<u>SOURCE OF VARIATION</u>	<u>SUMS OF SQUARES</u>	<u>d of f</u>	<u>MEAN SQUARES</u>	<u>'F' RATIO</u>	<u>SIGNIFICANCE</u>
C	79.21875	1	79.21875	40.108	1%
CS	77.03125	39	1.97516		
A	18.01875	1	18.01875	18.543	1%
AS	37.89792	39	0.97174		
T	40.13750	2	20.06875	16.023	1% *
TS	97.69583	78	1.25251		
CA	5.41875	1	5.41875	3.328	n.s.
CAS	63.49792	39	1.62815		
CT	4.16250	2	2.08125	3.358	n.s. *
CTS	48.33750	78	0.61971		
AT	10.38750	2	5.19375	5.834	5% *
ATS	69.44583	78	0.89033		
CAT	4.51250	2	2.25625	2.400	n.s. *
CATS	73.32083	78	0.94001		

* Based on a conservative test

KEY

- C Polarity of consequent (of conclusion)
- A Polarity of antecedent (of conclusion)
- T Conclusion type

TABLE OF ANOVA RESULTS. EXPERIMENT 2: 'TRUE' RESPONSES

<u>SOURCE OF VARIATION</u>	<u>SUMS OF SQUARES</u>	<u>d of f</u>	<u>MEAN SQUARES</u>	<u>'F' RATIO</u>	<u>SIGNIFICANCE</u>
C	121.00208	1	121.00208	77.259	1%
CS	61.08125	39	1.56619		
A	11.10208	1	11.10208	9.416	1%
AS	45.98125	39	1.17901		
T	85.8375	2	42.91875	20.455	1% *
TS	163.66250	78	2.09824		
CA	27.55208	1	27.55208	24.130	1%
CAS	44.53125	39	1.14183		
CT	7.52917	2	3.76458	5.799	5% *
CTS	50.63750	78	0.64920		
AT	0.80417	2	0.40208	0.648	n.s. *
ATS	48.36250	78	0.62003		
CAT	4.12917	2	2.06458	3.424	n.s. *
CATS	47.03750	78	0.60304		

* Based on a conservative test

KEY

- C Polarity of consequent (of conclusion)
- A Polarity of antecedent (of conclusion)
- T Conclusion type

TABLE OF ANOVA RESULTS. EXPERIMENT 2: 'FALSE' RESPONSES

<u>SOURCE OF VARIATION</u>	<u>SUMS OF SQUARES</u>	<u>d of f</u>	<u>MEAN SQUARES</u>	<u>'F' RATIO</u>	<u>SIGNIFICANCE</u>
C	74.93906	1	74.93906	52.075	1%
CS	56.12344	39	1.43906		
A	66.95156	1	66.95156	48.255	1%
AS	54.11094	39	1.38746		
T	66.32969	3	22.10990	21.583	1% *
TS	119.85781	117	1.02443		
CA	9.75156	1	9.75156	5.084	5%
CAS	74.81094	39	1.91823		
CT	18.76719	3	6.25573	9.216	1% *
CTS	79.42031	117	0.67881		
AT	5.55469	3	1.85156	3.156	n.s. *
ATS	68.63281	117	0.58661		
CAT	0.65469	3	0.21823	0.580	n.s. †
CATS	44.03281	117	0.37635		

* Based on a conservative test

KEY

- C Polarity of consequent (of conclusion)
- A Polarity of antecedent (of conclusion)
- T Conclusion type

CHAPTER 6

TRUTH STATUS EFFECTS ON SELECTION TASK BEHAVIOUR

EXPERIMENT 3

208

The second type of nonlogical bias discussed in the review is the effect of content that relates to the subjects' experience. Subjects tend to evaluate conclusions on the basis of their perceived truth or falsity. This bias has been much investigated in the syllogistic literature (see section 3 of Chapter 1) and has been suggested earlier as underlying the data of certain thematic conditional inference studies (e.g. those of Bucci, 1978 and Roberge, 1977). However, study of the bias has been almost wholly restricted to its effect upon conclusion evaluation and little attempt has been made to investigate its effect on other aspects of reasoning behaviour.

An exception to this is the study of Van Duyne (1976), who investigated the effect of truth status on selection task behaviour. This was a welcome extension as, due to the large amount of research on this task, it could be argued that the general influence of any bias on reasoning cannot be assessed until its role in mediating selection task responses is determined. Subjects were asked to evaluate test statements that had previously been generated by the subjects themselves, subjects having been asked to generate sentences that they believed to be always true ('necessity' condition) and sentences that they believed to be only sometimes true ('contingency' condition). Van Duyne reports that subsequent selection task performance was more 'logical' (i.e. biased towards falsification) on the contingent sentences and speculates that subjects' behaviour might to some extent be determined by a desire to fulfil their expectations - specifically, that, the less the subject believes the statement to be true, the more likely he is to attempt to falsify it.

The method employed differed from the usual presentation of the task in four main ways:

- 1) Subjects generated their own test statements.
- 2) Stimuli were not presented in card form. Instead, subjects were asked whether they thought it necessary to look for additional information if they only knew that TA, FA, TC or FC was the case. For example, given the statement 'all glucose is sweet', subjects were asked whether any additional information would be required in the case of a substance known to be: a) glucose; b) sweet; c) protein (i.e. not glucose) and d) bitter (i.e. not sweet).
- 3) Sequential presentation was used, whereas usually the subject is presented with all four stimuli simultaneously.
- 4) Subjects evaluated statements referring to potentially infinite universes of instances (e.g. an infinite number of samples of glucose, only one of which may be tested).

However, despite these variations, Van Duyne's finding that subjects appear to be more likely to attempt to falsify an (only) sometimes true statement, opened new possibilities. It is of interest, then, to take a closer look at the data from which this result was derived.

Van Duyne was specifically interested in 'logically correct' responses (i.e. selection of TA and FC and rejection of FA and TC). Noting the effect of matching bias on responses, he observed that subjects "sometimes make the correct selection for the wrong reason". Accordingly, he only scored a response as correct if the subject gave the logically correct response and, in addition, "gave the correct reasons for this selection". Subjects' selections were deemed to be 'for the correct reason' if they

explained them in terms of falsification. This procedure yielded the data shown in Table 6.1. As can be seen, the finding that subjects performed significantly 'better' on contingency sentences (i.e. sentences that were more likely to be false) is primarily based on response to TA, although subjects did perform 'better' under the contingency condition in all four cases.

Table 6.1

Data from Van Duyne (1976). Analysis of the 22 subjects

	<u>Table 6.1.A</u>				<u>Table 6.1.B</u>			
	<u>Number of subjects making logically valid decision, deemed to be for 'correct reason'</u>				<u>Actual frequency of selection</u>			
	<u>Selection of TA</u>	<u>Rejection of FA</u>	<u>Rejection of TC</u>	<u>Selection of FC</u>	<u>TA</u>	<u>FA</u>	<u>TC</u>	<u>FC</u>
NECESSITY CONDITION	8	16	6	13	21	5	14	20
CONTINGENCY CONDITION	17	19	8	14	22	3	14	19

However, given the results of Wason & Evans (1975), it is unsound to classify subjects' responses on the basis of their post hoc explanations. For this reason, Van Duyne's data were reanalysed in terms of the actual frequency of selections made. The results of this analysis are shown in Table 6.1.B and it is apparent from the table that actual selection behaviour was not affected by the experimental conditions. Van Duyne's data thus do not reveal an effect of truth status on selection task responses. His procedure modified not what subjects did on the task, but the

reasons why they thought they had done it. Analysed in this way, Van Duyne's experiment provides even better evidence for the dual-process hypothesis, than did that of Wason & Evans (1975). In the latter study, one condition (using the AN rule) produced a change in selection behaviour and a corresponding change in explanations. The authors argued that the selection differences (to those observed on the AA rule) were more likely to be due to the syntactic form of the AN rule than to the reasons proffered in the explanations, and thus they suggested that two processes were involved. However, that an experimental condition can affect explanations without affecting selection frequencies (as in Van Duyne's study), is very clear cut evidence that introspection does not reveal the causes of behaviour.

However, the main question of interest was whether differences in responses would be observed as a function of the differently perceived truth status of the statements to be evaluated. By using only 'always true' and 'sometimes true' statements, it was considered that Van Duyne's experiment had not fully tested this question. There was clearly only a small difference between the two types of statement and it was felt that this was due to the use of the word 'true' to describe both conditions. As this produced fairly similar statements (as regards truth status), it is not surprising that responses were also similar. Accordingly, the experiment was replicated with a major change: subjects being asked to generate four types of sentence, those that they believed always and usually true and those they believed always and usually false. This procedure allowed for both a truth/falsity comparison and a necessity/contingency comparison.

It was decided to again ask subjects for explanations of their responses as this would keep the procedure as close to that of Van Duyne as possible. Van Duyne's result was interpreted as showing that perceived falsity leads to justification of responses in terms of falsity. It was thus expected that falsity, not contingency per se, would facilitate verbal report scores.

There was no a priori reason to believe that either contingency or falsity would facilitate or inhibit selections. It was thus decided to test all behavioural measures two-tailed.

EXPERIMENT 3

METHOD

Design

Subjects were asked to generate their own (written) sentences to be used on four subsequent selection tasks. They were asked to generate five sentences that they considered 'always true', five that they considered 'usually true', five that they considered 'usually false' and five that they considered 'always false'. They were asked to generate one set of sentences at a time, with 24 subjects receiving all possible orders of presentation.

The experimenter selected one test statement from each category and presented these test statements on the four selection tasks in the same order as they had been generated by the subject. The selection tasks were presented in such a way that the subject had to answer questions about whether he would wish to investigate TA, FA, TC and FC instances of the rule. To keep order effects constant, the instances were always put to the subject in that order. Subjects' responses were recorded and their explanations scored as 'correct' or 'incorrect', in accordance with the scoring procedure.

Subjects

Twenty four undergraduates at Plymouth Polytechnic acted as paid volunteers. They had no previous selection task experience and were tested individually.

Apparatus

A tape recorder (visible to the subject) was used during the experiment.

Procedure

When asked to generate the sentences, subjects were handed a piece of paper with a heading relevant to the sentences required (e.g. 'usually true'). They were asked to generate only sentences of the form 'If P then Q' and asked not to include negatives or to construct sentences in which Q also implies P. After having finished the first set, subjects were immediately handed another sheet of paper and asked to generate the next set required; this procedure being repeated until the subject had generated all four sets of sentences.

The experimenter then selected a suitable sentence from each category for use on the subsequent selection tasks. The criteria used were whether or not the sentence could be used on the task, which required that the sentence conform to the instructions and that it be possible that the presence or absence of either antecedent or consequent could be independently observed. The first sentence meeting these criteria was selected. Some changes were made to certain sentences to make them grammatical or more formal.

After the sentence had been selected, it was presented to the subject as a statement that the experimenter would like to evaluate. The subject was asked to imagine that the experimenter did not know whether the statement was true or false and to imagine that the experimenter could observe only one fact, concerning either the first or second part of the statement, at a time. Four questions were posed on each of the four statements, constituting examples of (the knowledge of) TA, FA, TC and FC. After asking the subject to imagine each instance, the following question was put: "Would I need to find out anything else in order to find out whether the statement was true or false?". For example, for the sentence 'If a person drinks a bottle of whisky then he will be drunk', the subject was asked whether extra information would be required if it was known only that a person: had drunk a bottle of whisky (TA); had not drunk a bottle of whisky (FA); was drunk (TC); and was sober (FC). For each instance, after the subject had answered 'yes' or 'no', the experimenter asked "why?". Subjects were asked further questions as required until their justification was fully ascertained. This was scored as 'correct' or 'incorrect' in accordance with the scoring procedure used by Van Duyne.

Scoring Procedure

An explanation of a (logically correct) decision to select or reject was scored as 'correct' if it met the following criteria:

- TA The explanation was scored as 'correct' if the subject said that he selected TA in order to see whether the consequent might or might not happen.
- TC The explanation was scored as 'correct' if the subject

said either (i) that he rejected TC because the presence or absence of TA would not prove anything or (ii) that TC did not imply TA.

FA The explanation was scored as 'correct' if the subject said that he rejected FA because it was irrelevant or because the presence or absence of TC would not prove anything.

FC The explanation was scored as 'correct' if the subject said that he selected FC because it could (by discovering TA) show the statement to be false or reveal an exception.

RESULTS

Data were analysed by comparing necessity ('always') and contingency ('usually') statements and by comparing 'true' and 'false' statements. This involved pooling data from two statements for each condition.

The percentage frequency of selection of the four instances is shown in Table 6.2. There was no significant effect of contingency on these selections and no significant interactions between contingency and truth status.

Table 6.2

Percentage Frequencies of Selections (N = 24 for each Condition)

	<u>Experiment 3</u>			
<u>Sentence Type</u>	<u>TA</u>	<u>FA</u>	<u>TC</u>	<u>FC</u>
ALWAYS TRUE	96	50	96	54
USUALLY TRUE	92	42	87	67
USUALLY FALSE	92	46	87	87
ALWAYS FALSE	92	37	71	87

There was no significant effect of truth status on selection of TA or FA. However, there was strong evidence that falsity facilitated FC selection (sign test $p < 0.001$, two-tailed) and some suggestion (although not significant) that falsity inhibited TC selection (sign test, $p < 0.08$, two-tailed).

The percentage frequencies of subjects who made the logically correct response and whose verbal justifications were scored as meeting the criteria are shown in Table 6.3. (It should be noted that the basis of this table is logically correct decisions, not selection frequencies.) As for actual responses, there was no significant effect of contingency and no significant interactions between contingency and truth status.

Table 6.3
Percentage Frequencies of Subjects (N = 24) Making the Logically
Correct Response AND Giving a Verbal Explanation that met the
Scoring Criteria (Experiment 3)

<u>Sentence Type</u>	<u>Selection</u> <u>of TA</u>	<u>Rejection</u> <u>of FA</u>	<u>Rejection</u> <u>of TC</u>	<u>Selection</u> <u>of FC</u>
ALWAYS TRUE	25	46	4	37
USUALLY TRUE	37	58	4	42
USUALLY FALSE	58	46	8	71
ALWAYS FALSE	79	58	8	83

There was no evidence of an effect of truth status on verbal justifications of rejection of either TC or FA. However, falsity was found significantly to facilitate verbal report scores for both selection of TA (sign test, $p < 0.0001$) and selection of FC (sign test, $p < 0.0001$).

DISCUSSION

The findings apropos selection and verbal report are not independent as, to explain correctly a selection decision, that selection decision must first be correct. In the case of TA, there is no difference in selection frequency between the two conditions but, for FC, the facilitation of verbal report scores in the false condition could simply be due to the greater number of correct selections of this alternative. However, expressing correct verbal report as a percentage of correct selections (only), verbal explanations of FC selection were scored as 'correct' on 88% of occasions in the 'false' conditions, as opposed to only 65% of occasions in the 'true' conditions. It is thus apparent that truth status affected both of the 'dual processes' proposed by Wason & Evans (1975). It is worth noting in passing that, if introspections revealed the causes of behaviour, percentage of correct selections correctly explained should not vary between conditions. That they did so constitutes further evidence (as did Van Duyne's results) supporting the independence of the dual processes.

The crucial finding of this experiment, however, was that truth status affected the behavioural as well as introspective process, FC being selected significantly more frequently in the 'false' condition. It should be noted that, even if Van Duyne had asked subjects to generate 'false' statements, he could not have obtained this result, due to a 'ceiling effect' (although he may have observed an effect on verbal report scores for FC). 91% of his subjects selected FC on 'always true' statements. It is, perhaps, fortuitous that the present experiment was conducted on a sample of subjects who showed less overall tendency to select FC.

The two studies did, in fact, yield somewhat different response profiles. Subjects in the present study, when using the (directly comparable) 'always true' statements, selected FC significantly less often (Fishers exact, $p < 0.02$, two-tailed), and selected TC significantly more often (Fishers exact, $p < 0.02$, two-tailed), than did subjects in Van Duyne's experiment. However, there is sufficient variety of temporal, geographic and other dissimilarities between the two studies to render speculation as to the reasons for these differences injudicious. One noticeable similarity between the two studies was that in all conditions, the median number of selections per subject was three. In almost all other selection task experiments (both abstract and thematic), the median number of selections has been two. It is possible that sequential presentation of the four alternatives leads to higher selection frequencies.

Apart from the facilitation of FC selection, the possibility that falsity inhibits TC selection is also worthy of consideration, as it suggests that falsity has a general effect on consequent selections. The question arises as to the basis of this effect. Van Duyne, who believed that he had found such an effect, proposed that subjects were essentially seeking information that verified their own opinions, referring to this as 'cognitive self-reinforcement'.

One objection to this theory is that it is very task specific and does not contribute to the understanding of behaviour on the standard abstract selection task. As has been discussed earlier, the results of Manktelow & Evans (1979) suggest that there is no tendency to either verify or falsify on abstract selection tasks. Where, then, is the subject's motivation to prove his belief?

The best that could be said is that subjects have no beliefs about the truth or falsity of the rule on abstract tasks and therefore no motivation to prove it true or false.

A second objection is based on the fact that Evans (1977b), analysing a wide variety of selection task results, failed to find any evidence of an association between TC and FC selections. Similarly, there was no evidence of such an association in either the present study or that of Van Duyne. If subjects are verifying their own beliefs, they will select TC if they believe the rule to be true and FC if they believe the rule to be false. However, it follows from the apparent independence of TC and FC, that selections 'to prove the rule true' must be independent of selections 'to prove the rule false'; and thus (as these selections are aimed at verifying the subject's beliefs), that belief that the rule is true is independent of belief that the rule is false. This seems absurd, although it cannot be ruled out completely as there is ample evidence that subjects do not always behave consistently. For instance, Winthrop (1946) observed a high degree of inconsistency when subjects were asked to evaluate separately pairs of (attitude) propositions that were either contradictory or equivalent. However, the empirical independence of FC and TC selections would make it very difficult to construct a satisfactory theory based on the cognitive self-reinforcement principle.

Further, beliefs must depend on past experience. However, such past experience should not be confused with belief - belief is the affective consequence of that experience entering awareness. When a statement is evaluated, relevant past experience is retrieved and may then produce a conscious belief. Thus, a theory explaining responses in terms of belief is inadequate as it fails to account for the prior process of retrieval.

A possible basis of the retrieval process is the association between the constituent items of the statement. That is, one substrate of a belief about a conditional statement may be the strength of association between the antecedent and consequent. If subjects do evaluate conditionals on the basis of association strength, then this may well affect their behaviour on the selection task, leading to the selection of those alternatives that are associated. This is clearly a possible explanation of the effect of truth status. On false statements there is a stronger association between FC and TA than on true statements.

For instance, when thinking of a swan, the colour white is likely to come to mind and may be said to be an associate of swan. If the subject is given the opportunity of investigating various consequent values of a statement about the colour of swans, he may thus be likely to select white. This would be a TC case of the ('true') statement 'if it is a swan, then it is white' but would be an FC case of the ('false') statement 'if it is a swan, then it is black'. The increased selection of FC in the 'false' condition of experiment 3 could thus be due to FC alternatives tending to have been more associated with TA in that condition.

On an abstract selection task, the subject has no prior information available that will modify his consequent selections in the manner suggested above. However, when engaged on the task, he does have one relevant prior experience, as he has been exposed to the conjoint mention of P and Q in the rule to be evaluated. It is thus possible that an association is set up between P and Q that, in the absence of any other source of association, becomes the dominant nonlogical determinant of selections, leading to the selection of these 'matching' values.

The association theory can thus be extended to tasks using abstract materials. Further, in experiment 3, truth status affected both selections and verbal reports, which suggests that the associations not only modify selection behaviour but also give rise to conscious beliefs that subsequently modify verbal explanations. This is entirely consistent with the results of Wason & Evans (1975). If the subject 'believes' that Q occurs with P, he will expect that his selection will show 'If P then Q' to be true, and 'If P then not Q' to be false. Also he should tend to explain P selection in terms of finding (the expected) Q and Q selection in terms of finding P. Wason & Evans did in fact, notice such a tendency, referring to it as 'secondary matching bias'.

On the basis of the foregoing analysis, it is argued that the association theory has much to recommend it. Essentially, it has three distinct advantages:

- 1) It provides an explanation of truth status effects that generalises to standard tasks using abstract materials.
- 2) Although Evans (1975) suggested that the basis of matching bias may be that the subjects' attention is directed towards those values mentioned in the rule, no full attempt has been made to explain why subjects match. The association theory provides such an explanation.
- 3) Possibly the most important advantage is that it provides a consistent explanation of both selection behaviour and introspection on the selection task.

CHAPTER 7

THE EFFECTS OF EXPERIMENTALLY MANIPULATED TRUTH STATUS ON SELECTION

TASK BEHAVIOUR

EXPERIMENT 4	220
EXPERIMENT 5	226
APPENDIX	242

Experiment 3 used an unusual form of the selection task. Subjects generated their own (thematic) test statements, evaluated statements referring to potentially infinite universes of instances and selected or rejected sequentially presented alternatives. Thus, it is not necessarily the case that results obtained are generalisable to other selection tasks and the first problem, in the further investigation of the effect of truth status, must be to determine that a general effect has been discovered. Clearly, if responses are to some extent determined by the previously learned truth status of the conditional rule, it should be possible to produce response differences on abstract tasks by providing subjects with the relevant experimentally contrived experience.

In order to test this possibility, a pack of cards was developed that had letters (A or B) on one side and numbers (1 or 2) on the other, so arranged that most As had 1 on the back and most Bs had 2 on the back. It was decided to expose subjects to the contingencies of this pack via a probability learning task, conducted prior to selection task presentation. It was presumed that the effect of the probability learning task would be to induce more association between TA and FC, and less association between TA and TC, on the ('usually false') relationships, 'If A then 2' and 'If B then 1', than on the ('usually true') relationships, 'If A then 1' and 'If B then 2'.

It was expected that this procedure would produce results comparable to those obtained in Experiment 3. Specifically, it was predicted that subjects would make more FC, and less TC selections when reasoning with 'usually false' statements.

EXPERIMENT 4

METHOD

Design

Subjects were presented with a probability learning task, prior to a selection task, to familiarise them with the contingencies of the card pack (four of which were used in the selection task). In one selection task condition subjects evaluated a rule having a high probability of being true and, in the other condition, subjects evaluated a rule having a low probability of being true. The responses of those subjects who had shown significant evidence of (probability) learning were compared between the two conditions.

Materials

- 1) A pack of 20 record cards consisting of:
 - 8 cards with A on one side and 1 on the other
($p(1/A) = p(A/1) = 0.8$)
 - 2 cards with B on one side and 1 on the other
($p(1/B) = p(B/1) = 0.2$)
 - 2 cards with A on one side and 2 on the other
($p(2/A) = p(A/2) = 0.2$)
 - 8 cards with B on one side and 2 on the other
($p(2/B) = p(B/2) = 0.8$)
- 2) Four sheets of paper with a conditional rule typed on them as follows:

If there is an A on one side of a card, then there is a 1 on the other side.

If there is a B on one side of a card, then there is a 2 on the other side

If there is a B on one side of a card, then there is a 1 on the other side

If there is an A on one side of a card, then there is a 2 on the other side.

Given that four appropriate cards are drawn at random for the selection task, the first two rules have a probability of approximately 64% of being true¹ and were assigned alternately to subjects in the first condition, which was designated 'usually true'. The last two rules have a probability of only approximately 4% of being true² and were assigned alternately to subjects in the second condition, which was designated 'usually false'.³

Subjects

Thirty two undergraduates at Plymouth Polytechnic acted as paid volunteers. They were tested individually and assigned alternately to the two conditions.

Procedure

The card pack was shuffled and placed on the table, letter sides up. The following instructions were then given:

"These cards all have a letter on one side and a number on the other side. The letter is either an A or a B and the number is either a 1 or a 2. I am going to go through the pack one card at a time showing you the letter side of the card. For each card, I want you to guess whether the number on the other side is a 1 or a 2. I will turn each card over after your guess so that you can see whether or not you were right."

The subject then commenced to guess and, after each guess, the experimenter turned over the card, called out the number and called either "correct" or "wrong". At the same time, the experimenter recorded the subject's guess and marked it with a tick or a cross so that the subject could see that the experimenter was 'scoring' his answers. After the first run through, the pack was reshuffled and the subject asked to repeat the procedure. Four runs altogether were carried out and, at the end of each run, the subject's ticks were added up and his 'score' (out of 20) announced. The subject was then introduced to the selection task, the following instructions being given:

"I am now going to shuffle the cards again and ask you to select one 'A' card and one 'B' card without looking at the other side". (This was done and the two cards laid on the table, the pack was then turned over) "Now I am going to shuffle again and ask you to select one '1' card and one '2' card without looking at the other side." (these two cards were then placed on the table, producing the usual four card layout) "Now, we do not know what is on the other side of these four cards.....do you agree with that?" (the experimenter waited for the subject to indicate agreement) "but I want you to suppose that somebody makes the following statement about them." (The appropriate rule was then read twice to the subject who was then handed it on a typed sheet of paper and allowed to read it.) "Now, as we do not know what is on the other side of those cards, we do not know whether that statement is true or false. What I would like you to do is to decide which of the cards it would be necessary to turn over in order to find out whether that statement was true or false." (This instruction was repeated from the words 'which of the cards').

"You may choose any number of cards but take your time and don't give me an answer until you have made a decision about every card. That is, I want you to decide, for each card, whether or not it would be necessary to turn it over and then give me your answer about all the cards. Please don't actually turn any cards over." The experimenter then recorded the subject's selections.

The last run of the probability learning task was taken for each subject and a score (out of 20) assigned to it based not on correct answers but on 'optimal' answers (i.e. the most probable answer - a response of '1' to A or '2' to B). It was not expected that subjects would make wholly optimal responses as various studies (e.g. Grant et al, 1951) have reported that subjects match predictions to probabilities. A criterion was thus set at significantly above chance 'optimal' scoring. For 20 trials, this entailed that subjects had to achieve an optimal score of 15 to be considered as having properly learned the contingencies.

Results

Eight subjects failed to reach criterion in the probability learning task, and were excluded from the analysis. Of the remaining 24 subjects, 14 were in the 'usually true' condition and 10 were in the 'usually false' condition. The percentage frequency of selection of the four cards, in the two conditions, is shown in Table 7.1.

The prediction that more FC cards would be selected in the 'false' condition was confirmed, but was only near to significance ($p < 0.08$, Fishers exact, one-tailed). The prediction that less TC cards would be selected in the 'false' condition was not confirmed, the results being in the opposite direction. However, a post hoc test showed the latter finding to be not

significant ($p < 0.11$, Fishers exact, two-tailed). It is evident from Table 7.1 that no significant differences were observed on TA and FA selections (although this could possibly have been due to 'floor' and 'ceiling' effects).

Table 7.1

Percentage Frequency of Selections of the Four Cards, Experiment 4

<u>CARD</u>	<u>TA</u>	<u>FA</u>	<u>TC</u>	<u>FC</u>
<u>CONDITION</u>				
'USUALLY TRUE' (N = 14)	86	7	50	14
'USUALLY FALSE' (N = 10)	100	0	90	50

There was a noticeable tendency for subjects to select more cards generally in the 'false' condition and this tendency was found to be significant ($p < 0.02$, Kendall's S, two-tailed). An analysis of the number of cards selected in each condition is shown in Table 7.2.

Table 7.2

Frequencies of Cards Selected per Subject in the Two Conditions of Experiment 4

	<u>TRUE CONDITION</u>	<u>FALSE CONDITION</u>
3 CARDS	0	5
2 CARDS	9	4
1 CARD	4	1
NO CARDS	1	0
	14	10

DISCUSSION

The results for TC selection suggest that it was probably premature to interpret the near significant difference observed in Experiment 3 as a genuine effect. Experiment 4 did, however come near to replicating the one significant finding of Experiment 3, as subjects selected more FC in the 'false' condition ($p < 0.08$).

The basic question posed in this experiment was whether experimentally produced experience relevant to truth status can affect selection task responses. 'Falsity' did significantly affect responses, in that subjects tended to select more cards generally ($p < 0.02$). Unfortunately, this result makes interpretation of the FC result difficult. It could be regarded as further evidence of the specific effect of truth status on FC selections observed in Experiment 3, or alternatively, merely as an aspect of the general tendency. The latter interpretation would suggest that, although truth status effects may be produced as a function of either 'real life' or experimentally produced experience, these effects may well be qualitatively different.

Clearly, although Experiment 4 did demonstrate that experimentally manipulated truth status can affect responses, the nature of this effect was far from precisely determined. Accordingly, a further experiment was designed, with two notable alterations to improve the power of the design: a within subject design was adopted and a change made to the probability learning task. A set number of probability learning trials was presented to each subject in Experiment 4, so that each subject had an equal amount of prior exposure to the task materials. However, the loss of eight subjects on the probability learning task indicated that

it would be preferable to take all subjects up to a fixed criterion before presenting them with the selection task.

Various other changes were also made. Four conditions were used (an 'always false' and an 'always true' condition being added) and the whole task was presented via an on-line computer terminal. Predictions were that subjects would select specifically more FC in the 'false' conditions and, on the basis of the results of Experiment 4, that they would select more cards generally in the 'false' conditions.

EXPERIMENT 5

Method

Design

Subjects were first given a probability learning task on which they learned that:

- a) 'A' was always paired with '1'
- b) 'B' was usually paired with '2', but was occasionally paired with '1' or '3'
- c) 'C' was usually paired with '3', but was occasionally paired with '1' or '2'

The task was presented on a computer terminal and controlled by an on-line computer, as a somewhat complex 'running' criterion was employed.

After reaching criterion on the first probability learning task, the subject was presented with the first of four selection tasks. These four tasks were so formulated that, on the basis of their (probability learning) experience, subjects should perceive them as 'always true' ('If A then 1'), 'usually true' (If B then 2'), 'usually false' ('If B then 1') and 'always false' ('If A then 2'). Subjects were asked to select or reject each

of the four possible instances (of occurrence or non-occurrence of the components) sequentially. '1' and '2' were invariably used to represent TC and FC (or vice versa). TA was represented by either A or B (depending on the rule) and FA was always represented by C (thus giving a comparable FA instance for all four rules). The order in which the instances were presented was randomly determined.

After the selection task, subjects were returned to a subsequent probability learning task (identical to the first, but with a lowered criterion) three times. Each subsequent probability learning task was followed by another selection task such that each subject used each of the four types of rule. All 24 possible orders of presenting the four selection tasks were used, responses being recorded by the program.

A measure of each subject's probability learning performance (trials to criterion on the first task) was also taken.

Materials

There were no specific task materials as subjects received all instructions via the computer terminal. However, for the probability learning task, a set of letter/number pairs was stored in the computer and it is more convenient to describe this in a separate section. Fifteen pairs were stored as follows:

3 cases of A and 1 (p (1/A) = 1.0; p (A/1) = 0.6;

$$p (2 \text{ or } 3/A) = p (A/2 \text{ or } 3) = 0.0)$$

4 cases of B and 2 (p (2/B) = 0.67; p (B/2) = 0.8)

1 case of B and 1 (p (1/B) = 0.17; p (B/1) = 0.2)

1 case of B and 3 (p (3/B) = 0.17; p (B/3) = 0.2)

4 cases of C and 3 (p (3/C) = 0.67; p (C/3) = 0.8)

1 case of C and 1 (p (1/C) = 0.17; p (C/1) = 0.2)

1 case of C and 2 (p (2/C) = 0.17; p (C/2) = 0.2)

From the above conditional probabilities, the approximate probabilities of the four rules used in the selection tasks can be calculated. These are: 100% for 'If A then 1' ('always true'); 53% * for 'If B then 2' ('usually true'); 3% for 'If B then 1' ('usually false') and 0% for 'If A then 2' ('always false').

Subjects

Twenty four undergraduates, recent graduates and staff at Plymouth Polytechnic acted as paid volunteers. They were randomly assigned to the 24 possible orders of presenting the four selection tasks.

Procedure

Subjects were seated at the computer terminal and instructed in the operation of its controls. The only instruction about the task given verbally was: "If you try and do as sensibly as possible then you will finish the task quicker." This instruction was aimed at preventing subjects from 'playing' with the probability learning task (and thus never meeting the criterion). It was not presented as a requirement of the experiment but more as an 'inside tip' in the subjects' own interest (the 'tip' did, in fact, benefit subjects as they were paid a set amount to complete the task).

The subject then received instructions for the probability learning task via the terminal. He was told that the computer held a store of letter/number pairs (although he was not told how many), involving only As, Bs, Cs, 1s, 2s and 3s, and that

* See notes 1, 2 and 3

the computer would randomly select a pair, show him the letter and ask him to guess the number. The instructions then continued as follows:

"After you have made your guess, the computer will tell you what the number actually is so that you can see whether you were right or wrong.

Please note that, although the computer will present you with random examples of the letter/number pairs stored the association between letters and numbers within the pairs is not random.

Thus, for each letter, the three possible numbers are not equiprobable and one number is more likely to be paired with it.

If you try and work out what the relationships between the letters and numbers are, you will attain more correct answers and complete the task sooner."

The above 'clues' to the nature of the task were included as it was felt that subjects would find a computer an unfamiliar (and thus potentially confusing) medium of communication.

Although subjects were told that the selection of the pairs in store was randomly determined, this was not wholly true. The stored pairs were presented in an equivalent manner to a manual presentation. The store was presented sequentially and, at the end of a sequence, the store was randomly rearranged and the fifteen pairs presented again. Subjects were told that pairs were selected randomly so that the subject could never be (logically) certain that all As were paired with 1. If the subject could be certain that 'If A then 1' had a probability of 1.0, then this would affect the logic of the subsequent selection tasks.

After reaching criterion on the probability learning task, the computer presented the subject with the following instructions prior to the selection task:

"Suppose that the computer will set a logical reasoning task to another subject.

It has chosen at random a few of the stored pairs that you have been working with and placed them in a separate store. It will make a particular statement, about pairs in this new store, which will have the following general form:

If the letter is (A or B or C) then the number is
(1 or 2 or 3)

The subject (who has not seen any of the pairs) will be asked to attempt to find out whether the statement is true or false.

To help him do this the computer will offer him the chance of viewing a particular letter/number pair held in the new store. It will do this in one of two ways:

Either by the offer: The letter part of this pair
is (A or B or C), would you
like to know the number?

Or by the offer: The number part of this pair
is (1 or 2 or 3), would you
like to know the letter?

The problem is that the information may or may not be

helpful in finding out whether the statement is true or false and, because this is a logical reasoning task, the subject must be careful not to ask for information that does not, logically, help find out whether the statement is true or false.

Your task will be to advise the subject. You will be shown various possible offers that the computer may make and will be asked whether it would be logical to accept or reject them. Please only advise him to take up the offer if it is logically helpful from his point of view."

The subject was then presented with a conditional rule and with four 'offers' representing TA, FA, TC and FC instances (in a random order). For each offer he was asked: 'Would it be logical for the subject to accept this offer?'

It will be noted that the subject is asked to advise another person whether it is logical to accept or reject an offer. This was a further control for the effects of 'always' rules. If the subject forms an (inductive) conclusion that he 'knows' that all As are paired with ones, he may personally feel it unnecessary to accept any of the offers on rules concerning As. However, given that he is not allowed to tell the other person what he 'knows', he must help the other person find out for himself. The logically correct choice for the 'other person' is thus the same for all four rules.

After completion of the selection task, the subject was transferred back to a probability learning task prior to each subsequent selection task. Subsequent selection tasks were preceded by the following instructions:

"The computer has again fed a few of the pairs at random into a separate store and is once more going to make a

statement about this new store.

Suppose that another subject (who has never seen any of the letter/number pairs) is given the logical reasoning task.

As before, your task is to advise him whether, from his point of view, it would be logical to accept or reject the computer's offer."

Data for each subject were recorded by the program which presented a summary table once the sequence of tasks was completed.

Criteria for Probability Learning Tasks

To reach criterion on the first probability learning task, the subject was required to have:

- 1) Responded '1' to the last six As
- AND 2) Responded '2' to at least six of the last ten Bs and to have responded '1' or '3' to no more than three of the last ten Bs
- AND 3) Responded '3' to at least six of the last ten Cs and to have responded '1' or '2' to no more than three of the last ten Cs.

To reach criterion on subsequent probability learning tasks, the subject was required to have:

- 1) Responded '1' to the last four As
- AND 2) Responded '2' to at least four of the last seven Bs and to have responded '1' or '3' to no more than two of the last seven Bs
- AND 3) Responded '3' to at least four of the last seven Cs and to have responded '1' or '2' to no more than two of the last seven Cs.

RESULTS

The percentage frequencies (N = 24) of selection of the four alternatives, in the four conditions, are shown in Table 7.3.

Table 7.3

Percentage Selection Frequencies (N = 24) for the Four Conditions of Experiment 5.

	<u>TA</u>	<u>FA</u>	<u>TC</u>	<u>FC</u>
'IF A THEN 1' (ALWAYS TRUE)	37	62	54	62
'IF B THEN 2' (USUALLY TRUE)	46	37	42	58
'IF B THEN 1' (USUALLY FALSE)	46	50	62	71
'IF A THEN 2' (ALWAYS FALSE)	58	71	46	75

Analyses were performed on data 'pooled' between 'always' and 'usually' true conditions, and 'always' and 'usually' false conditions. The prediction that falsity would facilitate general selection was confirmed. Subjects accepted more 'offers' under 'false', than under 'true' conditions (sign test, $p < 0.04$, one-tailed). However, there were no significant differences in selection of any of the four alternatives individually and thus the prediction that subjects would select specifically more FC in the false conditions was not confirmed (although results were in the predicted direction).

There was a significant tendency to select more FA on

'always' statements (sign test, $p < 0.04$, two-tailed). There was no direct evidence of an interaction between this effect and that of truth status. However, the effect of falsity in increasing selection frequencies was significant on 'usually' statements alone (sign test, $p < 0.02$, two-tailed) but no way near significant on 'always' statements alone (sign test, $p > 0.5$, two-tailed).

The number of trials to criterion taken by each subject on the first probability learning task was taken as a measure of learning performance and associations were observed between this measure and selection task performance. Subjects taking a below median number of trials to criterion were categorised as 'fast learners' ($N = 10$) and subjects taking an above median number of trials to criterion were categorised as 'slow learners' ($N = 11$). Table 7.4 shows the data for each of these two groups under (pooled) 'true' and 'false' conditions.

Table 7.4

Percentage Frequencies of Selections of 'Fast' and 'Slow' Learners Under (Pooled) True and False Conditions in Experiment 5.

	<u>FAST LEARNERS</u>		<u>SLOW LEARNERS</u>	
	<u>(N = 10)</u>		<u>(N = 11)</u>	
	<u>'TRUE'</u> <u>CONDITIONS</u>	<u>'FALSE'</u> <u>CONDITIONS</u>	<u>'TRUE'</u> <u>CONDITIONS</u>	<u>'FALSE'</u> <u>CONDITIONS</u>
TA	65	80	23	32
FA	40	55	59	64
TC	60	55	41	55
FC	40	65	68	77

There was significant indication that fast learning was related to greater TA selection and lower FC selection and these differences are visually apparent from Table 7.4. Subjects who made an above median number of TA selections across the four tasks had taken significantly less trials to criterion than subjects who made a below median number of TA selections (Mann-Whitney, $p < 0.01$, two-tailed) and subjects who made a below median number of FC selections had taken significantly less trials to criterion than subjects who made an above median number of FC selections (Mann-Whitney, $p < 0.05$, two-tailed).⁴

Finally, it was noted that the high frequency of FA, and low frequency of TA selections provides an opportunity to test for correlations⁵ between various selections. Evans (1977b) tested for such correlations and, on the basis of the observed independence of TC and FC selections, concluded that correlations probably do not exist. However, as Evans points out, the high frequency of TA, and low frequency of FA, selections typically observed would not be expected to yield significant correlations. Thus correlations involving TA and FA selections are usually difficult to test but may well be present. Table 7.5 shows the analysis of all six possible pairs of selections and shows the direction of correlation and two-tailed significance level (if 5% or below).

As all 24 subjects performed a selection task under each of the four conditions, the data shown in Table 7.5 do not represent four independent tests of each correlation. Given this, and given the lack of significance in most cases, the data could be interpreted as consistent with an absence of any correlation between selections.

Table 7.5

Data Showing Relationships of Selections within Each of the six Possible Pairs of Alternatives under each Condition (Data from Experiment 5)

<u>COMPARISON</u>	<u>CONDITION</u>	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>DIRECTION OF CORRELATION</u>	<u>p</u>
TA v FA (P v \bar{P})	AT	3	6	12	3	-	n.s.
	UT	3	8	6	7	-	n.s.
	UF	2	9	10	3	-	< 0.02
	AF	10	4	7	3	+	n.s.
TA v TC (P v Q)	AT	8	1	5	10	+	< 0.03
	UT	8	3	2	11	+	< 0.02
	UF	9	2	6	7	+	n.s.
	AF	9	5	2	8	+	n.s.
TA v FC (P v \bar{Q})	AT	4	5	11	4	-	n.s.
	UT	6	5	8	5	-	n.s.
	UF	7	4	10	3	-	n.s.
	AF	9	5	9	1	-	n.s.
FA v TC (\bar{P} v Q)	AT	6	9	7	2	-	n.s.
	UT	5	4	5	10	+	n.s.
	UF	6	6	9	3	-	n.s.
	AF	6	11	5	2	-	n.s.
FA v FC (\bar{P} v \bar{Q})	AT	10	5	5	4	+	n.s.
	UT	8	1	6	9	+	n.s.
	UF	10	2	7	5	+	n.s.
	AF	14	3	4	3	+	n.s.
TC v FC (Q v \bar{Q})	AT	9	4	6	5	+	n.s.
	UT	4	6	10	4	-	n.s.
	UF	9	6	8	1	-	n.s.
	AF	9	2	9	4	+	n.s.

KEY

A - Both selected B - First only selected C - Second only selected
D - Neither selected

Alternatively, however, certain trends could be distinguished as worth further investigation. Firstly, although the data are consistent with the Evans (1977b) finding of lack of correlation between TC and FC selections, there is a suggestion of a negative correlation between TA and FA selections and, under one condition, this correlation reaches significance ($p < 0.02$, two-tailed). Secondly, the data are clearly consistent with a tendency to select TC, and not FC, when TA is selected, and to select FC, and not TC, when FA is selected. It is possible that card selections are positively correlated when they have the same matching status, and negatively correlated when they have opposite matching status. However, given the number of tests, any significant correlations in Table 7.5 would require replication before further discussion.

DISCUSSION

The profile of selections in Experiment 5 was somewhat unusual in comparison with results typically observed. In particular, FA selection was very high and it may well be that computer presentation of the task had more effect on responses than had been expected.

The associations observed between probability learning speed and selection task performance were unexpected and are difficult to interpret. It is possible that individual cognitive differences within the sample mediated both tasks. However, it should be noted that faster learners would have had less prior exposure to the letters and numbers used in the selection tasks and thus this factor could have been responsible for the differences in selection task responses.

The effect of truth status in this experiment was to produce a greater overall selection tendency in the 'false' condition. Although this finding replicates the results of Experiment 4 and again demonstrates an effect of truth status on an abstract selection task, it lends little apparent support to the association theory developed in the discussion of Experiment 3. The association theory would not predict that all selection frequencies would rise in the 'false' condition and, in fact, should predict that TC frequencies would fall, as TC is less associated with TA in the 'false' condition.

However, there is a possible explanation of this apparent inconsistency. It may well be that falsity did not reduce TC selection frequencies because such selections matched. Further, TA will be less associated, and FA comparatively (to TA) more associated, with TC in the 'false' condition, which may also have an effect on selection frequencies. In this case, falsity would be expected to increase both FC and FA selection and to decrease both TC and TA selection. However, if both these decreases were inhibited by matching bias, then the overall effect would be an increase in total selection frequencies. This interpretation is neither supported nor opposed by the data of Experiment 5, as evidence of a significantly higher overall selection frequency cannot be taken as evidence either for, or against, a higher selection frequency in any individual case. To test this interpretation, it would be necessary to control for matching bias by designing an experiment that utilises all four possible conditional rules, in which the presence of negative components is systematically manipulated. Falsity would then be expected, across the four rules, to increase FC and FA selections but to decrease TA and TC selections.

However, the data also lend themselves to an entirely different interpretation. In the probability learning phase of both Experiment 4 and Experiment 5, the contingency that subjects learned was, as far as likelihood was concerned, one of equivalence. That is, not only was (for instance) 2 the most likely number, given B, but also B was the most likely letter, given 2. The subjects may thus have been led to interpret rules presented on the selection task as implying equivalence. When the rule implies equivalence, all cards are potential falsifiers and the logically correct solution is to select all four cards. It is thus possible that falsity leads subjects to behave more logically. This explanation is also consistent with the data of Experiment 3 as, in that case, subjects were presented with (thematic) rules that depended on their past experience of material implication. Thus falsity could have caused subjects to behave more logically in all three truth status experiments:- selecting more FC (but not more TA, due to a 'ceiling effect') on the implication rules of Experiment 3 and selecting more cards generally on the equivalence rules of Experiments 4 and 5.

This explanation is testable by teaching subjects a contingency that implies implication rather than equivalence. If the effect of falsity on selection tasks is to lead subjects to behave more logically, then, if subjects learn an 'implication' contingency, they should select more TA and FC, but less FA and TC, in the 'false' condition. In order to observe effects on TA and FA, it would be necessary to utilise all four conditional rules as, when the 'If P then Q' form only is used, 'ceiling' and 'floor' effects are usually observed on TA and FA respectively.

An experiment that discriminates between the above alternative interpretations of the results of Experiments 3, 4 and 5 is reported in chapter 8.

FOOTNOTES

1. For the first rule, the probability of a '1' being on the back of the 'A' is 0.8 and the probability of an A not being on the back of the 2 is also 0.8. The probability of the rule being true for the four cards is thus $0.8 \times 0.8 = 0.64$. This analysis also applies to the 'If B then 2' relationship. However, certain complications arise from the presence of other cards on the table. (See further note in Appendix)
2. For the second two rules, a similar analysis yields a probability of the rule being true of $0.2 \times 0.2 = 0.04$.
3. Of course, it cannot be assumed that the subjective probabilities are the same as the objective probabilities quoted, or that the subjective probabilities perceived by one subject are the same as those perceived by another. In fact, subjects may well not perceive probabilities in mathematical terms, a possibility suggested by Phanér (1977).
4. It should be noted that, for clarity, Table 7.4 presents selection frequencies for subjects classified as above and below median on trials to criterion, whereas analyses were performed on numbers of trials to criterion between subjects classified as making an above or below median number of selections.
5. The term correlation is used here, although independence between card selections is tested by chi-squared. Such tests of association are essentially tests of correlation, and the latter term is used henceforth to distinguish statistical association (correlation) from psychological association, referred to elsewhere in the text.

APPENDIX TO CHAPTER 7

A FURTHER NOTE ON THE OBJECTIVE PROBABILITY OF 'TRUE'
AND 'FALSE' STATEMENTS USED IN EXPERIMENTS

4 and 5

In Experiments 4 and 5, various probabilities have been quoted, in the method sections, of the likelihood of the rule 'holding' for the four cards used in the selection task. The probabilities quoted have not been strictly accurate due to the distorting effects of other cards having been drawn from the pack. To avoid undue complication, this point will be illustrated only for the case of a 'true' rule in Experiment 4:

In Experiment 4, the pack was composed of 20 cards as follows:

8 cards having A on one side and 1 on the other
8 cards having B on one side and 2 on the other
2 cards having A on one side and 2 on the other
2 cards having B on one side and 1 on the other

For the 'true' rule 'if A then 1', the following four faces are displayed on the table:



The probability of this rule holding for these cards was given as 0.64. This is obtained from the multiplication of the probability of a 1 being on the back of an A (0.8) and the probability of a B being on the back of a 2 (0.8).

This calculation, however, ignores the presence of the other cards. For instance, if ten B cards were displayed then the probability of the rule holding would be zero (as the 2 card must

have an A on the back). Similarly, the rule would definitely be false if ten 1 cards were displayed. Even the presence of one B and one 1 card alters the probabilities (for instance, if a 2 is on the back of the B, then the probability of a B being on the back of the 2 is $7/9$, not $8/10$).

The probability of the rule holding is possibly a function of the order in which the cards are considered to have been drawn from the pack. The probability of the rule holding is 0.64 only if the A and 2 cards are drawn first. However, if the four cards are considered to have been drawn in the order (from left to right) depicted, then:

- 1) The probability of the A having a 1 on the back is 0.8
- 2) The probability of the 2 having a B on the back is:
 - a) 0.8, if the B card has a one on the back ($p = 0.2$)
 - b) $7/9 = 0.77$, if the B card has a 2 on the back ($p = 0.8$)And is thus: $(0.2)(0.8) + (0.8)(0.77) = 0.782$, approx.

The probability of the rule holding is thus (approximately)

$$(0.8)(0.782) = \underline{0.626}$$

However, the probability will not be the same if different orders are considered. Thus an exact probability of the rule holding cannot be given, as it depends upon the order in which the cards are drawn from the pack.

CHAPTER 8

THE EFFECT OF EXPERIMENTALLY MANIPULATED TRUTH STATUS ON SELECTION
TASK RESPONSES ACROSS ALL FOUR FORMS OF A CONDITIONAL RULE

EXPERIMENT 6	248
APPENDIX	268

In accordance with the points raised in the discussion of Experiment 5, a further experiment was designed that utilised all four forms of conditional rule, formed by systematic negation of the components, and ensured that subjects were exposed to an implication contingency during the probability learning phases.

Packs of cards were constructed that either did, or did not, have a given symbol on one side, and did, or did not, have another symbol on the other, and these packs were used to teach subjects the four types of contingent relationship. Subjects who had learned, for instance, the 'If P then Q' contingency, were presented with a selection task using a rule of this form in the 'true' condition, or a selection task using a rule of the form 'If P then Q' in the 'false' condition. Thus, the consequent of the contingency learned was negated in the 'false' condition. The composition of packs teaching an 'If P then Q' relation is listed below:

- 7 cards having P on one side and Q on the other
- 1 card having P on one side and \bar{Q} (i.e. blank) on the other
- 7 cards having \bar{P} (i.e. blank) on one side and Q on the other
- 7 cards having \bar{P} on one side and \bar{Q} on the other

The side of the card to which the antecedent referred was red and the other side was blue, to distinguish between blank cards. For the above contingency, subjects evaluated the rule 'If P then Q' in the 'true' condition or evaluated the rule 'If P then not Q' in the 'false' condition. That the contingency is one of implication should be clear from the relative numbers of cards. Although Q is on the back of most Ps (7 out of 8), it is on the back of only half the blank (\bar{P}) cards (7 out of 14) and P is on the back

of only half the Q cards (again, 7 out of 14). Put simply, whereas in Experiment 4, $p(P/Q) = p(Q/P) = 0.8$, in this experiment, $p(Q/P) = 0.875$ but $p(P/Q) = p(Q/\bar{P}) = 0.5$.

The central aim of this experiment was to distinguish between various possible theoretical interpretations of the effects of truth status observed in Experiments 3, 4 and 5. As was discussed in chapter 7, the association theory would predict an increase in FC and FA selections and a decrease in TC and TA selections, whereas the theory that falsity leads subjects to behave more logically would predict an increase in TA and FC selections and a decrease in FA and TC selections. A third possibility is that the effect of truth status on abstract tasks is qualitatively different from its effect on thematic tasks and that it does, in accordance with the results of Experiments 4 and 5, simply facilitate selection of all cards in the 'false' condition. This last hypothesis leads to a third set of predictions. The various predictions are compared in Table 8.1. It should be noted that all predictions about FC are the same and this prediction may thus be tested one-tailed, whereas differences in selections of other cards must be tested two-tailed.

The use of all four conditional rules in this experiment also allowed for a more detailed analysis of selection differences between 'fast' and 'slow' learners, as it was possible to determine whether such differences are observed on the basis of logical, or matching status.

Selection task rules were designed to be 'usually true' or 'usually false'. 'Always' conditions were not used because, as has been mentioned earlier, they disrupt the logic of the selection task (i.e. subjects know the rule to be true or false and need turn over no cards). This disruption is not easy to overcome when manual presentation is used.

Table 8.1

Predictions of Three Possible Explanations of the Effects of
Truth Status as to Increase or Decrease of Selections in the
'False' Condition.

	TA	FA	TC	FC
ASSOCIATION THEORY	LESS	MORE	LESS	MORE
MORE LOGICAL IN FALSE CONDITION	MORE	LESS	LESS	MORE
GENERAL FACILITATION OF ALL CARDS	MORE	MORE	MORE	MORE

EXPERIMENT 6

METHOD

Design

Subjects were presented with four probability learning tasks, to teach them the contingency relationships of four separate packs of cards. These represented all four possible contingencies that can be formed by making the presence or absence of a given symbol on the (blue) back of the cards contingent upon the presence or absence of another given symbol on the (red) front of the cards. Different symbols were used in each pack and the absence of a symbol was represented by a blank.

After the subject had reached criterion on each probability learning task, four cards of that pack were used for a selection task. Subjects in the 'true' condition were presented with a rule to evaluate that, on the basis of their probability learning experience, was likely to hold for the four cards (the rule

was a statement of the contingency that the subject had learned); whereas subjects in the 'false' condition were presented with a rule that was unlikely to hold (the rule was a statement of the contingency that the subject had learned, with the consequent negated). Thus, for instance, if the contingency learned was that cards with no symbol on the front usually had no symbol on the back, a subject in the 'true' condition would evaluate a rule of the form 'If not P then not Q' and a subject in the 'false' condition would evaluate a rule of the form 'If not P then Q'.

Subjects in both conditions performed one selection task on each of the four forms of the conditional rule. All 24 possible orders of presenting the probability learning/selection task pairs were used in each condition.

Materials

Packs of 22 6" x 4" cards were used, coloured red on the front and blue on the back. Four different types of symbol combination were used:

- 1) Packs having Triangles on the front and Stars on the back
- 2) Packs having Rectangles on the front and Ticks on the back
- 3) Packs having Diamonds on the front and Crosses on the back
- 4) Packs having Squares on the front and Circles on the back

Each subject was presented with one probability learning task followed by one selection task with each of the above sets of cards, each representing a different contingency. Each card set was used to present each of the four types of contingency relation to 12 of the 48 subjects. There were thus 16 packs of cards (four symbol types x four types of contingency). The composition of the packs to represent the four contingencies is shown overleaf for

the Squares and Circles set.

If Square then Circle

7 cards having a Square on the front and a Circle on the back

$$(p(C/S) = 0.875; p(S/C) = 0.5)$$

1 card having a Square on the front and a Blank on the back

$$(p(B1/S) = 0.125; p(S/B1) = 0.125)$$

7 cards having a Blank on the front and a Circle on the back

$$(p(C/B1) = 0.5; p(B1/C) = 0.5)$$

7 cards having a Blank on both sides

$$(p(\text{blue blank/red blank}) = 0.5; p(\text{red blank/blue blank}) = 0.875)$$

22 cards

If Square then not Circle

7 cards with Square and Blank

1 card with Square and Circle

7 cards with Blank and Circle

7 cards with both sides Blank

If not Square then Circle

7 cards with Square and Circle

7 cards with Square and Blank

7 cards with Blank and Circle

1 card with both sides Blank

If not Square then not Circle

7 cards with Square and Circle

7 cards with Square and Blank

7 cards with both sides Blank

1 card with Blank and Circle

Sixteen conditional rules were typed on separate sheets of paper (four forms of rule x four sets of symbols). These all had the same format, for instance:

"If there is a Square on the red side, then there is not a Star on the blue side."

The rules had a probability of being true, for the four cards used, (see notes 1, 2 and 3 of chapter 7) of approximately 77% in the 'true' condition and of approximately 6% in the 'false' condition.

Subjects

48 undergraduates at Plymouth Polytechnic acted as paid volunteers and were tested individually. They were assigned alternately to each condition and, within each condition, were randomly assigned to each of the 24 possible presentation orders.

Procedure

The procedure was essentially the same as in Experiment 3, with substitution of the relevant symbols for references to letters and numbers. The main difference was that the procedure was repeated four times, using a different pack (and contingency) on each occasion.

The other difference was in the criterion set for the probability learning task. Subjects were not given a set number of trials, but continued with the probability learning task until they reached criterion. Eight of the 22 cards were salient to the contingency relationship and thus, on each run through the pack, eight of the subject's guesses could be scored as to whether they conformed to the most probable outcome. The criterion set was that the subject should have scored either 13 (out of 16) on the last two trials or 17 (out of 24) on the last three trials. These are minimum scores that would be significant on a one-tailed test (although given that the subject may have many more than two or three trials and that 48 subjects are involved, attainment of the criterion does not necessarily represent significant evidence of learning in each case). It was decided that, if the subject had failed to reach criterion after 20 trials, he would be dropped from the sample and replaced. The same criterion was applied to all four probability learning tasks.

RESULTS AND DISCUSSION

Four subjects failed to reach criterion on the probability learning tasks and were replaced. Analysis of probability learning task results yielded no significant evidence that any contingencies were more difficult to learn than others.

The percentage frequency of selection ($N = 24$) of the four cards, for each of the four rules, under each of the two conditions, is shown in Table 8.2.

Table 8.2

Percentage Selection Frequencies of the Four Cards in the 'True' and 'False' Conditions of Experiment 6

<u>Logical Status</u> <u>of Card</u>	<u>IF P THEN Q</u>		<u>IF P THEN \bar{Q}</u>		<u>IF \bar{P} THEN Q</u>		<u>IF \bar{P} THEN \bar{Q}</u>		<u>ALL FOUR RULES</u>	
	<u>TRUE</u>	<u>FALSE</u>	<u>TRUE</u>	<u>FALSE</u>	<u>TRUE</u>	<u>FALSE</u>	<u>TRUE</u>	<u>FALSE</u>	<u>TRUE</u>	<u>FALSE</u>
	TA	96	75	96	79	100	88	88	71	95
FA	13	33	17	25	17	29	21	46	17	33
TC	83	58	67	58	83	71	75	38	77	56
FC	29	50	50	67	25	54	38	63	35	58

Effects of Truth Status

Subjects in the 'false' condition selected significantly more FC (Kendall's S, $p < 0.03$, one-tailed), less TC (Kendall's S $p < 0.02$, two-tailed) and less TA (Kendall's S, $p < 0.05$, two-tailed) than did subjects in the 'true' condition. Subjects in the 'false' condition selected more FA on all four rules, but this failed to reach significance.

Subjects selected significantly more FC in the 'false' condition and thus this experiment replicated, on abstract rules, the main effect of truth status observed on thematic rules in Experiment 3. This result accords with all three sets of predictions listed in Table 8.1. Subjects also selected significantly less TC in the 'false' condition and thus the effect of falsity is not a simple facilitation of card selection. There was, in fact, no overall facilitation of card selection in the 'false' condition.

Given that subjects learned 'implication' contingencies, the effect of truth status on consequent selections accords both with the association theory predictions and with those of the theory that falsity leads subjects to behave more in accordance with logical validity (see Table 8.1). However, on the basis of the antecedent selections, it is clear that subjects are not behaving more 'logically' in the 'false' condition. This hypothesis is disconfirmed by the significant decrease in TA selection ($p < 0.05$, two-tailed). This decrease in TA selection, and the increase in FA selection, does accord with the association theory predictions (see Table 8.1)

It had been expected that the 'matching bias' effect observed by Evans & Lynch (1973) and replicated by Manktelow & Evans (1979) would be further replicated on these data. Across both conditions, as predicted, significantly more TC was selected when the

consequent was affirmative (sign test, $p < 0.005$, one-tailed) and significantly more FC was selected when the consequent was negative (sign test, $p < 0.01$, one-tailed). Generally, across both conditions, more Q than \bar{Q} selections were made (sign test, $p < 0.001$, one-tailed), this result holding for both the 'true' (sign test, $p < 0.02$, one-tailed), and the 'false' (sign test, $p < 0.02$, one-tailed), conditions. There was thus good evidence of a matching effect on consequent selections. However, there was no significant evidence of an overall matching effect on antecedent selections.

Probability Learning Differences

The median number of trials to criterion on the first probability learning task was four. The data were split into a below median ('fast learner') group ($N = 21$) and an above median ('slow learner') group ($N = 20$). Twelve of the fast learners, and ten of the slow learners, were from the 'true' condition. Table 8.3 shows an analysis of selections for both of these groups and Table 8.4 presents a simplified version, with data pooled across all four rules.

In Experiment 5, fast learners selected significantly more TA and less FC. As only AA rules were used, they thus also selected more P and less \bar{Q} . However, these results were not replicated for either logical or matching values. Fast learners did select more TA than slow learners, but this result was not significant.

The only significant difference between the groups was on FA selection, fast learners selecting less FA than slow learners (Kendall's S, $p < 0.04$, two tailed).

Table 8.3

Separate Analysis of Percentage Selection Frequencies for Fast and Slow Learners (Data from Experiment 6)

	<u>IF P THEN Q</u>		<u>IF P THEN \bar{Q}</u>		<u>IF \bar{P} THEN Q</u>		<u>IF \bar{P} THEN \bar{Q}</u>	
	<u>TRUE</u>	<u>FALSE</u>	<u>TRUE</u>	<u>FALSE</u>	<u>TRUE</u>	<u>FALSE</u>	<u>TRUE</u>	<u>FALSE</u>

Table 8.3A Fast Learners (N = 12 for 'true', and N = 9 for 'false', condition)

TA	92	89	100	89	100	89	83	89
FA	8	22	17	11	8	11	25	11
TC	83	55	58	67	83	78	75	33
FC	33	67	50	89	25	55	50	78

Table 8.3B Slow Learners (N = 10 in each condition)

TA	100	60	90	70	100	80	90	50
FA	20	50	10	20	30	60	10	80
TC	90	80	80	40	90	70	80	40
FC	30	30	50	50	30	50	20	60

Table 8.4

Comparison Between Fast and Slow Learners of Percentage Selections

(Pooled Across the Four Rule Forms in Experiment 6)

	<u>'TRUE' CONDITION</u>		<u>'FALSE' CONDITION</u>	
	<u>FAST LEARNERS</u>	<u>SLOW LEARNERS</u>	<u>FAST LEARNERS</u>	<u>SLOW LEARNERS</u>
TA	94	95	89	65
FA	15	17	14	52
TC	75	85	58	57
FC	40	32	72	47

Differential Effects of Truth Status

It is apparent from Table 8.4 that truth status did not affect the antecedent selections of fast learners. Fast learners showed no statistical evidence of selecting less TA in the 'false' condition (Kendall's S, $p > 0.8$, two-tailed) and a nonsignificant tendency to select less FA (Kendall's S, $p > 0.5$, two-tailed). Slow learners however, selected significantly more FA (Kendall's S, $p < 0.03$, two-tailed), and significantly less TA (Kendall's S, $p < 0.04$, two-tailed), in the 'false' condition.

The effect of truth status on TA selections was thus significant for the slow learners but nowhere near significance for the fast learners, and the effect of truth status on FA selections was (marginally) reversed for fast learners. It may thus be concluded from this analysis that the effect of truth status on antecedent selections is both to increase FA, and decrease TA, selections in the 'false' condition, but that fast learners are

'immune' to this effect. It appears, however, that both fast and slow learners are equally affected by the effect of truth status on consequent selections.

Differential Effects of Matching Bias

Although slow learners showed a significant overall¹ effect of matching bias, ($p < 0.03$, one-tailed), this effect was only near significance for antecedent (sign test, $p < 0.09$, one-tailed) and consequent (sign test, $p < 0.08$, one-tailed) selections analysed separately. In contrast, fast learners showed no tendency to match on antecedent selections ($p = 1.0$) but a highly significant tendency to match on consequent selections (sign test, $p < 0.001$).

These results show that matching only affects consequent selections of fast learners but suggest that for slow learners, matching has an equal effect on antecedent and consequent selections. This differential effect can be tested directly by testing the interaction between matching bias and component (i.e. antecedent or consequent) of selection.² This interaction was found to be significant for fast learners (sign test, $p < 0.01$, two-tailed) but wholly nonsignificant for slow learners (sign test, $p = 1.0$, two-tailed). Thus matching bias only affected consequent selections of fast learners but had an equal effect on antecedent and consequent selections of slow learners.

Correlations³ Between Selections

A thorough analysis by Evans (1977b) has indicated that no correlations exist between TC and FC (or Q and \bar{Q}) selections. However, analysis of the data of experiment 5 suggested the possibility that there is a (negative) correlation between TA and

FA (or P and \bar{P}), and that antecedent selections are positively correlated with consequent selections of the same matching status, and negatively correlated with consequent selections of opposite matching status (see Table 7.5). However, few of the tests performed yielded a significant result and no conclusions may be drawn from those data. Clearly, a far more thorough test, of those correlations not tested by Evans (1977b), is required.

The problem, as Evans (1977b) pointed out, is that most selection task studies have used the AA rule only, on which P is almost always selected and \bar{P} very rarely selected. To test correlations involving these cards, it is thus preferable to consider data from all four forms of conditional rule. Such data are available from Experiment 6, but are also available from the experiment of Evans & Lynch (1973) and also from Experiments 1 and 2 of Manktelow & Evans (1979). It was thus decided that, to provide the strongest possible test of null hypotheses concerning correlations, an analysis of data from all these four studies would be conducted. Tests of correlation on the AA rule were included, but were restricted to these four studies to maintain comparability with other rule forms.

Tests of all five possible correlations not considered by Evans (1977b) were conducted, on each of the four rules, in the following manner:

- 1) Tests of correlation within each study were conducted by use of Fisher's Exact Tests (in accordance with Evans, 1977b). To provide the strongest possible test, data from 'true' and 'false' conditions in Experiment 6 were 'pooled' (as the groups were independent), as were the data from abstract and thematic groups in the two experiments of

Manktelow & Evans (1979), producing $N = 48$ in each case.

It might be objected that data from abstract and thematic groups should not be confused, but it should be noted that no differences between abstract and thematic conditions were observed by Manktelow & Evans.

- 2) For this purpose, the Fisher's Exact tests were conducted one-tailed, in the direction of the correlation present and a positive or negative Z value assigned to each result. Z was set equal to zero, whenever Fisher's Exact tests yielded $p > 0.5$ in either direction.
- 3) Tests of correlation were conducted across all four sets of data by the Stouffer method (also employed in chapter 5 - see Rosenthal, 1978). That is, the Z values from each set of data were used to produce a combined estimate of Z .

The results of these tests are shown in Table 8.5. When considering correlations between antecedent and consequent selections, it is possible to present them ordered as to logical or matching status. Presentation in terms of matching status has been adopted in Table 8.5 on the grounds that the pattern suggested by the data of Experiment 5 concerned matching, rather than logical, status. (The tests are of course the same - it is only a question of presentation.)

Table 8.5 shows the Z value obtained from each study, together with the combined estimate of Z . The last column gives the two-tailed significance level (if significant) of the 20 tests.

Tables of the raw data are given in the Appendix.

Table 8.5

'Z' Values Assigned to Results of Tests of Correlation Between Pairs of
Card Selections on the Four Rules. Data from Experiment 6,

Evans & Lynch, 1973, Manktelow & Evans, 1979

<u>COMPARISON</u>	<u>RULE</u> <u>FORM</u>	<u>STUDY FROM WHICH Z VALUE IS DERIVED</u>				<u>Combined</u> <u>'Z'</u> <u>Estimate</u>	<u>p</u> <u>2-tail</u>
		<u>E & L</u>	<u>M & E 1</u>	<u>M & E 2</u>	<u>Exp 6</u>		
P v \bar{P}	AA	- 0.70	- 0.60	- 1.84	- 1.75	- 2.44	0.02
	AN	- 1.38	0.00	- 0.34	- 1.30	- 1.51	n.s.
	NA	- 2.50	- 2.94	- 2.69	- 2.34	- 5.23	0.0001
	NN	- 2.06	- 4.00	- 4.00	- 3.83	- 6.94	0.0001
P v Q	AA	0.00	0.00	+ 0.87	+ 0.44	+ 0.65	n.s.
	AN	+ 0.98	0.00	0.00	0.00	+ 0.49	n.s.
	NA	- 0.15	+ 2.09	+ 0.32	- 0.81	+ 0.72	n.s.
	NN	0.00	+ 2.41	+ 2.65	+ 0.91	+ 2.98	0.005
P v \bar{Q}	AA	- 0.67	- 1.17	- 2.28	- 0.61	- 2.36	0.02
	AN	- 2.68	- 0.50	+ 0.30	0.00	- 1.44	n.s.
	NA	- 0.72	- 1.17	+ 1.01	+ 0.80	- 0.04	n.s.
	NN	- 0.89	- 2.38	- 2.91	- 1.54	- 3.86	0.0005
\bar{P} v Q	AA	0.00	- 0.12	- 1.24	- 0.24	- 0.80	n.s.
	AN	- 0.21	- 0.65	- 1.07	+ 0.46	- 0.73	n.s.
	NA	0.00	- 1.40	0.00	+ 1.13	- 0.13	n.s.
	NN	0.00	- 2.46	- 2.29	- 1.06	- 2.90	0.005
\bar{P} v \bar{Q}	AA	+ 1.27	+ 0.87	+ 2.66	+ 0.80	+ 2.80	0.01
	AN	+ 1.38	+ 2.26	+ 1.60	+ 0.91	+ 3.07	0.005
	NA	+ 0.89	+ 1.59	+ 1.67	0.00	+ 2.07	0.04
	NN	+ 1.82	+ 1.35	+ 3.98	+ 2.25	+ 4.70	0.0001

Key

E & L - Evans & Lynch (1973)

M & E - Manktelow & Evans (1979)

Before detailed discussion of Table 8.5, one general point should be made. When performing 20 analyses, there is a problem with significance levels as, mathematically, it would be 'expected' that one result would prove significant at the 5% level. (i.e. that one 'type 1' error would occur). Even four or five such results, in 20 tests, would be very difficult to interpret. However, this difficulty does not greatly apply to interpretation of the results of the tests shown in Table 8.5, for two reasons. Firstly, it would certainly not be 'expected' that eleven out of 20 analyses would prove significant and, secondly, it is noticeable that many of the significant results have a probability which is far lower than 5%. Three results yield a probability of a type 1 error of less than one in 10,000 and four other results yield a probability of a type 1 error of less than one in 200.

The most clear cut finding shown in Table 8.5 is that \bar{P} selections are positively correlated with \bar{Q} selections. This result is significant on all rules and it is particularly noticeable that, within each individual study, no test yields a negative correlation. As this correlation proved significant on the AA rule, the result was further investigated, with respect to the data from all the other selection task experiments presented in this thesis, and it was found that \bar{P} and \bar{Q} selections on AA rules were invariably positively correlated in all conditions of each experiment (although not always significantly correlated). This correlation is thus very consistent.

It is also reasonable to conclude from Table 8.5 that there is a negative correlation between P and \bar{P} selections, although this correlation was not significant for the AN rule and produced noticeably better levels of significance on the NA and NN rules

(which produce lower frequencies of P, and higher frequencies of \bar{P} , selections. The correlation was consistently negative across the four studies. However, analysis across the (AA rules of) other experiments presented in this thesis did not reveal a consistent negative association between P and \bar{P} , although in some cases, such tests were possibly meaningless, due to the very high frequencies of P selection.

Generally, the sign of the combined Z estimates is consistent with the idea that antecedent and consequent selections are positively correlated when they have similar matching status and negatively correlated when they have dissimilar matching status. However, apart from the $\bar{P} \text{ v } \bar{Q}$ comparison, these correlations rarely approached significance, except in the case of the NN rule. Although it is possible that these tests may prove significant across a large number of studies, it is probable that, even if this general tendency exists, it plays a very minor role in influencing the data. Only in the case of the NN rule does this pattern play a clear role.

In conclusion, the results shown in Table 8.5 may be summarised in three ways:

- 1) There is a highly consistent positive correlation between \bar{P} and \bar{Q} selections.
- 2) There is a negative correlation between P and \bar{P} selections although this correlation appears to be much stronger on the NA and NN rules and did not reach significance on the AN rule. Further tests on the data of other experiments presented in this thesis suggested that, on the AA rule, this correlation is not as consistent as that between \bar{P} and \bar{Q} selections.
- 3) On the NN rule there is a highly significant tendency

for selection of cards of similar matching status to be positively correlated and for selection of cards of dissimilar matching status to be negatively correlated. It is possible that this general tendency extends to other rule forms but, even if this were the case, it is unlikely that it would play an important role in determining responses.

CONCLUSIONS

In the 'false' condition of Experiment 6, subjects selected significantly more FC and significantly less TA and TC. They also selected more FA, although this result did not reach significance. However, that the increase in FA selection failed to reach significance may be ignored, as further analysis revealed that truth status did not affect the antecedent selections of fast learners, but that slow learners selected significantly less TA, and significantly more FA, in the 'false' condition.

The results of Experiment 6, concerning truth status effects, were thus all in accord with the association theory predictions shown in Table 8.1. The other sets of predictions were not confirmed. There was no overall facilitation of card selection in the 'false' condition and the hypothesis that 'falsity' leads subjects to behave more in accordance with logical validity was disconfirmed by the observed effects of truth status on antecedent selections.

Fast learners selected significantly less FA, and non-significantly more TA, than slow learners. These results are consistent with the Experiment 5 finding that fast learners selected

significantly more TA and nonsignificantly less FA than slow learners (see Table 7.4). Further, in Experiment 6, fast learners appeared to be 'immune' from the effect of both matching and truth status on antecedent selections. These results suggest that fast learners display response differences, to slow learners, on antecedent selections in general, such that their behaviour is more in accordance with logical validity and less susceptible to the effect of nonlogical factors. It is of interest to consider why this should be the case.

The most apparent explanation is that fast learners have a stronger appreciation of the (logical) necessity to test modus ponens. In the discussion of Experiment 5 it was pointed out that an alternative possibility is that selection differences are a result of the fact that slow learners had a longer period of exposure to the task materials. For instance, it is possible that longer exposure may enhance the effect of truth status on antecedent selections. However, it cannot be argued that the shorter prior exposure to the materials inhibited the fast learners from making antecedent matching responses. All subjects in Experiment 6 received considerably more prior exposure to the task materials than has been the case in other experiments which have produced matching effects. Thus differences in the effects of matching bias, between fast and slow learners, indicate that response differences between them are not a function of differential prior exposure to the task materials. It is thus not unreasonable to conclude that individual differences do play a role in selection task behaviour, such that some subjects are 'better' at both probability learning and making logically valid antecedent selections. There was no evidence of academic differences between fast, and

slow, learners, who were sampled from the same year of the same course. It is thus reasonable to assume that differential abilities were specifically related to behaviour concerning conditional relationships.

The implications of the correlations between certain selections, shown in Table 8.5, will be considered in relevant parts of the general discussion.

FOOTNOTES

1. Analysed across both conditions and across antecedent and consequent. (i.e., for each subject, subtracting \bar{P} and \bar{Q} , from P and Q selections.

2. This may be done by computing $(P - \bar{P}) - (Q - \bar{Q})$ for each subject, where P, \bar{P} , Q and \bar{Q} stand for the total number of that type of selection. A negative result indicates that the subject matched more on consequent selections and a positive result indicates that the subject matched more on antecedent selections

3. See Note 5 of Chapter 7.

APPENDIX TO CHAPTER 8

SOURCE TABLES FOR TABLE 8.5

Table 8.A1

Tests of Association between P and \bar{P}

<u>RULE</u>	<u>SOURCE</u>	<u>BOTH</u>	<u>P ONLY</u>	<u>\bar{P} ONLY</u>	<u>NEITHER</u>	<u>Direction of</u>	<u>Z value</u>
<u>FORM</u>						<u>Association</u>	
AA	E & L	1	20	1	2	-	0.70
	M & E 1	6	40	1	1	-	0.60
	M & E 2	3	38	3	4	-	1.84
	Expt 6	7	34	4	3	-	1.75
AN	E & L	0	22	1	1	-	1.38
	M & E 1	4	42	0	2	+	0.00
	M & E 2	4	40	1	3	-	0.34
	Expt 6	7	35	3	3	-	1.30
NA	E & L	1	7	12	4	-	2.50
	M & E 1	6	6	34	2	-	2.94
	M & E 2	9	11	24	4	-	2.69
	Expt 6	8	3	37	0	-	2.34
NN	E & L	3	9	9	3	-	2.06
	M & E 1	5	9	33	1	-	4.00
	M & E 2	7	15	24	2	-	4.00
	Expt 6	7	9	31	1	-	3.83

Key

E & L - Evans & Lynch (1973)

M & E 1 - Manktelow & Evans (1979) Expt 1

M & E 2 - Manktelow & Evans (1979) Expt 2

Table 8.A2

Tests of Association between P and Q

<u>RULE</u> <u>FORM</u>	<u>SOURCE</u>	<u>BOTH</u>	<u>P ONLY</u>	<u>Q ONLY</u>	<u>NEITHER</u>	<u>Direction of</u> <u>Association</u>	<u>Z value</u>
AA	E & L	10	11	2	1	-	0.00
	M & E 1	24	22	1	1	+	0.00
	M & E 2	28	13	3	4	+	0.87
	Expt 6	30	11	4	3	+	0.44
AN	E & L	14	8	0	2	+	0.98
	M & E 1	34	12	2	0	-	0.00
	M & E 2	31	13	3	1	-	0.00
	Expt 6	24	18	4	2	-	0.00
NA	E & L	4	4	10	6	-	0.15
	M & E 1	12	0	24	12	+	2.09
	M & E 2	18	2	23	5	+	0.32
	Expt 6	7	4	30	7	-	0.81
NN	E & L	9	3	8	4	+	0.00
	M & E 1	13	1	18	16	+	2.41
	M & E 2	19	3	12	14	+	2.65
	Expt 6	10	6	14	18	+	0.91

Key

E & L - Evans & Lynch (1973)

M & E 1 - Manktelow & Evans (1979) Expt 1

M & E 2 - Manktelow & Evans (1979) Expt 2

Table . A3

Tests of Association between P and \bar{Q}

<u>RULE</u> <u>FORM</u>	<u>SOURCE</u>	<u>BOTH</u>	<u>P ONLY</u>	<u>\bar{Q} ONLY</u>	<u>NEITHER</u>	<u>Direction of Z value</u> <u>Association</u>	
AA	E & L	6	15	2	1	-	0.67
	M & E 1	15	31	2	0	-	1.17
	M & E 2	8	33	5	2	-	2.28
	Expt 6	15	26	4	3	-	0.61
AN	E & L	0	22	2	0	-	2.68
	M & E 1	7	39	1	1	-	0.50
	M & E 2	10	34	0	4	+	0.30
	Expt 6	26	16	4	2	-	0.00
NA	E & L	2	6	8	8	-	0.72
	M & E 1	3	9	18	18	-	1.17
	M & E 2	7	13	5	23	+	1.01
	Expt 6	6	5	13	24	+	0.80
NN	E & L	2	10	5	7	-	0.89
	M & E 1	2	12	19	15	-	2.38
	M & E 2	3	19	15	11	-	2.91
	Expt 6	6	10	21	11	-	1.54

Key

E & L - Evans & Lynch (1973)

M & E 1 - Manktelow & Evans (1979) Expt 1

M & E 2 - Manktelow & Evans (1979) Expt 2

Table 8.A4

Tests of Association between \bar{P} and Q

<u>RULE</u> <u>FORM</u>	<u>SOURCE</u>	<u>BOTH</u>	<u>\bar{P} ONLY</u>	<u>Q ONLY</u>	<u>NEITHER</u>	<u>Direction of</u> <u>association</u>	<u>Z value</u>
AA	E & L	1	1	11	11	=	0.00
	M & E 1	3	4	22	19	-	0.12
	M & E 2	2	4	29	13	-	1.24
	Expt 6	7	4	27	10	-	0.24
AN	E & L	0	1	14	9	-	0.21
	M & E 1	2	2	34	10	-	0.65
	M & E 2	2	3	32	11	-	1.07
	Expt 6	7	3	21	17	+	0.46
NA	E & L	8	5	6	5	+	0.00
	M & E 1	28	12	8	0	-	1.40
	M & E 2	28	5	13	2	-	0.00
	Expt 6	36	9	1	2	+	1.13
NN	E & L	9	3	8	4	+	0.00
	M & E 1	21	17	10	0	-	2.46
	M & E 2	16	15	15	2	-	2.29
	Expt 6	17	21	7	3	-	1.06

Key

E & L - Evans & Lynch (1973)

M & E 1 - Manktelow & Evans (1979) Experiment 1

M & E 2 - Manktelow & Evans (1979) Experiment 2

Table 8.A5

Tests of Association between \bar{P} and \bar{Q}

<u>RULE</u>	<u>SOURCE</u>	<u>BOTH</u>	<u>\bar{P} ONLY</u>	<u>\bar{Q} ONLY</u>	<u>NEITHER</u>	<u>Direction of Association</u>	<u>Z value</u>
AA	E & L	2	0	6	16	+	1.27
	M & E 1	4	3	13	28	+	0.87
	M & E 2	5	1	8	34	+	2.66
	Expt 6	6	5	13	24	+	0.80
AN	E & L	1	0	1	22	+	1.38
	M & E 1	3	1	5	39	+	2.26
	M & E 2	3	2	7	36	+	1.60
	Expt 6	8	2	22	16	+	0.91
NA	E & L	7	6	3	8	+	0.89
	M & E 1	20	20	1	7	+	1.59
	M & E 2	11	22	1	14	+	1.67
	Expt 6	18	27	1	2	+	0.00
NN	E & L	6	6	1	11	+	1.82
	M & E 1	19	19	2	8	+	1.35
	M & E 2	18	13	0	17	+	3.98
	Expt 6	25	13	2	8	+	2.25

Key

E & L - Evans & Lynch (1973)

M & E 1 - Manktelow & Evans (1979) Experiment 1

M & E 2 - Manktelow & Evans (1979) Experiment 2

CHAPTER 9

AN EXPERIMENT ON THE EFFECT OF IMPLICIT TRUTH STATUS,

USING THEMATIC RULES

EXPERIMENT 7

277

In the three abstract experiments on truth status (Experiments 4, 5 and 6), the truth status was always implicit, whereas it was explicit in the thematic Experiment 3 (i.e. subjects were specifically asked to generate 'true' or 'false' rules). It was thus decided to extend investigation to thematic rules having implicit truth status. For this purpose, it was decided that an implicit relationship exists between certain words and colours and that rules such as 'if the word on one side, is grass, then the colour on the other side is 'green', would serve to introduce an implicit truth status to the selection task.

The experiments carried out to date have used 'true' conditions in which TC is related to TA and 'false' conditions in which FC is related to TA. For clarity, an example of these two conditions is shown in Table 9.1A and 9.1B, using thematic terms of the type used in this experiment:

Table 9.1

Table showing Possible Thematic Content of a 'True' Rule (9.1A) and Various Possible Alternative Contents of 'False' Rules (9.1B, 9.1C and 9.1D) to be used in Experiment 7

<u>LOGICAL STATUS</u>	<u>'TRUE' RULE:</u>	<u>'FALSE' RULE: 'IF GRASS THEN RED'</u>			
	<u>'IF GRASS THEN GREEN'</u>				
	<u>9.1A</u>	<u>9.1B</u>	<u>9.1C</u>	<u>9.1D</u>	
<u>TA</u>	GRASS	GRASS	GRASS	GRASS	GRASS
<u>FA</u>	COAL	COAL	COAL	COAL	BLOOD
<u>TC</u>	GREEN	RED	RED	RED	RED
<u>FC</u>	RED	GREEN	BLUE	BLUE	BLUE

It can be seen from Table 9.1A and 9.1B that the four cards used are the same in each condition but that the logical status of the 'green' and 'red' cards is dependent on the rule to

be evaluated. In explaining the effect of truth status, it has been argued that FC is more frequently selected in the 'false' condition due to its association with TA (i.e. that subjects will tend to select the 'green' card in both 'true' and 'false' conditions). However, there are other forms of 'false' condition such as that depicted in Table 9.1C.

In condition C, neither TC nor FC is associated with TA. Grass is neither red nor blue. Thus, if truth status effects are mediated by association, FC selection should be facilitated in 'false' condition B, but not in 'false' condition C.

A further possible 'false' condition is that depicted in Table 9.1D. As in condition C, TA (grass) is associated with neither TC (red) nor FC (blue). However, TC (red) is associated with FA (blood) and this greater association with the consequent value mentioned in the rule should facilitate FA selection.

Accordingly, an experiment was designed which presented subjects with four selection tasks, one of each of the above four types. The 'true' condition A was used as a comparison condition for the three 'false' conditions and the following predictions were made:

A - B Comparison: That subjects would select less TC and more FC in condition B (in accord with previous results).

A - C Comparison: That subjects would select less TC, but would not select more FC in condition C.

A - D Comparison: That subjects would select more FA in condition D.

Finally, a further feature was introduced into this

experiment. Subjects were asked to write down the cards they wished to select and the cards they wished not to select (in separate columns of a sheet of paper). This modification has two advantages. Firstly, it minimises experimenter/subject interaction during the task and, secondly, it allows for an analysis of the order in which cards are selected or rejected.

Due to the difficulty of deriving NA conditional rules from nonbinary stimuli, only affirmative (AA) rules were used in this experiment.

EXPERIMENT 7

METHOD

Design

Subjects were presented with four selection tasks, all concerned with the evaluation of a (thematic) AA conditional rule. One task was presented under each of four conditions, as follows:

- 1) A 'true' condition (A) in which there was a natural relationship between TA and TC.
- 2) A 'false' condition (B) in which there was a natural relationship between TA and FC.
- 3) A 'false' condition (C) in which TA was related to neither TC nor FC.
- 4) A 'false' condition (D) in which TA was related to neither TC nor FC but in which FA was related to TC.

These four conditions were presented to each subject in a random order.

Materials

1) A pack of 3" x 4" cards was constructed, all having one coloured side and a word written in brown ink on the other (white) side. The following words and colours were used:

<u>Words</u>	<u>Colours</u>
Grass	Green
Blood	Red
Canary	Yellow
Sky	Blue
Coal	Black

The pack consisted of 50 cards, comprising 2 each of all possible ($5 \times 5 = 25$) combinations of the above.

The cards to be used in the selection task were pre-assigned such that:

- a) Each word was used as TA and FA, and each colour was used as TC and FC, an equal number of times in each of the four conditions.
- b) Each of the ten possible pairs of words, and ten possible pairs of colours, (5C_2), were used on at least one problem.
- c) Each word was used as TA no more than once for each subject.
- d) For each TA, the colour used as the 'false' TC in conditions B, C and D was also used as FC in condition A. Thus, for a given TA word, the same two colours were used in conditions A and B. For instance, for rules about the word 'sky', condition A used (an expansion of) the rule 'sky then blue' and used blue as TC and green as FC, whereas condition B used the rule 'sky then green', and used blue as FC and green as TC.

2) Ten sheets of paper with conditional rules typed on them as follows:

For use in Condition A

If the word on one side of the card is SKY, then the colour on the other side is BLUE.

If the word on one side of the card is GRASS, then the colour on the other side is GREEN.

If the word on one side of the card is BLOOD, then the colour on the other side is RED.

If the colour on one side of the card is COAL, then the colour on the other side is BLACK.

If the word on one side of the card is CANARY, then the colour on the other side is YELLOW.

For use in Conditions B, C and D

If the word on one side of the card is SKY, then the colour on the other side is GREEN.

If the word on one side of the card is GRASS, then the colour on the other side is RED.

If the word on one side of the card is BLOOD, then the colour on the other side is YELLOW.

If the word on one side of the card is COAL, then the colour on the other side is BLUE.

If the word on one side of the card is CANARY, then the colour on the other side is BLACK.

3) Sheets of duplicated paper of the following form:

NECESSARY TO TURN OVER	NOT NECESSARY TO TURN OVER

Subjects

Thirty undergraduates at Plymouth Polytechnic acted as paid volunteers.

Procedure

The selection task presentation was essentially the same as for Experiments 4 and 6. Changes in the presentation are listed below:

- 1) The subject was not asked to choose the cards to be laid on the table, the experimenter selecting the first cards that were the preassigned designates for that condition.
- 2) At no time during the experiment was a subject allowed to see both sides of the same card.
- 3) Instead of verbally indicating their selections, subjects were asked to categorise all four cards as either necessary, or not necessary, to turn over, on the piece of paper illustrated in the Materials section.

As analysis was to be conducted on the order in which the subject wrote down his selections or rejections, the order in which the cards were laid on the table was of particular importance. Two orders were used:- A) word, word, colour, colour; and B) colour, colour, word, word. These orders were alternated across the four tasks, with fifteen subjects being presented with order A first. The position (within a pair) of TA or FA, and of TC or FC, was determined randomly - the first required word or colour encountered being placed on the right of the pair. The pack was shuffled after each selection task.

RESULTS

The percentage frequencies of selection of the four cards, in the four conditions, is shown in Table 9.2.

Table 9.2

Percentage Frequencies of Selections of the Four Cards in the Four
Conditions of Experiment 7

<u>CARD</u>	<u>C O N D I T I O N</u>			
	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>
TA	87	90	87	93
FA	10	7	13	17
TC	57	57	53	50
FC	27	27	23	23

There is clearly no difference between any of the conditions and no statistical analysis was carried out.

Order Effects

Analysis of the order in which subjects write down their evaluations is best carried out by analysing separately each of the six possible pairs of cards. For each pair, a data point is thus obtained whenever a subject selects or rejects both members of the pair (if he selects one, and rejects the other, no data is obtained, as the cards will be written in different columns). Each subject may thus contribute between 0 and 4 data points to each comparison.

Table 9.3 shows the distribution of these data points, together with the significance level obtained from two-tailed sign tests. It should be noted that the figures in the table contain more than one measure from some subjects and thus they are not the figures on which the sign tests are based.

Table 9.3

Analysis of the Order in which Pairs of Cards were Selected or

<u>COMPARISON OF PAIR SELECTED OR REJECTED TOGETHER</u>	<u>Rejected in Experiment 7</u>		
	<u>FIRST NAMED</u>	<u>SECOND NAMED</u>	<u>p</u>
	<u>WRITTEN DOWN</u>	<u>WRITTEN DOWN</u>	<u>2-tailed</u>
	<u>FIRST BY SUBJECT</u>	<u>FIRST BY SUBJECT</u>	
TA v TC	38	16	0.01
TA v FA	15	4	0.04
TA v FC	29	6	0.03
FA v FC	63	37	0.02
TC v FC	32	23	n.s.
FA v TC	31	30	n.s.

For the TC v FC comparison, TC was written first significantly more often ($p < 0.04$) when TC and FC were selected but there was no difference when TC and FC were rejected. There was also some suggestion in the data that TC is selected before FA - all five subjects who selected both TC and FA, on one or more occasions, wrote TC down first. It appears reasonable to draw the following conclusions:

- 1) There is a significant tendency to respond to TA before all other cards.

- 2) There is a significant tendency to respond to FA before FC.
- 3) There is a possibility that the position of TC in this hierarchy varies. It may be selected, but not rejected, before FA and FC.

One further test was carried out. It was considered of interest to test whether comparative association strength had any effect on order. In conditions A and B there is one associated, and one nonassociated (with TA), consequent stimulus, the associated stimulus acting as TC in condition A and as FC in condition B. As there were no selection differences between the conditions, associated and nonassociated stimuli can thus be compared across conditions A and B, with logical status controlled for.

Whenever TC and FC were selected or rejected together in conditions A and B, the associated stimulus tended to be written down first. This tendency, however, just failed to reach significance (sign test, $p < 0.06$, two-tailed). However, the tendency was weakened in the case of selections by the strong tendency to select TC before FC. When rejections alone were considered, subjects did show a significant tendency to write down the associated card before the nonassociated card (sign test, $p < 0.05$, two-tailed).

DISCUSSION

Comparisons between the various 'false' conditions cannot be made as no effect of truth status was observed.

One possible explanation of this is that sequential presentation of four selection tasks (without another task in

between , as was the case in Experiment 5 and 6) will inhibit differential effects. In fact, 17 of the 30 subjects made the same selection on each of the four tasks. However, when initial tasks only are considered, there is still no evidence of an effect of truth status. For instance, FC was selected by two of the eight subjects who received condition B first and by one of the ten subjects who received condition A first. There is thus no evidence, from initial selections, that condition B facilitated FC selection and it would appear that truth status does not affect behaviour with these materials.

That subjects attended to the associated consequent before the nonassociated consequent is suggested by the fact that they showed some tendency to write this value down first. This 'cueing' of the associated value has been argued in Experiments 3, 4, 5 and 6 to lead to selection. In Experiment 7, however, subjects correctly reevaluated the association as irrelevant. One possible explanation is that the association between, for instance, 'grass' and 'green' is too available. It may be that the association is so obvious to the subject that he becomes suspicious of the experimenter's motives and discounts it.

There were, however, other more clear cut findings of Experiment 7. It is of interest that TA is responded to before the other cards and that FA is responded to before FC. Clearly, knowledge of the order in which the cards are responded to can contribute to the understanding of behaviour on the task. For instance, the priority of TA selection suggests that subjects find this decision more 'obvious', and is consistent with the typical finding that most subjects select it. It is also consistent with the finding of Evans & Newstead (1977) that, on a truth table task,

the least variable responses across subjects (e.g. evaluation of the TT case as 'true') tended to have a lower latency.

PART 3

GENERAL DISCUSSION

Truth status effects are discussed in Chapter 10 with particular reference to the 'availability' theory of Tversky & Kahneman (1973). Theoretical implications of this discussion are extended, in Chapter 11, to other areas of reasoning research. Chapter 12 focusses upon the position of Evans (1977b), which it was concluded, at the end of Part 1, was the best available published explanation of reasoning behaviour. The structure of the mathematical model is considered in some detail and the Evans (1977b, 1980a, 1980b) position is evaluated, both in terms of the theoretical discussion contained in Chapters 10 and 11, and in the light of various experimental findings of Part 2.

CHAPTER 10

TRUTH STATUS EFFECTS

SECTION 10.1	Main Experimental Findings	288
SECTION 10.2	The 'Availability' Theory of Tversky & Kahneman (1973)	297
SECTION 10.3	Some Applications of the 'Availability' Heuristic to Thematic Reasoning Studies	302
SECTION 10.4	Some Applications of the 'Availability' Heuristic to Abstract Reasoning Studies	306
SECTION 10.5	Functional Utility of Availability	311
SECTION 10.6	Broader Implications	315

SECTION 10.1 MAIN EXPERIMENTAL FINDINGS

A significant effect of truth status on selection task behaviour was observed in all relevant experiments, except Experiment 7. Any full explanation of truth status must take account of the findings of this latter study but they will be disregarded in this initial discussion.

The significant results obtained in experiments 3, 4, 5 and 6 are shown in Table 10.1, which for each card, shows whether more, or less, selections were observed in the 'false' conditions relative to the 'true' conditions. If there was no significant difference, this is indicated in the table by 'N.S.D.'. The last column of the table shows the results of analyses across all four cards.

Table 10.1

Significant Effects of Truth Status Observed in Experiments 3, 4, 5 and 6. Differences Expressed in Terms of Selection Frequencies in the False Condition Relative to the True Condition.

<u>CARD</u>	<u>TA</u>	<u>FA</u>	<u>TC</u>	<u>FC</u>	<u>ALL 4 CARDS</u>
EXPERIMENT 3	NSD	NSD	NSD	<u>MORE</u>	NSD
EXPERIMENT 4	NSD	NSD	NSD	NSD	<u>MORE</u>
EXPERIMENT 5	NSD	NSD	NSD	NSD	<u>MORE</u>
EXPERIMENT 6	<u>LESS</u>	<u>MORE</u> *	<u>LESS</u>	<u>MORE</u>	NSD

* This result was not significant across all subjects, but it was significant for slow learners. The effect of truth status on both FA and TA selections was not significant for fast learners.

The effect of truth status on FC selection is reasonably clear cut. Subjects selected more FC in the 'false' condition, than in the 'true' condition, of Experiments 3, 4, 5 and 6 and this result reached significance in Experiments 3 and 6. It has been suggested that this result is due to the greater associational bond between TA and FC in the 'false' condition. In the probability learning phases of Experiments 4, 5 and 6, subjects had learned an association that connected TA and TC in the 'true' conditions and TA and FC in the 'false' conditions.

It follows from this theory that there should also be a decrease in TC selection in the 'false' condition, as this card will be less associated with TA than it is in the 'true' condition. However, results regarding TC selection have been somewhat variable. Subjects did select less TC in the 'false' conditions of Experiments 3 and 6, this result reaching significance in the latter case and being near significance in the former. In Experiments 4 and 5, however, subjects selected more TC in the 'false' condition, although this result was not significant in either case. In Experiments 4 and 5, subjects selected significantly more cards overall in the 'false' condition, but this result cannot be taken as significant evidence of a tendency to select more of one particular card. The frequencies of selections in Experiment 5, in fact, suggest that there was little difference in TC selections between the conditions, although, in Experiment 4, the difference in TC selection was near significance. As subjects also selected (nonsignificantly) more FA in Experiment 5, the results of Experiments 4 and 5 suggested an alternative explanation of the effect of truth status. If the probability learning phase of these experiments did induce the subjects to interpret the rule as implying equivalence (see discussion of Experiment 5, in Chapter 7), then selection of TC and FA would

be logically valid responses. TC and FA are potential falsifiers if the rule implies equivalence but not if the rule implies implication, whereas FC is a potential falsifier on both interpretations. It was thus possible, on the basis of the results of Experiments 3, 4 and 5, that 'falsity' led subjects to behave more in accordance with logical validity: producing a tendency to select more TC and FA on equivalence rules, less TC and FA on implication rules, and more FC on all rules. Subjects also selected (nonsignificantly) more TA in the false condition of Experiment 5, although generally differences in TA selection on AA rules would not be expected due to a ceiling effect.

However, this explanation does not fit the data of Experiment 6 with regard to antecedent selections. Subjects selected less TA, and more FA, in the 'false' condition, and thus this condition clearly did not produce behaviour more in accordance with logical validity. However these results would be expected if subjects' behaviour is mediated by associations between the cards. If consequent selections are mediated by their association with the antecedent value mentioned in the rule (i.e. TA), then antecedent selections will be mediated by their association with the consequent value mentioned in the rule (i.e. TC). In Experiment 6, FA was more associated, and TA less associated, with TC in the 'false' condition, than in the 'true' condition.¹

Given that subjects do not appear to be behaving more in accordance with logical validity in the 'false' condition, an explanation of antecedent and consequent responses in terms of their association with TC and TA, respectively, appears to be the best available. Given this, the results of Experiment 6 could be taken as revealing the full effect of truth status on responses. That is, that 'falsity' leads subjects to select more FA and FC, and less TA and TC.

However, this position is presented with two problems that require explanation. Firstly, no decrease in TA and TC selections was observed in the 'false' conditions of Experiments 4 and 5. A possible explanation of these results has already been proposed in Chapter 7. On AA rules, a decrease in TA and TC selections may not be observed because these values match. Clearly, if FA and FC selections were facilitated in the 'false' conditions of Experiments 4 and 5, and the decrease in TA and TC selections was inhibited by matching bias, then the net effect would be an increase in overall selection frequencies in the 'false' condition. However, this hypothesis must be treated with reservations. It could not explain an increase in TA or TC selections in the 'false' condition. The results of Experiments 4 and 5, regarding these selections, although not significant, are thus in the 'wrong' direction. Further, it is noticeable that, in Experiment 6, less TA and TC selections were made in the 'false' condition even on AA rules.

The second problem concerns the findings of Experiment 3. In this experiment, subjects did select less TC in the 'false' condition, but no effects of truth status were observed on antecedent selections. That they selected less TC may have been due to matching having had less effect on responses in this experiment, due to the thematic content and the mode of presentation employed. The near significant reduction in TC selection in the 'false' condition may thus be seen as consistent with the association theory in a situation where truth status effects were not inhibited by matching bias. The main problem, for interpretation of the results of this experiment, appears to be that truth status had no effect on FA selections. The lack of effect on TA selections may be explained as due to a 'ceiling' effect but the lack of effect on FA selections

cannot be explained as due to a 'floor' effect, as a reasonable frequency of FA selections was observed in Experiment 3.

The lack of effect of truth status on FA selections in Experiment 3 is probably due to the content of the rules. In Experiment 6, FA was far more likely, given TC, in the 'false' condition, than in the 'true' condition. In the 'false' condition, seven of the eight cards having TC on one side would have FA on the other side, whereas, in the 'true' condition, only half the cards having TC on one side would have FA on the other side. The higher frequency of FA selection in the 'false' condition has thus been argued to be due to its greater association with the consequent value mentioned in the rule. However, in experiments based on the Van Duyne (1976) paradigm, this is not the case.

For instance, in the example given by Van Duyne, using 'sugar', 'protein', 'sweet' and 'bitter', the FA value ('protein') is no more related to one consequent value than another. Further, it was realised in retrospect that there is an unavoidable experimenter bias in this paradigm, that makes interpretation of results concerning FA selections very difficult. Consider, for instance, the following examples of 'true' and 'false' statements used in Experiment 3:

TRUE 'If a person has a bath then he gets wet'

FALSE 'If a person cuts his finger, then he dies'

Most subjects generated such statements, in which the consequent value, but not the antecedent value, is essentially binary. When constructing an FC instance, it is clear that, if a person has not died, then he is alive, and that, if a person is not wet, then he is dry. However, when constructing an FA instance, there is no such clear cut alternative. For instance, for the 'false'

statement, there are a variety of FA alternatives, such as:

- A) A person has not cut his finger
- B) A person has broken his leg
- C) A person is in perfect health
- D) A person has contracted bubonic plague

Clearly, the association between FA and the consequent values will be determined by the nature of the FA alternative adopted. If alternative D above were adopted, then this would have a strong association with 'dies', but the opposite of this association would be likely to result (to varying degrees) from the adoption of alternatives A, B or C.

Bearing the above points in mind, three reasons may be advanced for the absence of an effect of truth status on FA selections in Experiment 3:

- 1) During the conduct of the experiment, it was always attempted to construct 'neutral' instances of FA to avoid influencing the subject's decisions. In retrospect, this procedure led to a lack of any associative effects and thus was possibly the basis of the absence of a truth status effect on FA selections. However, although in some cases (such as the 'sugar' example), a neutral FA instance may be constructed, in other cases the nearest alternative to a neutral instance is the explicit negation of TA (alternative A above). This may lead to some association between FA and TC or between FA and FC.
- 2) If, however, some effects of association were present, they did not necessarily consistently lead to a stronger association between FA and TC on 'false' statements. The examples given above show that an FA instance that is associated with either TC or FC may be constructed (in any condition).

3) If any associative effects were present they could not have had an opposite effect on one condition to another, as different thematic rules were used in each condition.

Thus, due to the attempt to construct 'neutral' instances, it is unlikely that, in most cases, FA had any differential association with either TC or FC in any condition. If such associations did exist in some cases, there is no reason why the association with TC should have been always stronger in the 'false' condition, and thus no reason why subjects should have selected more FA in the 'false' conditions of Experiment 3.

In summary, evaluation of the association theory as an explanation of the results of Experiments 3, 4, 5 and 6, leads to the following conclusions:

- 1) The most consistent effect of truth status has been the increase in FC selection in the 'false' condition and the association theory provides a very good explanation of why this card is 'cued'.
- 2) The effect of truth status on antecedent selections in experiment 6 was consistent with the association theory but inconsistent with the hypothesis that 'falsity' leads subjects to behave more in accordance with logical validity.
- 3) The data lend support to the idea that FA selections will be more frequent in the 'false' condition when FA is more likely, given TC, than in the 'true' condition. That is, that truth status affected FA selections in Experiment 6, but not in Experiment 3, is consistent with the association theory.
- 4) The only significant results concerning TA and TC selection have revealed a decrease in frequency of these

selections in the 'false' condition (see Table 10.1). These results were consistent with the association theory predictions.

The association theory is very similar to the 'availability' theory of Tversky & Kahneman (1973) and, in fact, may be regarded as a 'special case' of their theory, as applied to reasoning problems. Although primarily concerned with subjects' frequency judgements, it is of interest to consider the work on 'availability' in some detail, as it will provide a broader perspective within which to view the possible effects of association in a range of problem environments.

SECTION 10.2 THE 'AVAILABILITY' THEORY OF TVERSKY & KAHNEMAN (1973)

Tversky & Kahneman (1973) propose that frequency judgements are made on the basis of the information that the subject can retrieve, referring to this as 'availability'. They present a variety of experimental evidence for this. For instance, they found that the majority of subjects report that words beginning with K, L, N, R or V are more frequent than words having these letters in the third position, although, in fact, the latter type of words are more frequent. They point out that the former type of words are more easily 'brought to mind'. In another experiment, subjects were presented with lists of names, 19 of them famous people of one sex and 20 of less famous people of the other sex. One group of subjects recalled more of the famous names and another group of subjects judged that sex to have occurred more frequently. The authors argue that, as more examples of one sex were 'available', that sex was judged to have occurred more frequently. As general support for their findings, they quote Leicht (1968) and Underwood, Zimmerman & Freund (1971) as showing that items that are better recalled are judged more frequent. A less artificial effect of availability is to be observed in the data of Lichtenstein et al (1978). They found that, in estimating relative frequencies of lethal events, subjects tended to overestimate the frequency of those events likely to attain sensational media coverage (e.g. hurricanes and murder). Subjects were far more accurate in judging the relative frequency of common words, for which the subjects' prior exposure would have 'matched' actual frequency (see also Shanteau, 1978).

Tversky & Kahneman propose that "availability provides a natural explanation for" the 'illusory correlation' findings of Chapman (1967). Chapman presented subjects with lists of pairs of words, some pairs repeated, and found that subjects overestimated the frequency of cooccurrence of associated pairs. For instance, subjects tended to estimate the word pair 'bacon-eggs' as having occurred more often than the word pair 'lion-eggs'. In subsequent experiments, subjects were exposed to a list of clinical diagnoses and test results and asked to evaluate the cooccurrence of various symptoms and diagnoses. Subjects overestimated the cooccurrence of pairs that were judged to be natural associates by an independent group of subjects. Subjects perceived correlations between such pairs, even when the actual correlation was negative (Chapman & Chapman, 1967, 1969). Tversky & Kahneman replicated these findings with minor modifications to ensure that the effect was due to the presence of associations.

Tversky & Kahneman argue that, if stimuli occur together, then they tend to become associated and thus that subjects are "likely to conclude" that high association strength indicates a high frequency of cooccurrence. They point out, however, that there are other sources of association, besides repetition, and conclude: "thus, the various sources of illusory correlation can all be explained by the operation of a single mechanism - the assessment of 'availability' or associative strength".

On the basis of this analysis, the 'availability' of associations between antecedent and consequent values will determine a subject's evaluation of a conditional statement. There will be more available TA-TC examples in the case of 'true' statements and more available TA-FC examples in the case of 'false' statements.

On the selection task, the FC card will thus be more 'available', and the TC card less 'available' on 'false' statements. Similarly, on standard abstract tasks, the matching values will be more 'available' as they have been mentioned in the rule, in the same way as lethal events are more 'available' if they have been mentioned in the media. Of course, although availability may explain the mechanism underlying subjects' belief, there is no logical reason why the subject should select 'available' cards and thus, if 'availability' does mediate selections, its effect in 'cueing' certain cards must be seen as a nonlogical influence.

That experimentally manipulated availability can affect subjects' responses is well demonstrated by the work of Kubovy (Kubovy, 1977, Kubovy & Psotka, 1976). When asked to report the "first digit that comes to mind", 28.4% of subjects chose 'seven' and only 2.2% chose 'one', whereas, when asked to report the "one-digit number that comes to mind", 18% chose 'one'. Similarly when asked for a number between 1,000 and 9,999, only 4.3% of subjects chose a number beginning with four, whereas, when the first four-digit number that 'comes to mind' was requested, 27.4% of subjects chose a number beginning with four. Subjects tend to respond with the 'available' number that is embedded in the task instructions. Although it is not suggested that subjects, performing a selection task, are asked to select the 'first card that comes to mind', it should be remembered that data, from both the selection task and conditional inference tasks, suggest that logical validity has little effect on responses to problems concerning the consequent of a conditional rule. Thus the selection task presents subjects with two consequent cards, whose relevance, or irrelevance, to the logical structure of the problem they are

apparently incapable of comprehending. The problem, for the subject, is thus very similar to being asked to select any card that 'comes to mind', and it is perhaps not surprising that his consequent selections are determined by the 'available' cues.

A further finding of Kubovy (1977) is of relevance to the null results of Experiment 7. Although embedding the number 'one' in the task instructions increased its frequency of choice by the subjects, explicit mention of 'one', as an example of a number between zero and nine, did not affect its frequency of choice. In fact, Kubovy & Psocketka found that explicit mention of 'seven' actually reduced its frequency of choice. Kubovy suggests that "the availability heuristic does not operate in a context in which the most available response is made to appear externally caused". A reasonable explanation of the results of Experiment 7 is thus that the associations present in the materials were, in effect, 'too available'. There was some suggestion in the data that, when a consequent value was associated with TA (TC in condition A and FC in condition B), a decision was made about this value, before the other consequent value. When both values were rejected, the associated value was rejected significantly more often before, than after, the nonassociated value. It was suggested in the discussion of Experiment 7 that the associated card thus was 'cued' but that the 'cue' was so obvious that the subject may well have ignored it, due to suspicion of the experimenter's motives. Given the Kubovy results, a more sophisticated form of this explanation may be put forward. The subjects may have been well aware that the available cue was a function of the experimental design (i.e. that it was 'externally caused') and avoided making the response cued by availability, not because they were suspicious of the experimenter's

motives, but simply because they are capable of avoiding the bias when they are made aware of it.

The lack of effect of 'too available' cues, in the Kubovy experiments and in Experiment 7, appears to represent an interaction between the type 1 and type 2 processes, proposed by Wason & Evans (1975). Many subjects in Experiment 7 spontaneously reported that they had noticed, and purposely ignored, associations in the cards, and it was often obvious that they were aware of the associations during performance of the task. In contrast, subjects in Experiment 6 made no spontaneous reports about the truth status of the rules.

SECTION 10.3 SOME APPLICATIONS OF THE 'AVAILABILITY' HEURISTIC TO THEMATIC REASONING STUDIES

When subjects are presented with logically structured problems having thematic content relating to their experience, a variety of available cues are likely to be embedded in this content.

The experiments discussed in Chapter 1.3 presented subjects with thematic syllogistic arguments and required them to evaluate both the truth status and the logical validity of the conclusion. Most of these studies reported that subjects' evaluation of the logical validity of a conclusion was affected by the perceived truth status of that conclusion, such that subjects tended to accept conclusions with which they agreed, and reject conclusions with which they disagreed. It was concluded in Chapter 1.3 that, in particular, the experiments of Feather (1964) and Kaufman & Goldstein (1967) provide reasonable evidence of the nonlogical effect of truth status on the evaluation of the logical validity of a syllogism's conclusion. If Tversky & Kahneman are correct, evaluation of the truth status of a conclusion is mediated by the availability of that conclusion. However, there is no reason to assume a sequential chain such that subjects evaluate a conclusion as 'valid' because it is 'true' and as 'true' because it is available.

The Kubovy experiments indicate that availability may mediate other responses, besides frequency assessment (which underlies truth status evaluations). An alternative explanation, of the correlation between subjects' evaluations of truth status and logical validity, is thus that both evaluations are directly mediated by availability. This interpretation suggests that subjects

do not have a motivation to accept statements that accord with their beliefs, as no cause and effect relationship is postulated between evaluation of truth status and evaluation of logical validity. Both are held to be determined by availability, which produces a pseudo-association between the two.

Some indication, that apparent truth status effects are not a resultant of subjects' motivations to accept statements consistent with their beliefs, is provided by research on the so-called 'knew it all along effect'. This paradigm is of particular interest as it represents another area of research in which truth status has been experimentally manipulated, although not by probability learning tasks but simply by the experimenter telling subjects whether particular statements are true or false. Subjects have been asked to predict the likelihood of various outcomes of past historical events (Fischhoff, 1975) and future historical events (Fischhoff & Beyth, 1975) or to predict the likelihood of alternative answers to general knowledge problems (Fischhoff, 1977, Wood, 1978). Subjects have been told the correct 'outcome', at various stages of the experiment, and the following results have typically been observed:

- 1) Given outcome knowledge, subjects rate that outcome as more likely than do subjects without outcome knowledge.
- 2) Subjects given outcome knowledge, after having made their likelihood estimates, tend to overestimate their original estimate of the correct outcome, when asked to recall it.
- 3) Subjects given outcome knowledge, overestimate the likelihood they would have assigned to that outcome, and overestimate the likelihood that hypothetical peer group subjects would have assigned to that outcome. Subjects without outcome knowledge attribute lower likelihood estimates (of the correct outcome) to hypothetical peer group subjects, than do subjects with outcome knowledge.

Various controls indicate that these results are not due to subjects attempting to present a 'good' impression of themselves or of hypothetical peers. For instance, Slovic & Fischhoff (1977) found that subjects, given the result of an experiment, considered replication to be more likely than did subjects asked the likelihood of replication, if that particular result was obtained (in the future).² Clearly, the perceived likelihood of replication (logically) should be the same whether the subject knows, or merely assumes, that a given result has been obtained.

The explanation most consistently advanced for the 'knew it all along effect' is that outcome knowledge raises the availability of that outcome at the cost of all other possible outcomes. (Experimentally inducing an increase in the availability of other outcomes reduces the bias). When asked to recall his original judgements, or asked to give estimates of the judgements of hypothetical peers not having outcome knowledge, the subject is biased by the increased availability of the actual outcome. Thus, although judgements are influenced by the truth status imposed by the experimenter, this judgemental bias is derived directly from the availability of the outcome, rather than from any intention on behalf of the subject to take truth status into account. This interpretation parallels the interpretation of apparent truth status effects on syllogism evaluation, as being mediated by the availability of the conclusion, rather than by any bias of the subject towards accepting the validity of conclusions because he believes them.

Experiments which have presented subjects with thematic selection tasks were discussed in Chapter 3.6. In certain of these experiments use of thematic content facilitated FC selection and it was concluded that such an effect is obtained only if the

subject is presented with a 'believable' problem that relates to his experience of 'real life'. For instance, subjects have experience of the necessity to stamp letters correctly and thus the context of the Johnson-Laird, Legrenzi & Legrenzi (1972) experiment did relate to their experience. In the case of those thematic experiments that have reported a facilitation of FC selection, the effect was explained in Chapter 3.6 as due to the FC card having been 'cued' by the content or context of the problem. This interpretation is easily transposed into the hypothesis that the effect is mediated by the availability of the FC card.

In the first part of the Legrenzi (1971) study, the falsifying (negative) instance was most salient to the subject's original derivation of the rule and thus was more available on the subsequent selection task. In the Johnson-Laird, Legrenzi & Legrenzi (1972) study, the availability of the TA-FC relation was enhanced by the context of the role playing exercise. Thus, in these studies, as in the 'false' conditions of the truth status experiments of Part 2, the availability of the FC card is enhanced. Such an effect will not be obtained, however, if thematic content is used that does not enhance the availability of FC. The drinks and foodstuffs content used by Manktelow & Evans (1978) produced no availability effects as wholly arbitrary pairings, of particular foods and drinks, were used.

If subjects cannot reason 'backwards' from a conditional rule, then consequent selections will be particularly susceptible to any subtle influence or bias and these selections may well be solely 'cued' by availability. Any experimental modification of the relative availability of consequent values will thus also modify the subjects' actual selections. There is no reason to believe that such experimental modification has any other effect.

SECTION 10.4 SOME APPLICATIONS OF THE 'AVAILABILITY' HEURISTIC TO
ABSTRACT REASONING STUDIES

On the basis of the 'cueing' effect of availability it would be expected that, if the subject is presented with a problem whose premisses employ particular terms, then he will exhibit a preference for conclusions employing those same terms. However, this possibility does not apply to syllogistic, and conditional inference, problems, as all possible conclusions contain terms that have appeared in the premisses. It is, of course, possible that the availability of the terms in the premisses accounts for what Revlis (1975a, 1975b) has referred to as a 'bias against nonpropositional conclusions'. That is, on syllogistic tasks, subjects may be biased against nonpropositional conclusions, simply because they do not contain any terms used in the premisses.

It has been pointed out in Section 10.3 that matching bias readily lends itself to interpretation in terms of availability, as it is simply a bias towards those terms that are available in the given rule. In Chapter 6, it was suggested that the concurrent mention of the antecedent and consequent values produced an (available) association between the two, which led to their selection on the selection task.

However, if subjects select the matching stimuli for this reason, then matching effects on P and Q should be related, as subjects are held to be focussing on the PQ case. The more a subject focusses upon this available case, the more likely he should be to select both P and Q. A correlation between selection of these cards would thus be expected. However, the analyses given in Chapter 8 (see Table 8.5) yield little support for this, except

in the case of the NN rule, although they do yield strong support for the existence of a correlation between \bar{P} and \bar{Q} selections.

It appears, from these results, that matching bias should be more properly regarded as a 'mismatching bias'. That is, if a subject's behaviour is influenced by the bias, it appears likely to lead to avoidance of selection of mismatching cards, but not necessarily to selection of both matching cards. This is consistent with the results of truth table experiments, in which, although the $\bar{P}Q$ and $P\bar{Q}$ cases are evaluated as 'irrelevant' more often than the PQ case, the bias is most noticeable in the large increase in 'irrelevant' responses to the $\bar{P}\bar{Q}$ case (e.g. Evans & Newstead, 1977).

On the basis of the above, it is reasonable to conclude that it is the terms that are available and not the connection between them. If matching selections are explained as due to the availability of the individual terms, rather than the availability of the association between them, then this is clearly similar to the attentional explanation of Evans (1975). However, it does not appear to be the case that subjects select those cards that they attend to, but rather that they fail to select those cards that they do not attend to. The subject will not necessarily select all available terms, but clearly will fail to select all nonavailable terms.

While on the subject of availability effects on abstract tasks, it is worth noting that availability provides a reasonable explanation of the Wason '2, 4, 6' experiments (Wason, 1960, 1968b, 1971). Wason (1971) explains the task used as follows: "subjects were told that the three numbers, 2, 4, 6, conformed to a simple relational rule, which they were to try to discover by generating successive triads of numbers. Each time they were told whether their

numbers conformed to the rule." They were warned to announce the rule "only when they were highly confident that they had discovered it." The rule was, in fact simply 'numbers increasing in order of magnitude'.

Typical results from this research may be summarised as follows:

- 1) Most subjects announced an (incorrect) rule prematurely on at least one occasion and, apparently, "they all appeared bewildered when told that their announcements were wrong."
- 2) That the given instance was an arithmetical progression strongly influenced subjects' early generation of triads and many subjects incorrectly announced a rule that depended on arithmetical progression.
- 3) Subjects' generation of triads prior to an announcement appeared to be consistent with that announcement. That is, subjects appeared to be obtaining verifying support for their supposed hypothesis, before announcing it as the rule. There was also some tendency for subjects who had announced an incorrect rule to generate immediately a further triad that was consistent with the rule that they had announced.

The possible effects of availability in these results are self-evident. The available arithmetic progression had a strong effect on responses and, once a subject had developed an hypothesis about the rule, the availability of this influenced his responses both before, and after, announcement of this hypothesis. A very good example of the strong effect of the availability of the arithmetical progression is shown in protocol 2 of Wason (1971).

This subject generated '4, 6, 8' and '6, 8, 10' and announced the rule 'arithmetical progression by two'. Subsequent to this, he generated three triads, each followed by an incorrect announcement of the rule, as follows: '8, 10, 12' ('progression of even integers'), '13, 15, 17' ('any three positive integers') and '3, 5, 7' ('any three numbers'). It will be noted that, although each of the three generated triads is consistent with the subsequent announcement, they are also consistent with the (original) hypothesis 'arithmetical progression by two'. There was no reason for this in relation to the hypotheses announced. The subjects' responses thus appear to be influenced by the progression made available by the original example, even though (on the basis of his announcements) he appears to have accepted that this progression is irrelevant to the rule. Wason & Evans (1975) suggest that the '2, 4, 6' results indicate that the generation of triads is influenced by different factors from those that influence the announcement of the rule. They suggest that generation of triads is a type 1 process, unconsciously influenced by an earlier hypothesis, whereas the rule announcement is a type 2 process which rationalises previous generations by formulating a rule which is consistent with them.

It should be noted that the influence of the available progression by two, in the above protocol, does not indicate a verification bias, as the subject's second, third and fourth announcements did not concern progression by two. There is thus no reason to believe that the first two triads generated prior to his first announcement ('arithmetical progression by two') were intentionally aimed at verification. The subject's responses were influenced by the available progression, irrespective of his subsequent announcements.

As in the case of selection tasks using the AA rule, the '2, 4, 6' experiments may produce the illusion of a verification bias, due to the influence of certain available features of the problem.

SECTION 10.5 FUNCTIONAL UTILITY OF AVAILABILITY

The availability heuristic has been shown to be a possible source of bias in a variety of experimental situations. However, it is unlikely that decisions would be mediated by availability if the heuristic had no functional utility.

In general, the heuristic clearly has utility with regard to its effect on frequency judgements, although the work of Chapman & Chapman shows that it can lead to error. However, in 'real life' judgements, the heuristic most probably leads to the best estimate possible, on the basis of the subject's experience, and the heuristic may, in fact, only prove wholly useless on artificially structured experimental tasks. In the light of this possibility, it is of interest to consider whether availability, although producing inappropriate bias on experimental reasoning tasks, may have functional utility in 'real life' reasoning.

When considering whether the availability heuristic aids reasoning in 'real life' situations, it may be inappropriate to ask whether it facilitates the drawing of logically valid conclusions. The subject will learn behaviours that are likely to produce correct decisions and the assessment of logical validity has no intrinsic utility in 'real life', except in as much as it can aid the subject in making such correct decisions. Further, there are drawbacks to assessment of logical validity as a method of assessing whether statements are correct. If arguments are evaluated as invalid, then this yields no information whatsoever as to whether the conclusion is right or wrong. Even if the argument is evaluated as valid, this yields only the information that the conclusion is true, if the premisses are true. An

appreciation that the argument is valid is thus only of use in those situations in which evaluation of the truth status of the premisses is easier than evaluation of the truth status of the conclusion.

Direct evaluation of conclusions, on the basis of experience, is thus an advantageous 'real life' behaviour that expedites acceptance of correct, and rejection of incorrect, conclusions. If an argument were advanced that concluded that 'London is the capital of France', it would be a highly inefficient method of evaluation to evaluate the logical validity of the argument. It is irrelevant whether the argument is invalid, or whether the argument is valid and the premisses incorrect. The essential difference between experimental and real life situations is that, in the latter, logical validity is not the salient dimension and it matters not at all whether an invalid conclusion, or inference, is accepted, as long as it is true, or whether a valid conclusion is rejected, as long as it is false. In certain cases, assessment of logical validity may be of potential use in deciding whether a statement is correct but, in most cases, there will be a more useful 'shortcut' method.

Thus, in 'real life' problem solving, available cues will have a beneficial effect in aiding correct decision making. When evaluating arguments, any available cues from the subject's experience will provide a shortcut method of deciding whether a conclusion is correct. This is likely to be so automatic that, even when a subject is specifically asked to judge the logical validity of a thematic argument, available cues will affect his responses. In fact, it is debatable whether all subjects share the experimenter's definition of 'logically valid'. In real life, the

phrase 'valid argument' may sometimes be taken to mean an argument with a true conclusion, and it may be unusual to apply the word 'valid' to an argument with a false conclusion.

The utility of availability in evaluating conditional statements may be that it leads to the accessing of relevant (frequent) instances. A logician will discount any number of verifying instances of a conditional rule as irrelevant, and accept one falsifying instance as proving the rule false. However, this is unlikely to be the procedure of the nonlogician, who is more likely to evaluate conditional statements in an inductive, rather than deductive, manner. The nonlogician's need is to produce sense and order in his environment,³ and the prime determinant of the perception of a conditional relationship will be the relative amount of verifying and falsifying instances (this is, in fact, the basis of decisions about contingent relationships - Jenkins & Ward, 1965, Smedslund, 1963, Ward & Jenkins, 1965). If there are a greater number of verifying, than falsifying cases, then the rule will prove a useful tool in decision making, even though some exceptions may exist. Thus, subjects will treat information that is readily available as having potential relevance to evaluation of a conditional rule, and information that is not readily available will tend to be perceived as not relevant (at best, it can only reveal an exception). This behaviour will have 'real life' utility, as available cues will lead to recall of, or attention to, verifying instances of statements that are (generally) true, and falsifying instances of statements that are (generally) false.

Availability can thus mediate correct evaluation of real life conditional statements. That it serves no useful purpose on the selection task, is due to the artificial nature of the task. Subjects are very unlikely to have encountered such a problem in real life and will thus have had no opportunity to learn relevant behaviours, or 'unlearn' irrelevant behaviours.

On the basis of the above arguments, subjects may have learned to use the availability heuristic, as it aids the making of a variety of 'real life' decisions. That the heuristic leads to error on experimental tasks may be interpreted, not as revealing the essential irrationality of the subject, but more as revealing the artificial nature of the experimental situation. A similar argument has been advanced by Navon (1978) with regard to 'conservatism' in Bayesian probability revision. He argues that, in experimental tasks, the given probabilities are reliable but that in 'real life' situations, there is an essential unreliability in any probabilistic estimate, and thus that 'conservatism' may be an adaptive heuristic that the subject has learned to utilise to offset the unreliability of 'real world' probabilities. He points out that "since that (probability assessment) system is adapted to everyday 'real world' judgement situations,it may fail when confronted with artificial tasks which lack those (real life) characteristics."

Thus observed bias on a range of experimental tasks may be interpreted as resulting from the application of behaviours that yield optimal performance in 'real life'. This gives rise to the question of whether other typically observed responses to reasoning tasks, not so far considered, reflect the application of such behaviours. Chapter 11 will explore the possibility that all the main preferred responses to abstract reasoning tasks are the result of learned behaviour patterns that facilitate 'real life' decision making.

FOOTNOTES

1. In the 'false' condition, given TC, FA was far more likely than TA, whereas in the 'true' condition, given TC, TA and FA were equally likely.

2. This result has implications for research on conditional statements. It suggests, for instance, that the contrapositive is not psychologically equivalent to the MT inference. For the AA rule, the MT inference is ' \bar{Q} is the case, therefore \bar{P} is the case', whereas the contrapositive inference is essentially 'if \bar{Q} were the case, then \bar{P} would be the case'. The Slovik & Fischhoff results suggest that there is a psychological distinction between arguing from the knowledge of \bar{Q} and arguing from the assumption of \bar{Q} .

3. Some possible empirical support for the idea that subjects have a preference for general laws about their environment is the preference for universal, rather than particular, statements reported by Revlis (1974), Revlis & Hayes (1972) and Revlis, Lipkin & Hayes (1971).

CHAPTER 11

ABSTRACT REASONING TASKS: SOME POSSIBLE EFFECTS OF PRIOR EXPERIENCE

SECTION 11.1	Nonlogical Factors	318
SECTION 11.2	Logical Factors	322
SECTION 11.2.1	Behaviour Concerning the Antecedent of Conditional Rules	325
SECTION 11.2.2	Conditional Inference Tasks	327
SECTION 11.2.3	Disjunctive Inference Tasks	329
SECTION 11.2.4	Syllogistic Reasoning	331
SECTION 11.3	Conclusions	335

SECTION 11.1 NONLOGICAL FACTORS

Several nonlogical factors have been discussed in Part 1 and these are summarised below:

- 1) Matching bias. This bias is known to affect conditional truth table tasks (see Chapter 2.2) and the selection task (see Chapter 3.4).
- 2) Atmosphere bias. Although there are alternative explanations this has been argued to affect responses to syllogistic problems (see Chapter 1.2 and 1.6). This bias has also been argued to be preferable to alternative explanations on the grounds that such explanations are inapplicable to selection task behaviour (see Chapter 4.2). An explanation in terms of a 'feature matching' bias thus extends across paradigms.
- 3) Conclusion bias. On conditional inference tasks, subjects have a preference for affirming MT, AC and DA inferences having negative conclusions (see Chapter 2.3.1) Related to this is the preference observed in Experiments 1 and 2 for affirmation of contrapositive, converse and inverse arguments having negative consequents (see Chapter 5).
- 4) The effect of truth status of conclusion on evaluation of the logical validity of that conclusion (see Chapter 1.3).

That effects of truth status of conclusion, and effects of matching bias, are the resultant of a learned behaviour (responding on the basis of availability) that has 'real life' utility, has been argued in Chapter 10. It thus only remains to consider the possible utility of atmosphere and conclusion bias.

In considering the functional utility of atmosphere bias,

it is worth noting the following two features of determinate premiss pairs:

1) When both premisses are affirmative, the valid conclusion is always affirmative, otherwise it is always negative.

2) When both premisses are universal, the conclusion is always universal in the case of figure 2 and figure 1/4 premiss pairs having 'forwards' conclusions, (although never universal in the case of figure 3 or figure 1/4 premiss pairs having 'backwards' conclusions) otherwise (for all figures) the valid conclusion is always particular.

The above features of determinate premiss pairs have a striking resemblance to the principles of atmosphere. If it was known that a premiss pair was determinate then matching the mood of the conclusion to the atmosphere of the premisses, and allowing negative and particular atmospheres to dominate when the atmospheres were 'mixed', would almost always produce the correct mood of conclusion. It would only fail in the case of two universal premisses in figure 3, or figure 1/4, allowing only a 'backwards' conclusion, which are probably rarely, if ever, used. Thus responding in accordance with atmosphere is a very useful behaviour pattern, when solving determinate premiss pairs, or evaluating valid syllogisms.

However, the utility of atmosphere bias depends upon the proportion of determinate arguments that the subject encounters in 'real life'. If the subject primarily encounters indeterminate arguments, then atmosphere responses have little utility and it is impossible to tell how the subject acquires them. However, it is quite possible that subjects primarily encounter determinate arguments in 'real life', and thus that they have learned to respond on the basis of atmosphere, which almost always produces correct

behaviour. In this case, atmosphere bias may be seen as a learned behaviour pattern which has functional utility in 'real life', but which is inappropriate to those 'artificial' experimental situations in which the subject is presented with primarily indeterminate arguments.

In Chapter 5, a preference was observed for affirmation of contrapositive, converse and inverse conclusions having negative consequents and affirmative antecedents. This was interpreted as an effect of 'caution' such that subjects are biased against any specific evaluation (i.e. 'true' or 'false') of statements having broad (negative) referents, and in favour of making nonspecific (negative) conditional predictions. The latter bias led to a greater tendency to affirm conclusions having negative, and to deny conclusions having affirmative, consequents. The results regarding affirmation of the contrapositive, converse and inverse paralleled results concerning affirmation of the MT, AC and DA inferences (e.g. Evans, 1977a). It is possible that all the effects observed in Experiments 1 and 2 would be paralleled by a relevant study of the MT, AC and DA inferences and their opposites. It can thus be argued that observed nonlogical biases on such tasks have 'real life' utility, in raising the likelihood of subjects making correct decisions. Unless the consequent dimension is binary, a negative statement, or conditional prediction, is more likely to be less specific than an affirmative statement or prediction. That subjects may have learned to avoid making specific predictions about uncertain things, may now be seen as a viewpoint consistent with the above suggestions about the possible utility of other 'nonlogical' biases.

However, this behaviour is likely to be only observed in conditions of maximum uncertainty. The thematic content of most

conditional statements used in 'real life' will allow the likelihood of inferences to be directly evaluated on the basis of their availability (e.g. Bucci, 1978). Only when the subject has no relevant available experience, need he rely solely on the cautious adoption of statements making vague, nonspecific predictions.

SECTION 11.2 LOGICAL FACTORS

There are a variety of typically observed responses to reasoning tasks that accord with logical validity. However, that they accord with logical validity, is not necessarily, in itself, an adequate explanation of these responses. However, most workers in the area accept (either explicitly or implicitly) that behaviour in accordance with logical validity is the result of an appreciation of logical structure. Differences of opinion tend to concern those responses not in accordance with validity. Indeed, most workers in the area appear to believe that it is only 'errors' that require explanation (e.g. Chapman & Chapman, 1959, Dickstein, 1978a, 1978b, Johnson-Laird, 1975, Revlis, 1975a, 1975b, Roberge, 1971b, 1974, 1976a, 1976b, 1978, Taplin, 1971, Taplin & Staudenmayer, 1973, Wason, 1968a, 1969a, Wason & Johnson-Laird, 1969b, 1970).

There thus appears to be a general assumption that subjects do have some intrinsic ability to appreciate the logical structure of a problem ('valid' responses being tacitly regarded as evidence for this), and interest is thus centered on 'invalid' responses. However, there is an alternative point of view. If no such assumption is made about subjects' abilities to reason in a logical manner, then all common responses to reasoning tasks are of equal interest. For instance, if subjects apply learned behaviours to abstract reasoning tasks, and these behaviours facilitate correct decision making in 'real life', then it would not be unexpected that some of these behaviours would produce responses in accordance with logical validity. These behaviours would be of equal interest, whether or not they accorded with validity.

There are, in fact, (at least) three possible explanations of preferred responses that accord with logical validity:

- 1) As suggested above, the subject may be applying some learned behaviour that happens to produce an identical response to one based on an analysis of the structure of the problem. This behaviour may relate to parts of the structure, superficial features of the problem, or to the learned manipulation of syntactic expressions.
- 2) The subject may produce a logically correct answer because he has been trained to apply the sequence of behaviours necessary to analyse the problem in a formal way. That this is possible, is clearly attested to by the fact that many people successfully complete logic courses.
- 3) The subject's basic cognitive structure may be such that he can appreciate the logical structure of the problem without any specific prior experience.

Alternative 2 is an unlikely explanation, as subjects with training in formal logic are specifically excluded from most reasoning task experiments.

As has been noted, most workers, by concentrating on the explanation of 'errors', implicitly assume that responses in accordance with logical validity are explainable as under (3) above. The question to be considered is whether there is any evidence for this position, or whether there is not at least equally good indication that explanation (1) above is a viable alternative.

This question will be pursued with regard to four main categories of logically valid responses:

- 1) Behaviour concerning the antecedent of conditional statements. Of all tasks, such behaviour is most consistent in being frequently in accordance with logical validity, and would thus be expected to yield the best indication that subjects are capable of appreciating logical structure.
- 2) Certain logical behaviours observed on conditional inference tasks, with reference to the findings of Chapter 5
- 3) Disjunctive inference tasks.
- 4) Syllogistic reasoning tasks.

SECTION 11.2.1 BEHAVIOUR CONCERNING THE ANTECEDENT
OF CONDITIONAL RULES

Subjects make few errors in affirming the MP inference or in denying its opposite (see Chapter 2.3, particularly with reference to Evans, 1977a, and Taplin, 1971), and show some tendency to avoid affirmation of DA. Similarly, on the selection task, subjects show a strong tendency to select TA and to avoid selection of FA (see Chapter 3). On truth table tasks, subjects almost invariably report TT to be a verifying case and there is a very strong tendency to report TF to be a falsifying case (see Chapter 2.2). Thus, on the basis of selection task responses, subjects appear to appreciate the need to test modus ponens and, on the basis of behaviour on both inference and truth table tasks, they are usually capable of evaluating the outcome of such a test.

All the above behaviours accord with logical validity. However, that they reflect the fact that subjects' thinking processes are isomorphic with logical structure, is by no means demonstrated. To say that the MP inference follows logically from a statement of the form 'if P then Q', may well cloud the issue, as it can equally well be seen as a verbal equivalent of the conditional. The MT inference on the AA rule can be verbally expressed as (the contrapositive) 'if not Q then not P', but the equivalent verbal expression of MP, as was shown in Chapter 5, is the conditional (see Table 5.1). It would thus be impossible to learn to use conditional statements without acquiring both the ability to draw MP and the ability to evaluate the TT and TF cases. A subject's abilities, concerning the antecedent of a conditional statement, are thus inextricably conjoined with his acquisition of the verbal use of

such statements, and need not be based in any formal structuring of his thinking processes.

SECTION 11.2.2 CONDITIONAL INFERENCE TASKS

There is very little evidence for logical performance on conditional inference tasks except for those behaviours considered in Section 11.2.1. In Experiments 1 and 2 (see Chapter 5) no evidence was found for the hypothesis that subjects affirm the contrapositive, or deny its opposite, due to an appreciation of its logical status. Although this hypothesis has not been tested concerning affirmation of the MT inference, or denial of its opposite, there is no reason to believe that subjects would appreciate the logical validity of MT any more than the contrapositive.

Although frequencies of AC affirmation can be manipulated by the content or context of the problem, (see Chapter 2.3.2) the changes most probably stem from the application of the subject's available past experience, rather than from differential acceptance of the validity of the inference. That AC and DA affirmation on abstract tasks is not mediated by a perception of the major premiss as implying equivalence, is indicated by the differential frequencies with which these inferences are affirmed and by the fact that AC and DA responses do not 'move together'. That is, those rule forms that yield the highest frequencies of AC affirmation are not necessarily those that yield the highest frequencies of DA affirmation.

The only apparently 'logical' behaviour on inferences concerning the consequent of a conditional rule is that subjects usually avoid the evaluation of inferences as 'false' and the evaluation of 'opposite' inferences as 'true'. That is, if inferences are not affirmed, or if their 'opposites' are not denied, then they tend to be evaluated as 'indeterminate'. However, if the

subject takes account of the minor premiss, he need only be attending to the superficial structure of the problems. The major premiss will either have an odd or even number of negatives and, taking account of both minor premiss and conclusion, the MP, DA, MT and AC inferences will match the major premiss in this respect, whereas there will be a mismatch in the case of 'opposite' inferences. The inference forms used in Experiments 1 and 2 may be evaluated in a similar way. Instead of counting the negatives in the minor premiss and conclusion, the subject would simply need to count the number of negatives in the conditional statements offered and compare this count with the key statement. Subjects' experience of negatives may simply lead them to assume that statements with an even number of negatives cannot mean the same as statements with an odd number of negatives, and vice versa.

Thus subjects need perform no more than a simple negative count to avoid denying inferences or to avoid affirming 'opposite' inferences. In the case of the AC inference, the comparison is made particularly easy as the negatives are attached to the same terms as they are in the major premiss. Admittedly, subjects do not always avoid the denial of inferences. For instance, in Experiment 2, the contrapositive was denied on over 10% of occasions. However, subjects would be expected to miscount occasionally.

SECTION 11.2.3 DISJUNCTIVE INFERENCE TASKS

There is very little behaviour in accordance with logical validity on disjunctive problems involving rules with negative components and there appears to be some tendency to evaluate all rule forms as if they were AA disjunctions. In the case of problems involving AA disjunctions, error rates are very low (see Chapter 2.4). For abstract AN and NA rules, the data of Roberge (1978) show that across all problems (which were all determinate) subjects drew the valid conclusion on 38% of occasions, drew the opposite of the valid conclusion on 35% of occasions and rated the argument as 'indeterminate' on 26% of occasions. There appears to be no effect of logical validity on this response distribution, although 'error' rates are much lower on exclusive NN rules.

Typical responses on disjunctive tasks readily lend themselves to an explanation in terms of the subject's experience. It would surely be impossible for a subject to learn the use of the word 'or' without, at the same time, acquiring the ability to handle an AA disjunction. Subjects' experience of the use of disjunctives cannot be questioned. However, it is unlikely that subjects have encountered NA or AN disjunctive statements in 'real life', except on the most infrequent occasions. That subjects do not have an appreciation of the concept of disjunction, other than that acquired from the learned use of the word 'or', is indicated by the fact that high frequencies of correct responses are only observed on the (AA) rule form with which they are familiar.¹ Presented with a rule form that does not relate to their 'real life' experience, subjects' frequencies of correct responses drops to little above chance level. If subjects had an intrinsic concept

of disjunction, this concept would be expected to generalise to rule forms with which they were unfamiliar.

Finally, it should be noted that, in 'real life', the content of the disjunction will aid inferences from it. The data of Roberge (1977) suggest that semantic factors are at least as important as the structure.

Valid responses to syllogistic arguments are difficult to analyse as there are a variety of syllogisms, each having its own structure. That behaviour in accordance with logical validity is observed is indicated by two findings. Firstly, that atmosphere bias appears strongest when it accords with the logically valid conclusion and, secondly, that subjects often affirm the valid nonpropositional conclusion to indeterminate premiss pairs (this tendency was particularly noticeable in the data of Johnson-Laird & Steedman, 1978). However, certain syllogisms rarely produce valid responses.

Reference to typical data from syllogistic studies reveals that those problems yielding high frequencies of valid nonpropositional conclusions are those having EE, EO, OE and OO; and II, OI and IO premiss pairs. The reader will note that these constitute the 'category 1' and 'category 2' premiss pairs discussed by Dickstein (1978b) (see Chapter 1.2). Frequencies of valid nonpropositional conclusions are much lower on other premiss pairs, and it is thus these two categories that yield the best evidence of subjects' ability to derive a valid conclusion from indeterminate premiss pairs. In fact it is also these premiss pairs that underlie the finding that more atmosphere responses are made when they accord with validity. The latter frequency is not particularly higher than atmosphere responses on indeterminate premiss pairs other than those in category 1 and 2. Behaviour on these problems thus constitutes the main evidence for an effect of logical validity on responses to syllogisms.

Dickstein (1978b) attributes a particular error to each of these two categories, to explain why not all subjects accept a nonpropositional conclusion. He presents an interesting analysis showing within subject consistency in performance within, but not between, the two categories. On the basis of this analysis, Dickstein concluded that a subject susceptible to one type of error would tend to make this error on all problems within one category, but would not be more, or less, likely, than other subjects, to be susceptible to the other type of error. Dickstein interpreted this as evidence for the operation of two distinct errors. However, it is possible that Dickstein was viewing his data from the wrong perspective.

As the premiss pairs in these two categories yield the highest frequencies of logically valid nonpropositional conclusions, then surely the most interesting question is why subjects avoid errors, rather than make errors on these problems. Given the simple statement, 'no valid conclusion may be derived from two negative premisses', all category 1 premiss pairs can be validly solved. Similarly, all category 2 premiss pairs can be validly solved, given the simple statement, 'no valid conclusion may be derived from two particular premisses'. Thus no complex analysis of these problems is required for their solution. This does not necessarily imply that such complex analysis is not performed, but it at least suggests that subjects may have acquired one or both of the necessary rules as a 'shortcut' method. Two points may be advanced in favour of the latter hypothesis:

- 1) The hypothesis explains why more valid nonpropositional conclusions are affirmed on these, than on other, indeterminate, premiss pairs. However, if it were held that these premiss pairs were 'solved' by logical analysis,

this would not explain why other indeterminate premiss pairs are not frequently solved, unless it was argued that category 1 and 2 arguments are easier. It is not apparent why this should be the case.

- 2) If acceptance of valid nonpropositional conclusions is a function of some intrinsic logical ability, then performance should correlate across, (as well as within) categories.

The two rules given above may be combined with the atmosphere principles to produce the following two simple rules which will yield the correct solution to 28 of the 36^2 possible premiss pairs.

- A) If no premiss is negative, conclusion affirmative; if one premiss negative, conclusion negative; if two premisses negative, no valid conclusion.
- B) If no premiss is particular, conclusion universal; if one premiss particular, conclusion particular; if two premisses particular, no valid conclusion.

It could be objected that the above rules do not define the direction of conclusion, but this is usually cued by the premisses. For instance, for determinate premiss pairs, if the word 'some' is attached to a conclusion term in one of the premisses, then it is attached to the same term in the conclusion.

The important point about the above is that, if subjects use such rules, then no formal logical analysis is required. That subjects' behaviour may possibly be a function of these rules is suggested by the fact that 'error' rates are typically very high on those 8 problems for which the above rules yield an invalid solution.

Rule A above appears to have a stronger effect on responses than Rule B. The dominance of negative atmosphere when one premiss is negative is more marked than the dominance of particular atmosphere when one premiss is particular, and the (invalid) acceptance of a propositional conclusion to two negative premisses is typically less frequent, than the acceptance of a propositional conclusion to two particular premisses. The differences in effect are thus consistent across both the atmosphere and the 'added' parts of Rules A and B. However, the important point is that, in all cases, 'error' rates are lower on those problems for which the above rules yield a valid solution. Thus a simple extension of the atmosphere principles can explain behaviour on those (category 1 and category 2) premiss pairs that appear to yield the best evidence for subjects' 'logical' behaviour. The revised atmosphere theory given in Rules A and B above is, of course, very speculative. It does explain a considerable amount of the data but it is clearly debatable whether subjects would have had the opportunity of acquiring these rules. However, if subjects respond on the basis of these rules, they need attend only to the superficial features of the problem and, given the possibility that these rules have been acquired, data from syllogistic reasoning tasks cannot be taken as definitive evidence that subjects attend to, and perform operations upon, the formal structure of the problems.

SECTION 11.3 CONCLUSIONS

It is not claimed that the issues raised in this, and the previous, chapter, have conclusively demonstrated that subjects do not have an intrinsic logical competence. The intention has been, however, to demonstrate that valid responses are not necessarily a function of such competence and that alternative explanations are possible.

The alternative explanation presented here is that subjects' responses on abstract tasks are a resultant of behaviours that they have learned to apply to 'real life' problems. It is suggested that a strength of this interpretation is that it sees no essential difference between those behaviours that accord with logical validity and those that do not. Both are held to reflect behaviour patterns that have functional utility in 'real life'. This position essentially needs to demonstrate the following two things:

- 1) That 'nonlogical' behaviours can have utility in real life problem solving.
- 2) That 'logical' behaviours are not necessarily mediated by a logical analysis of the structure of the problem and that they may be a resultant of aspects of the problem that the subject has learned to respond to.

It is considered that there is reasonable indication that 'nonlogical' behaviours may have utility in 'real life' (for instance in direct evaluation of conclusions) and that the subject has acquired them prior to the experimental situation (see Chapters 10 and 11.1). The position regarding (2) above is more complicated. Valid behaviours concerning the antecedent of conditional rules, and

concerning AA disjunctive rules, cannot be taken as evidence of subjects' appreciation of logical structure, as these behaviours must be acquired at the same time as the subject acquires the use of the relevant language. However he need acquire nothing other than the use of the language. That subjects do not acquire a concept of disjunction when learning the use of the word 'or' is indicated by the high error rate on negated rules. The learning is stimulus-specific and does not generalise. There is little other evidence for logical appreciation on propositional reasoning tasks. However, the evidence for logical competence on syllogistic reasoning tasks cannot be dismissed lightly and the suggestions advanced in Chapter 11.6 are very speculative.

The essential point is that it is quite possible that the labels 'logical' and 'nonlogical' reflect an arbitrary distinction made by the experimenter on the basis of an irrelevant dimension (logical validity). For instance, if subjects have acquired the two simple rules suggested in 11.6, then the same form of behaviour 'looks like' a nonlogical bias on some occasions and like a 'logical' response on others. The rules would have utility in real life because they yield correct behaviour on most problems, but they do not relate to logical structure.

It cannot be denied that the subjects' experience has a strong effect upon responses. There are two areas of clear cut evidence for this:

- 1) Use of meaningful thematic content that relates to a subject's experience affects his responses. Experiments 4, 5 and 6 have demonstrated that even a short period of prior experience can affect responses.

- 2) High frequencies of logically valid responses are observed on those problems whose solution depends on the subjects' learned use of the syntactic structure (see 11.3 and 11.5).

That some 'logical' and 'nonlogical' behaviours are a function of the subjects' experience is thus not in question. Further, there is no conclusive evidence that any particular behaviour is not a function of the subjects' experience. However, it is not simply the point here to argue that logical behaviours are learned. The point is that there is a distinction between the term 'learning' and the term 'understanding'. It is argued that subjects learn behaviours and that some of these behaviours happen to accord with logical validity but that there is no reason why acquisition of a response that accords with validity need be taken as having induced a state of appreciation of that validity. At least in the case of propositional reasoning, those valid responses that are frequently observed are notably related to a basic learning of the use of language (i.e. MP affirmation or evaluation of the TT case). There is no reason to believe that the subject acquires anything other than this usage.

FOOTNOTES

1. Error rates are low on exclusive rules of the NN form but this is consistent with the argument. Response to an exclusive NN rule as though it is an AA rule does not lead to error.

2. There are only 36 possible premiss pairs, when no conclusion is presented for evaluation (as their order is irrelevant to the argument): 10 each in figures 2 and 3 and 16 in figure 1/4.

CHAPTER 12

A REEVALUATION OF THE POSITION OF EVANS (1972a, 1977b)

SECTION 12.1	Introduction	340
SECTION 12.2	A fit of the Data of Experiment 6 to the Model of Evans (1977b)	343
SECTION 12.3	Further Examination of the Underlying Assumptions of the Model	365
SECTION 12.4	The Use of the Four Conditional Rules in the Disambiguation of 'Logical' and 'Nonlogical' Factors	368
SECTION 12.5	Introspection and the Stochastic Nature of Responses	373
SECTION 12.6	The Distinction Between Interpretational and Operational Factors	378
SECTION 12.7	Conclusions	383

During Part 1, considerable emphasis was placed upon the distinction between apparent 'logical' and 'nonlogical' effects on reasoning task responses and, at the end of that review, it was concluded that the Evans (1972a, 1977b) position was the best published interpretation of reasoning task data. In succeeding chapters, however, various experimental findings and theoretical issues have been discussed which have some bearing upon this conclusion. It is thus necessary to reevaluate the Evans position in the light of these developments.

The following three central tenets may be identified in the Evans position:

- 1) That responses are not a simple function of interpretation, as 'operational' factors also influence the data and must be taken into account. This tenet is crucial to all other theories that Evans has proposed.
- 2) That the processes underlying responses are not necessarily available to introspection. Thus protocol data, although possibly of interest in itself, cannot reveal the basis of responses to the task. Evans (1980a) cites Byrne (1977) as pointing out that verbal reports are a form of data to be explained in their own right, and not a means of explaining some other form of behaviour. Thus Evans is not without support for his position, as the Nisbett & Wilson (1977) paper also shows.
- 3) That processes underlying reasoning task responses are stochastic (rather than deterministic) within individual subjects. Thus differences in responses between subjects

are not seen as due to some subjects being in a different state of insight, or adopting a different (conscious) strategy, to other subjects. To clarify this position, suppose that 70% of subjects make a particular response. If behaviour is stochastic within subjects, then all subjects have a probability of making, or not making, this response. The observed result may thus be taken as an estimate of subjects' mean probability of making the response.

A detailed exposition of these three tenets, and of their implications for other theories of reasoning, is given by Evans (1980b).

Before consideration of these general tenets, it is of interest to consider the more specific extension of the position to the selection task in the form of the mathematical model of Evans (1977b). The development of this model entailed the adoption of further, subsidiary, assumptions, to those given above, and certain of the experimental findings of Part 2 bear a direct relation to these assumptions.

There are two ways to evaluate a mathematical model - to attempt to fit the model to a set of data, or to evaluate its assumptions on the basis of theoretical considerations or experimental findings. Both of these approaches will be adopted here, as one specific purpose of utilising all four conditional rules, in Experiment 6, was to provide a set of data on which the model could be tested.

In the following sections, the Evans position will thus be evaluated in three ways. Firstly, the results of Experiment 6 will be used as a direct test of the model's ability to fit

selection task data and, secondly, the assumptions of the model will be evaluated, both on the basis of the results of this fit. and on the basis of relevant experimental findings of Part 2. Finally, the three central tenets of the Evans position, given above, will be evaluated on a more general basis.

SECTION 12.2 A FIT OF THE DATA OF EXPERIMENT 6 TO THE MODEL
OF EVANS (1977b)

On the basis of the theory that reasoning responses are a function of both logical (interpretational) and nonlogical (operational) factors, and that such responses are probabilistic within subjects, Evans (1977b) proposed that such responses could be modelled in mathematical terms. He suggested a formulation which, although theoretically a general model, was related specifically to the selection task. He proposed that selection probabilities for each card are determined by a weighted addition of interpretational (I) and response bias (R) tendencies. Formally, the probability of a response, $p(r)$, is given by:

$$p(r) = \alpha I + (1 - \alpha) R$$

where $0 < \alpha < 1$

$$0 < I < 1$$

and $0 < R < 1$

The I values refer to the subjects' "interpretation of the sentences comprising the logical premisses". Thus, for instance, as subjects apparently 'interpret' conditionals as requiring the test of TA, I values would be expected to be higher (i.e. be a stronger influence towards selection) in the case of TA than in the case of FA. Similarly, R values would be expected to be higher in the case of 'matching' cards.

The weighting factor, α , determines the relative effect of the I and R tendencies. The lower the α value, the more dominant the influence of nonlogical factors. Evans (1977b) proposes that α (values) are "a function of the situation (e.g. presence of realistic or abstract materials) and certain gross logical

distinctions (in this application, the direction of reasoning between antecedent and consequent)". The latter is a reference to the apparent greater effect of nonlogical influences on consequent selection.

The selection probability of each card may be represented in terms of a particular expression of the general model (i.e. including the specific I and R terms for that card). An estimate of this expression can then be obtained from the selection frequency of this card observed in a sample of subjects. It should be noted that, as Evans points out, this procedure implicitly assumes that all subjects have the same parameter values. However, in an homogeneous sample, little variation in individual parameter values would be expected. (This point will be returned to in section 12.3.)

The approach is of little value when applied only to affirmative rules as the true cases of the antecedent and consequent (TA and TC, respectively) always match and the false cases (FA and FC) always mismatch. Interpretation and matching (response) bias are thus confused. However, the two factors can be distinguished by use of all four rules formed by systematic negation of the antecedent and consequent components. Thus, for instance, TA matches in the case of 'If P then Q' but mismatches in the case of the rule 'If not P then Q'. It was by this method that Evans & Lynch (1973) demonstrated that the tendency to select P and Q on the affirmative rule is a function of matching and not a function of a (verifying) tendency to select TA and TC.

The results of Evans & Lynch (1973) appeared to suggest that more attention is directed to matching on consequent selections than on antecedent selections and thus Evans (1977b) suggested that there should be two values of α (the weighting factor); α_a and α_c .

I values will differ for the different logical values of the cards and thus separate I values must be assigned to true and false instances of the antecedent and consequent (ITA, IFA, ITC and IFC). As the only known response bias on the selection task was matching bias, Evans proposed that R should take just two values: one when the card matches (RM) and another when the card mismatches ($\bar{R}\bar{M}$).

The results of an experiment, using all four forms of conditional rule, yield sixteen data cells (selection of four cards on four rules). From these, two estimates of each of eight expressions derived from the model may be obtained. The nature and source of these expressions are shown in Table 12.1.

It is apparent from Table 12.1 that, although numerical values can be assigned to each of the eight expressions, estimates of individual parameter values cannot be obtained. However, Evans (1977b) adopted prior assumptions of certain parameter values in order to test whether consistent estimates of the remaining parameters could be obtained. The basis of these assumptions is best explained by quoting from the paper:

"Now, in the limiting case of realistic materials where α approaches 1, behaviour approaches that of being perfectly logical on this task (assuming a truth table for material implication). Thus it is assumed that the I parameter has the value 1 for correct responses (p and \bar{q}) and 0 for incorrect responses (\bar{p} and q)".

Specifically, then, Evans assumed that $ITA = IFC = 1$, and that $IFA = ITC = 0$. Inspection of Table 12.1 shows that these assumptions greatly simplify the eight expressions, the αI values all reducing to either α or 0. Evans (1977b) fitted the model, with these assumptions, to the data of Evans & Lynch (1973) and produced impressively consistent estimates of the remaining four parameters. The manner in which these were derived is shown in Table 12.2, which is a reprint of the relevant table in the Evans (1977b) paper.

Table 12.1

Source of the Eight Fundamental Expressions Derivable from the Model when Fitted to Selection Task Data

<u>INTERPRETATION AND MATCHING COMBINATION</u>		<u>EXPRESSION</u>	<u>SELECTION FREQUENCIES FROM WHICH THE TWO ESTIMATES ARE DERIVED</u>
TA,	MATCH	$\alpha a \text{ ITA} + (1 - \alpha a) \text{ RM}$	P selections on (i) 'If P then Q' (ii) 'If P then not Q'
FA,	MATCH	$\alpha a \text{ IFA} + (1 - \alpha a) \text{ RM}$	P selections on (i) 'If not P then Q' (ii) 'If not P then not Q'
TA,	MISMATCH	$\alpha a \text{ ITA} + (1 - \alpha a) \text{ RM}$	\bar{P} selections on (i) 'If not P then Q' (ii) 'If not P then not Q'
FA,	MISMATCH	$\alpha a \text{ IFA} + (1 - \alpha a) \text{ RM}$	\bar{P} selections on (i) 'If P then Q' (ii) 'If P then not Q'
FC,	MATCH	$\alpha c \text{ IFC} + (1 - \alpha c) \text{ RM}$	Q selections on (i) 'If P then not Q' (ii) 'If not P then not Q'
TC,	MATCH	$\alpha c \text{ ITC} + (1 - \alpha c) \text{ RM}$	Q selections on (i) 'If P then Q' (ii) 'If not P then Q'
FC,	MISMATCH	$\alpha c \text{ IFC} + (1 - \alpha c) \text{ RM}$	\bar{Q} selections on (i) 'If P then Q' (ii) 'If not P then Q'
TC,	MISMATCH	$\alpha c \text{ ITC} + (1 - \alpha c) \text{ RM}$	\bar{Q} selections on (i) 'If P then not Q' (ii) 'If not P then not Q'

Table 12.2

Table III of Evans (1977b).

Predicted equalities from provisional model based on probability estimated from the data of
Evans & Lynch (1973) (N = 24)

A. Probability Estimates

<u>Model</u>	<u>ESTIMATE 1</u>	<u>ESTIMATE 2</u>	<u>MEAN ESTIMATE</u>	<u>STANDARD ERROR</u>
a) $\alpha_a + (1 - \alpha_a) \cdot RM$	0.875	0.917	0.896	0.044
b) $(1 - \alpha_a) \cdot RM$	0.292	0.458	0.375	0.070
c) $\alpha_a + (1 - \alpha_a) \cdot R\bar{M}$	0.583	0.542	0.563	0.072
d) $(1 - \alpha_a) \cdot R\bar{M}$	0.083	0.042	0.063	0.035
e) $\alpha_c + (1 - \alpha_c) \cdot RM$	0.583	0.750	0.667	0.068
f) $(1 - \alpha_c) \cdot RM$	0.500	0.583	0.542	0.072
g) $\alpha_c + (1 - \alpha_c) \cdot R\bar{M}$	0.333	0.417	0.375	0.070
h) $(1 - \alpha_c) \cdot R\bar{M}$	0.083	0.292	0.188	0.056

Table 12.2 (Continued)

B. Predicted equalities of model

<u>Parameter</u>	<u>Estimate 1</u>	<u>S.E.</u>	<u>Estimate 2</u>	<u>S.E.</u>	<u>Diff.</u>	<u>S.E.</u>	<u>Z</u>
α_a	(a) - (b) 0.521	0.094	(c) - (d) 0.500	0.080	0.021	0.123	0.17
α_c	(e) - (f) 0.125	0.099	(g) - (h) 0.187	0.090	0.062	0.134	0.46
RM *	(b)/(1 - α_a) 0.765	0.173	(f)/(1 - α_c) 0.642	0.099	0.123	0.199	0.62
$\bar{R}\bar{M}$ *	(d)/(1 - α_a) 0.129	0.073	(h)/(1 - α_c) 0.223	0.068	0.094	0.098	0.96

* Using best estimates of α_a and α_c

C. Best estimates of parameters

<u>Parameter</u>	<u>Estimate</u>	<u>Standard Error</u>
α_a	0.511	0.062
α_c	0.156	0.067
RM	0.704	0.100
$\bar{R}\bar{M}$	0.176	0.049

This table is best explained by the following quote from Evans (1977b): "It will be seen, from Table IIIA (see Table 12.2) that estimates of the parameter value α_a may be obtained in two independent ways from the data: by the subtraction (a) - (b) and the subtraction (c) - (d). Similarly, two estimates of α_c are obtained. In both cases the two estimates are impressively close, and the difference between them far from significant. By calculating best estimates for α_a and α_c (see Table IIIC) it is also possible to calculate two independent estimates for each of the Parameters RM and $\bar{R}\bar{M}$. The predicted equalities here depend on ratios and also fail to differ significantly. The relatively large standard errors of these differences, however, suggest that the sample size is insufficiently large for a powerful test of these equalities."

Having explained the method, the data of Experiment 6 may now be fitted to the model. The 'true' and 'false' conditions of Experiment 6 used different subjects and thus there are, in fact, two sets of data which may be fitted. The estimates of the eight expressions shown in Table 12.1 are shown in Table 12.3A, for the 'true' condition, and in Table 12.3B, for the 'false' condition. Table 12.3C shows a comparison of the estimates between the conditions.

It is apparent from Table 12.3 that, for each condition, the two estimates of each expression do not differ significantly from each other. However, Table 12.3C shows that all but one of the expressions differ significantly between the conditions. In the case of the $\alpha_a IFA + (1 - \alpha_a) \bar{R}\bar{M}$ expression, the difference between the conditions is close to significance.

Table 12.3 A

Estimates of the Eight Expressions of Table 12.1, Derived from the Data of Experiment 6

'True' Condition

	<u>EXPRESSION</u>	<u>ESTIMATE 1</u>	<u>ESTIMATE 2</u>	<u>MEAN</u> <u>ESTIMATE</u>	<u>S.E.</u>	<u>Z</u>
a)	$\alpha_a \text{ ITA} + (1 - \alpha_a) \text{ RM}$	0.958	0.958	0.958	0.029	0.00
b)	$\alpha_a \text{ IFA} + (1 - \alpha_a) \text{ RM}$	0.167	0.208	0.188	0.056	0.37
c)	$\alpha_a \text{ ITA} + (1 - \alpha_a) \text{ RM}$	1.000	0.875	0.938	0.034	1.84
d)	$\alpha_a \text{ IFA} + (1 - \alpha_a) \text{ RM}$	0.125	0.167	0.146	0.051	0.41
e)	$\alpha_c \text{ IFC} + (1 - \alpha_c) \text{ RM}$	0.500	0.375	0.438	0.071	0.88
f)	$\alpha_c \text{ ITC} + (1 - \alpha_c) \text{ RM}$	0.833	0.833	0.833	0.054	0.00
g)	$\alpha_c \text{ IFC} + (1 - \alpha_c) \text{ RM}$	0.292	0.250	0.271	0.064	0.33
h)	$\alpha_c \text{ ITC} + (1 - \alpha_c) \text{ RM}$	0.667	0.750	0.709	0.065	0.64

Table 12.3 B

Estimates of the Eight Expressions of Table 12.1, Derived from the data of Experiment 6

'False' condition

	<u>EXPRESSION</u>	<u>ESTIMATE 1</u>	<u>ESTIMATE 2</u>	<u>MEAN</u> <u>ESTIMATE</u>	<u>S.E.</u>	<u>Z</u>
a)	$\alpha_a \text{ ITA} + (1 - \alpha_a) \text{ RM}$	0.750	0.792	0.771	0.060	0.35
b)	$\alpha_a \text{ IFA} + (1 - \alpha_a) \text{ RM}$	0.292	0.458	0.375	0.069	1.20
c)	$\alpha_a \text{ ITA} + (1 - \alpha_a) \text{ R}\bar{\text{M}}$	0.875	0.708	0.792	0.058	1.44
d)	$\alpha_a \text{ IFA} + (1 - \alpha_a) \text{ R}\bar{\text{M}}$	0.333	0.250	0.292	0.065	0.64
e)	$\alpha_c \text{ IFC} + (1 - \alpha_c) \text{ RM}$	0.667	0.625	0.646	0.069	0.30
f)	$\alpha_c \text{ ITC} + (1 - \alpha_c) \text{ RM}$	0.583	0.708	0.646	0.069	0.91
g)	$\alpha_c \text{ IFC} + (1 - \alpha_c) \text{ R}\bar{\text{M}}$	0.500	0.542	0.521	0.072	0.29
h)	$\alpha_c \text{ ITC} + (1 - \alpha_c) \text{ R}\bar{\text{M}}$	0.583	0.375	0.479	0.071	1.46

Table 12.3 C

Estimates of the Eight Expressions of Table 12.1, Derived from the Data of Experiment 6

		<u>Comparison Between Conditions</u>				
	<u>EXPRESSION</u>	<u>ESTIMATE FROM</u>	<u>ESTIMATE FROM</u>	<u>DIFFERENCE</u>	<u>S.E. OF</u>	<u>Z</u>
		<u>'TRUE' COND.</u>	<u>'FALSE' COND.</u>		<u>DIFFERENCE</u>	
a)	$\alpha_a \text{ ITA} + (1 - \alpha_a) \text{ RM}$	0.958	0.771	0.187	0.067	2.79
b)	$\alpha_a \text{ IFA} + (1 - \alpha_a) \text{ RM}$	0.188	0.375	0.187	0.089	2.10
c)	$\alpha_a \text{ ITA} + (1 - \alpha_a) \text{ RM}$	0.938	0.792	0.146	0.067	2.18
d)	$\alpha_a \text{ IFA} + (1 - \alpha_a) \text{ RM}$	0.146	0.292	0.146	0.083	1.76
e)	$\alpha_c \text{ IFC} + (1 - \alpha_c) \text{ RM}$	0.438	0.646	0.208	0.099	2.10
f)	$\alpha_c \text{ ITC} + (1 - \alpha_c) \text{ RM}$	0.833	0.646	0.187	0.088	2.12
g)	$\alpha_c \text{ IFC} + (1 - \alpha_c) \text{ RM}$	0.271	0.521	0.250	0.096	2.60
h)	$\alpha_c \text{ ITC} + (1 - \alpha_c) \text{ RM}$	0.709	0.479	0.230	0.096	2.40

If the assumptions adopted by Evans (1977b), (i.e. that $ITA = IFC = 1$, and that $IFA = ITC = 0$), are employed to simplify these expressions, two estimates of α_a ($a - b$ and $c - d$) and α_c ($e - f$ and $g - h$) may be derived. The resulting estimates of these parameters are shown in Table 12.4 A for the 'true' condition, and in Table 12.4 B for the 'false' condition. Table 12.4 C shows a comparison of best estimates between the conditions.

It can be seen from Table 12.4 that the two estimates of each α value are not significantly different within each condition but that the best estimates of α_a and α_c are significantly different between the conditions. However, the more crucial finding shown in Table 12.4 is the negative value of α_c obtained in the 'true' condition. All parameter values should vary between 0 and 1 and thus α_c cannot be negative and it is apparent from this result that there is something fundamentally wrong, either with the model itself or with the method used to fit it to the data. Inspection of the method used by Evans (1977b) makes it apparent that a negative value of α_c will always be obtained whenever overall TC selections are more frequent than FC selections, and thus that it is an assumption of the method that subjects will always select more FC than TC. This assumption underlies the prior assumptions about I values and thus these are clearly not valid general assumptions.

The problem posed by the conclusion that the assumptions about I values are invalid is that, if no assumptions are made, it is impossible to derive estimates of any individual parameter values as each of the eight basic expressions contains all three types of parameter. However, it is possible to obtain assumption-free estimates of expressions that contain only two parameters. Of particular interest is an expression that contains no I parameter.

Table 12.4

Estimates of α and αc from the Data of Experiment 6

<u>PARAMETER</u>	<u>ESTIMATE 1</u>	<u>S.E.</u>	<u>ESTIMATE 2</u>	<u>S.E.</u>	<u>DIFFERENCE</u>	<u>S.E. OF</u>	<u>Z</u>	
						<u>DIFF.</u>		
<u>12.4 A</u>								
<u>'TRUE'</u>	α	0.770	0.063	0.792	0.061	0.022	0.088	0.25
<u>CONDITION</u>	αc	-0.395	0.089	-0.438	0.091	0.043	0.127	0.34
<u>12.4 B</u>								
<u>'FALSE'</u>	α	0.396	0.091	0.500	0.087	0.104	0.126	0.83
<u>CONDITION</u>	αc	0.000	0.098	0.042	0.101	0.042	0.141	0.30
<u>12.4 C</u>								
<u>PARAMETER</u>	<u>Best Est.</u>	<u>S.E.</u>	<u>Best Est.</u>	<u>S.E.</u>	<u>DIFFERENCE</u>	<u>S.E. OF</u>	<u>Z</u>	
<u>COMPARISON</u>	<u>'True' Cond</u>		<u>'False' Cond</u>			<u>DIFF.</u>		
<u>BETWEEN</u>	α	0.781	0.044	0.448	0.063	0.333	0.077	4.32
<u>CONDITIONS</u>	αc	-0.417	0.064	0.021	0.071	0.438	0.096	4.56

Subtraction of c from a (see Table 12.3) subtracts αa ITA from itself and subtraction of d from b involves the subtraction of αa IFA from itself. The result in both cases is an estimate of the expression $(1 - \alpha a) RM - (1 - \alpha a) R\bar{M}$ which can be simplified as:

$$(1 - \alpha a) (RM - R\bar{M})$$

A similar procedure can be followed to obtain two estimates of the expression $(1 - \alpha c) (RM - R\bar{M})$. An interesting result is obtained if estimates of these two expressions are calculated for the data of Evans & Lynch (1973). The result of this procedure is shown in Table 12.5 (letters refer to the data shown in Table 12.2).

It is somewhat surprising that the two mean estimates are identical. This arises from the fact that subjects selected exactly as many more P than \bar{P} cards as they did Q than \bar{Q} cards. However, the point of interest is that, for the data of Evans & Lynch, the following result obtains:

$$(1 - \alpha a) (RM - R\bar{M}) = (1 - \alpha c) (RM - R\bar{M})$$

There are only two possible explanations of this:

- 1) αa does not equal αc , and $(RM - R\bar{M})$ varies between antecedent and consequent selections in just such a way that the above equality is obtained. This possibility is difficult to accept if α and R are assumed to be independent.
- 2) $(RM - R\bar{M})$ is invariant across antecedent and consequent selections and, in this experiment, $\alpha a = \alpha c$. This is certainly the simplest and most likely explanation.

Table 12.5

Calculation of $(1 - \alpha) (RM - R\bar{M})$ Values from the Data of Evans & Lynch (1973). Estimates taken from
Evans (1977b)

<u>EXPRESSION</u>	<u>ESTIMATE 1</u> <u>a - c</u>	<u>S.E.</u>	<u>ESTIMATE 2</u> <u>b - d</u>	<u>S.E.</u>	<u>DIFFERENCE</u>	<u>S.E. OF</u> <u>DIFF.</u>	<u>Z</u>	<u>MEAN</u> <u>ESTIMATE</u>
$(1 - \alpha a) (RM - R\bar{M})$	0.333	0.084	0.312	0.078	0.021	0.115	0.18	0.323
	<u>ESTIMATE 1</u> <u>e - g</u>		<u>ESTIMATE 2</u> <u>f - h</u>					
$(1 - \alpha c) (RM - R\bar{M})$	0.292	0.098	0.354	0.091	0.062	0.134	0.46	0.323

The crucial point is that the Evans method of fitting the model cannot cope with the above equality as, when the data were fitted to the model assuming an invariant $(RM - \bar{RM})$, they produced widely differing values of α_a and α_c . Given the above equality, these results are clearly erroneous.

It should be noted that, if $(RM - \bar{RM})$ is presumed invariant¹ then the conclusion that $\alpha_a = \alpha_c$ (from the above equality) is independent of assumptions about I values, whereas the actual calculation of α_a and α_c by Evans (1977b) was not independent of his assumptions about I values.

However, Evans' calculations, of $\alpha_a = 0.511$ and $\alpha_c = 0.156$, although wrong, do have some meaning. If no assumptions are made about I values, then it can be seen that the calculations of α_a and α_c are, in fact, calculations of the expressions:

$$\alpha_a (ITA - IFA)$$

$$\text{and } \alpha_c (IFC - ITC)$$

(where, in the case of the Evans & Lynch data, $\alpha_a = \alpha_c$)

From the finding that $\alpha_a = \alpha_c$, it may thus be concluded that:

$$(ITA - IFA) > 3 (IFC - ITC)$$

This result is of some interest, as it suggests that consequent selections are not typified by a lack of attention to the logical structure. Evans argues that the relative influence of the interpretational factors was stronger on antecedent selections, but this was clearly not the case, as $\alpha_a = \alpha_c$. It may be inferred from the above expression that subjects' tendency to behave more in accordance with logical validity on antecedent, than on consequent, selections is a function of their interpretation of the sentence. In fact, given the Manktelow & Evans (1979) results,

it is possible that, over a range of studies, the mean finding would be that $IFC = ITC \approx 0.5$. That is, that subjects have no appreciation of the logical status of TC and FC and that response to these cards is random, unless influenced by matching,

The above analyses have used the model to obtain two estimates of each of the following expressions:

$(1 - \alpha_a)(RM - \bar{RM})$ which may be referred to as an index of antecedent matching

$(1 - \alpha_c)(RM - \bar{RM})$ which may be referred to as an index of consequent matching

$\alpha_a (ITA - IFA)$ which may be referred to as an index of antecedent interpretation

$\alpha_c (IFC - ITC)$ which may be referred to as an index of consequent interpretation.

These indices contain only two terms and may vary between -1 and 1. There is no a priori reason why the two estimates of these expressions should, in fact, be equalities and thus their calculation constitutes a valid test of the model itself. Inspection of Tables 12.2 B and 12.5 shows that the two estimates were in all cases impressively close. Thus, although it has been shown that Evans (1977b) was somewhat premature in his assumptions of parameter values, the above analysis supports his conclusion that the model fits the data of Evans & Lynch (1973).

An assumption-free method having been derived, the model can be fitted to the data of Experiment 6. Table 12.6 shows the estimates of the $(1 - \alpha) (RM - \bar{RM})$ expressions, Table 12.6 A for the 'true' condition and Table 12.6 B for the 'false' condition (letters refer to the data of Tables 12.3 A and 12.3 B respectively). Table 12.6 C shows a comparison of the best estimates between the conditions.

Table 12.6

Estimates of $(1 - \alpha) (RM - R\bar{M})$ Values from the Data of Experiment 6

	<u>EXPRESSION</u>	<u>ESTIMATE 1</u>	<u>S.E.</u>	<u>ESTIMATE 2</u>	<u>S.E.</u>	<u>DIFFERENCE</u>	<u>S.E. OF</u> <u>DIFF.</u>	<u>Z</u>
<u>12.6 A</u>								
	<u>$(1 - \alpha_a) (RM - R\bar{M})$</u>							
<u>'TRUE'</u>	(a - c and b - d)	0.020	0.045	0.042	0.076	0.022	0.088	0.25
<u>CONDITION</u>	<u>$(1 - \alpha_c) (RM - R\bar{M})$</u>							
	(e - g and f - h)	0.167	0.096	0.124	0.085	0.043	0.128	0.34
<hr/>								
<u>12.6 B</u>								
	<u>$(1 - \alpha_a) (RM - R\bar{M})$</u>							
	(a - c and b - d)	-0.021	0.083	0.083	0.095	0.104	0.126	0.83
<u>'FALSE'</u>	<u>$(1 - \alpha_c) (RM - R\bar{M})$</u>							
<u>CONDITION</u>	(e - g and f - h)	0.125	0.100	0.167	0.099	0.042	0.141	0.30

Table 12.6 (Continued)

Estimates of $(1 - \alpha) (RM - \bar{RM})$ Values from the Data of Experiment 6

	<u>EXPRESSION</u>	<u>'TRUE'</u>	<u>S.E.</u>	<u>'FALSE'</u>	<u>S.E.</u>	<u>DIFFERENCE</u>	<u>S.E. OF</u> <u>DIFF.</u>	<u>Z</u>
<u>12.6 C</u>								
<u>COMPARISON</u>	$(1 - \alpha_a) (RM - \bar{RM})$	0.031	0.044	0.031	0.063	-	0.077	0.00
<u>BETWEEN</u>								
<u>CONDITIONS</u>	$(1 - \alpha_c) (RM - \bar{RM})$	0.146	0.064	0.146	0.071	-	0.096	0.00
<u>(BEST ESTIMATES)</u>								

12.6 D

<u>MEAN ESTIMATE OF</u> <u>$(1 - \alpha_a) (RM - \bar{RM})$</u>	<u>S.E.</u>	<u>MEAN ESTIMATE OF</u> <u>$(1 - \alpha_c) (RM - \bar{RM})$</u>	<u>S.E.</u>	<u>DIFFERENCE</u>	<u>S.E. OF</u> <u>DIFF.</u>	<u>Z</u>
0.031	0.039	0.146	0.048	0.115	0.062	1.85

The two estimates of each expression are very close in each condition. The crucial finding, however, is shown in Table 12.6 C, which strongly suggests that these expressions are invariant across conditions.² It is possible that α_a and α_c are covarying with $(RM - \bar{RM})$ in just such a way between conditions as to keep the $(1 - \alpha) (RM - \bar{RM})$ values constant, but this is highly unlikely, and it is reasonable to conclude that α_a , α_c and $(RM - \bar{RM})$ are, in fact, invariant across conditions. Thus it may be concluded from this analysis that truth status affects only the I values.

Thus two of the Evans assumptions about I values must be rejected. Firstly, as was demonstrated earlier, his assumption of the actual numerical values were unwarranted and, secondly, it appears that Evans was also incorrect in his assumption that I values cannot vary. This latter finding has important implications which will be discussed later.

On the basis of the Evans & Lynch data, the difference between $(1 - \alpha_a) (RM - \bar{RM})$ and $(1 - \alpha_c) (RM - \bar{RM})$ estimates must be tested two-tailed and, as can be seen in Table 12.6 D, the observed difference in Experiment 3 does not quite reach significance ($p < 0.07$). However, the result suggests that the equality of these values observed on the Evans & Lynch data may not hold in all cases. It is worth noting at this point that any difference observed would not necessarily be due to differences between α_a and α_c , as $(RM - \bar{RM})$ may change between antecedent and consequent. Thus, to avoid making any assumptions whatsoever about the source of potential differences, these two expressions should properly be written as:

$$(1 - \alpha_a) (RM_a - \bar{RM}_a)$$

and $(1 - \alpha_c) (RM_c - \bar{RM}_c)$

Table 12.7

(α x) Differences in the Four I Values, 'False' Condition Relative to 'True' Condition

	<u>ESTIMATE 1</u> <u>OF</u> <u>DIFFERENCE</u>	<u>S.E.</u>	<u>ESTIMATE 2</u> <u>OF</u> <u>DIFFERENCE</u>	<u>S.E.</u>	<u>MEAN EST OF</u> <u>DIFFERENCE</u> <u>(D)</u>	<u>S.E. OF</u> <u>DIFFERENCE</u>	<u>Z</u> <u>(D/S.E.)</u>
αa (Difference in ITA)	-0.187	0.067	-0.146	0.067	-0.167	0.047	3.55
αa (Difference in IFA)	0.187	0.089	0.146	0.083	0.167	0.061	2.73
αc (Difference in IFC)	0.208	0.099	0.250	0.096	0.229	0.069	3.31
αc (Difference in ITC)	-0.187	0.088	-0.230	0.096	-0.209	0.065	3.21

The strong effect of truth status on I values is apparent from Table 12.4, in which estimates of α_a and α_c are, in fact, estimates of:

$$\alpha_a \text{ (ITA - IFA)}$$

and

$$\alpha_c \text{ (IFC - ITC)}$$

However, it cannot be inferred from changes in (ITA - IFA) and (IFC - ITC) that all four I values differ between the conditions. Fortunately, use can be made of the $(1 - \alpha) (RM - \bar{RM})$ invariance across conditions to investigate this further. Due to this invariance, subtraction of expressions a and c in the 'true' condition from expressions a and c in the 'false' condition respectively yields two estimates of the expression α_a (ITA 'false' - ITA 'true') which can be written as α_a (Difference in ITA) and is an index of the difference in ITA in the 'false' condition (relative to the 'true' condition). A similar procedure yields estimates of all four expressions shown in Table 12.7.

The last column of Table 12.7 shows the Z values derived from a test of the (null) hypothesis that there is no difference between the conditions. As can be seen, there is a highly significant difference in each case. The table shows that differences in ITA are similar to those of IFA and that differences in IFC are similar to those of ITC. The absolute values of these differences cannot be computed as the values of α_a and α_c are not known. In summary, analysis in terms of the model yields three pieces of useful information:

- 1) α_a , α_c , RM and \bar{RM} values are invariant between conditions.
- 2) ITA, IFA, IFC and ITC are all significantly different between the conditions.
- 3) In the 'true' condition, $ITC > IFC$ whereas these values

are approximately the same in the false condition.

$ITA > IFA$ holds for both conditions.

The primary importance of these findings is that truth status affects I values only, and that such effects can be demonstrated to occur by an analysis in terms of the model. Such analysis can thus play a useful role in identifying significant effects. More generally, the reader's attention is drawn to the fact that the two estimates of the four indices were in all cases impressively close in both the 'true' and 'false' conditions. Thus, together with the analysis of the Evans & Lynch data, the model itself has been tested on three sets of data. It is thus argued that this section has constituted a powerful test which the model can be considered as having passed.

However, the assumptions of I values adopted by Evans (1977b) have clearly failed this test and it has been shown that the specific parameter values derived from the Evans (1977b) fit were incorrect.

SECTION 12.3 FURTHER EXAMINATION OF THE UNDERLYING ASSUMPTIONS
OF THE MODEL

Although the model has proved quite useful in fitting the data of Experiment 6, these data, in fact, violate certain assumptions, regarding α values and the independence of responses, that were adopted to obtain the fit.

The analysis shown in Table 8.5 reveals that there are correlations between various selections on the selection task, notably a negative correlation between P and \bar{P} selections and a positive correlation between \bar{P} and \bar{Q} selections. This suggests, for instance, that \bar{Q} selections will be more frequent on rules having negative antecedents and thus that the initial two estimates of the eight expressions are not estimates of the same probability. However, an analysis (not presented here) across the data of Experiment 6, Evans & Lynch (1973) and both abstract experiments of Manktelow & Evans (1978) revealed no evidence that the two estimates, of each of the eight expressions, differed significantly. Thus, although correlations exist in the data, it is unlikely that they badly distort the fit.

The problem regarding α values is that, for the purpose of fitting the model, Evans assumes that there will be little variability between subjects "within an homogeneous sample such as undergraduate students". However, within such a sample, significant differences in selection frequencies were observed, in Experiments 5 and 6, between 'fast' and 'slow' learners. In Experiment 6, it was observed that fast learners were 'immune' to the effects of matching bias on antecedent selections and, in terms of the Evans model, this would imply that 'fast' learners

had higher values of α than 'slow' learners. Of course, in theory, α values may be expected to show some variability across subjects (although individual α values cannot be computed), and calculation of the overall α value is an average. However, the utility of an average is inversely related to the variability of the data. If there is little variability across subjects, the computed average α value has psychological relevance but, if there is a large amount of variability across subjects (as the significant differences between 'fast' and 'slow' learners suggest), then the 'representativeness' of the average α value is debatable.

However, the effect of a wholly unrepresentative α value would be to produce a bad fit to the data. The good fit to the data obtained in the last section, and the invariance of α values across the two conditions of Experiment 6, indicates that differences in α values between 'fast' and 'slow' learners did not seriously distort the calculations. Further, the existence of differences between subjects is no essential problem for the Evans model. Evans (1977b) points out that parameter values may vary between groups of subjects "who vary in intelligence, logical experience, etc." Thus, the more subgroups of subjects that are analysed individually, the more sophisticated will be the fit to the model, but the results of section 12.2 indicate that the model is sufficiently robust to cope with a certain amount of variability within the data.

In conclusion, it may be fairly said that the violation of assumptions, discussed above, would be expected to distort a fit to the model. Thus, the fact that these assumptions were to some extent violated cannot be used as an explanation of the good fit achieved in section 12.2. However, the fact that these

assumptions were violated to some extent suggests that, although a simple model has its uses, a more complex model would allow a more sophisticated analysis.

One further (implicit) assumption, adopted in fitting the model, is that use of all four conditional rules disambiguates 'logical' and 'nonlogical' factors. This assumption warrants detailed consideration, as it is the basis of the methodology of some crucial experiments, both in selection task research (notably Evans & Lynch, 1973 and Wason & Evans, 1975), and in other research on conditional rules (e.g. Evans, 1972b, 1977a). This assumption is sufficiently crucial to be awarded its own section.

SECTION 12.4 THE USE OF THE FOUR CONDITIONAL RULES IN THE
DISAMBIGUATION OF 'LOGICAL' AND 'NONLOGICAL' FACTORS

The matching bias results, in particular, depend on the use of all four conditional rules to control for interpretation. For instance, on the selection task, across the four rules P and \bar{P} both have the status of TA, and FA, on 50% of occasions. Higher selection frequencies of P , than \bar{P} , may thus be attributed to matching, rather than logical status. Similarly, TA and FA both match, and mismatch, on 50% of occasions. Higher selection frequencies of TA, than FA, may thus be attributed to logical status, rather than matching.

The use of the four rules is thus a very neat technique which controls for one dimension, whilst assessing the effects of the other. Most of the important findings, from research on propositional reasoning, depend upon this technique. For instance all the following findings depend, to at least some extent, on a control in one or other dimension:

- A) That matching bias mediates responses on both the selection task and truth table tasks.
- B) That conclusion bias mediates responses on inference tasks.
- C) That subjects have a strong tendency to select TA, and reject FA, on the selection task.
- D) That subjects apparently have no preferential tendency to select either FC or TC, on the selection task.
- E) That subjects have some tendency to evaluate the FT (truth table) case as 'false' and the FF case as 'true', and a very strong tendency to evaluate the TF case as 'false'. (Findings with respect to the TT case are so clear cut that they cannot be said to depend on a control for

matching, however the control does indicate the dominance of the tendency to evaluate the TT case as 'true').

The use of all four conditional rules has thus garnered an important body of information. Unfortunately, to interpret the data as definitely showing the above findings, the following two assumptions need to be adopted:

- 1) That the effect of logic, or interpretation, is consistent across the four rules.
- 2) That the effect of nonlogical factors is consistent across the four rules.

For instance, if subjects considered that testing MP was less important, and testing DA more important, on rules having negative antecedents, then this would inflate the apparent size of the matching bias effect on antecedent selections, and could even be responsible for the entire effect. Similarly, if, for some (unknown) reason, subjects had a stronger tendency to match on rules having affirmative antecedents, then this would inflate the size of the preference for TA selection, or could even create an apparent preference for TA selection.

The analysis performed in Chapter 8 (see Table 8.5) showed that, on selection tasks using the NN rule, there is a strong positive correlation between selection of cards of similar matching status (P and Q, and \bar{P} and \bar{Q}) and a strong negative correlation between selection of cards of dissimilar matching status (P and \bar{Q} , and \bar{P} and Q). However, on other rules, except in the case of the correlation between \bar{P} and \bar{Q} , there was little evidence for this pattern. It is thus possible that matching effects are not consistent across the four rules, although these results clearly cannot be considered as good evidence for this. However, it is a remote possibility that observed effects of 'logical' factors are a function of differential matching across the four rules, as similar effects (i.e. a tendency to behave in accordance

with logical validity on the antecedent but not the consequent) are also observed on conditional inference tasks, on which matching bias (presumably) has no effect.

A far more realistic hypothesis, one that presents some challenge to the validity of the methodology, is that there may be a differential effect of interpretation across the four rules. In particular, it is considered that there is something 'unusual' about the NA rule for the following two reasons:

- 1) It has proved very difficult, and sometimes impossible, to construct meaningful NA rules, implying implication, comparable to the other three rule forms. (see relevant comments in part 2). Unless the antecedent dimension is binary, it is very difficult to construct a thematic statement such that everything that is not P, is Q and even some Ps are Q. Even with a binary antecedent, the task is not easy, although it is simple to construct binary NA statements implying equivalence (e.g. 'If a number is not even, then it is odd'). Thus, as in 'real life', NA rules will almost always be used to imply equivalence, subjects may well interpret such rules as implying equivalence on an abstract task.
- 2) The NA rule states that everything is either P or Q, whereas the other conditional rules, when implying implication, do not make such an assertion. That is, if implication rules are presented in diagram form, then the diagrams of AA, AN, and NN rules all have a blank space in them, allowing for a (possible) state of the world in which neither P nor Q is the case. The NA rule is thus unlike other rules in that it is a disjunctive

statement about P and Q, a point made by the Greek philosopher, Galen. Given the ease with which subjects make inferences from AA disjunctives, and the difficulty incurred by all other (inclusive) disjunctives, it is probable that, in 'real life', conditionals are used preferentially (to disjunction), except in the case of the AA disjunction, which is used in preference to the NA conditional. There are grounds for arguing that disjunctions are usually interpreted as exclusive (thus leading to an equivalence interpretation of the NA conditional) but it is unnecessary to pursue this point, as the very possibility that the NA conditional may be interpreted as an AA disjunction, in itself, suggests that its interpretation may well not be equivalent to other conditionals. However, the purpose here is not to suggest that apparent 'nonlogical' effects are due to interpretational differences across the four rules...The above points have been made simply to introduce a note of caution. There are, in fact, very good reasons for assuming that nonlogical effects are not a function of differential interpretation of the rules, for instance:

- 1) The consistent lack of significant difference, within the pairs of equalities predicted by the model, suggests that the NA rule does not yield (selection task) responses inconsistent with other rule forms. For instance, if interpretation was held to underlie responses, leading to more ('logical') 'tests' of AC and DA on rules implying equivalence (i.e. by selection of TC and FA, respectively), then it would be difficult to explain why AC is 'tested' no more often on NA rules than on AA rules (on which DA is rarely 'tested') and why DA is 'tested' no more often

on NA rules than on NN rules (on which AC is rarely 'tested').

- 2) The above findings, and other findings discussed earlier, show that behaviour that can be logically related to the AC inference does not 'move' with behaviour that can be logically related to the DA inference. Clearly, these behaviours would be expected to 'move together' if they were a function of interpretation.
- 3) Behaviour that can be related to the MT inference is equally susceptible to response bias, although MT is logically valid on all known interpretations.
- 4) As Evans has pointed out (e.g. Evans, 1980b), subjects appear to interpret the conditional as implying equivalence more often on inference, than on truth table, tasks. This is inconsistent with an 'interpretational' view, but is not inconsistent with the view that factors other than interpretation determine responses.
- 5) On the selection task, the significant positive correlation between \bar{P} and \bar{Q} selections, on all four rules (see Table 8.5), would be impossible to explain as based in differential interpretation of the rules, and may be regarded as firm evidence that the matching status of the cards has psychological relevance.

Finally, it is worth noting that, although in 'real life' the NA rule may be more often used to imply equivalence, there is no need for the subject to learn differential interpretations of different syntactic structures. In 'real life', the extent to which a conditional statement implies equivalence is wholly determined by its semantic content.

SECTION 12.5 INTROSPECTION AND THE STOCHASTIC NATURE OF RESPONSES

Certain errors in the fit to the model obtained by Evans (1977b) have been pointed out and corrected in Section 12.2 and other possible problems for the model have been discussed in 12.3 and 12.4. In general, the model has stood up reasonably well to detailed analysis, and the more general tenets of the Evans position may now be evaluated.

His position regarding introspection, and the probabilistic nature of responses, will be discussed in one section partly because they are, in one sense, related. It is more natural to view a response as deterministic if it is introspectable, if only because subjects report what appears to be a deterministic process. It is only when introspection is not regarded as revealing the causes of behaviour that the possibility that behaviour is stochastic may be considered. That is, as subjects report a deterministic process, then behaviour can only be stochastic if it is not introspectable. (This is not to suggest that, if behaviour is not stochastic, it is introspectable).

There is no evidence that reasoning responses are introspectable and, further, for introspection to be of any use in interpreting reasoning data, responses would have to be introspectable on all occasions. If not, then the experimenter will never know whether any particular response is introspectable or not. As Evans (1980b) has pointed out, "we can only assess the validity of a subjective report in a situation where we have an independent objective method of determining the cause of the subject's action. In such a situation, however, we do not need his introspective report".

That at least some introspective reports are not valid is clearly indicated by the data of Wason & Evans (1975), and the 'rationalisation' hypothesis adequately accounts for the content of the subjects' protocols. The subject is apparently unaware³ of the causes of his behaviour and, when asked for an explanation, generates the best explanation available. This effect is particularly noticeable in the data of Experiment 3 and of the Van Duyne (1976) study (see Chapter 6). In the Van Duyne study, significant differences were observed in the protocol data, across conditions which produced no significant differences in selection behaviour. That explanations of selections can be manipulated (by the experimental conditions) independently of the selections themselves, is definitive evidence that those explanations do not mirror the causes of behaviour.

Fellows (1976) has argued that "if Evans dismisses the subjects' reports as rationalisations, then logically he must also dismiss his own explanations in the same way" and claims that an artificial distinction is drawn "between 'the experimenter' who is all wise and 'the subject' who is a fool". This is a valid point as any interpretation of data, must be, at least to some extent, a personal rationalisation. If this were not the case, there would be no theoretical controversy within the psychological literature. When selection task data from AA rules only was considered, responses were explained in the same way by experimenters and subjects (as due to a verification bias). This suggests that subjects' rationalisations should be interpreted in terms of the theories of Bem (1965) rather than those of Festinger (1957). That is, the subject observes his own behaviour and attempts to construct the best available explanation of it, as does the experimenter.

However, the experimenter has the advantage of a broader perspective in cases where, either he is in possession of data from a wide range of studies or, in cases of a single study, he manipulates a variable of which the subject is apparently unaware and which he predicts will have an effect upon responses. Thus, for instance, in the case of the Evans & Lynch (1973) study, the experimenters may be said to have been 'wiser' than the subjects.

The question of whether behaviour is stochastic or deterministic extends across the whole of experimental psychology, which essentially infers aspects of human behaviour from group data. It is a question of whether results are interpreted as showing that (all) subjects are most likely to behave in a particular way, or as showing that most subjects will behave in a particular way. A preference must be admitted for the former alternative and there is no apparent reason why data from reasoning tasks should not be interpreted in the same way as data from other areas of experimental psychology. However, certain findings reported in Part 2 indicate that a simple stochastic approach is unsatisfactory.

The selection differences between 'fast' and 'slow' learners show that there is greater variability within an homogeneous group of subjects than might have been expected. This does not invalidate a stochastic approach but it does indicate that, even in an homogeneous sample, the group average probability of making a response is a poor indicator of individual probabilities of making a response. Further, the Dickstein results concerning the independence of responses between, but not within, category 1 and 2 premiss pairs, indicate that individual differences between subjects may apply to some, but not other, problems. Thus individual parameter values within a stochastic model may vary with

regard both to individual differences and the type of problem. However, the Dickstein finding, that some subjects behave more in accordance with validity on some problems, whereas other subjects behave more in accordance with validity on other problems, would presumably be as much of a problem for a deterministic, as for a stochastic, approach.

The associations observed (in Chapter 8) between card selections on the selection task also suggest that a more complex stochastic approach is required. Given that certain selections appear to be made prior to other selections or rejections, it is possible that the probability of subsequent selections is contingent upon prior selections. It is possible, for instance, that, on the AA rule, prior selection of TA increases the tendency to reject FA (as these selections are negatively associated) and that this rejection of FA increases the tendency to subsequently reject FC (as these selections are positively associated on this rule). However, no causal relation need be postulated. The point is that, given the associations between the selections, certain of the probabilities involved must be conditional probabilities.

Thus, although no objections are made to a stochastic approach, it is suggested that any sophisticated stochastic approach would need to take account of the following two factors:

- 1) The role of individual differences (such as probability learning speed) even within an apparently 'homogeneous' group of subjects.

- 2) The contingent relationship between various responses.

For instance, subjects who select \bar{P} are more likely to select \bar{Q} . Similarly, it was reported in Chapter 5 that subjects have differential tendencies to affirm inferences,

and thus that subjects who affirm one inference are more likely to affirm another.

SECTION 12.6 THE DISTINCTION BETWEEN INTERPRETATIONAL AND
OPERATIONAL FACTORS

The central theme of this thesis has been that there are preferred responses that accord with logical validity, and other preferred responses that do not, and consistent support has been advanced for the Evans viewpoint that the latter are not the result of error in a process of 'logical reasoning', but are separate behaviours to be explained in their own right. Thus the position adopted in previous chapters must be regarded as, at least, similar to that adopted by Evans. Certainly, the points made in previous sections of this chapter have been entirely consistent with the Evans position regarding introspection and have, with certain modifications, supported a mathematical approach. However, when the core feature of the Evans position, his conception of the nature of interpretational and operational factors, is evaluated, important differences emerge between his position and that adopted here.

Before discussing these differences, it is important to be quite clear about Evans' definition of interpretational and operational factors. Evans (1977b) states that "generally speaking, interpretational influences take the form of logical tendencies resulting from the subjects' comprehension of the sentence, while operational influences consist of nonlogical response tendencies such as matching bias." Operational tendencies have been described as (due to the fact that) "subjects.....ignore the logical structure of the problems" (Evans, 1972a) and simply as "a nonlogical response bias" throughout Evans (1978).

Evans has thus tended to describe, rather than explain, nonlogical effects. However, there is no wish to imply here that such accounts of observed responses are in any way invalid. One does not have to explain an effect to argue that it exists. Thus Evans' arguments that nonlogical effects exist, and must be taken account of, have constituted a valuable contribution to the understanding of reasoning task behaviour. However, it is clearly an extension of his position that is advanced in Chapters 10 and 11.1, in which some attempt was made to investigate the substrate of apparent response biases, and to explain them as possibly learned behaviours that have utility in 'real life', but that are inappropriate to artificial experimental situations. In fact, as long ago as 1972, Evans (1972a) has suggested that operational effects may be products of the experimental situation, but has not attempted to explain why this may be the case.

The Evans position regarding interpretational factors is that they relate to the logical structure of the problem, dependant upon the subject's interpretation of the sentence. Although this position appears to assume that subjects are capable of a valid logical analysis in certain situations, it is not inconsistent with the arguments developed in Chapter 11.2. Evans has himself suggested that prior use of the language may influence subjects' responses on formally structured tasks. Specifically, Evans (1977a) has suggested that, in 'real life', subjects use conditionals of the form 'P only if Q' in situations where Q precedes P in time, whereas the usual form of the conditional is used more often in situations where P precedes Q in time. On the basis of this suggestion, he predicted that this prior usage would influence responses on a formal reasoning task, and Evans & Newstead (1977)

have presented some empirical support for this prediction. Thus, in this case, behaviour in accordance with an 'interpretational' influence is not necessarily seen as leading to behaviour in accordance with logical validity. This is similar to arguments about the learned use of the word 'or' advanced in Chapter 11.2.3 and, more generally, is consistent with the point of view that 'logical' behaviours may result from a learned use of the language. However, it is not, of course, suggested that Evans himself has argued that all interpretational effects result from such learned usage.

The view that interpretational effects do not necessarily lead to valid behaviour is supported by the data of Experiment 6 (see Chapter 8). In this experiment, truth status affected responses with regard to their logical, rather than matching, status. As was seen in section 12.2, it was the I values of the Evans model that varied with truth status. However, this cannot be regarded as a consistent effect of logical validity, as behaviour in accordance with logical validity was facilitated on antecedent selections in the 'true' condition and on consequent selections in the 'false' condition. Thus this appears to be an effect that is related to the logical structure of the problem but can not be held to be related to the subject's appreciation of the logic of the problem.

However, if the effect of truth status is taken as being an interpretational influence, then other findings of Experiment 6 pose a problem for the Evans position. Evans assumes that interpretational and operational factors are independent (although Evans, 1977, has accepted that there must be some interaction between nonlogical influences and rule form as, for instance, matching bias

is not observed on disjunctive tasks). In Experiment 6, those subjects ('fast learners') who were 'immune' to the (interpretational) effects of truth status on antecedent selections were also 'immune' to the (operational) effects of matching bias on antecedent selections. As these subjects also selected more TA, and less FA, than slow learners, the following associations were observed in the data:

- 1) An association between the (operational) effect of matching bias and the (interpretational) effect of behaviour in accordance with logical validity on antecedent selections.
- 2) An association between the above interpretational effect and the effect of truth status.
- 3) An association between truth status and matching bias.

Thus, the effect of truth status was associated with both an interpretational tendency (to select TA and reject FA), and an operational (matching) tendency, and these two tendencies were associated with each other. These associations are difficult to account for from the Evans point of view, but they are entirely consistent with the position developed in this discussion. If all preferred responses are a result of behaviour acquired in the solution of 'real life' problems, then there is no essential difference between the influence of truth status and the influence of matching (both being 'available' influences), even though one happens to relate to the logical structure and the other does not. Similarly, there is no essential difference between the influence that leads subjects to test conditional rules by evaluating TA, and the influence of truth status, even though one happens to correlate with logical validity and the other does not. The subject will behave in accordance with these influences, dependent upon

their strength. If fast learners have had greater experience of use of conditional relationships, as their faster learning of such relationships implies, then they will have a stronger tendency, than other subjects, to evaluate TA. Clearly, a strong dominant response will be less susceptible to other influences.

In conclusion, although there are typically observed responses that may be categorised as relating, or not relating, to the logical structure of the problem, and responses that may be categorised as according, or not according, with logical validity, there is no reason to believe that these behaviours are fundamentally different from one another.

SECTION 12.7 CONCLUSIONS

At the end of Part 1, it was concluded that data from reasoning tasks could not be explained without taking account of 'nonlogical' factors, unrelated to the logical structure of the problem, in addition to logically valid behaviour. However, the truth status findings, particularly those of Experiment 6, demonstrate that responses cannot simply be categorised into those depending on an appreciation of logical validity, which relate to structure, and those that result from a response bias, independent of structure. The effect of truth status in Experiment 6 related to the logical structure of the task, but did not have a consistent effect in terms of logical validity. This result appears to 'blur' the distinction between 'logical' and 'nonlogical' factors.

In Chapters 10 and 11, it has been argued that there is, in fact, no essential difference between behaviours that accord with logical validity and those that do not. It was pointed out that logical validity is not the salient dimension in 'real life' problem solving (which requires correct, rather than valid, decisions) and that behaviours that have been learned to solve 'real life' problems would be expected to produce logically valid responses on some, but not necessarily all, occasions. From this point of view, the designation of responses as 'interpretational' or 'operational' (or as 'related, or not related, to logical structure') is seen as an essentially irrelevant categorisation imposed by the experimenter.

If typically observed responses are held to have utility in real life, then it is of particular interest to consider the apparent response biases observed on reasoning tasks, as it may at first

be thought that such responses could have no 'real life' utility. However, deeper analysis of these apparent response biases does indicate that they may serve a useful function in 'real life' decision making. The functional utility of behaviours that lead to valid conclusions is, presumably not in question. The more controversial suggestion, advanced in this discussion, is that subjects have acquired responses to specific aspects of problems and that, although such responses may sometimes accord with logical validity, there is no reason to assume that an appreciation of that validity is acquired with the response.

Whether or not the reader is convinced, from the points raised in Chapters 10 and 11, that there is, at least, a case for arguing that such responses are acquired, it cannot seriously be denied that subjects more frequently make valid responses on problems of which they have some experience. The decision process is facilitated if the subject is familiar with the problem. For instance, subjects are experienced at using the word 'or' with relation to AA disjunctions.

Another important variable is the semantic content of the problem. Subjects' experience of the content of 'real life' problems will have an important influence upon behaviour. For instance, many 'real life' inferences may be 'pragmatic', rather than logical, (Harris & Monaco^O, 1978) and direct evaluation of the conclusion has utility when it is as easy to evaluate the conclusion as it is to evaluate the premisses, of an argument. Experiments 5 and 6 indicate that even a small amount of prior experience will influence responses.

It has been the contention of this discussion that subjects will learn useful responses to problems that they encounter

in 'real life', but that, when they encounter experimental problems in which neither the content nor the structure relates to their prior experience, it is not surprising if their previously learned responses are inappropriate. On the basis of this argument, if it is required to investigate how subjects actually do solve problems in 'real life', then they must be given 'real life' problems to solve.

It was intended to conclude this discussion with some reference to early work on memory, but this has been succinctly referred to by Evans (1978) who points out: "it is now generally acknowledged that the attempt to investigate human memory via meaningless nonsense syllables was misguided and largely unproductive: the study of meaning or semantics is central to the understanding of memory." In the same way, 'reasoning' cannot be investigated by the use of problems devoid of relevance to the subjects' experience, as the ability to 'reason' is an inseparable function of that experience.

FOOTNOTES

1. In fact, this is a very reasonable assumption as, if $(RM - \bar{RM})$ were not invariant across antecedent and consequent selections, it would have to interact with α values in a highly coincidental manner to produce the equality between $(1 - \alpha_a) (RM - \bar{RM})$ and $(1 - \alpha_c) (RM - \bar{RM})$.

2. Surprisingly, there is no difference between conditions in the best estimates of each expression. Subjects in the 'true' condition selected exactly as many more P, than \bar{P} , cards and exactly as many more Q, than \bar{Q} cards, as did subjects in the 'false' condition.

3. It is, of course, possible that the subject is aware of matching, but supresses this information in his verbal report.

REFERENCES

- ASCH, S.E. (1956) Studies of independence and conformity.
Psychological Monographs, 70, no. 9 (Whole no. 416)
- AYER, A.J. (1936) Language, Truth and Logic. London: Gollancz
- BEGG, I. & DENNY, J.P. (1969) Empirical reconciliation of
atmosphere and conversion interpretations of syllogistic
reasoning errors. Journal of Experimental Psychology,
81, 351-4
- BEM, D.J. (1965) An experimental analysis of self-persuasion.
Journal of Experimental Social Psychology, 1, 199-218
- BRACEWELL, R.J. (1974) Interpretation factors in the four card
selection task. Paper presented at the Trento conference
on the Selection Task
- BRACEWELL, R.J. & HIDI, S.E. (1974) The solution of an inferential
problem as a function of stimulus materials. Quarterly
Journal of Experimental Psychology, 26, 480-8
- BRAINE, M.D.S. (1976) On the meaning of if-then. Paper presented
at the Psycholinguistic Circle of New York
- BRAINE, M.D.S. (1978) On the relation between the natural logic
of reasoning and standard logic. Psychological Review,
85, 1-21
- BREE, D.S. (1973) The interpretation of implication. In A. Elithorn
and D. Jones (Eds): Artificial and Human Thinking.
Amsterdam: Elsevier
- BREE, D.S. (1974) If the context is open the inference is uninvited.
Working paper 18 Delft: Interfakulteit Bedrijfskunde

- BREE, D.S. & COPPENS, G. (1976) The difficulty of an implication task. British Journal of Psychology, 67, 579-86
- BREE, D.S. & COPPENS, G. (1978) A reply to Moskman. British Journal of Psychology, 69, 373
- BUCCI, W. (1978) The interpretation of universal affirmative propositions. Cognition, 6, 55-77
- BUDNER, S. (1962) Intolerance of ambiguity as a personality variable. Journal of Personality, 30, 29-50
- BYRNE, R. (1977) Planning meals: problem solving on a real data base. Cognition, 5, 287-332
- CAMILLI, G. & HOPKINS, K.D. (1978) Applicability of chi-square to 2 x 2 contingency tables with small expected cell frequencies. Psychological Bulletin, 85, 163-7
- CERASO, J. & PROVITERA, A. (1971) Sources of error in syllogistic reasoning. Cognitive Psychology, 2, 400-10
- CHAPMAN, L.J. (1967) Illusory correlation in observational report. Journal of Verbal Learning and Verbal Behaviour, 6, 151-5
- CHAPMAN, L.J. & CHAPMAN, J.P. (1959) Atmosphere effect re-examined. Journal of Experimental Psychology, 58, 220-6
- CHAPMAN, L.J. & CHAPMAN, J.P. (1967) Genesis of popular but erroneous psychodiagnostic observations. Journal of Abnormal Psychology, 73, 193-204
- CHAPMAN, L.J. & CHAPMAN, J.P. (1969) Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. Journal of Abnormal Psychology, 74, 271-80
- COHEN, M.R. & NAGEL, E. (1934) An Introduction to Logic and Scientific Method. London: Routledge & Kegan Paul

COPI, I.M. (1972) Introduction to Logic. New York: Macmillan

DICKSTEIN, L.S. (1975) Effects of instructions and premise order on errors in syllogistic reasoning. Journal of Experimental Psychology, 104, 376-84

DICKSTEIN, L.S. (1978a) Error processes in syllogistic reasoning. Memory and Cognition, 6, 537-43

DICKSTEIN, L.S. (1978b) The effect of figure on syllogistic reasoning. Memory and Cognition, 6, 76-83

EVANS, J.St.B.T. (1972a) On the problems of interpreting reasoning data: logical and psychological approaches. Cognition, 1, 373-84

EVANS, J.St.B.T. (1972b) Interpretation and matching bias in a reasoning task. Quarterly Journal of Experimental Psychology, 24, 193-9

EVANS, J.St.B.T. (1972c) Reasoning with negatives. British Journal of Psychology, 63, 213-19

EVANS, J.St.B.T. (1974) On the origin of selections in the selection task. Paper presented at the Trento conference on the Selection Task.

EVANS, J.St.B.T. (1975) On interpreting reasoning data.....a reply to Van Duyne. Cognition, 3, 387-90

EVANS, J.St.B.T. (1976) A critical note on Quinton & Fellows' observation of reasoning strategies. British Journal of Psychology, 67, 517-18

EVANS, J.St.B.T. (1977a) Linguistic factors in reasoning. Quarterly Journal of Experimental Psychology, 29, 297-306

EVANS, J.St.B.T. (1977b) Toward a statistical theory of reasoning. Quarterly Journal of Experimental Psychology, 29, 621-35

- EVANS, J.St.B.T. (1978) The psychology of deductive reasoning.
In A. Burton and J. Radford (Eds): The Study of Thinking
London: Methuen
- EVANS, J.St.B.T. (1980a) Thinking: Experiential and information
processing approaches. In G. Claxton (Ed) New Directions
in Cognitive Psychology (in press) London: RKP
- EVANS, J.St.B.T. (1980b) Current issues in the psychology of
reasoning. British Journal of Psychology (in press)
- EVANS, J.St.B.T. & LYNCH, J.S. (1973) Matching bias in the
selection task. British Journal of Psychology, 64, 391-7
- EVANS, J.St.B.T. & NEWSTEAD, S.E. (1977) Language and reasoning:
A study of temporal factors. Cognition, 5, 265-83
- EVANS, J.St.B.T. & WASON, P.C. (1976) Rationalisation in a
reasoning task. British Journal of Psychology, 67, 479-86
- FEATHER, N.T. (1964) Acceptance and rejection of arguments in
relation to attitude strength, critical ability and in-
tolerance of inconsistency. Journal of Abnormal and
Social Psychology, 69, 127-36
- FELLOWS, B.J. (1976) The role of introspection in problem solving
research: A reply to Evans. British Journal of
Psychology, 67, 519-20
- FESTINGER, L. (1957) A Theory of Cognitive Dissonance. Stanford
Univ. Press
- FHANER, S. (1977) Subjective probability and everyday life.
Scandinavian Journal of Psychology, 18, 81-4
- FILLENBAUM, S. (1975) If: Some uses. Psychological Research, 37, 245-50
- FILLENBAUM, S. (1976) Inducements: On the phrasing and logic of
conditional promises, threats and warnings. Psychological
Research, 38, 231-50

- FISCHHOFF, B. (1975) Hindsight \neq foresight: The effect of outcome knowledge on judgement under uncertainty. Journal of Experimental Psychology: Human Perception and Performance, 1, 288-99
- FISCHHOFF, B. (1977) Perceived informativeness of facts. Journal of Experimental Psychology: Human Perception and Performance 3, 349-58
- FISCHHOFF, B. & BEYTH, R. (1975) "I knew it would happen" - remembered probabilities of once-future things. Organisational Behaviour and Human Performance, 13, 1-16
- FRASE, L.T. (1966a) Validity judgements of syllogisms in relation to two sets of terms. Journal of Educational Psychology, 57, 239-45
- FRASE, L.T. (1966b) Belief, incongruity and syllogistic reasoning. Psychological Reports, 18, 982
- FRASE, L.T. (1968a) Effects of semantic incompatibility upon deductive reasoning. Psychonomic Science, 12, 64
- FRASE, L.T. (1968b) Associative factors in syllogistic reasoning. Journal of Experimental Psychology, 76, 407-12
- GEIS, M.L. & ZWICKY, A.M. (1971) On invited inferences. Linguistic Inquiry, 2, 561-6
- GENTZEN, G. (1964) Investigations into logical deduction. American Philosophical Quarterly, 1, 288-306
- GILHOOLY, K.J. & FALCONER, W.A. (1974) Concrete and abstract terms and relations in testing a rule. Quarterly Journal of Experimental Psychology, 26, 355-9
- GOODWIN, R.Q. & WASON, P.C. (1972) Degrees of insight. British Journal of Psychology, 63, 205-12

- GORDON, R. (1953) The effect of attitude toward Russia on logical reasoning. Journal of Social Psychology, 37, 103-11
- GRANT, D.A., HAKE, H.W. & HORNSETH, J.P. (1951) Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. Quarterly Journal of Experimental Psychology, 42, 1-5
- HARRIS, R.J. & MONACO, G.E. (1978) Psychology of pragmatic implication: Information processing between the lines. Journal of Experimental Psychology: General, 107, 1-31
- HENLE, M. (1962) On the relation between logic and thinking. Psychological Review, 69, 366-78
- HENLE, M. & MICHAEL, M. (1956) The influence of attitudes on syllogistic reasoning. Journal of Social Psychology, 44, 115-27
- HULL, C.L. (1943) Principles of Behaviour: An Introduction to Behaviour Theory. New York: Appleton - Century - Crofts
- JANIS, I. & FRICK, F. (1943) The relationship between attitudes toward conclusions and errors in judging logical validity of syllogisms. Journal of Experimental Psychology, 33, 73-7
- JENKINS, H.M. & WARD, W.C. (1965) Judgement of contingency between responses and outcomes. Psychological Monographs, 79, no 1. (Whole no. 594)
- JOHNSON-LAIRD, P.N. (1975) Models of deduction. In R.J. Falmagne (Ed): Reasoning: Representation and Process. New York: Wiley
- JOHNSON-LAIRD, P.N., LEGRENZI, P & LEGRENZI, M.S. (1972) Reasoning and a sense of reality. British Journal of Psychology, 63, 395-400

- JOHNSON-LAIRD, P.N. & STEEDMAN, M. (1978) The psychology of syllogisms. Cognitive Psychology, 10, 64-99
- JOHNSON-LAIRD, P.N. & TAGART, J. (1969) How implication is understood. American Journal of Psychology, 82, 367-73
- JOHNSON-LAIRD, P.N. & TRIDGELL, J.M. (1972) When negation is easier than affirmation. Quarterly Journal of Experimental Psychology, 24, 87-91
- JOHNSON-LAIRD, P.N. & WASON, P.C. (1970) A theoretical analysis of insight into a reasoning task. Cognitive Psychology, 1, 134-48
- KAUFMAN, H. & GOLDSTEIN, S. (1967) The effects of emotional value of conclusions upon distortions in syllogistic reasoning. Psychonomic Science, 7, 367-8
- KNEALE, W. & KNEALE, M. (1962) The Development of Logic
London: Oxford Univ. Press
- KUBOVY, M. (1977) Response availability and the apparent spontaneity of numerical choices. Journal of Experimental Psychology: Human Perception and Performance, 3, 359-64
- KUBOVY, M. & PSOTKA, J. (1976) The predominance of seven and the apparent spontaneity of numerical choices. Journal of Experimental Psychology: Human Perception and Performance, 2, 291-4
- LEFFORD, A. (1946) The influence of emotional subject matter on logical reasoning. Journal of General Psychology, 34, 127-51
- LEGRENZI, P. (1970) Relations between language and reasoning about deductive rules. In G.B. Flores D'Arcais & W.J.M. Levelt (Eds): Advances in Psycholinguistics Amsterdam: North-Holland
- LEGRENZI, P. (1971) Discovery as a means of understanding. Quarterly Journal of Experimental Psychology, 23, 417-22

- LEICHT, K.L. (1968) Recall and judged frequency of implicitly occurring words. Journal of Verbal Learning and Verbal Behaviour, 7, 918-23
- LEWIS, D. (1973) Counterfactuals. Oxford: Basil Blackwell
- LICHTENSTEIN, S., SLOVIC, P., FISCHHOFF, B., LAYMAN, M. & COMBS, B. (1978) Judged frequency of lethal events. Journal of Experimental Psychology: Human Learning and Memory, 4, 551-78
- LIPPMAN, M.Z. (1972) The influence of grammatical transform in a syllogistic reasoning task. Journal of Verbal Learning and Verbal Behaviour, 11, 424-30
- LUNZER, E.A., HARRISON, C. & DAVEY, M. (1972) The four card problem and the generality of formal reasoning. Quarterly Journal of Experimental Psychology, 24, 326-39
- MANKTELOW, K.I. (1978) Realistic reasoning: it's not that easy. Paper presented at the London Conference of the British Psychological Society
- MANKTELOW, K.I. & EVANS, J.St.B.T. (1979) Facilitation of reasoning by realism: effect or non-effect? British Journal of Psychology (in press)
- MORGAN, J.J.B. & MORTON, J.T. (1944) The distortion of syllogistic reasoning produced by personal convictions. Journal of Social Psychology, 20, 39-59
- MOSHMAN, D. (1978) Some comments on Brée & Coppens' The difficulty of an implication task. British Journal of Psychology, 69, 371-2
- NAVON, D. (1978) The importance of being conservative: Some reflections on human Bayesian behaviour. British Journal of Mathematical and Statistical Psychology, 31, 33-48
- NISBETT, R.E. & WILSON, T.D. (1977) Telling more than one can know: Verbal reports on mental processes. Psychological Review, 84, 231-59

- PEZZOLI, J.A. & FRASE, L.T. (1968) Mediated facilitation of syllogistic reasoning. Journal of Experimental Psychology, 78, 228-32
- POPPER, K.R. (1959) The Logic of Scientific Discovery London: Hutchinson
- REVLIN, R. & LEIRER, V.O. (1978) The effects of personal biases on syllogistic reasoning: Rational decision from personalised representations. In R. Revlin and R.E. Mayer (Eds) Human Reasoning Washington: Winston/Wiley
- REVLIS, R. (1974) Prevarication: Reasoning from false assumptions. Memory and Cognition, 2, 87-95
- REVLIS, R. (1975a) Two models of syllogistic reasoning: Feature selection and conversion. Journal of Verbal Learning and Verbal Behaviour, 14, 180-95
- REVLIS, R. (1975b) Syllogistic reasoning: Logical decisions from a complex data base. In R.J. Falmagne (Ed): Reasoning: Representation and Process. New York: Wiley
- REVLIS, R. & HAYES, J.R. (1972) The primacy of generalities in hypothetical reasoning. Cognitive Psychology, 3, 268-90
- REVLIS, R. LIPKIN, S.G. & HAYES, J.R. (1971) The importance of universal quantifiers in a hypothetical reasoning task. Journal of Verbal Learning and Verbal Behaviour, 10, 86-91
- 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 New York: Wiley
- ROBERGE, J.J. (1970) A reexamination of errors in formal syllogistic reasoning. Psychonomic Science, 19, 331-3
- ROBERGE, J.J. (1971a) Further examination of mediated associations in deductive reasoning. Journal of Experimental Psychology, 87, 127-9

- ROBERGE, J.J. (1971b) Some effects of negation on adults' conditional reasoning abilities. Psychological Reports, 29, 839-44
- ROBERGE, J.J. (1974) Effects of negation on adults' comprehension of fallacious conditional and disjunctive arguments. Journal of General Psychology, 91, 287-93
- ROBERGE, J.J. (1976a) Effects of negation on adults' disjunctive reasoning abilities. Journal of General Psychology, 94, 23-8
- ROBERGE, J.J. (1976b) Reasoning with exclusive disjunction arguments. Quarterly Journal of Experimental Psychology, 28, 419-27
- ROBERGE, J.J. (1977) Effects of content on inclusive disjunction reasoning. Quarterly Journal of Experimental Psychology, 29, 669-76
- ROBERGE, J.J. (1978) Linguistic and psychometric factors in propositional reasoning. Quarterly Journal of Experimental Psychology, 30, 705-16
- ROSENTHAL, R. (1978) Combining results of independent studies. Psychological Bulletin, 85, 185-93
- SELLS, S.B. (1936) The atmosphere effect: An experimental study of reasoning. Archives of Psychology, 29, 3-72
- SHANTEAU, J. (1978) When does a response error become a judgemental bias? Commentary on "Judged frequency of lethal events". Journal of Experimental Psychology: Human Learning and Memory, 4, 579-81
- SIEGAL, S. (1956) Nonparametric Statistics. New York: McGraw-Hill
- SIMPSON, M.E. & JOHNSON, D.M. (1966) Atmosphere and conversion errors in syllogistic reasoning. Journal of Experimental Psychology, 72, 197-200

- SLOVIC, P. & FISHHOFF, B. (1977) On the psychology of experimental surprises. Journal of Experimental Psychology: Human Perception and Performance, 3, 544-51
- SMALLEY, N.S. (1974) Evaluating a rule against possible instances. British Journal of Psychology, 65, 293-304
- SMEDSLUND, J. (1963) The concept of correlation in adults. Scandinavian Journal of Psychology, 4, 165-73
- SMEDSLUND, J. (1970) Circular relation between understanding and logic. Scandinavian Journal of Psychology, 11, 217-19
- STAUDENMAYER, H. (1975) Understanding conditional sentences with meaningful propositions. In R.J. Falmagne (Ed): Reasoning: Representation and Process. New York: Wiley
- STOUFFER, S.A., SUCHMAN, E.A., DEVINNEY, L.C., STAR, S.A. & WILLIAMS, R.M. Jnr. (1949) The American Soldier: Adjustment During Army Life (Vol 1). Princeton, N.J: Princeton Univ. Press
- STRAWSON, P.F. (1952) Introduction to Logical Theory. London: Methuen
- TAPLIN J.E. (1971) Reasoning with conditional sentences. Journal of Verbal Learning and Verbal Behaviour, 10, 219-25
- TAPLIN, J.E. & STAUDENMAYER, H. (1973) Interpretation of abstract conditional sentences in deductive reasoning. Journal of Verbal Learning and Verbal Behaviour, 12, 530-42
- THOULESS, R. (1959) The effect of prejudice on reasoning. British Journal of Psychology, 50, 289-93
- TSAL, Y. (1977) Symmetry and transitivity assumptions about a nonspecified logical relation. Quarterly Journal of Experimental Psychology, 29, 677-84
- TVERSKY, A. & KAHNEMAN, D. (1973) Availability: a heuristic for judging frequency and probability. Cognitive Psychology, 5, 207-32.

- UNDERWOOD, B.J, ZIMMERMAN, J. & FREUND, J.S. (1971) Retention of frequency information with observations on recognition and recall. Journal of Experimental Psychology, 87, 149-62
- VAN DUYNE, P.C. (1973) A short note on Evans' criticism of reasoning experiments and his matching response hypothesis. Cognition, 2, 239-42
- VAN DUYNE, P.C. (1974) Realism and linguistic complexity in reasoning. British Journal of Psychology, 65, 59-67
- VAN DUYNE, P.C. (1976) Necessity and contingency in reasoning. Acta Psychologica, 40, 85-101
- WARD, W.C. & JENKINS, H.M. (1965) The display of information and the judgement of contingency. Canadian Journal of Psychology, 19, 231-41
- WASON, P.C. (1960) On the failure to eliminate hypotheses in a conceptual task. Quarterly Journal of Experimental Psychology, 12, 129-40
- WASON, P.C. (1964) The effect of self-contradiction on fallacious reasoning. Quarterly Journal of Experimental Psychology, 16, 30-4
- WASON, P.C. (1965) The contexts of plausible denial. Journal of Verbal Learning and Verbal Behaviour, 4, 7-11
- WASON, P.C. (1966) Reasoning. In B. Foss.(Ed): New Horizons in Psychology. London: Penguin
- WASON, P.C. (1968a) Reasoning about a rule. Quarterly Journal of Experimental Psychology, 20, 273-81
- WASON, P.C. (1968b) "On the failure to eliminate hypotheses" - a second look. In P.C. Wason and P.N. Johnson-Laird (Eds): Thinking and Reasoning. Harmondsworth: Penguin

- WASON, P.C. (1969a) Regression in reasoning? British Journal of Psychology, 60, 471-80
- WASON, P.C. (1969b) Structural simplicity and psychological complexity: Some thoughts on a novel problem. Bulletin of the British Psychological Society, 22, 281-4
- WASON, P.C. (1971) Problem solving and reasoning. British Medical Bulletin. No. 3 (Cognitive Psychology), 206-10
- WASON, P.C. & EVANS, J.St.B.T. (1975) Dual processes in reasoning? Cognition, 3, 141-54
- WASON, P.C. & GOLDING, E. (1974) The language of inconsistency. British Journal of Psychology, 65, 537-46
- WASON, P.C. & JOHNSON-LAIRD, P.N. (1969a) Proving a disjunctive rule. Quarterly Journal of Experimental Psychology, 21, 14-20
- WASON, P.C. & JOHNSON-LAIRD, P.N. (1969b) Expectation, contradiction and insight. Paper presented at the XIX International Congress of Psychology.
- WASON, P.C. & JOHNSON-LAIRD, P.N. (1970) A conflict between selecting and evaluating information in an inferential task. British Journal of Psychology, 61, 509-15
- WASON, P.C. & JOHNSON-LAIRD, P.N. (1972) Psychology of Reasoning: Structure and Content. London: Batsford
- WASON, P.C. & SHAPIRO, D. (1971) Natural and contrived experience in a reasoning problem. Quarterly Journal of Experimental Psychology, 23, 63-71
- WINTHROP, H. (1946) Semantic factors in the measurement of personality integration. Journal of Social Psychology, 24, 149-75

WOOD, G. (1978) The knew-it-all-along effect. Journal of Experimental Psychology: Human Perception and Performance, 4, 345-53

WOODWORTH, R.S. & SELLS, S.B. (1935) An atmosphere effect in formal syllogistic reasoning. Journal of Experimental Psychology, 18, 451-60