University of Plymouth

PEARL **https://pearl.plymouth.ac.uk**

04 University of Plymouth Research Theses 01 Research Theses Main Collection

1991

LINGUISTIC AND SEMANTIC FACTORS IN CONDITIONAL REASONING

ELLIS, MATTHEW CHARLES

http://hdl.handle.net/10026.1/1831

http://dx.doi.org/10.24382/3285
University of Plymouth

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

LINGUISTIC AND SEMANTIC FACTORS IN CONDITIONAL REASONING

MATTHEW CHARLES ELLIS

A thesis submitted in partial fulfilment of the requirements of the Council for National Academic Awards for the degree of Doctor of Philosophy

February 1991

Polytechnic South West
Plymouth

POLYTECHNIC SOUTH WEST
LIBRARY SERVICES

Item No.	900066835-5
Class No.	T 160 ELL
Contl No.	X702400296

TELEPEN

90 0066835 5

REFERENCE ONLY

MATTHEW CHARLES ELLIS

•

LINGUISTIC AND SEMANTIC FACTORS IN CONDITIONAL REASONING

•

ABSTRACT

Two complementary sets of experiments were carried out to investigate how people understand, and reason with, truth-functional conditional forms. The first four experiments employed abstract materials and sought primarily to investigate the notion that people interpret conditionals like 'if P then Q' a 'fuzzy' way - that is, as if such a statement means something like 'if P then, generally speaking, Q'. Evidence was found that people will tolerate, or even expect, a certain degree of fuzziness in a conditional scenario such that they will be prepared to rate such a conditional statement as true despite the existence of a (small) proportion of falsifying counterexamples. Systematic effects of various extra-logical factors such as syntactic form and availability were also identified, and these were shown to be at their strongest when features of the reasoning task rendered it more difficult to perform.

The second four experiments employed thematic materials to investigate systematically the effect upon the understanding and use of the conditional of its presentation in a range of linguistic contexts. A tentative taxonomy of such contexts is proposed, in which eight linguistic contexts are variously distinguished in terms of such factors as temporality, causality, speaker's intention, direct control of speaker and hearer over the truth of the propositions and the non-contingent likelihood of the consequent. The systematic effects of linguistic context identified include the reliable observation of two non-traditional truth tables. It is argued that the 'natural' interpretation of the conditional in everyday language corresponds to material equivalence, and a mechanism is proposed whereby the suggestion by the scenario of various distinct forms of fuzziness might cause this interpretation to 'degrade' into variously defective forms. In these terms, the truth table traditionally referred to as the 'defective implication' truth table is argued to be more properly considered as a 'wholly defective' version of the material equivalence truth table, the two novel tables being 'partially defective' versions.

For Emma

Acknowledgements

I am extremely, and forever, grateful to Professor Steve Newstead, my Director of Studies, for his very considerable support, both academic and personal, over the six long years of struggle which underlie this work. This dissertation would not have seen the light of day without his invaluable advice and his constant readiness to re-embrace the project with enthusiasm and encouragement following each of the several occasions upon which I approached terminal inertia.

I am indebted also to Professor Jonathan Evans for his guidance in the early stages of the work and for his many illuminating suggestions as to how various aspects of my data might be interpreted, and to Dr Ian Dennis for his many incisive observations throughout and for steering me away from being drawn into a maelstrom of unworkable statistical analyses.

Finally, I should like to record my undying gratitude to my wife, Emma, for her unwavering support and encouragement over the last four years, for the many privations which she has cheerfully suffered in the cycle of frenetic euphoria and dark despond which is a PhD, for being a constant source of inspiration and delight and for never letting me forget her confidence that I would get there eventually (which might have flagged on occasions but, unlike my own, never evaporated completely).

Outline Contents

		page
Chapter 1		
Introduction and a Review of the Conditional Reasoning Literature		1
1.1	General Introduction	2
1.2	Research Using Abstract Materials	65
1.3	Three More Theoretical Frameworks	95
1.4	Research Using Thematic Materials	112
1.5	The Major Approaches Revisited	150
Chapter 2		
Experiments 1 to 4: The 'Fuzzy' Experiments		156
2.1	General Introduction	157
2.2	Experiment 1	158
2.3	Experiment 2	180
2.4	Experiment 3	207
2.5	Experiment 4	223
2.6	General Discussion	238
Chapter 3		
Experiments 5 to 8: The 'Contexts' Experiments		248
3.1	General Introduction	249
3.2	Experiment 5	268
3.3	Experiment 6	304
3.4	Experiment 7	327
3.5	Experiment 8	359
3.6	General Discussion	373
Chapter 4		
Synthesis and Conclusion		388
4.1	Fuzziness	389
4.2	Familiarity and Facilitation	390
4.3	Logical Competence and Rationality	396
4.4	Conclusion	398
References		399
Appendices		414

Detailed Contents

page

Chapter 1
Introduction and a Review of the Conditional Reasoning Literature 1

1.1 General Introduction 2

1.1.1 What is Propositional Reasoning? 9
1.1.2 What is Conditional Reasoning? 12
1.1.3 Conditionals and Truth Table Tasks 17
1.1.4 Conditionals and Inferences 20
1.1.5 Logic and Logical Competence 26
1.1.6 The Rationality Debate 29
1.1.7 Natural Deduction 34
1.1.8 Conditionals and 'Possible Worlds' Semantics 38
1.1.9 Representing Conditionals: Mental Models 41
1.1.10 Counterexamples and Fuzziness 55

1.2 Research Using Abstract Materials 65

1.2.1 Some Typical Paradigms 65
1.2.2 Inference Tasks - Some Typical Phenomena 66
1.2.3 Negativity Effects 70
1.2.4 Inference Tasks and Underlying Truth Tables 77
1.2.5 Truth Table Tasks- Some Typical Phenomena 77
1.2.6 Linguistic Factors and Directionality Effects 83
1.2.7 The Wason Selection Task 88

1.3 Three More Theoretical Frameworks 95

1.3.1 The Heuristics and Biases Approach 96
1.3.2 Evans' 'Heuristic and Analytic' Framework 98
1.3.3 Cheng and Holyoak's Pragmatic Reasoning Schemas 102

1.4 Research Using Thematic Materials 112

1.4.1 Abstract *versus* Thematic Materials 113
1.4.2 Language as Behaviour - Speech Acts & Implicature 115
1.4.3 Effects of Perceived Context 125
1.4.4 Effects of Belief 133
1.4.5 Thematic Facilitation in the Wason Selection Task 138

1.5 The Major Approaches Revisited 150

page

Chapter 2
Experiments 1 to 4: The 'Fuzzy' Experiments 156

2.1 General Introduction 157

2.2 Experiment 1 158

2.2.1 Design and Methodology 159
2.2.2 Results 162
2.2.3 Discussion 174

2.3 Experiment 2 180

2.3.1 Design and Methodology 181
2.3.2 Results 183
2.3.3 Discussion 198

2.4 Experiment 3 207

2.4.1 Design and Methodology 208
2.4.2 Results 212
2.4.3 Discussion 220

2.5 Experiment 4 223

2.5.1 Design and Methodology 224
2.5.2 Results 227
2.5.3 Discussion 235

2.6 General Discussion 238

Chapter 3
Experiments 5 to 8: The 'Contexts' Experiments 248

3.1 General Introduction 249

3.1.1 A Taxonomy of Usage 250
3.1.2 Overview of the four Experiments 262

3.2 Experiment 5 268

3.2.1 Design and Methodology 269
3.2.2 Results 276
3.2.3 Discussion 289

page

3.3 Experiment 6 304

3.3.1 Design and Methodology 305
3.3.2 Results 311
3.3.3 Discussion 319

3.4 Experiment 7 327

3.4.1 Design and Methodology 329
3.4.2 Results 332
3.4.3 Discussion 347

3.5 Experiment 8 359

3.5.1 Design and Methodology 361
3.5.2 Results 364
3.5.3 Discussion 369

3.6 General Discussion 373

3.6.1 The four 'Inducement' Contexts 373
3.6.2 The subjective Non-contingent Likelihood of the Consequent 374
3.6.3 Speaker's and Hearer's Control 375
3.6.4 X-tables and Y-tables - genesis and reality 376
3.6.5 E-tables - the canonical interpretation? 383
3.6.6 Uncertainty and Control 385

Chapter 4
Synthesis and Conclusion 388

4.1 Fuzziness 389
4.2 Familiarity and Facilitation 390
4.3 Logical Competence and Rationality 396
4.4 Conclusion 398

References 399

Appendices 414

Chapter One

Introduction and Review of the Conditional Reasoning Literature

1.1 : General Introduction

This thesis is concerned with the way that people reason with, and about, conditional statements - that is, statements of a form such as 'if *some-antecedent* then *some-conclusion*', in which the truth of the conclusion appears to be conditional upon the truth of the antecedent. It is appropriate to begin, therefore, by saying something about the nature of reasoning and about the kinds of conditional statement with which this investigation will concern itself. According to Craik (1943, as summarised by Johnson-Laird, 1983a), human reasoning comprises three distinct processes:

> " 1 A 'translation' of some external process into an internal representation in terms of words, numbers or other symbols.
>
> 2 The derivation of other symbols from them by some sort of inferential process.
>
> 3 A 'retranslation' of these symbols into actions, or at least a recognition of the correspondence between these symbols and external events, as in realizing that a prediction is fulfilled." (Johnson-Laird, 1983a, pp.2-3)

The most quintessential of these activities to a definition of 'reasoning' is the drawing of inferences, so as to generate new information which goes beyond that explicitly presented or known already. The other two processes stress the importance of an abstract internal representation to the current cognitive psychological view of man as a 'symbol processor'.

The conditional statements with which this investigation will be concerned are, necessarily, restricted in kind in terms of logical status, linguistic type, grammatical mood and tense. In terms of logical status, the investigation is restricted to conditional statements which are truth-functional - that is, those

in which (at least) the truth or falsity of the consequent is logically dependent upon the truth of the antecedent. An example should make this distinction clear: in the conditional statement "If you poke Fido then he will bite you", one might expect Fido's biting the hearer to be in some way logically connected to whether or not the hearer pokes him. This conditional is, therefore, truth-functional. On the other hand, in the conditional "If you're hungry then there's some cheese in the fridge", one would not expect whether or not there actually is some cheese in the fridge to be affected by whether or not the listener is hungry. This kind of conditional is not truth-functional and will not concern us further in this thesis.

In terms of linguistic type, the investigation will concern itself only with declarative sentences (that is, those which state something) as opposed to interrogative or imperative sentences. Our restrictions in terms of mood and tense fall into one of Kaluza's six kinds. Kaluza (1986) proposed a binary systematization of conditional sentences by dichotomising in terms of hypothetical *vs*indicative, past *vs* non-past and (for hypothetical conditionals only) factual *vs* non-factual. In terms of Kaluza's classification system, the investigation is restricted to conditional statements which are non-past (that is, those which use present or future tenses) and indicative (that is, those which express a fact, as in "If you poke Fido then he will bite you"), as opposed to subjunctive (as in "If you poked Fido then he would bite you"). In summary, then, the investigation is restricted to conditionals which are truth-functional, declarative, indicative and non-past.

Two separate sets of experiments will be presented, each of which belongs to a relatively distinct research tradition. The first four experiments employ 'abstract' materials, which are intended to be as devoid as possible of any

intrinsic 'real world' meaning. The remaining four experiments, in contrast, use 'thematic' materials, which are intended to reflect 'real' ideas, entities and relations - the sort of things about, and with, which people reason in their everyday life.

A number of features and concerns are common to both sets of experiments, however. At a superficial level, all eight experiments are similar in that they require subjects either to make or to evaluate conditional inferences based upon statements made in English. At a deeper level, all eight experiments are crucially concerned with the internal representations which underlie the inference processes. Most importantly, both sets of experiments have a common aim: the search for a quintessential underlying meaning of the logical connective 'if'[1] which might form the basis of conditional reasoning behaviour - irrespective of the type of task, or of whether the materials are abstract or thematic. Such a search might be expected, however, to be fraught with difficulty. As will be seen, it is essentially the very fickleness of 'if' which has given rise to the vast body of conditional reasoning research in which this present study has its roots. In the much-quoted words of Wason & Johnson-Laird (1972, p.92):

> "... conditionals are not creatures of a constant hue. Like chameleons ... they take on the colour suggested by their surroundings."

By using abstract materials in the first set of experiments, it was intended that should any 'colour' be suggested by the surroundings, then it would at least be constant. The use of thematic materials in the second set of experiments, in contrast, was intended quite explicitly to suggest certain rather particular 'colours', so that the chameleon-like quality of the

1 (in truth-functional, declarative, indicative, non-past conditionals, at least)

conditional might be examined.

One crucial issue underlies the whole of reasoning research, particularly (as will usually be the case) where the task involves reasoning with, or about, verbal statements of a 'logical' nature, such as 'If *such-and-such* then *some-conclusion*'. The point about tasks such as these is that they will in general be capable of being represented in terms of some formal logical system (such as the propositional calculus, which will be introduced shortly) and, more importantly, that they will have a *solution* in terms of that logical system. To the extent that a subject's solution accords with this 'logical' solution, it may, therefore, be tempting to say that they[2] have performed in a 'rational' way - and *vice versa*. The question of rationality is at issue throughout the whole of cognitive psychology and thus will appear in a number of guises. In so far as it arises in the field of reasoning, however, the 'Rationality Debate', as it has become known, is essentially about whether, given their particular individual interpretations of the situation, people reason 'logically' (that is, in accordance with the rules of some given logical system) or, alternatively, their behaviour is determined by certain *extra-logical* factors.

Before considering the question of 'logical' behaviour in verbal reasoning, it is necessary to introduce in some detail the concepts of propositional reasoning in general, and of conditional reasoning in particular, within the framework of the propositional calculus, a formal system of logic which underlies the tasks which subjects are typically asked to perform in this

2 The universal use 'he' and its masculine bedfellows will be avoided wherever practicable in this thesis. This will not be achieved by the use of bizarre neologisms such as 's/he', nor by the distracting device of sprinkling the two genders more or less evenly but, rather, by the use of the ungrammatical - but idiomatically quite acceptable - 'they'.

field. A crucial distinction will then be drawn between logical *competence* and logical *behaviour*. It will then be possible to return to the Rationality Debate in rather more detail.

Whilst this introduction to propositional logic and conditional reasoning will consider both the mental representation of the relationships between propositions and the processes which act upon those representations, it will in the main be rather more concerned with the *effect* of particular representations rather than their essential *nature*. In the section which follows our consideration of the Rationality Debate, a qualitatively rather different system of logic is described - 'natural deduction' - which is concerned with the sorts of mental representations of 'logical' relationships which people 'possess', and the procedures which they bring to bear upon them, in virtue of which they behave as they do in reasoning tasks.

This is followed by a brief consideration of an important alternative approach to the problem of finding an unequivocal semantics for 'if' which has its roots in linguistic philosophy - the use of 'possible worlds' semantics. This, in turn, leads into a description of Johnson-Laird's (1983a) theory of 'mental models' - a major theoretical framework which proposes a system of symbolic representation and procedural semantics whereby 'reasoning' might proceed without recourse to any sort of logic whatever.

Central to the mechanics of reasoning under Johnson-Laird's theory of mental models is the notion of the search for falsifying counterexamples to a conclusion. The four abstract experiments were directly involved with the manipulation of visual arrays which contained varying proportions of falsifying items, such that a conditional statement about the array was at

times generally, but not absolutely, true. The four thematic experiments investigated various effects of linguistic context and content upon the interpretation of 'realistic' conditional statements. As will be argued later, it may be shown that conditional statements in the real world are usually only *probabilistically* true, in that they will generally be considered perfectly reasonable and earnest statements to make even when the speaker knows that the consequent only follows from the antecedent most, rather than all, of the time. Whilst subjects in these thematic experiments are not, of course, presented with arrays of positive and negative exemplars, it is nevertheless reasonable to assume that (in some particular circumstances) they will have a certain tolerance of counterexamples, since that accords with what naturally occurs in the real world. In both sets of experiments, then, the investigation is, in some sense, concerned with 'fuzziness'. In the abstract experiments, fuzziness is explicitly represented as particular proportions of positive exemplars and counterexamples; in the thematic experiments, it may be expected to manifest itself indirectly as a tolerance of counterexamples given certain contextual conditions. It is appropriate, therefore, to round off the general introduction to conditional reasoning with a consideration of research concerned with reasoning about situations and/or assertions which are more or less 'fuzzy'.

Following this general introduction to conditional reasoning research, the typical paradigms and phenomena to be found in the 'abstract' literature are described. It will become clear from this section that whilst subjects do appear to behave relatively systematically on reasoning tasks, only sometimes will their behaviour accord with formal logic - so that their behaviour must be at least partially determined by 'extra-logical' factors.

Before moving on to the 'thematic' literature, it will be appropriate to introduce briefly three further theoretical frameworks proposed to account for the complex determination of human reasoning behaviour: the 'heuristics and biases' approach (originating from the work of Tversky and Kahneman, 1971, *et seq.*), Evans' (1984) 'heuristic and analytic framework' and Cheng & Holyoak's (1985) 'pragmatic reasoning schemas'.

Each of these accounts will be referred to from time to time in the section which follows, which reviews the 'thematic' conditional reasoning literature, which is concerned with the effects of content and context. The typical phenomena will be described and contrasted with those from the 'abstract' literature. Additionally, the notion of 'language as behaviour' will be introduced as the basis for a review of the particular literature concerned with the manipulation of linguistic context.

The chapter concludes with a comparative summary of the various theoretical approaches described and a consideration of the extent to which each is able to account for the phenomena identified.

1.1.1 : What Is Propositional Reasoning?

Propositional reasoning might be characterised as a set of methods (rules of inference) for going from given items of information (premises) to another item of information (the conclusion). Early reasoning research (for example, Chapman & Chapman, 1959; Henle, 1962) tended to be based upon Aristotelean logic. As Strawson (1952) has pointed out, however, there are certain quite ordinary and unremarkable inferences which are not supported under Aristotelean logic[3]. More recent research has tended to be characterised by its use of tasks based on 'modern' propositional logic and first-order predicate logic, such as that formalised by the analytic philosopher Gottlōb Frege (1879). A diversity of propositional logics have followed in Frege's wake (see, for example, McCawley, 1981) but common to all such logics (and to propositional reasoning) is that they are about *propositions*. For our present purposes, a proposition will either be a well-formed, grammatical statement in English or else a variable symbol (typically 'P' or 'Q') which will stand for such a statement. Formally, a proposition will be what an analytic philosopher would call "an internal sentence token in logical form". A crucial requirement of a proposition is that it be capable of having a truth value. At the very least, a proposition must be able to be said to be 'true' or 'false': as McCawley (1981, p.3) puts it:

> " Since propositions are whatever logic is done on, if logic is to provide an account of truth and inference, propositions must be things that can be said to be true or false and must also be things that can serve as premises or conclusions of inferences."

It will be seen later, however, that there are circumstances in which at least

[3] Braine (1978) offers the example of the inference from *'All F are either G or H'* to *'All F that are not G are H'*

one further truth value - 'indeterminate' (neither true nor false) - must be admitted. Put simply, the sentence

[1] 'Emma likes the cat'

is a statement of a single proposition, which in this case is (trivially)

[2] *Emma likes the cat*

which is clearly capable of being either true or false. A simple proposition such as this is known as an 'atomic proposition'. At the lowest level of analysis, propositions are always affirmative, so that the sentence

[3] 'Emma doesn't like the cat'

should be seen as containing the same atomic proposition,

[2] *Emma likes the cat*

but now it is explicitly negated by the use of 'not'. Note, however, that whether or not an atomic proposition is negated is quite independent from whether the sentence containing it is true or false, since both [1] and [3] are clearly equally capable of being true or false.

Finally, atomic propositions might be merged with each other so as to form compound propositions, so that the compound proposition

[4] *Emma likes the black cat*

may be analysed into two atomic propositions,

[2] *Emma likes the cat*

(as before) and

[5] *The cat is black*

For our purposes, propositions may be either simple or complex and, at the same time, either negated or unnegated. Propositions may be related to each other in particular ways. The six most important relations in propositional logic are shown in Table 1.1, together with the notation which will be used to denote them throughout this thesis. In a formal system of propositional logic, the meaning of these relations - and of the rules of inference concerned with their interrelationship - might be precisely defined as a set of *axioms* which are, in effect, templates which state precisely what constitutes a well-formed formula of that logical system and what follows given certain premises expressed in the language of that formal system. Table 1.1 offers for each relation a putative natural language definition of its meaning. As will be seen, however, there is not always a one-to-one mapping between the natural language definition and this precise definition in terms of axioms in the formal calculus.

relation	notation	putative definition
negation	~P	not P
conjunction	P & Q	P and Q
inclusive disjunction	P v Q	P or Q *(or both)*
exclusive disjunction	P / Q	P or Q *(but not both)*
material implication	P -> Q	*at-least-if* P then Q
material equivalence	P <-> Q	if-*and-only-if* P then Q

Table 1.1 : The six basic relations of propositional logic

This is particularly true in the case of material implication and material equivalence, both of which will often be indistinguishably expressed rather more loosely in natural language as 'if P then Q', and it is this ambiguity

that is of central concern in conditional reasoning research.

1.1.2 : What Is *Conditional* Reasoning?

Essentially, conditional reasoning is that class of propositional reasoning in which the truth of one proposition is dependent - or conditional - upon the truth of the other. In natural language usage, the most commonly encountered expression of this sort of contingent relationship is the linguistic form 'if P then Q' (where P and Q each stand for some proposition).

In propositional logic, there are two different ways in which the truth values of P and Q might be related to each other conditionally. The first such relation is known formally as 'material implication'. In this case, the truth of the antecedent proposition (P) implies the truth of the consequent proposition (Q), so that whenever P is true, we know that Q is true also. This is not, however, to say that Q cannot be true quite independently of P - so that knowing that Q is true cannot tell us anything definite about the truth of P. Furthermore, since Q might be true independently of P, knowing that P is false tells us nothing about the truth of Q. Note, however that the *falsity* of Q does tell us something about P: knowing that Q is not true must tell us that P cannot be true either, since if P were to be true, then Q would be true also - and if P can't be true, then it (P) must be false.

The second such relation is known formally as 'material equivalence', although it is sometimes referred to in the literature as the 'biconditional'. In material equivalence, the truth of each proposition (P and Q) is dependent upon the truth of the other, so that the truth of P implies that Q is true and, at

the same time, the truth of Q implies that P is true. Put another way, it is always either the case that both P and Q are true, or else that neither of them is true. Under this relation, then, knowing that either proposition is true tells us that the other is true also: furthermore, it is now also the case that knowing that either proposition is false tells us that the other is false too.

An alternative way of distinguishing between these two interpretations of the conditional is in terms of necessity and sufficiency. Under material implication, the antecedent, P, is *sufficient* for the consequent, Q - that is, knowing that P is true is sufficient information for us to know that Q is true also. P is not, however, *necessary* for Q, since we allow situations in which Q is true yet P is false. Under material equivalence, on the other hand, P *is* necessary for Q, since we do not allow a situation in which Q is true but P is false - note, however, that P is still sufficient for Q so that, strictly, we should now say that P is *necessary and sufficient* for Q.

In natural language usage, 'if P then Q' is far from the only way of expressing a conditional relationship. Whilst it may not be immediately obvious in every case, it turns out that every one of the following sentential forms is a logically consistent expression of material implication: 'P only if Q'; 'never P without Q'; 'not both P and not-Q'; 'either not-P or Q'; 'not-P unless Q'.

What may be less obvious still is that each of these expressions is also a logically consistent expression of material *equivalence* - that is, each of these expressions is *ambiguous* with respect to the particular relation which obtains between P and Q. Whilst unambiguous natural language expressions do exist (for example, 'if, and only if, P then Q' to express material equivalence)

a moment's reflection will confirm that in everyday discourse such phrases will generally be avoided as being somehow rather too clumsy and pedantic. This is not, however, to say that they sound wholly unnatural. Rather, they tend to have a somewhat 'legalistic' feel to them - and the typical use of such phrases in legal documents and scientific papers is precisely so as to avoid the ambiguity of their less cumbersome counterparts.

Whilst each of these sentential forms is, strictly speaking, ambiguous as to whether it expresses material implication or material equivalance, it will be rare indeed for this ambiguity to be apparent to the hearer of a given utterance. Rather, a given conditional statement will tend quite naturally and smoothly to suggest that a particular relationship holds between P and Q - and just what that suggested relationship might be will typically depend upon such things as the content of the propositions, the sentential form in which the sentence is phrased and the linguistic context in which it is uttered.

The importance of the sentential form, and the various effects of content and context, are addressed at length later in this chapter. It may be helpful to offer a number of typical illustrations of these effects at this point, however, so as to demonstrate the fact that conditional statements do indeed appear to be far from consistent in their meaning and applicability[4].

The first example uses the sentential form 'if P then Q' to express the material implication relationship between sparkling wine and country of origin in the conditional statement "if the sparkling wine is genuine

4 Note that just how a given subject will interpret a given conditional statement is an empirical question: the suggested interpretation offered in these examples is for the purpose of illustration only.

Champagne then it's made in France". In natural language usage, this quite natural-sounding statement would indeed appear to suggest a material implication relation between P and Q, in so far it seems to express the fact that a given sparkling wine can only be genuine Champagne (P) if it comes from France (Q), but this is not to say that it might not equally well be a méthode Champenoise wine (not-P) from another region of France (Q). In terms of necessity and sufficiency, being genuine Champagne (P) is sufficient to say that it comes from France (Q), but it is not necessary that a sparkling wine be genuine Champagne for it come from France (so P is sufficient, but not necessary, for Q).

In the example given, the form 'if P then Q' appears to stand quite naturally (and unambiguously) for material implication. Changing the content of the propositions might, however, give rise to quite a different interpretation. Take, for example, a (logically isomorphic) statement issued by a mother to her son "if you do the washing-up then I'll let you borrow the car". It should be clear that in natural language usage, this equally natural-sounding statement may now tend to be interpreted as a material equivalence, in so far as it would now seem to suggest that whilst doing the washing-up (P) will mean that the son gets to borrow the car (Q), yet failing to do it (not-P) will now mean that the loan of the car will be withheld (not-Q). In terms of necessity and sufficiency, doing the washing-up (P) is now necessary (and sufficient) for the loan of the car (Q).

Whilst each of these statements appears to suggest quite a different relationship between P and Q, it is clear that neither of them seems at all ambiguous about that relationship. The ambiguity is rather more apparent in the statement "if we use aerosols containing CFCs then the ozone layer will be irreparably damaged". In this case, it is by no means clear what the

statement has to say about what will happen if we stop using aerosols containing CFCs (not-P) - that is, we are not told enough to feel certain about whether this would prevent irreparable damage (not-Q) or, alternatively, whether other potential causes would continue to operate such that the ozone layer would be irreparably damaged anyway (Q). In terms of necessity, we simply do not know whether the use of aerosols containing CFCs (P) is a necessary condition for irreparable damage to the ozone layer (Q).

In none of the above examples has the particular sentential form used ('if P then Q') seemed to be in any way unnatural. This universal applicability will not be the case for all of the sentential forms presented above, however. For example, it seems totally unnatural (or even nonsensical) to say "we don't use aerosols containing CFCs unless the ozone layer will be irreparably damaged", although this is logically equivalent[5] to "if we use aerosols containing CFCs then the ozone layer will be irreparably damaged". Part of the bizarreness of this sentence comes, of course, from the inappropriate retention of the verb tenses, in as much as the original statement presumably has the use of aerosols as a *cause* of the damage, and causes must obviously precede their effects. Even with the tenses suitably massaged, however, the sentence appears to be a pretty odd thing to say: "we won't have used aerosols containing CFCs unless the ozone layer has been irreparably damaged".

It has been demonstrated above that the relationship between P and Q which is suggested by a particular sentential form of the conditional may depend (at least) upon the content of the propositions and, furthermore, that a given

5 (in so far as the sentential forms are equivalent and the contents of each proposition are identical in each form)

sentential form may be somehow more or less natural for the expression of a given relationship in given circumstances. This is precisely what Wason & Johnson-Laird (1972) had in mind in likening the conditional to a 'chameleon' and it is the search for the elusive nature of this chameleon which motivates the present study and, indeed, much of the conditional reasoning research which has gone before it.

1.1.3 : Conditionals and Truth Tables

One way of looking at how a person interprets a given conditional statement is to infer from a set of responses the *truth table* which that subject in some sense 'possesses' for that statement. A conditional statement 'if P then Q' contains two propositions, P and Q. As has been said, propositions are things which can be said to be either true or false, so that there are four possible combinations of P and Q in terms of their truth - P alone might be true, Q alone, both or neither. These situations are traditionally denoted by the letters T (for true) and F (for false), so that TT denotes an instance in which both P and Q are true, TF an instance where P is true but Q is false, FT an instance where P is false but Q is true, and FF an instance where both P and Q are false.

In the simplest type of 'truth table task', a subject is given a conditional statement such as 'if P then Q', and then presented with each of these four combinations in turn and asked whether it shows the conditional statement to be true, or false[6]. For example, a subject might be given the statement "if

6 This is the 'evaluation' type of 'truth table task'. In the alternative, 'construction', type of truth table task, a subject is asked to identify one (or many) instances which would show a given conditional statement to be true and/or one (or many) which would show the statement to be false.

the symbol is red then it is triangular" together with a picture of a coloured symbol - which might perhaps be green and triangular - and asked whether the picture shows the statement to be true or false. In this case, then, the instance presented is the FT case, where P (red) is false, and Q (triangular) is true. From the set of a subject's responses to all four combinations (TT, TF, FT and FF) the truth table which they 'possess' for a given statement might be inferred. The truth table will consist of the ('T' or 'F') responses given by the subject to the TT, TF, FT and FF cases presented, as in Table 1.2. So far, only two possible responses have been considered - and, indeed, this is all that some studies have permitted the subject to make. Other studies, however, have permitted a third type of response (such as '?', 'indeterminate', 'irrelevant', 'don't know' or 'can't say') which thus corresponds to the third 'indeterminate' truth value mentioned earlier.

Wason (1966), following Quine (1952), argues that this third truth value (which Wason terms 'irrelevant') is required in virtue of the fact that there are occasions upon which subjects wish to indicate that they are unable to *apply* the statement to the instance (or *vice versa*) and offers by way of example the statement "if it is a table then it has four legs". Where the instance presented is indeed a table (in the TT and TF cases), then the statement is clearly applicable, and in order to establish its truth the subject merely has to observe whether the table does indeed have four legs. The problem arises where the instance isn't a table (in the FT and FF cases). Wason argues that the statement now no longer applies - it says nothing *relevant* about this instance as it only talks about things which are tables. On this basis, he proposes that instead of one of the truth tables which correspond to material implication or material equivalence (which are determinately true or false for each truth case), a subject might have what he terms a 'defective' truth table (which includes the indeterminate

response usually denoted as '?').

given : "if P then Q"			I-table	E-table	D-table
P	Q	truth case	material implication P -> Q	material equivalence P <-> Q	Wason's 'defective' *(non-logical)*
true	true	TT	T	T	T
true	false	TF	F	F	F
false	true	FT	T	F	?
false	false	FF	T	T	?

Table 1.2 : Truth tables : I-table - material implication
E-table - material equivalence
D-table - Wason's (1966) 'defective' table

Table 1.2 shows the two 'logical' truth tables together with Wason's 'defective' table. Since these - and other - truth tables will be the subject of considerable discussion throughout this thesis, they will often hereafter be referred to more readably as I-tables, E-tables and D-tables respectively. In addition, Table 1.2 also shows the propositional logic notation applying to I-tables (P->Q) and E-tables (P<->Q); D-tables do not correspond directly to any relation in propositional logic.

It will be noted that there is no distinction whatever between any of the tables in terms of the TT and TF cases - that is, the two cases where the antecedent (P) is true. Furthermore, the 'logical' tables, I-tables and E-tables, are distinguished only on the FT case.

1.1.4 : Conditionals and Inferences

All that has been addressed so far is the sort of interpretation that subjects might arrive at given a conditional statement - at their *representation* of the situation which it describes. In considering conditional reasoning, however, we are also concerned with processes of *inference*. As has been stated above, rules of inference in formal propositional logic might be stated as axioms, which might be seen as templates which specify precisely what follows from certain given information. The four 'classical' inferences of formal logic each take two pieces of given information (premises) and from them deduce a conclusion. In an 'inference task', a subject is given a first premise (a conditional statement - in Table 1.3 this is always of the form 'if P then Q', although other forms may, of course, be used), a second premise (an assertion about the truth of P or the truth of Q) and a conclusion, and is asked to say whether or not the conclusion follows from the premises[7]. The four 'classical' inferences, Modus Ponens (MP), Modus Tollens[8] (MT), Denial of the Antecedent (DA) and Affirmation of the Consequent (AC) are set out in Table 1.3. Table 1.3 also shows the propositional logic notation for each inference rule. For example, the notation for Modus Tollens, [**->(->PQ, ->(~Q,~P))**] may be read as "the fact that P implies Q [**->PQ**] implies that not-Q implies not-P [**->(~Q,~P)**]".

---

7 This is the 'evaluation' type of 'inference task'. In the alternative, 'construction', type of inference task, a subject is given a first premise (a conditional statement), a second premise (an assertion about the truth of P or the truth of Q) and asked what (if any) conclusion follows. A 'hybrid' inference task is also reported, in which the argument is presented with a range of possible conclusions from which a subject must select one (or all) which follow (if any).

8 More properly, this inference is known as *Modus Tollendo Tollens*, but the shorter term is rather more widely found in the literature and so will be used throughout this thesis.

inference		procedure	
MP:	->(->PQ, ->(P,Q))		
	'Modus Ponens'	*given the conditional*	if P then Q
		and the assertion	P
		infer	Q
MT:	->(->PQ, ->(~Q,~P))		
	'Modus Tollens'	*given the conditional*	if P then Q
		and the assertion	not Q
		infer	not P
DA:	->(->PQ, ->(~P,~Q))		
	'Denial of the Antecedent'	*given the conditional*	if P then Q
		and the assertion	not P
		infer	not Q
AC:	->(->PQ, ->(Q,P))		
	'Affirmation of the Consequent'	*given the conditional*	if P then Q
		and the assertion	Q
		infer	P

Table 1.3 : The four 'classical' inferences of propositional logic

Before considering the validity of each of these inferences under each of the three truth tables presented so far, it is necessary to clarify what shall be meant by 'valid'. According to McCawley (1981, p.70), a proposition is valid if it is "true in all states of affairs", whilst an inference is "valid if its conclusion follows from its premises": unfortunately, this appears to beg the question of what is meant by "follows" in this context. Johnson-Laird (1983a, p.28) is rather more helpful:

> " The fundamental semantic principle governing all logic is that an argument is warranted provided that there can be no counterexample to it - that is, an inference is valid if there is no interpretation of the premises that is consistent with a denial of the conclusion".

This definition of 'valid' is consonant with, and perhaps helpfully illustrated by, one of the classical methods of proof in logic - *reductio ad absurdum.* A *reductio ad absurdum* proof requires that we temporarily suppose the negation of the conclusion and then use that negated conclusion in an attempt to derive a contradiction. If we can indeed derive a contradiction, then the negated conclusion is false, and so the original (unnegated) conclusion must be true (since if a thing is not permitted to be false, then it must be true). Given, for example, the conditional statement 'if P then Q' and the fact that Q is false, we are able to make the Modus Tollens inference:

given if P then Q
and not-Q
———
infer not-P

In order to use a *reductio ad absurdum to* prove this inference valid, we temporarily suppose that P (the negation of the conclusion, not-P) is true. From this, we can make the Modus Ponens inference:

given if P then Q
and P
———
infer Q

We know, however, that not-Q is actually true (because it was given to us as a fact at the outset), and so this (derived) Q is clearly a contradiction. The (temporarily supposed) P must therefore be false: if P is false, then (almost tautologically) not-P must be true, and so the original inference has been

proven valid.

The need to spell out what shall be meant by 'valid' arises out of the presence in certain truth tables of some indeterminate truth values. So far, it is only the D-table to which this observation applies: later in this thesis, however, two further partially indeterminate truth tables will be identified.

For the present purposes, the validity of an inference will only be considered with respect to a given truth table, so that McCawley's "all states of affairs" will refer merely to each of the four possible truth cases under that truth table. An inference will be considered to be 'valid' if *and only if* the counterexample to its conclusion is expressly forbidden by that truth table. Note that it is not necessary that a truth table expressly *permit* a given conclusion, since to expressly forbid the counterexample is to insist that the conclusion itself be true - that is, a conclusion about which a given truth table is indeterminate will be taken to be valid provided that the truth table is determinate about the falsity of the counterexample to the conclusion. Wherever the counterexample to the conclusion is permitted, the inference will be considered to be invalid. Where a truth table is indeterminate about the counterexample to the conclusion of an inference, then that inference will be considered to be of *indeterminate* validity under the truth table in question (irrespective of whether the conclusion itself is expressly permitted).

The validity of each of the four 'classical' inferences under each of the three truth tables presented so far is now considered. The first of these, Modus Ponens (MP), might appear so trivially simple as not to require any *reasoning* at all: indeed the formal symbolic expression of this inference, [->(->PQ, ->(P,Q))] could be said to be tautological in so far as its conclusion is

little more than a restatement of its antecedent. MP does not distinguish between any of the truth tables set out in Table 1.2, in that each table has TT as true and TF as false - that is, none of them allows us to have a true antecedent with a false consequent. If the consequent can never be false when the antecedent is true, then it must in those cases be true, and so the MP inference is demonstrably valid for all three truth tables.

On similar grounds, Modus Tollens (MT) is also valid for all three tables. This can again be demonstrated by considering the forbidden TF case - if the antecedent can never be true when the consequent is false, then it must in those cases be false, and so the MT inference is also demonstrably valid.

The Denial of the Antecedent (DA) is clearly valid for E-tables, which hold FF to be true and forbid FT, so that a false antecedent can only ever occur with a false consequent (so that knowing that the antecedent is false tells us that the consequent must be false also). DA is equally clearly *invalid* for I-tables, which permit FT as well as FF, so that a false antecedent may now be accompanied by either a true or a false consequent (and so knowing that the antecedent is false cannot now tell us anything about the truth of the consequent).

For D-tables, however, the validity of the DA inference is less clear cut, since the D-table is *indeterminate* about the truth of the consequent whenever the antecedent is false (that is, it records no determinate truth value for either the FT or the FF case). Under the definition of validity set out above, we must say that the DA inference is of *indeterminate* validity under the D-table, since the D-table is silent with respect to the truth of the counterexample (FT).

Finally, the Affirmation of the Consequent (AC) is clearly valid for E-tables (where TT is supported and FT is forbidden) and invalid for I-tables (in which the true consequent is allowed to occur with both true and false antecedents in TT and FT). Here again, the D-table does not expressly forbid the (FT) counterexample to the conclusion and so, here again, we must say that the AC inference is of *indeterminate* validity under the D-table.

P	Q	truth case	I-table material implication	E-table material equivalence	D-table Wason's 'defective'
true	true	TT	T	T	T
true	false	TF	F	F	F
false	true	FT	T	F	?
false	false	FF	T	T	?

inference	true antecedents	valid for I-tables	valid for E-tables	valid for D-tables
MP	TT, TF	√	√	√
MT	TF, FF	√	√	√
DA	FT, FF	x	√	?
AC	TT, FT	x	√	?

Table 1.4 : Relationship between truth tables and inferences

The relationship between the truth tables and the inferences supported by each are shown in Table 1.4, which thereby makes explicit predictions about the sort of inferences which a subject with a given truth table ought to be expected to endorse[9]: 'I-table subjects' should endorse only the MP and MT inferences, 'E-table subjects' should endorse all four inferences and 'D-table

9 (or *vice versa* - that is, the truth table which a subject might be expected to 'possess' given a particular pattern of inference endorsement)

subjects' should endorse MP and MT but classify the DA and AC inferences as 'indeterminate' (where such a response is permitted) or else (where their responses are constrained to 'yes' or 'no') fall somewhere between rejection and endorsement of these inferences overall.

1.1.5 : Logic and Logical Competence

Logic and reasoning are long-established, if not always comfortable, bedfellows. In the nineteenth century, logicians such as Boole (1854) and Mill (1874) considered logic to provide the very laws of reasoning. As will be seen when the Rationality Debate is described, however, later philosophers (such as Ayer, 1936) strongly rejected this notion and preferred to see logic as only normative.

In the context of reasoning research, questions of logic and logical competence usually refer to some system of logic which subjects in some sense 'possess'. The possession of such a system implies the logical *competence* to solve any task amenable to solution within that particular logical system. This is not to say, however, that all such problems will be solved: solutions are in general seen as subject to some performance system, by virtue of which extra-logical factors may influence the solution. The distinction between competence and performance has its origins in the field of linguistics, in the work of Noam Chomsky (1957), who saw (linguistic) competence as being expressed by a grammatical theory 'possessed' by a subject.

In the field of human reasoning, the notion of logical competence is somewhat problematic, as it would appear to defy any form of non-circular

test except in circumstances where the performance system can be fully specified (and performance factors tend to be highly task-specific). The problem is not merely that failure to solve a problem does not imply that the subject is not competent to solve it: Pollard (1979) stresses that a correct solution does not necessarily imply logical competence either, as such a solution could conceivably arise through the application of extra-logical heuristic strategies, or simply be a correct guess - we must strive, therefore, to explain *how* a subject reaches a correct solution.

The doctrine of mental logic, in so far as it concerns psychology rather than philosophy, has its roots with the Piagetian school. Piaget (for example, in Inhelder and Piaget, 1958) proposed the existence of a competence system based in formal logic: an understanding of logic was seen as growing in the child through various stages, culminating in the *concrete operational stage* and, finally, the *formal operational stage.* In the concrete operational stage, the child is seen as learning to apply logical operations in particular contexts, such learning resulting in the growth of a general abstract system, so that in the formal operational stage, the child is able to represent a given situation propositionally and to apply formal logical rules to such a representation. There are, however, a number of problems with this doctrine. Firstly, there is the question of whether even adult human reasoners are in fact competent in terms of a formal logical system: as Wason (1977) points out, adults performing reasoning tasks tend not to appreciate such formal notions as *reversibility* - neither do they apply the combinatorial analysis which is required within a formal system.

Secondly, Piaget's doctrine predicts a *general* competence: Wason (1977) points out, however, that this is rather at odds with the highly task-specific

nature of reasoning performance and the fact that - as will be seen later in this chapter - performance is generally only particularly good where thematic, rather than abstract, materials are employed. To be fair to Piaget, however, he does seem to retreat somewhat on the generality question later in his work (for example, Piaget, 1972), where he still holds that we possess a general logical competence but admits that it is only likely to be manifested where we are interested, and experienced, in the content of the problem, in which case abstract tasks will clearly tend to underestimate competence.

Thirdly, as Evans (1982) points out, formal logic does not correspond with the everyday use of natural language: for example, in formal propositional logic the statement

(2 + 2 = 5) -> (the world is square)

is true, but it is nonsensical and devoid of meaning in everyday language. Fourthly, there is the need to account fully for the *acquisition* of such a system. Falmagne (1980), for example, has suggested that learning logic is akin to learning language (but, as Johnson-Laird, 1983a, points out, adults simply do not display consistent logical behaviour from which a child might learn). Piaget proposes that children internalise their own actions and reflect upon them (but this is circular and totally untestable). Fodor (1980) suggests - not unlike Chomsky's (1957) notion of a *universal grammar* - that the system is innate (but such a claim could only ever gain support from the failure of any other attempt so far to account for the process of acquisition).

The notion of logical competence, then, is fraught with difficulties: most damning of all, perhaps, are the simple facts that people *do* make 'errors' in

their reasoning performance and that such 'errors' are highly task-specific, context-specific and content-specific. The extent to which such 'errors' are properly so termed, however, is a major question in the Rationality Debate, a description of which now follows.

1.1.6 : The Rationality Debate

The field of conditional reasoning research might be seen as one of the bloodiest arenas of battle between two distinct schools of thought about the rationality of man. In a broader psychology, the rationalist approach has its modern roots in the work of such social psychologists as Kelly (1955), who espouses the view of 'man as scientist'. The basic notion underlying this view is that in going about our daily life, we are continually involved in constructing theories about the world, which we then test by making predictions and collecting evidence. The opposing camp is that of behaviourism - as epoused, for example, by Skinner (1938; 1972). Underlying this view is the idea that all our behaviour is determined by our environment and our particular personal history of reinforcement. In the rationalist approach, then, 'man as theory-tester' would require some internal system of deductive logic. The behaviourist view, however, requires no such system.

In the psychology of reasoning, each of these two extreme positions underlies a distinct view of behaviour on reasoning tasks. As was seen in the previous section, it has been argued that performance which is consonant with some formal system of logic is no certain indication that the subject possesses that formal system, and neither is failure so to perform on a particular occasion any indication that they do *not* possess such a system.

The fundamental difficulty in deciding whether a subject is behaving rationally was set out succinctly by Smedslund (1970), and is essentially concerned with the circular relationship between logic and understanding. If, on the one hand, we assume that a subject *understands* a problem, then we can determine whether or not they are behaving logically by looking at the *accuracy* of their responses. Conversely, we can assume that a subject will behave logically and thereby use the responses to measure their understanding of the problem. One way out of this dilemma was suggested by Evans (1972a) - the use of a multi-paradigm approach. The idea underlying this approach is that if a phenomenon is due to *logical interpretation*, then it should appear on each of a range of tasks where the same linguistic form is used. If, on the other hand, it is due to the way in which that subject carries out a particular reasoning operation, then it should appear whenever that particular operation is required, irrespective of the linguistic form involved.

Originally espoused by Henle (1962; see also Falmagne, 1975; Revlin and Leirer, 1978; Revlin and Mayer, 1978), the extreme rationalist view interprets this circularity in a rather different way. For the rationalists, reasoning processes are seen as fundamentally logical, so that variations in behaviour should be seen as arising solely out of subjects' logical interpretation of the situation. It is proposed that individual subjects' interpretations vary by virtue of such erroneous operations as *restatement* of a premise or of the conclusion, *omission* of a premise or *addition* of a premise.

Henle (1962) proposes one additional source of variation in behaviour - subjects might "fail to accept the logical task". This refers to a situation where the subject does not actually respond on the basis of any reasoning

operation at all, but rather, for example, evaluates a conclusion on its own merits, irrespective of the premises which are presented along with it. In this particular example, a subject's belief in the truth of the conclusion might effect the response.

The reader may have noticed that what subjects do in reasoning experiments has thus far been referred to as 'behaviour' and 'performance' - and that where the word 'errors' has been unavoidable, it has generally been placed in quotation marks.It is traditional in the reasoning literature to use the term 'errors' to refer to responses which fail to fit a particular representation within a particular formal system of logic. In the strictest case, where the I-table is held to be 'correct', then not only will various *random* response patterns be classed as 'errors', but also E-table and D-table patterns. In a rather more relaxed position, where I-tables and E-tables might be seen as perfectly reasonable alternative interpretations, D-tables are still likely to be referred to as being in some way erroneous. It is argued that to do so is nonsense. That a reasoning problem is *designed* on the basis of some formal logical system is not to say that a non-logician will prefer to go about its solution in a way which is consonant with that system. Formal systems of logic are artifacts manufactured by analytic philosophers and mathematicians for the description of particular formal relationships and operations - most importantly, they are largely intended as languages in which colingual practitioners might converse about matters close to their hearts. Psychologists might find such systems convenient for the description and specification of their reasoning problems, but there is no necessary basis for imagining that such systems will in any way be entailed in the behaviour of non-logicians in arriving at their solutions.

This view is broadly similar to that espoused by Cohen (1979; 1981; 1982), who holds that it is incorrect from a philosophical point of view to describe man as 'irrational', since to do so is to describe the behaviour of a subject in terms of some preconceived norm. Cohen's essential objection is that to do so necessarily (but indefensibly) presupposes that the preconceived normative rule held by the subject is the same as that held by the experimenter. Quite apart from the impossibility of ascertaining in a non-circular way just what each of these norms might be, there would appear to be grounds upon which to suspect that they would, in fact, differ, since the subject will be carefully selected so as to be as näive as possible, whilst the experimenter (in virtue of having designed such a task) must, perforce, be comparatively well versed in propositional logic.

What, then, do these various viewpoints say about the possible sources in variation in behaviour on reasoning tasks? The rationalists say that variation occurs in the way in which subjects interpret the relation which obtains between antecedent and consequent - once interpreted, the *operations* of reasoning should be consistent across subjects and across tasks. Performance for the rationalists, then, is a direct reflection of the subject's interpretation of the situation. Additionally, it is allowed that there will be occasions when no reasoning occurs at all, but rather that subjects merely respond on the basis of the conclusion in its own right.

Non-rationalists typically admit idiosyncratic variations in interpretation, but hold in addition that various factors affect the operations carried out upon those interpretations. Some of these factors will have to do with the operations themselves (for example, essential variations in difficulty across the various rules of inference), whilst others will have nothing whatever to

do with the logic of the task - for example, Evans' (1972c) 'matching bias'[10].

Braine (1978) offers an interesting perspective upon the extreme rationalist position by identifying two essential additional demands of reasoning formally in the laboratory as contrasted with reasoning practically in everyday life. He argues that in everyday reasoning, a person will use all the information which they have at their disposal, whereas in formal laboratory reasoning, they are asked merely to say whether (or what) conclusions follow from premises. The first additional demand made by the formal task is that subjects must "compartmentalize" information - that is, restrict the information used in reasoning to just that which is explicitly contained in the premises. Secondly, they must "discover the minimum commitments of the premises as they are worded" (Braine, 1978, p.2). This means that they must actively disregard any and all of the implicit additional information which might have been available to them had the premises been part of an instance of 'everyday reasoning' - that is, information arising out of such mechanisms as 'conversational implicature' (Grice, 1967)[11]. Furthermore, they must resist making or endorsing 'invited inferences' (Geis & Zwicky, 1971)[11] - that is, inferences which are not necessarily logically entailed, but which are somehow 'invited' by the context in which the premises are presented.

Braine's point seems to cast further doubt upon the propriety of labelling a human reasoner "irrational" simply because their behaviour is not in

10 This phenomenon will be described in some detail later in this chapter. Suffice it to say at this stage that matching bias is a non-logical bias which manifests itself as a tendency to attend to, or to select, items which match those named in a rule, irrespective of their logical status.

11 These mechanisms are considered in some detail later, when we address the effects of context upon reasoning behaviour and consider the notion of 'language as behaviour'.

accordance with some formal system of logic: but this is a question which can never be resolved absolutely, since it is ultimately a matter for philosophical, rather than scientific, argument. What is indisputable, however, is that the effects of certain determinants of behaviour are at least *concordant* with formal logic, whilst those of other determinants are clearly *extra-logical*.

1.1.7 : Natural Deduction

So far, only one system of logic has been considered - the propositional calculus. However firmly rooted in antiquity such a system might be, this does not alter the fact that it is an artifact devised by mathematicians. Uneasy at the disparity observed between human reasoning behaviour and that dictated by propositional logic, a number of writers have been motivated to find an alternative to standard logic which is psychologically, rather than merely mathematically, valid.

An alternative approach, generically termed 'natural deduction' was initially proposed by Gentzen (1935/64), who sought "to set up a formal system that came as close as possible to actual reasoning" (Gentzen, 1964, p.288). In general terms, systems of natural deduction can lead to similar conclusions to the inference rules of a formal propositional calculus, but without the need for explicit knowledge of any formal proof procedure.

It has been shown that an inference is valid provided that there is no interpretation of the premises that is consistent with a denial of the conclusion. The approach which has been taken so far to decide whether a

given inference is valid on this basis is to apply the inference rule axioms of a formal system of propositional logic on the basis of some particular truth table. In natural deduction, validity is determined by the application of some method (known as a 'decision procedure') by which a systematic search might be made for some critical interpretation of the premises that is consistent with a denial of the conclusion. If the decision procedure fails to find such an interpretation, then the inference is valid (assuming, of course, that the decision procedure is correct and complete and that it is applied correctly).

	(1)	(2)	(3)	(4)
given:	A, B	A or B, not-A	A->B, A	A->B, not-B
infer:	A & B	B	B	not-A
	(5)	(6)	(7)	(8)
given:	A & B	A	not-A	not-not-A
infer:	A	A or B	A->B	A

Table 1.5 : Some inference schemata for propositional logic (after Johnson-Laird, 1983a, Table 2.2, p.29)

Typically, natural deduction is based on the use of *inference schemata,* a set of 'templates' which provide a conclusion by a process of direct matching and variable substitution. Some such schemata appear to correspond relatively closely to certain of the rules of inference, but share little of their economy, in that each logical connective will typically require the introduction of a separate set of schemata. Table 1.5 offers some examples of inference schemata for propositional logic - note that (3) corresponds to Modus Ponens

and (4) to Modus Tollens.

A number of researchers have used natural deduction as the basis for psychological theories of reasoning (for example, Johnson-Laird, 1975; Osherson, 1975; Ennis, 1976; Braine, 1978; Rips, 1983; Braine, Reiser and Rumain, 1984). Rips' (1983) ANDS computer model of deduction is representative of such approaches. ANDS (A Natural Deduction System) is a formal artificial intelligence model of propositional reasoning using "natural deduction rules" that manipulate propositions in a hierarchically structured working memory. The rules work in either a forward (from premises to conclusion) or backwards (from conclusion to premises) direction and allow suppositions to be introduced during the deduction process: formally, the rules are descended from the natural-deduction procedures pioneered by Gentzen (1935/64) and Jaskowski (1934), which typically produce proofs without axioms such that the conclusion is deduced directly from the premises. The central assumption is that

> " ... deductive reasoning consists in the application of mental inference rules to the premises and conclusions of an argument ... the sequence of applied rules forms a mental proof or derivation of the conclusion from the premises, where these implicit proofs are analogous to the explicit proofs of elementary logic ... both the routines and the proofs are claimed to correspond to those used intuitively by subjects who have not received formal training in logic." (Rips, 1983, p.40)

Additionally, it is assumed that the rules are available to the reasoner only on a probabilistic basis: an assumption in line with Newell's (1980) prediction of ranges of possible behaviour by a given subject on a given task. Rips' model bears a resemblance to Newell's (1980) problem-solving approach in so far as the deduction system is seen as being at least isomorphic to a production system of condition-action pairs and also in that a

complete specification of any piece of reasoning behaviour is seen as being provided by the trace (the sequential list of operations and variable bindings) of the rules used therein.

Rips argues that the use of suppositions is central to conditional reasoning: he argues that there is an "intimate connection" between his model's supposition-creating rules and conditional propositions in that sentences of the form 'if P then Q' can often be paraphrased as 'suppose that P; then Q'. He points out that Mackie (1973) and Rips & Marcus (1977) have given accounts of the ability of the word 'if' to evoke suppositions and that a supposition is required by a *reductio ad absurdum* argument.

The rules themselves are not claimed to be exhaustive in that there are arguments in classical sentential logic which the system simply cannot handle (and which it will declare to be invalid due its being written under the 'negation by failure' principle). Additionally, Rips suggests that there are "undoubtedly" deductive procedures which subjects use, but which are not part of the system as they do not happen to have been included in the subject protocols from which the rules were gleaned. Rips (1983) does set out to test his model empirically, by comparing the behaviour of the system with that of a human subject. This test of the model is, however, somewhat weakened by his having used the very same subject whose protocols were used to build the rule set in the first place - and so, not surprisingly, the system does behave in a largely consonant 'human' fashion. Nevertheless, the observed correspondence does at the very least demonstrate the *plausibility* of such a rule set as a basis for human conditional reasoning in so far as it is capable of producing typical reasoning behaviour which is consonant with at least one human subject's actual behaviour.

1.1.8 : Conditionals and 'Possible Worlds' Semantics

The search for an unequivocal semantics of 'if' began, in fact, not in psychology but rather in linguistic philosophy, with Ramsey's (1931) proposal of a procedure for evaluating conditional expressions. Ramsey's proof procedure took the form of a 'thought experiment'. First, add the antecedent of the conditional to your stock of beliefs; next, assess whether the consequent is true. If you believe there to be a necessary connection from the antecedent to the consequent then you should evaluate the consequent as true and thus the entire conditional as true. If, prior to adding the antecedent to your stock of beliefs, you already believe the consequent to be true, then it should remain a part of your stock of beliefs and, again, the conditional will be evaluated as true.

Stalnaker (1968) pointed out that as it stands, this procedure only works if you have no prior belief about the truth of the antecedent. Where you believe the antecedent to be true, then your evaluation of the conditional is simply equivalent to your belief about the truth of the consequent. Where, on the other hand, you believe the antecedent to be false, then in order to add it to your stock of beliefs, you need (temporarily) to modify some of that stock to maintain consistency . The conditional may then be evaluated as before - that is, in terms of the truth of the consequent. The modified procedure (which has become known as the 'Ramsey-Stalnaker thought experiment') thus runs as follows: first, add the antecedent of the conditional to your stock of beliefs; next, evaluate the conditional on the basis of whether the consequent is true.

Stalnaker's account of the semantics of 'if' involves the notions of 'possible worlds'. A conditional 'if P then Q' is true if, and only if, Q is true in that 'possible world' in which P is true but which otherwise differs minimally from the world in question (the actual world[12]). Where P is true in the actual world, then this 'possible world' is precisely similar to the actual world. Where the truth of P is either unknown to the reasoner or is known to be false, then the possible world should differ from the actual world just as little as is necessary for the addition of P to the world while maintaining consistency. In the special case that P is self-contradictory, then Q is evaluated in a special 'absurd' possible world in which everything, including the contradictions of everything, is true - and thus Q is true.

Stalnaker's theory has been heavily criticised for the way in which it deals with *counterfactual* conditionals - that is, those for which the antecedent, P, is actually false. According to Stalnaker, such a conditional should be evaluated in the (unique) possible world which is as close as possible to the actual world but differs (minimally) in that P is true and any resulting contradictions are avoided by additional modifications. A number of theorists (for example, Lewis, 1973; Kutschera, 1974; Pollock, 1976) take issue with Stalnaker's insistence that there is a unique possible world in which the counterfactual antecedent is true, and offer alternative accounts. Lewis (1973), for example, proposes a universe (without limit) of possible worlds arranged akin to a Ptolemaic universe of concentric spheres, such that the closer a sphere is to the world at its centre, the more similar the worlds comprising the sphere are to that world. Kutschera (1974), meanwhile, proposes a model in which the worlds which are possible under the

12 Strictly speaking, this is not the actual world but rather the world in which the conditional is being evaluated - frequently, of course, these will be one and the same.

condition that P is true are included in the worlds which are possible under the condition that Q is true.

Quite apart from the fact many theorists (including Lewis, 1973) have argued that counterfactuals are a distinct species of conditional which should not be compounded with indicative conditionals, all of these accounts suffer from a number of difficulties in consequence of their assumption of a 'possible-worlds' semantics. This assumption does no violence to their usefulness in elucidating conditionals from the analytic philosopher's point of view, in which endeavour they have made a significant contribution. Where they run into difficulty is in their suitability for the basis of a psychologically plausible account of the way in which human reasoners understand conditionals.

One major such difficulty is that in order for many of these accounts of possible worlds to work, the set of possible worlds must be infinitely large in size, such that it could not conceivably fit into a person's mind. Furthermore, certain of the theories require that the reasoner have an understanding of some relatively sophisticated philosophical or mathematical notions - for example, Braine (1978) points out that Kutschera's (1974) account would require that a person should have a concept of the power set of the set of possible worlds, which is clearly somewhat implausible.

Johnson-Laird (1983c, pp.23-24) rather neatly captures the irreconcilability of these 'possible worlds' accounts in the following *reductio ad absurdum* argument:

> " ... if Stalnaker, Lewis, or their colleagues, have defined the truth conditions of conditionals, then no-one can ever grasp them, and *a fortiori,* no-one can ever properly evaluate any conditional. Yet,

even little children are able to decide that certain conditionals are true (or false)."

1.1.9 : Representing Conditionals: Mental Models

Whilst formal logic, natural logics and 'possible worlds' semantics all have something important to contribute to our understanding of the conditional relation, none of them is able to offer anything approaching a complete specification of what people actually do in solving reasoning tasks in terms of the underlying knowledge, the representation of the task and the processes which act upon that representation. Formal logic certainly represents the structure of the conditional relation elegantly and unambiguously, but people simply do not always behave in a way which is consonant with the underlying logical structure and the inferential axioms of any formal system of logic. Furthermore, people with no logical training at all are able to solve certain reasoning problems with ease. Natural logics do take account of what people actually do, since they tend to be based largely on behavioural observations: they tend, however, to have little more basis for their schemata than that they make accurate (but *post hoc)* predictions about the behaviour from which they were derived in the first place. In addition, they tend to be far from complete and are often unable to make predictions about behaviour outside the operations which they embody explicitly. As was shown in the previous section, approaches based on 'possible worlds' semantics tend to be at best psychologically implausible and, at worst, psychologically impossible.

An alternative approach which may be sufficiently powerful to account for the range of observed behaviour, and yet flexible enough to be able to

overcome most of these objections, is Johnson-Laird's (1983a) theory of mental models, first espoused by Johnson-Laird (1980; 1981) and developed fully in the eponymous book "Mental Models" (Johnson-Laird, 1983a).

Rouse and Morris (1986) review the many and various guises in which the term "mental models" has appeared throughout cognitive science and elsewhere - for example, in the field of manual control (Conant & Ashby,1970; Wickens, 1984; Jagacinski & Miller, 1978) and in considering how people understand systems and devices (Norman, 1983; Brown & de Kleer, 1981; de Kleer & Brown, 1983; Williams, Hollan & Stevens, 1983; Kieras & Bovair, 1984). Such approaches are typified by Hayes' (1979) 'näive physics' and by many of the contributions to Gentner & Stevens' (1983)[13] - where the major concern is to explain common sense notions about such physical phenomena as electricity, heat and motion. Rips (1986) argues that such researchers belong to what he terms the 'figurative' school of mental models, in which the central concern is the use and development of domain-specific knowledge or common-sense beliefs in problem-solving and comprehension, rather than their representational format.

Of more relevance to this thesis are those approaches which Rips (1986) terms 'literal' rather than 'figurative'. In Rips' terms, a 'literal' mental model refers to a conception in which the model is held in a unique type of representation, which is quite unlike the usual propositions or networks and which promotes a kind of reasoning that is different from that under standard systems of logic or under probability theory[14]. As examples of

13 (which, confusingly, shares with Johnson-Laird's book the title "Mental Models")

14 Rips (1986) is at pains to point out that he does not suggest that the phenomena studied by 'figurative' approaches are illusory - rather, merely that no representational uniqueness is implied in this approach.

endeavours which have been held to necessitate such special-purpose types of representation, Rips (1986) cites the provision of a psychological semantics for natural language (Johnson-Laird, 1983a) and approaches to probabilistic reasoning (Kahneman & Tversky, 1982) and deductive reasoning (Johnson-Laird, 1983a; Johnson-Laird & Bara, 1984).

This consideration of mental models approaches now turns to Johnson-Laird's (1983a) theory of mental models in particular, and considers what it might have to say about how people reason with conditionals. Johnson-Laird seeks to bring together ideas from experimental psychology, logic, linguistics and artificial intelligence into a broad theoretical framework within which such cognitive activities as thinking, reasoning and language might be considered. His approach represents a synthesis of two extant notions: (1) Craik's (1943) notion that a representation is built up not only from information given, but also from perception and imagination, and (2) the concept of procedural semantics, which has its origins in artificial intelligence research (for example, Woods, 1967; Davies and Isard, 1972; Longuet-Higgins, 1972; Wilks, 1972). He suggests that a procedural semantics is used in the construction, manipulation and interrogation of mental models and offers experimental evidence (for example, Johnson-Laird, 1980; 1981) that in addition to representing discourse in a superficial linguistic format, subjects will typically build representations in the form of mental models. The theory is offered as a basis of an account of a range of cognitive activity of which propositional reasoning is but part. The fundamental principle of the theory so far as reasoning in general is concerned is that

> " ... reasoning consists in the construction of mental models on the basis of the premises, and the search for alternative models that

might render putative conclusions false ..."

(Johnson-Laird, 1983a, p.177)

It is clear from this that the mental models approach characterises 'man as scientist', in that putative models are actively tested by an explicit, active search for equally valid alternative models which might render the current conclusions false - a sort of *reductio ad absurdum.*

The theory is based on six principal assumptions, as follow:

1 the premises (or discourse) are represented by the construction of an integrated mental model in which the logical properties are not directly represented, but rather "emerge naturally as a consequence of the use of the expression in the search process" *(ibid.)* - it is clear from this assumption that in order to use the expression, subjects must be aware of its normal linguistic function

2 the conclusion drawn is based on novelty and parsimony; this is in line with Johnson-Laird's (1975) principle of semantic informativeness - that is, that a proposition is more informative the greater the number of contingencies it rules out

3 the order of information in the conclusion is based on a FIFO (first-in-first-out) model of human working memory

4 if the figure of the premises does not permit immediate integration then extra operations are necessary

5 the greater the load on working memory, the harder it will be to make an inference - for example, as a result of the extra operations posited in assumption (4)

6 (most controversially of all) ordinary individuals who have not been taught logic do not use inference rules. Instead, they

have

> "... one essential piece of semantic knowledge. They know that an inference is valid if the conclusion is true in every state of affairs in which the premises are true. In other words, a putative conclusion follows validly from a set of premises if it is true when the premises are true and there is no way of interpreting the premises so as to render it false ..." (Johnson-Laird, 1983a, p.180)

How then might the departure of actual reasoning behaviour from canonical solutions be characterised within this framework? For Johnson-Laird, 'errors' are attributed to limitations in the working memory system, which not only has to hold the current mental model, but also to provide space for the construction and comparison of alternative models, so that systematic patterns of error are predicted amongst problems of alternative sentential and logical forms. Such an approach might, then, provide an account for the various effects of linguistic form, and for the varying apparent difficulty amongst the four classical inferences.

Another source of error for Johnson-Laird is variation in mastery of the skills involved. The theory assumes performance on a verbal inference task to involve three component skills: (1) the ability to form an integrated mental model of the premises, (2) an appreciation of the fact that an inference is only valid in the absence of counterexamples and (3) the ability to put into words the common characteristics of a set of mental models.

Having described the theory in general terms, it is now appropriate to consider what it has to say about conditional reasoning in particular. Johnson-Laird (1983a; 1986) argues that the ideal solution to the problem of 'if' would be neither a single unequivocal logic, nor a variety of different

meanings, but rather a

> " ... single uniform semantics from which both the diversity of the interpretations of conditionals and the vagaries in their logical behaviour will emerge." (Johnson-Laird, 1986, p.63)

Johnson-Laird's (1986) mental models account of conditionals rests crucially on five assumptions:

1 people can understand the true and proper nature of conditionals

2 the linguistic notion of compositionality applies to the semantic interpretations of conditionals, so that the interpretation of the whole conditional can be built up from the interpretations of its constituents

3 given (2), the interpretation of the consequent of a conditional can be done in isolation provided that it is done in a context which is known to satisfy the antecedent of the conditional[15]

4 it is a corollary of (3) that the function of the antecedent is to establish a context in terms of a state of affairs which is to be taken into account when considering the consequent

5 the Ramsey-Stalnaker (1968) notion of a 'thought experiment' is an ideal - people do not really evaluate a conditional by adding its antecedent to their stock of beliefs (with minimal

---

15 Note that it follows from (3) that if both speaker and hearer are conscious of the content of an antecedent (for example, of the imminence of an impending state of affairs), then the antecedent need not even be stated. The example offered by Johnson-Laird (1986, p.64) will illustrate this point: a mother who observes her child about to grab a forbidden cake can make a conditional assertion merely by saying "I'll smack you!". Incidentally, the antecedent may even be omitted where the conditional is counterfactual, so that if the mother observes that the child has overcome the temptation of the cake then she might say "I'd have smacked you".

adjustment to maintain consistency) and then evaluating the consequent

Given these assumptions, the basic mental models procedure for evaluating a conditional is as follows:

> "Step 1. Construct a mental model based on the superficial linguistic representation of the antecedent and on those beliefs triggered during this process.
>
> Step 2. Interpret the consequent in the context of the model and general knowledge." (Johnson-Laird, 1986, p.65)

Before we can do conditional reasoning, we need to set up an interpreted model of the conditional and the context. Under the mental models approach, we interpret conditionals by setting up a mental model based both on the meaning of the antecedent, and on our beliefs and knowledge of the context in which the conditional is presented. We then determine the nature of the relationship between antecedent and consequent, which may lead to the antecedent model being refined recursively. Finally, we set up a mental model of the consequent and then set up a scenario which relates a model of the consequent to the model of the antecedent.

Having interpreted and modelled the conditional, we are in a position to reason with it. For Johnson-Laird (1986), conditional reasoning consists in using the integrated mental model of the conditional to formulate 'informative conclusions' and then searching for alternative interpretations (alternative models) which would refute these putative conclusions: in the absence of any such refutation, the conclusions are proven.

On the essential meaning of 'if', it is proposed that

> " *If* is a cue to consider a possible or hypothetical state of affairs ... the conditional is true if and only if its consequent is true in the antecedent model and there is no alternative model in which it is false. The majority of conditionals, however, lack clear-cut truth conditions because their antecedents and the beliefs they trigger place insufficient constraints on the set of possible antecedent models." (Johnson-Laird, 1986, p.73)

Byrne and Johnson-Laird (1989) argue that their findings offer evidence for the relationship between the form of such models and the words typically chosen to express them. Subjects in that study were given sets of co-referential assertions - such as ('s') 'Lisa can have a fish supper', ('c') 'Lisa catches some fish' and ('g') 'Lisa goes fishing' - and asked to construct a single sentence which related the assertions together. In discussing this study, Byrne and Johnson-Laird (1990) argue that a task such as this requires a subject to find "a plausible causal or global framework" for the events described and then to organise the events within that framework.

They argue that subjects' knowledge about the usual relations between events enables them to identify ('s') 'Lisa can have a fish supper' as an obvious outcome, such that their task then becomes one of interrelating the other two sentences so as to lead to that outcome. They propose that these three assertions would give rise to two initial models as follows:

(1) [g c] s

(2) . . .

The use of brackets in model (1) denotes that 'g' and 'c' have been exhaustively represented in the model, whereas 's' has not. Model (2) denotes that alternative possibilities may exist (such as Lisa not being able to have a fish supper) but that such alternatives have, as yet, to be represented.

In terms of the model-based theory of sentential connectives proposed by Johnson-Laird, Byrne and Schaeken (1989), the two models described above correspond to a conditional sentence with 'conjunctive antecedents', in that the antecedents derive from a conjunction of events within the model. Byrne and Johnson-Laird (1989) also gave subjects sets of assertions corresponding to 'disjunctive antecedents' in Johnson-Laird, Byrne and Schaeken's (1989) terms: for example, ('c') 'Lisa catches some fish' was replaced with ('m') 'Lisa goes to the fish market'. In this case, Byrne and Johnson-Laird (1990) suggest that the initial models are as follows:

(1) ⌈ g ⌉ s
    ⌊ m ⌋ s

(2) ...

Byrne and Johnson-Laird (1989) predicted that in connecting these assertions into a single sentence, the initial models both for conjunctive and disjunctive antecedents would lead to subjects choosing a hypothetical construction rather than a merely conjunctive one. As predicted, they found that 69% of subjects used hypothetical or causal terms (such as 'if', 'when' or 'because'), whilst only 21% of subjects used simple conjunctions throughout.

It might be argued, however, that subjects' preference for terms such as 'if' to represent conditional scenarios is quite independent from - and can tell us little about - the nature of the underlying representation of such scenarios. Whilst the findings of Byrne and Johnson-Laird (1989) are certainly consonant with the particular models which they suggest, there seem to be no grounds upon which to argue that the choice of terms results from the use of this particular type of representation rather than from the simple recognition of a contingent situation, however it might be represented.

Its lack of amenability to empirical test is, however, far from the only shortcoming of the mental models account. One of the most damning arguments against any assertion that people use rules of inference (of either standard or natural logic) is the extreme difficulty in accounting for the way in which they acquire such rules in the first place. Whilst Johnson-Laird only requires that people possess "one essential piece of semantic information", he must still show how such possession comes about. The principle whereby he proposes that complex procedures are acquired rests crucially upon Turing's (1936) theory of computatability, wherein any complex function can be expressed (albeit in a somewhat cumbersome fashion) in terms of just three primitive functions. Johnson-Laird argues that such a notion of compositionality offers a mechanism whereby very simple procedures might be built into more complex ones, including those which are able to carry out inferences without recourse to any formal system of logic. The problem for Johnson-Laird is that simplify the procedures as he may, he still needs eventually to admit "an innate armamentarium of data and procedures" (Johnson-Laird, 1983a, p.144) such that at least primitive functions and data are present *ab initio* of which the more complex procedures might be composed.

Rips (1986), in a critique of mental models sharply entitled "Mental Muddles", sets out to show that 'literal' mental models (that is, those which imply a unique, special-purpose, type of representation) are unnecessary for any of the various endeavours for which they have been proposed - and thus that the 'figurative' approach is the more reasonable. He identifies three major "muddles". Firstly, Rips argues that Johnson-Laird appears to violate Fodor's (1980) 'formality constraint'[16] by claiming that standard cognitive

psychological accounts of comprehension are inadequate in that they are deficient in semantic properties such as truth and reference (which his mental models approach is able to supply).

The examples offered by Johnson-Laird (1983a) in support of the inadequacy of any account which is constrained in this way are persuasive. Rips (1986) argues equally persuasively, however, that Johnson-Laird's objections can, in fact, be overcome by the provision of a mechanism which can compare representations derived from discourse with representations derived from perception - and a mechanism such as this is (trivially) available within a standard cognitive psychological framework.

It may be argued, however, that to the extent that we are only concerned with explaining and predicting human behaviour on conditional reasoning tasks, then objective 'real world' truth is not of any great consequence to us. Rather, the only truth with which we need to concern ourselves is that embodied in a subject's individual belief system - and this sort of truth is, indeed, adequately provided for in a more traditional cognitive psychological framework, since the reasoning process only requires access to that internal representation. If this argument is correct, then it matters not whether Rips' first objection is valid: the fact that the 'traditional' approach can deal adequately with truth for our purposes does not mean that we are obliged to reject alternative approaches which are equally adequate.

---

16 Put simply, the 'formality constraint' says that cognitive psychology is about the description of computational processes which act upon mental representations and that the only computationally relevant properties of those representations are those internal to the representations themselves: in particular, cognitive psychology is constrained to do without notions of truth and reference, which have to do with the relationship between internal representations and the 'real world'.

The second "muddle" identified by Rips (1986) concerns the use of mental models as simulations, and appears to be as much a criticism of 'figurative' mental models as of 'literal' ones. Since, however, Johnson-Laird's (1983a) account of mental models in deductive reasoning does not involve simulation in the sense intended by Rips, this second objection will not be discussed further here.

Rips' third objection to mental models - aimed particularly at 'literal' models - is of particular relevance to this thesis as it addresses directly Johnson-Laird's claim that mental models are superior to both inference rules and the schemata of natural deduction in accounting for human behaviour on deductive reasoning tasks. The larger part of Rips' defence of inference rules and schemata concerns the particular application of the theory of mental models to syllogistic, rather than conditional, reasoning - primarily because it is Johnson-Laird's (1983a) theory of syllogistic reasoning in particular which is used as the central example in his book and which has been the subject of most empirical studies of his theory of mental models since (see for example, Johnson-Laird & Bara, 1984; and, for a critique, Boolos, 1984). The four essential points of Rips' defence of inference rules and schemata do, however, apply *mutatis mutandis* to conditional reasoning.

Firstly, Rips argues that the dismissal of traditional inference rules is only at the expense of the introduction of some special-purpose rules about mental models - for example, how the models should be used to represent the premises, how they should be combined and transformed into other models and how putative conclusions should be evaluated. It may be argued, however, that the word 'rule' is being used in a rather different sense in each case. 'Inference rule' is a term specific to formal systems of logic

(whether of the standard or of the natural deduction kind) and denotes a rigid template, containing such things as connectives and variable symbols, to which statements expressed in the formal language supported by the logical system might be matched, so as to generate a new statement in the formal language. The rules used in Johnson-Laird's theory of mental models, on the other hand, are effective procedures for the construction and manipulation of mental representations. One might wish to argue, therefore, that Rips is here playing somewhat fast and loose with different senses of the word 'rule' and thus that the objection is somewhat less than fair.

Secondly, Rips notes that although the tokens in a mental model are supposed to be the *actual* referents for premises and conclusions (as opposed to just standing for them), certain rules very like those of standard logic still apply. For example, two given tokens in a given model may be connected so as to insist that they refer to the same individual - or separated so as to insist that they cannot stand for the same individual. On this basis, it might be argued that to the extent that the rules of mental models are sensitive to such logical constraints then they are not really distinct from inference rules - so that "manipulation of mental models isn't fundamentally different from manipulation of mental propositions". This observation seems unobjectionable - but it may also be argued to be somewhat tangential to the question of whether mental models are a useful framework for our present purposes.

Thirdly, Rips takes exception to Johnson-Laird's claim that mental models will be better able to account for the effects of thematic content, in so far as mental models "wear their content on their sleeves" - and so if deduction is performed by manipulating mental models, then content will be taken into

account automatically. Rips argues that because in Johnson-Laird's theory of the syllogism the terms of the syllogism are no more than tokens, then there is no reason to suppose that content will be better handled here than under alternative approaches. Whilst this is undoubtedly true of the particular details of that particular theory, it is strongly implied by Johnson-Laird (1983a) that the models are sufficiently abstract and flexible to accommodate very much more information than that, for example, a given token is a beekeeper or an artist (which was all that was necessary for illustration of the theory of syllogisms). Moreover, the ability of mental models (in principle) to hold the necessary diversity of information to take account of content and context is made clearer still in Johnson-Laird (1986).

Finally, Rips asks "what's so bad about inference rules anyway?". He identifies a selection of the phenomena which inference rules are able to predict and argues that the theory of mental models fails to make many specific predictions at all, other than about behaviour on classical syllogisms and transitive inferences. Rips is, however, silent about the 'non-logical' phenomena which inference rules are unable to predict (and which, if only because of the extreme flexibility of the theory, mental models are, in principle at least, able to accommodate).

Despite these objections, the theory is an attractive one for a number of reasons. It is absolutely independent of any formal system of logic, so that there is no question of the resulting representations being 'correct' other than by way of coincidental concordance with the formal truth tables: there is, accordingly, no necessary reason to restrict the set of possible representations to those strictly concordant with just I-tables, E-tables and D-tables[17]. By explicitly providing that the interpreted model of the

[17] As will be seen later, this will be of vital importance to the current thesis,

conditional should include beliefs and knowledge of the context in which the conditional is presented, it automatically anticipates, and provides a mechanism for, the effects of such things as content, linguistic context and sentential form.

Additionally, the theory offers a number of plausible sources of both systematic and unsystematic 'error' and behavioural variation, both within and between subjects. Within subjects, systematic variation is predicted amongst the four classical inferences, and amongst problems utilising alternative linguistic and logical forms, due to limitations in the working memory system. Systematic between-subject variation is also predicted due to variation in mastery of the skills involved and, presumably, performance and capacity differences amongst subjects' working memory systems. Variation is also predicted by the proposal that (typically) "antecedents and the beliefs they trigger place insufficient constraints on the set of possible antecedent models". Given that the process is unconstrained, but that people do produce solutions in a finite (and typically rather short) period of time, there is reason to suspect that this variation will tend to be largely unsystematic[18] since it will often result from the (more or less orderly) premature termination of a process.

1.1.10 : Counterexamples and Fuzziness

The search for falsifying counterexamples to a conclusion has been shown

since two additional types of (non-logical) truth table will be identified.

18 To suggest that the prediction of unsystematic variation is attractive in a theory might be considered heresy. Nevertheless, the psychological literature of conditional reasoning is so very rich in just this sort of variation (both within and between subjects) that it might be argued that this feature of the theory actually increases its psychologically plausibility.

to be an operation which is central to Johnson-Laird's (1983a; 1986) account of conditional reasoning within his theory of mental models. Since the first four experiments were directly involved with the manipulation of visual arrays which contained varying proportions of falsifying counterexamples, it is appropriate to round off this general theoretical introduction with a brief review of relevant research concerned with reasoning about situations and/or assertions which are more or less *fuzzy*. Throughout the remainder of this thesis, the descriptor 'fuzzy' will be used from time to time both with respect to the truth of a conditional statement and with respect to the consistency of state of affairs about which the statement is made. A statement may, for example, be described as 'fuzzily true' or 'fuzzily false' with respect to a given set of instances. The assertion that a statement is 'fuzzily true' with respect to a given set will be used to mean that the set contains some certain (small) proportion of (logically falsifying) TF cases and that a substantial proportion of the remaining items are (logically verifying) TT cases. Saying that a statement is 'fuzzily false' with respect to such a set, on the other hand, is intended to convey that the set contains some certain (small) proportion of (logically verifying) TT cases and that a substantial proportion of the remaining items are (logically falsifying) TF cases.

It is important, however, to be clear that in using the term 'fuzzy' it is by no means intended to imply that conditional statements should be viewed within the framework of any particular formal theory based on fuzzy sets, such as Zadeh's (1979) theory of 'approximate reasoning'. One fundamental idea is, however, borrowed directly from Zadeh's theory for the purposes of this thesis - the notion of degrees of truth, whereby statements such as conditionals might be not merely true or false but, alternatively, very true,

more or less true and so on.

It is appropriate to point out at this stage that the ascription of the term 'fuzzy' is assymmetrical with respect to formal logic. Where a statement is 'fuzzily true' with respect to a set of instances (or a set is to be constructed such that a given statement about it is shown to be true), then even a single TF case in the set renders the statement *logically* false. Where, on the other hand, a statement is 'fuzzily false' with respect to a set of instances (or a set is to be constructed such that a given statement about it is shown to be false), then the inclusion of any number of TT cases does not in any way make a logically false statement logically true, since so long as at least one TF case is present, the statement is still logically false. In terms of formal logic, then, something 'fuzzily true' is, in fact, logically false, whilst something 'fuzzily false' is logically just as false as something 'absolutely' false.

As has been seen, the conditional statement 'if P then Q' appears sometimes to be interpreted by subjects as standing for the material implication relation (P -> Q : that is, 'at least if P then Q') and sometimes for the material equivalence relation (P <-> Q : that is, 'if, and only if, P then Q'). Whilst the factors which decide which of these interpretations will be made constitute the very *raison d'être* of conditional reasoning research, these two alternative interpretations themselves are fully and formally specified, such that to assign one or other of these labels to a situation is fully to specify the relationship that obtains between P and Q. It should be noted, moreover, that in assigning one of these two formal labels, we make the relationship between P and Q utterly determinate - we refuse to admit any uncertainty in that counterexamples are explicitly forbidden under both interpretations.

Before considering the extent to which it might be desirable to provide for uncertainty, it is necessary to consider just how uncertainty might manifest itself in the context of a reasoning experiment. As has been stated already, reasoning experiments may be distinguished in terms of whether they use thematic or abstract materials. In using thematic materials, we are attempting to capture how people reason in 'real world' situations - and so we use materials and relationships with which they will be familiar in their everyday experience. It is intuitively obvious that the hearer of a conditional statement uttered in reality (that is, outside the reasoning laboratory) will often be aware that the truth of that statement is less than 100% certain, but will nonetheless act upon it *as if* it were absolutely true. There are a number of reasons why this might be the case. For example, a statement such as "if you go through to the Public Bar then you'll be served straight away" seems to be a 'shorthand' for "if you go through to the Public Bar then *it is, in my opinion, statistically probable that* you'll be served straight away". The important point is that the hearer is perfectly well aware that there is some uncertainty involved, so that given the antecedent, the consequent will be expected to be true only probabilistically: nevertheless, in following the advice offered, the hearer is still acting *as if* they had interpreted the relationship as canonical.

Conditional reasoning in the real world is an inherently fuzzy business. If we are to capture how people reason in 'real world' situations, therefore, it may be argued that we must provide for the effects of perceived uncertainty or, to put it another way, for 'fuzziness' in the interpretation of the relationship between P and Q.

Fuzziness manifests itself rather differently where we are using abstract materials. In abstract investigations, the conditional statements are all about

totally abstract materials (such as symbols of various colours and shapes) and are uttered in a laboratory context which is quite deliberately devoid of any other context. The utterance of a statement such as "if the symbols are red then they are square" is clearly a statement about the relationship between redness of symbols and squareness of symbols. More specifically, in the particular context of a reasoning experiment - where the only entities that exist are coloured symbols - it is quite clearly a statement about the relationship which obtains between the set of all red symbols and the set of all square symbols. The conditional statement "if the symbols are red then they are square" in such a context is, therefore, merely an alternative expression of the universal generalisation "all red symbols are square symbols". The interpretation of this statement to which we would assign the label 'material implication', then, would have the set of all red symbols contained within the set of all square symbols, such that there were no non-square red symbols, but there were (or, at least there may be) some non-red square symbols. The interpretation to which we would assign the label 'material equivalence', on the other hand, would have the set of all red symbols identical with the set of all square symbols, such that not only were there no non-square red symbols but neither were there any non-red square symbols.

References to 'fuzziness' in the psychological literature abound, but tend in the main to have been concerned with the meaning of quantifiers (such as 'all' or 'some') rather than of connectives (such as 'if ... then'). In his review of the literature on the 'fuzziness' of quantifiers, Newstead (1988) identifies studies from three distinct areas: syllogistic reasoning, rating scales and psycholinguistics. Studies on the interpretation of quantifiers in rating scales (for example, Bass, Cascio & O'Connor, 1974; Newstead & Collis, 1987) have,

in the main, been concerned with quantifiers of frequency (such as 'often' or 'never') or amount (such as 'few' or 'many'). Psycholinguistic studies of fuzzy quantifiers have generally been concerned with the search for a suitable theoretical framework within which to describe the interpretation of quantifiers. Typical of these studies have been Just's (1974) featural theory, Holyoak & Glass' (1978) analogue theory and Anderson's (1981) propositional theory. In addition, a number of approaches founded in fuzzy set theory have been proposed, as typified by Hersh & Caramazza's (1976) study motivated by advances in semantic theory such as the work of Labov (1973) and Lakoff (1973).

Whilst an exhaustive search of the particular literature of conditional reasoning fails to reveal any comparable investigations in which 'fuzziness' is directly manipulated, a number of studies do concern themselves with the availability of alternative 'routes' to the consequent in addition to the antecedent, and so are of interest here. For example, Bucci (1978) and Rumain, Connell & Braine (1983) have shown that the logically erroneous AC (Q, therefore P) and DA (not-P, therefore not-Q) inferences are more likely to be rejected for the 'if P then Q' conditional when subjects are aware of specific examples for which something other than P also implies Q (that is, examples of the logical case ~PQ). Similar results are reported by Markovits (1984; 1985; 1986) and further support comes from Markovits' (1988) formally equivalent (but non-verbal) concrete task.

In a thematic task, where P and Q are replaced by something with which subjects are familiar, then it should be easier for subjects to generate such alternatives to P and thus performance on these two inferences should improve - and, as will be seen in the review of the thematic literature - this is

indeed the case. Markovits (1984;1985;1986) argues, however, that a more abstract capacity might be involved, such that subjects' awareness of the possible existence of (unspecified) ~PQ cases is independent of their ability to generate specific examples of such cases. In support of this notion, Markovits *(ibid)* has demonstrated that for such facilitation to occur, subjects need to be aware not merely of the existence of ~PQ examples but also of their pertinence to the reasoning task - and he argues that differences in the ability to recognise and use such examples should be seen as competence differences between subjects.

In so far as it might be argued that (in abstract tasks at least) a conditional statement might be interpreted as an alternative linguistic expression of a universal generalisation, then it is appropriate to describe also some of the phenomena typically reported in the 'fuzziness' literature concerned with syllogistic reasoning.

Syllogistic reasoning involves reasoning from two premises to a conclusion, as in

All cognitive scientists eat quiche
Some psychologists are cognitive scientists

Therefore Some psychologists eat quiche

In classical Aristotelean syllogisms, each premise will incorporate a quantifier such as 'all', 'some', 'no' or 'some ... not'. The precise nature of the relationship between quantified assertions and conditional statements is, of course, an empirical question - and is outside the scope of this thesis. To the extent that a conditional statement might be interpreted as an alternative expression of a universal generalisation, however, it seems appropriate to consider what the syllogistic reasoning literature has to say about fuzziness

in the interpretation of (at least) the quantifiers 'all' and 'no'.

Once an abstract conditional is viewed as a universal generalisation, then it becomes clear that uncertainty must be admitted. Kyburg (1983), for example, argues that in people's real world experience, there are few genuine universal generalisations, such that 'all' will typically mean 'almost all'. Specifically, he argues that in the "natural ecology" (Einhorn and Hogarth, 1981), the generalisations with which people mainly deal, and which they habitually confirm, are essentially statistical, such that they have a natural tendency to treat a conditional generalisation such as 'if something is an X then it is a Y' as representing 'almost all Xs are Ys'. Certainly, it does seem intuitively plausible that the way in which we habitually understand and use generalisations does involve fuzziness. Newstead and Griggs (1984) offer the example "All Germans are hardworking" and argue that such a statement does seem to permit exceptions and so seems to constitute a useful generalisation rather than a statement of something that is true of each and every German.

Empirical studies of the interpretation of 'all' and 'no' in reasoning experiments appear relatively rarely in the psychological literature, but those few that are reported suggest that people do indeed interpret generalisations in this way. For example, Newstead and Griggs (1984), report that subjects given a fact about a population such as "There is one person over 100 years old, 110,933 under 100 years old" still rated the quantified statement "All the people are aged under 100" as moderately appropriate.

In common with Newstead and Griggs (1984), the four 'abstract'

experiments to be reported in this thesis also involve asking subjects to make subjective ratings on a continuous scale (in this case, subjective ratings of the truth of a conditional statement). It is appropriate at this stage, therefore, to identify a potentially confounding demand effect which is unavoidable in such studies. Where subjects are presented with scenarios which vary along some continuum and are asked to rate the scenarios along a continuous scale, there is inevitably a danger that rather than using the scale to indicate what it purports to measure (for example, degree of appropriateness) they might feel obliged to use it to reflect their perception of where on the continuum a particular scenario falls. In the Newstead & Griggs (1984) study, such a demand effect would manifest itself as an appropriateness rating which varied inversely with the proportion of counterexamples - which is, of course, precisely confounded with their predictions. To the extent that such an effect might be in operation, however, Newstead and Griggs (1984) argue that it should be expected to be just as strong for 'no' as for 'all'. The observation that even a single positive instance resulted in the rejection of the "No ..." statement almost as strongly as when there were many such instances, however, permitted Newstead and Griggs (1984) to conclude that 'all' is interpreted more fuzzily than 'no'. It must be noted, however, that the potential confound is still there: whilst 'all' does appear to be more fuzzy than 'no', this is by no means to say that some part of the apparent fuzziness for both quantifiers might be no more than an artifact of this demand effect[19].

In conditional reasoning with abstract materials in the laboratory context, then, it has been argued that to utter a conditional is, effectively, to utter a

[19] In the extreme, it may even be argued that the *whole* of the apparent fuzziness of 'no' is an artifact of the demand effect, the only 'true' fuzziness being the *additional* apparent fuzziness found for 'all'.

universal generalisation. Since it appears that, in people's natural experience, universal generalisations are inherently fuzzy, it seems reasonable to suppose that a certain (small) proportion of counterexamples will be tolerated - and it may well be that such a proportion will even be *expected.*

In so far as a 'real world' conditional statement might be interpreted as a universal generalisation, it seems not unreasonable to expect to observe a certain level of tolerance of counterexamples in thematic experiments also. Furthermore, it has been suggested that certain 'real life' conditional statements (such as the advice given above about getting served in the pub) are inherently fuzzy in that they actually reflect the speaker's (fallible) belief and, furthermore, that our knowledge of the world often tends to be statistical rather than categorical.

1.2 : Research Using Abstract Materials

This section reviews the typical paradigms and findings of studies which employed abstract materials - that is, those selected to be as devoid as possible of any 'real-world' meaning. As has been stated, the rationale behind the use of such materials is to avoid any possible effects of such factors as the content of the propositions, the perceived linguistic context in which the statement is uttered and individual differences between subjects in terms of their experience with the materials and their beliefs about the truth of the conclusion as an assertion in its own right. Typical questions of interest in abstract tasks are the effects of factors such as the sentential form and the presence of negative components.

1.2.1 : Some Typical Paradigms

As has been seen, two major kinds task have been used in conditional reasoning research - inference tasks and truth table tasks - each of which might be presented in one of two forms - construction and evaluation. In an inference construction task, a subject is given a conditional statement, together with an assertion about one of the propositions in the conditional, and asked to say what (if anything) follows from those two pieces of information. In an inference evaluation task, a subject is given a conditional statement and assertion as before, but now asked to say whether or not a particular conclusion follows on that basis - or, alternatively, that it is not possible to say whether it follows or not[20]. Additionally, some studies have

[20] This indeterminate response is not always made available - in some studies, subjects are constrained to answer "yes" or "no" (or "true" or "false", depending on the task).

used a 'hybrid' inference task in which subjects have to select one (or more) conclusions from amongst a given set of options.

In a truth table construction task, a subject is given a conditional statement and is asked to construct one (or many) instances which show the statement to be true (or false). In a truth table evaluation task, a subject is given a conditional statement as before, but is now additionally given a particular state of affairs to which the statement purportedly applies, and asked to say whether the statement is true or false with respect to that state of affairs - or, alternatively, to say that neither of these determinate responses is appropriate[20].

1.2.2 : Inference Tasks - Some Typical Phenomena

A number of studies in the early 1970's were concerned with subjects' endorsement of the four "classical" inferences described above. These studies were primarily of the evaluation type, so that subjects were required to evaluate a given argument. The following argument is typical of the *genre:*

given	if the letter is C then the number is 4
and	the number is not 4
	———————
conclude	the letter is not C

In this particular case, the materials are letter-number pairs, the most common type of abstract material used in these studies. The inference represented in this example is Modus Tollens (MT), so that the first premise of the argument is a conditional statement of the form 'if P then Q', where P

(the antecedent proposition) is 'the letter is C' and Q (the consequent proposition) is 'the number is 4', whilst the second premise of the argument - 'the number is not 4' - is the negation of the consequent. The conclusion of the argument - 'the letter is not C' - is the negation of the antecedent.

Subjects would typically be asked to state whether the conclusion *necessarily* follows *assuming the truth of the premises*. As has already been seen, it is no easy matter to decide just what "necessarily follows" means for these purposes. For the purposes of this thesis, a strict view has been taken that a conclusion follows from an argument (that is, an inference is valid) if, and only if, any contradiction of the conclusion can be seen to be explicitly ruled out. In using terms such as "necessarily follows" in their instructions, investigators such as Evans (1977) intend that subjects should take a similarly strict view of what it is for a conclusion to follow from an argument, such that only "yes" and "no" responses are required[21]. As will be seen, studies have varied in terms of the response options available to the subject - Taplin (1971), for example, used "true" and "false" rather than "yes" and "no", whereas Taplin and Staudenmayer (1973) provided a third 'indeterminate' response.

As was seen in Table 1.4, a subject who interprets the conditional as expressing material equivalence (and who responds purely on the basis of 'possession' of the E-table) should endorse all four inferences, whilst a subject who interprets it as expressing material implication (and who responds purely on the basis of 'possession' of the I-table) should endorse just

[21] In that particular study, the written instructions to subjects said merely that the conclusion "may or may not follow logically from the two statements" and that they should respond "yes" if they "think the conclusion follows". It is reported, however, that subjects had the strict interpretation of "follows" explained to them by the experimenter.

MP and MT, rejecting DA and AC.

Where a subject responds on the basis of a 'defective' truth table (such as the D-table), they should still endorse MP and MT, but may now wish to respond 'indeterminate' to DA and AC on the grounds that their representation of the relation is indeterminate with respect to the falsifying counterexample (which for both these inferences is the FT case). Where an 'indeterminate' response is available, this is straightforward. There is also no problem where (as in Evans, 1977) the available responses are "yes" ('it does necessarily follow') and "no" ('it doesn't necessarily follow') since, on a strict interpretation of "necessarily follows", even though they would not like to say that the DA and AC inferences were necessarily false, they (crucially) do not rate them as necessarily true, and so "no" is a suitable response. There is a problem for these subjects where the only available responses are "true" and "false", however (as was the case in Taplin, 1971), since the preferred 'indeterminate' response is now at odds with both these options.

For group data - where we might reasonably expect a mixture of E-table, D-table and I-table interpretations[22] - the expected pattern on this basis would be 100% endorsement of MP and MT, and some intermediate (but more or less equal) level of endorsement for DA and AC. This predicted pattern of responses does, of course, rest on various assumptions: for example, that every subject will be consistent in their interpretation of the conditional statement irrespective of the inference of which it forms a part - and that the materials are quite genuinely devoid of meaning in any way which might

[22] (at the very least - later in this thesis, two further types of 'defective' truth table are proposed, and there is no necessary reason to suppose that even this will constitute an exhaustive set).

influence the subject differentially for different inferences.

Table 1.6 shows summary data from three typical studies of this type - Taplin (1971), Taplin and Staudenmayer (1973) and Evans (1977).

study	MP	MT	DA	AC
Taplin (1971)	92	63	52	57
Taplin and Staudenmayer (1973)	99	87	82	84
Evans (1977)	100	75	69	75

Table 1.6 : Percentage of endorsements of inferences involving a conditional statement of the form 'if P then Q'
(adapted from Evans, 1982)

It may be seen from Table 1.6 that the predictions about MP, DA and AC are pretty well borne out: MP endorsement is virtually 100%, whilst the rates for DA and AC are roughly equal and at an intermediate level. The rates for MT, however, are nothing like the 100% which should be expected if subjects respond purely on the basis of their underlying truth tables - indeed, the MT endorsement rates are virtually indistinguishable from those for DA and AC.

Typical explanations for the observed superiority of MP over MT revolve essentially around the observation that MP is so trivially easy that it hardly involves reasoning at all, whilst the MT argument is really jolly hard to work out. Since it has been argued that to endorse an argument as "valid" we must be able to convince ourselves that the counterexample to its conclusion is explicitly denied, difficulty in working the argument out might be expected to result in lower endorsement rates as a result of subjects

'losing track' of the argument and thus failing to find that explicit denial. Explanations in terms of difficulty are, however, obliged to specify the source of the difficulty, and such candidate features have been proposed as the direction in which the inference is made (that is, from the antecedent to the consequent, or *vice versa)* and the presence of negative components.

In terms of direction, working through the MP inference [->(->PQ,->(P,Q))] involves proceeding from antecedent to consequent, as does the DA inference [->(->PQ,->(~P,~Q))], whereas both MT [->(->PQ,->(~Q,~P))] and AC [->(->PQ,->(Q,P))] involve proceeding 'backwards' from consequent to antecedent. Directionality will turn out to be an important feature in the explanation of behaviour on conditional reasoning tasks and will be seen often to be dependent upon the sentential form in which the conditional is stated: this question is considered again later when syntactic factors are addressed. In terms of the presence of negative components in the argument, MP requires nothing to be negated, DA and AC one item each, and MT two items. It is the general question of the presence of negative components that is considered next.

1.2.3 : Negativity Effects

It has been seen that the involvement of negative components is a possible source of extra difficulty in making the MT inference (where the second premise and the conclusion both involve a negative) as opposed to the MP (which does not involve negatives at all). What has not so far been addressed is the effect of including negative components in the conditional statement itself - that is, the discussion so far has only been concerned with

statements where both P and Q are *affirmative.*

A particular problem with restricting the investigation to double- affirmative statements is that the presence of negative components is confounded with the logical structure, so that certain inferences might be easier to make simply because they do not involve negation - in as much as statements including negative components are syntactically more complex, have more words and thus are more likely to be misread, and so on.

As will be seen, this confound may be removed by the systematic permutation of the presence of negative components. Table 1.7 sets out what sentences expressing the various truth cases might look like for each of the four possible combinations of affirmative and negative propositions in the conditional statement. Traditionally, the corresponding conditional statements are referred to in terms of whether each of their components is affirmative ("A") or negative ("N"), so that 'if P then Q' would be referred to as an AA statement since both P and Q are affirmative, whilst 'if P then not Q' is an AN statement, 'if not P then Q' an NA statement and 'if not P then not Q' an NN statement.

conditional	case	sentence
AA : <u>if P then Q</u>		
eg:	*if the letter is X then the number is 4*	
	TT:	the letter is X; the number is 4
	TF:	the letter is X; the number is not 4
	FT:	the letter is not X; the number is 4
	FF:	the letter is not X; the number is not 4
AN : <u>if P then not Q</u>		
eg:	*if the letter is B then the number is not 6*	
	TT:	the letter is B; the number is not 6
	TF:	the letter is B; the number is 6
	FT:	the letter is not B; the number is not 6
	FF:	the letter is not B; the number is 6
NA : <u>if not P then Q</u>		
eg:	*if the letter is not Z then the number is 9*	
	TT:	the letter is not Z; the number is 9
	TF:	the letter is not Z; the number is not 9
	FT:	the letter is Z; the number is 9
	FF:	the letter is Z; the number is not 9
NN :<u>if not P then not Q</u>		
eg:	*if the letter is not W then the number is not 2*	
	TT:	the letter is not W; the number is not 2
	TF:	the letter is not W; the number is 2
	FT:	the letter is W; the number is not 2
	FF:	the letter is W; the number is 2

Table 1.7 : Conditional statements including negative components and the sentences which express the corresponding truth cases

In addition to looking at how the presence of negatives affects the way in which the various truth cases are expressed, it is necessary to consider how the four "classical" inferences will now be expressed, and this is set out in

Table 1.8. Table 1.8 makes it clear that the introduction of negative components into the conditional statements systematically alters the occurrence of negative components in the conclusions that subjects are being asked to endorse.

conditionals		inferences	
AA :<u>if P then Q</u>			
eg: if the letter is X then the number is 4			
MP:	the letter is X	*therefore*	the number is 4
AC:	the number is 4	*therefore*	the letter is X
DA:	the letter is not X	*therefore*	the number is not 4
MT:	the number is not 4	*therefore*	the letter is not X
AN :<u>if P then not Q</u>			
eg: if the letter is B then the number is not 6			
MP:	the letter is B	*therefore*	the number is not 6
AC:	the number is not 6	*therefore*	the letter is B
DA:	the letter is not B	*therefore*	the number is 6
MT:	the number is 6	*therefore*	the letter is not B
NA :<u>if not P then Q</u>			
eg: if the letter is not Z then the number is 9			
MP:	the letter is not Z	*therefore*	the number is 9
AC:	the number is 9	*therefore*	the letter is not Z
DA:	the letter is Z	*therefore*	the number is not 9
MT:	the number is not 9	*therefore*	the letter is Z
NN :<u>if not P then not Q</u>			
eg: if the letter is not W then the number is not 2			
MP:	the letter is not W	*therefore*	the number is not 2
AC:	the number is not 2	*therefore*	the letter is not W
DA:	the letter is W	*therefore*	the number is 2
MT:	the number is 2	*therefore*	the letter is W

Table 1.8 : Conditional statements including negative components and the sentences expressing the relevant classical inferences

Various studies have looked at the effect that the presence of negative components has upon the patterns of endorsement of each of the four classical inferences. One of the earliest such studies to do so systematically was that of Evans (1977), the data from which are summarised in Table 1.9.

conditional	inference :	MP	MT	DA	AC
AA : if P then Q		100	75	69	75
AN : if P then not Q		100	56	12	31
NA : if not P then Q		100	12	50	81
NN : if not P then not Q		100	25	19	81
	overall:	100	42	38	67

Table 1.9 : Evans (1977) : Percentage of endorsements of inferences *(adapted from Evans, 1982)*

One way in which negative components might make the arguments more difficult to work out is in that they will sometimes involve the subject in having to negate something which is already negated. Neither MP or AC involve negation as part of their argument. For DA, however, whenever the consequent is already negative (AN and NN), the subject is obliged to deal with a double negative in order to reach the conclusion, so that denial of the consequent 'the number is not 6' leads to 'the number is 6', but only via the intermediate step 'it is not the case that the number is not 6'. Similarly, the conclusion to the MT argument will involve a double negative where the antecedent is negative (on NA and NN).

If it is this double negative which makes it harder for the subject to prove the argument, then we should expect lower endorsement rates of DA on AN

and NN, and for MT on NA and NN. Table 1.9 appears to support this prediction. DA is far less frequently endorsed when the consequent is negative (on AN and NN); similarly, MT is far less frequently endorsed when the antecedent is negative (on NA and NN) and this is found to be a reliable effect over a number of other studies (for example, Roberge, 1971; 1974; 1978; Evans, 1972b; Pollard and Evans, 1980).

If dealing with a double negative is, indeed, the source of the trouble then MP and AC endorsement rates should not be affected by the presence of negative components. This is demonstrably true for MP endorsement in this study, which remains at 100% throughout (and this is in line with the typical findings on AA statements, as might be seen by comparison with Table 1.6). The AC inference, however, is endorsed considerably less often on the AN form than on the other three. Since the AC inference never involves a double negative, it may be that an explanation of the effect of negatives in terms of additional difficulty is rather too simplistic.

A very different explanation of the effect of negatives is proposed by Evans (1977) and Roberge (1978), who suggest that subjects are more likely to accept a conclusion which is negative - irrespective of whether it arises by affirmation of a negative component or by negation of an affirmative component. How, then would such a bias manifest itself here? It may be seen from Table 1.8 that negative conclusions are associated with affirmative consequents for DA (AA and NA) and with affirmative antecedents for MT (AA and AN). For these two inferences, then, it is clear that 'negative conclusion bias' makes precisely the same predictions as the suggestion that double negatives cause difficulty, in that the cases preferred under the former are precisely those which the latter does not make more difficult. Since, for these two inferences, the two accounts make identical predictions,

the data in Table 1.9 obviously support each account to the same extent.

The 'negative conclusion bias' proposal does, however, make a distinctive prediction for the AC inference, in so far as those cases with negative antecedents (NA and NN) should be preferred, whereas the 'double negative' account predicts no effects whatever on AC - and it is indeed the case that the highest AC endorsement rates are found on NA and NN in Table 1.9. It will be noted, however, that those on AA are only marginally lower than those on NA and NN. It appears, then that both these accounts of the effects of negativity are somewhat embarrassed by the AC inference, and so perhaps neither tells the whole story.

Before moving on to consider further studies, it is appropriate to distinguish these two accounts in terms of their mechanism. This consideration of the effects of negativity started from the simple premise that the presence of negative components may make it harder to work out the arguments of certain of the inferences - and the 'double negative' account is this vein. It should be noted that the 'negative conclusion bias' account is a rather different kettle of fish altogether, however, in so far as it has nothing whatever to do with difficulty. Pollard and Evans (1980) see 'negative conclusion bias' as a 'caution heuristic': more specifically, Evans (1982) suggests that it might be seen as a bias towards the least falsifiable statement (and Popper, 1959, points out that a conditional is more falsifiable the more general its antecedent and the more specific its consequent). As this review continues, further such 'extra-logical' biases will be identified .

1.2.4 : Inference Tasks and Underlying Truth Tables

In the studies described so far, subjects were asked to evaluate, or to construct, conclusions to arguments corresponding to the four "classical" rules of inference. What was investigated, then, was the way in which subjects reason with conditional information. Clearly, in order to reason, subjects must have some underlying mental representation of the situation which obtains - but the studies considered so far can only address the question of representation indirectly - by asking something like "given that a subject argues thus (or endorses this line of argument), what kind of mental representation must they have in order to do so?". A number of researchers (for example, Taplin, 1971; Taplin and Staudenmayer, 1973; Staudenmayer, 1975; Staudenmayer and Bourne, 1978; Rips and Marcus, 1977; Marcus and Rips, 1979) have argued that we may use the results of studies such as these to infer the nature of the underlying representations.

The major problem with such studies, as Evans (1982) points out, is that when subjects' behaviour is compared across tasks, it turns out to be highly inconsistent. By changing the task to a 'truth table task', however, we are able to get rather more directly at the underlying representation, and it is these tasks which are considered next.

1.2.5 : Truth Table Tasks - Some Typical Phenomena

Truth table tasks attempt to get more directly at the truth tables which subjects 'possess' by asking them to classify instances as true, false or irrelevant, either by evaluation of given instances or by the construction/generation of instances classifiable in one of these ways.

Typically, such studies systematically manipulate the presence of negative components in the same sort of way as has been described for inference tasks.

Evans (1972c) is typical of the *genre*. Subjects were presented with an array of coloured shapes together with a conditional statement about coloured shapes, such as

> IF THE SHAPE ON THE LEFT IS A RED CIRCLE
> THEN THE SHAPE ON THE RIGHT IS A GREEN SQUARE

Subjects were given two tasks for each conditional statement: a 'verification' task, in which they had to select two coloured shapes to make the statement true, and a 'falsification' task, in which they had to select two coloured shapes to make the statement false. Subjects were asked to provide an exhaustive set of solutions to each task, but were allowed to generalise verbally, for example by saying "A red circle on the left, together with anything except a green square on the right, makes it true" (Evans, 1982, p.138). Evans argues that since the subject was asked for an exhaustive set of solutions (or at least a specification of what an exhaustive set would be), any truth case not constructed on either task could be inferred to be 'irrelevant'.

It has already been shown in Table 1.7 that systematically manipulating the presence of negative components in the conditional statements results in a deconfounding of truth case with the presence of negative components in the sentences describing the relevant instances. There is, however, another confound which is removed in addition. From the example statement above, it will be noticed that the TT case (red-circle & green-square) contains both the items which are named in the statement. The FF case (which might, for

example, be blue-triangle and yellow-star) contains neither of these items, whilst TF and FT will contain exactly one each. What is confounded with truth case, then, is the degree to which the pair of items selected *matches* the items mentioned in the statement. Table 1.10 shows how the manipulation of negative components deconfounds logical status and matching status.

conditional	truth case : TT	TF	FT	FF
AA : if P then Q	PQ	P~Q	~PQ	~P~Q
AN : if P then not Q	P~Q	PQ	~P~Q	~PQ
NA : if not P then Q	~PQ	~P~Q	PQ	P~Q
NN : if not P then not Q	~P~Q	~PQ	P~Q	PQ

Table 1.10 : The distribution of matching cases (P, Q) and mismatching cases (~P, ~Q) over truth cases (TT, TF, FT, FF) and statement types (AA, AN, NA, NN)

When referring to whether an item matches a proposition named in a conditional statement, it is traditional to refer to those that match as P (for the antecedent) and Q (for the consequent), and to those that mismatch as ~P and ~Q respectively. For the example given, then, the pair of items which match both propositions in the statement (red-circle & green-square) is referred to as PQ, a pair which matches the antecedent only (for example, red-circle & yellow-star) as P~Q, and so on. Given that each possible matching status occurs exactly once within each truth case, the data can be pooled over the four statements to give a measure either of the effect of truth

case (unbiased by matching status) or of the effect of matching status (unbiased by truth case). Table 1.11 classifies the data from Evans (1972c) in both these fashions.

Evans (1972c) claims that the classification by truth case is evidence for the predominance of the D-table, since virtually all the responses to TT are 'true', 80% of those to TF are 'false' and the modal response to FT and to FF is 'irrelevant'. It should however, be noted that this is not an analysis of individual subjects' truth tables, but rather a pooling of data from all 24 subjects.

The classification by matching status reveals that when there is a double match (PQ), then very few subjects (14%) fail to select that item. Conversely, where there is double mismatch (~P~Q) then the level of non-selection is at its highest (41%), and in the middle fall the semi-matching P~Q and ~PQ items. Put another way, there is a tendency for subjects to select items which are explicitly named in the conditional statement, and to reject items which are not named. Evans (1972c) terms this bias 'matching bias': like 'negative conclusion bias', 'matching bias' is extra-logical and is not related to the complexity of the task.

To the extent that 'matching bias' might be seen as an example of 'feature-matching', its reality is corroborated by a related phenomenon in the field of syllogistic reasoning: a number of researchers in that field (for example, Woodworth and Sells, 1935) have reported the occurrence of 'atmosphere bias', a rather similar phenomenon whereby the syntactic form of the premises can induce a preference for a conclusion of a similar type.

classification		verification task [true]	falsification task [false]	item not selected [irrelevant]
truth case :	TT	99	0	1
	TF	3	80	17
	FT	14	34	52
	FF	33	23	44
matching status :	PQ	34	52	14
	P~Q	41	33	26
	~PQ	40	27	33
	~P~Q	34	25	41

Table 1.11 : Evans (1972c) : Percentage selections of items classified by truth case and by matching status

(adapted from Evans, 1982)

It might be argued, however, that in order for a descriptive feature of the data to be raised to the status of a 'bias', we must be able to show how such a determinant of behaviour might come about - and for what it might be used outside the reasoning laboratory. As will be seen shortly when the 'heuristics and biases' approach to reasoning is considered, Pollard (1982) has applied to deductive reasoning Tversky & Kahneman's (1983) notion of the 'availability heuristic', which was originally proposed to explain certain phenomena in intuitive statistical inference. For the present purposes, 'availability' means something like the perceived salience of an item for use in a particular situation. Pollard (1979) has suggested that in an abstract task, the primary determinants of what elements will be available are the features of the task itself - so that the appearance of items in a rule increases their availability and thus produces 'matching bias'.

In a not dissimilar vein, Evans (1984), proposes that items mentioned in the rule are judged as 'relevant' due to the operation of linguistic mechanisms such as the topic-comment effect and the effect of the presence of the linguistic marker 'if' in inducing the use of presuppositions, so that such items come to 'dominate' a subject's thinking.

Support for both these explanations comes from Evans (1983b). This study used a truth table evaluation (as opposed to construction) task, a type of task on which 'matching bias' has been shown to be relatively reliable (see, for example, Evans, 1975; Evans and Newstead, 1977). In Evans (1983b), however, items were negated either implicitly (as is generally the case in such experiments) or else explicitly. For example, given a rule 'if the letter is K then the number is 3', the implicit AN case might be given as 'the letter is K and the number is 5', whilst the explicit AN case might be given as 'the letter is K and the number is not 3'. As predicted by both the 'availability' and the 'relevance' accounts, 'matching bias' was, indeed, significantly reduced by the use of explicit negation (by which, in this example, both K and 3 are equally 'available' and 'relevant').

Evans (1984) argues, however, that the 'availability' account is in some difficulty on evaluation tasks since all possible logical cases are presented to the subject "thus ensuring the 'availability' of each" (Evans, 1984, p.453), and thus that his 'relevance' account is the better explanation. It seem clear that the correctness of this argument rests upon the mechanism by which an item becomes 'available'. If it is merely necessary that a subject be exposed to an item, then Evans' argument is unobjectionable. If, on the other hand (as seems likely) the conditional rule enjoys some sort of special status in the whole presentation (in as much as it is the rule - and the items in it - which

are being reasoned about, rather than the logical cases), then the 'availability' account is as applicable to evaluation tasks as it is to construction tasks, since the superior availability of the items named in the rule is assured.

The 'heuristics and biases' approach is considered in rather more detail later in this chapter, but now the focus now returns as promised to the question of directionality, which has been suggested as a potential source of difficulty on certain inferences.

1.2.6 : Linguistic Form And Directionality Effects

The question of directionality arises in considering whether, for a given linguistic form of expression of P -> Q , it might be more natural (or easier, or more usual) to reason from antecedent to consequent, or *vice versa*. As was mentioned at the outset, the basis of the psychological interest in conditional reasoning is that such natural language expressions of the conditional relation as 'if P then Q' are ambiguous about whether the relation is one of material equivalence or material implication - we want to be able to specify what it is that determines how people will interpret a given statement in such a linguistic form. This is not, however, to say that all linguistic forms of the conditional are ambiguous. The putative natural language definition of material implication offered in Table 1.1, for example, ("*at-least-if* P then Q") whilst clumsy, is unambiguous - but only if the hearer pays careful attention to the "*at-least*". Similarly, that of material equivalence ("*if-and-only-if* P then Q") is unambiguous, but again ungainly and requiring careful attention. The linguistic forms that people actually use, however, tend by contrast to be easy to listen to, but potentially ambiguous. As has

been shown, 'if P then Q' is not by any means the only way of expressing the conditional - there are numerous others, for example, 'never P without Q', 'not P if not Q' and so on - but the alternative form of primary interest here is 'P only if Q' (which will be referred to as the "OI" form), which at first hearing does not appear to mean the same thing at all as 'if P then Q' (the "IT" form). That it does express the same logical relation as 'if P then Q' might be demonstrated by the examples given by Evans (1977) as follows:

IT: If he is a policeman then he is over 5'9" in height

OI: He is a policeman only if he is over 5'9" in height

It should be clear that these sentences are logically equivalent. The IT form appears to be about the set of all policemen (and tells us that all policemen are over 5'9") - it does not, however, deny the existence of people who are over 5'9" but who do not happen to be policemen. The OI form, on the other hand, appears to be about some attribute which a person must have in order also to have the attribute of being a policeman. Here again then, it follows that all policemen are over 5'9" but, as for the IT form, it does not deny the existence of people who are over 5'9" but who do not happen to be policemen.

Logically equivalent these sentences might be - but it should be clear that they are not at all 'saying the same thing'. The important question here is the circumstances in which a speaker would choose to utter one of these in preference to the other. One way of illustrating this point is to imagine these statements to be answers to a hypothetical question - and to ask what that question might be in order to attract one answer as opposed to the other. It is suggested that the IT form might be offered in reply to a question like

(a) Is there any way we can find out how tall this chap is without actually meeting him?

whereas the OI form would seem more suited to a question such as

(b) Is there any way we can find out if he really is a policeman?

Evans (1977) characterises the difference in terms of necessity and sufficiency. In the IT form, it is emphasised that the antecedent is sufficient for the consequent - that is, knowing that he is a policeman is sufficient to know also that he is over 5'9". In the OI form, on the other hand, the emphasis is on the fact that the consequent is necessary for the antecedent - that is, unless the consequent is true, the antecedent cannot be true. On this basis, Evans (1977) proposes that the IT form ('if P then Q') is concerned with reasoning from the antecedent to the consequent, whilst the OI form ('P only if Q') is concerned with reasoning from the consequent to the antecedent. If this is the case, then IT forms should lead more easily to MP and DA inferences, whilst the OI form should lead more easily to MT and AC inferences - and with the exception of DA endorsements, which were equal (and rather low), each of these predictions was significantly supported in that study.

Evans & Beck (1981) observe that such a characterisation leads to the observation that

> "The use of the word 'if' appears to direct attention to the proposition which it modifies, irrespective of the presence of the logically critical word 'only'." (Evans & Beck ,1981, p.111)

That the OI form appears to 'invite' inferences from consequent to antecedent is endorsed by a number of authors (for example, Rips and Marcus, 1977; Braine, 1978). Rips and Marcus (1977) suggest that

directionality effects might best be explained within the framework of a 'suppositional' theory of conditional reasoning, wherein subjects evaluate 'if P then Q' by supposing the truth of the antecedent P and assessing the effect of doing so against what they term their "current database". In as much as no such assessment is made of scenarios in which P is false, Rips and Marcus (1977) predict the D-table; furthermore, they specifically predicted that MP would be endorsed more often than MT on the grounds of the extra difficulty arising out of the presence of negative components. Subjects were given a diamond-shaped presentation of letters and numbers contained in squares together with a statement either in IT or OI form. In addition to evaluating the statement, they were required to report the order in which they checked the squares, the prediction being that for IT statements the antecedent would be checked first, whilst for OI statements the consequent would be checked first. For IT statements, subjects did, indeed, behave as predicted, but for OI statements there was no consistent order of checking.

One way in which directionality differences might arise is by way of temporal or causal factors. Evans (1977) suggests that the normal linguistic function of the two forms might be to express implication according to the time order of antecedent and consequent events. This notion was tested by Evans and Newstead (1977) using a truth table task involving the comparison of sentences such as

[IT1] if the first letter is B then the second letter is R
vs [OI1] the first letter is B only if the second letter is R

and

[IT2] if the second letter is Z then the first letter is T
vs [OI2] the second letter is Z only if the first letter is T.

Evans & Newstead suggested that [IT1] was linguistically more 'natural' than [IT2] in that it had a 'forwards' time order, whilst [OI2] was more natural than [OI1] in virtue of its 'backwards' time order - [IT1] and [OI2] should thus be easier to evaluate correctly than their counterparts. This was indeed found to be the case.

Since the phenomena of reasoning research often turn out to be paradigm-specific, it has been argued that a multi-paradigmatic approach is appropriate (see, for example, Evans, 1982). With this in mind, Evans & Beck (1981) sought support for Evans & Newstead's (1977) findings on a truth table evaluation task, but now using an inference construction task in which negative components were systematically permuted. The effect of directionality was as predicted, in that both 'forward' (MP and DA) inferences were endorsed significantly more often with IT rules, whilst both 'backward' (AC and MT) inferences were endorsed significantly more often with OI rules. The effect of negativity was in line with an explanation in terms of increased task difficulty, in that for both IT and OI forms, there was an overall reduction in all inferences endorsed as the presence of negative components increased (so that, for both IT and OI rules, all four inferences were endorsed most often on AA rules and least often on NN rules).

Interestingly, Evans & Beck (1981) found no sign of 'negative conclusion bias' (as reported, for example, by Evans, 1977). Evans & Beck attribute the non-appearance of this bias to the change in the nature of the task, in as much as their study employed a construction task, whereas Evans (1977) used an evaluation task. Their argument, which if *post hoc* is nevertheless unobjectionable, is essentially that 'negative conclusion bias' is a caution

heuristic (as suggested by Pollard and Evans, 1980) which manifests itself as acceptance of a negative conclusion in situations where a subject has failed to generate a positive conclusion: in a construction task, however, all responses (including those corresponding to negative conclusions) involve the active generation of some actual item.

The typical effects on reasoning behaviour of the presence of negative components and the sentential form of the conditional statement have now been described, and various explanations of these effects have been considered in terms of task difficulty and in terms of heuristics and biases. Before taking stock of the various determinants of behaviour which have been identified so far - and of the theoretical issues raised thereby - one final paradigm will be described.

1.2.7 : The Wason Selection Task

The 'Wason selection task', or 'four card problem', (Wason, 1966) has, arguably, been the single most useful paradigm in conditional reasoning research. Somewhat perversely, perhaps, its major impact is probably not so much in terms of explanations springing from its use so much as in terms of the considerable amount of research which it has spawned - as often by its shortcomings as by its virtues. In this section, a number of abstract selection task studies are reviewed and some of the issues to which these have given rise are considered. The majority of research using the selection task has, however, used thematic materials and thus will not be considered until rather later in this chapter.

What, then, is the Wason selection task? To the uninitiated, the task appears

almost facile. Subjects are shown four cards, on two of which is a number, and on the other two a letter, and are told that each card has a number on one side and a letter on the other. They are then given a conditional statement such as

> If there is a B one side of the card
> then there is a 7 on the other

and are asked to indicate the card or cards (and only those cards) which need to be turned over in order to establish the truth or falsity of the conditional statement. The correct answer is, in fact, that which falsifies the statement, so that if, for example, the cards B, G, 3 and 7 were given with this particular statement, then the only cards which need to be turned over are B and 3, as no other card combination is capable of showing the statement to be false.

Despite its simple appearance, most adults fail to solve the task. This deceptive simplicity has attracted a number of derogatory comments, several of which are quoted with almost perverse pride by Wason (1983): Wetherick (1970) calls it a "laboratory game", Cohen (1981) a "cognitive illusion" - but perhaps the least disputable description of all is Vuyk's (1981) "irritating".

The task was originally conceived as a laboratory analogue of the Popperian hypothetico-deductive approach to scientific theory testing (see, for example, Popper, 1959), in which the central tenet is that in order to be 'scientific', a theory has to be capable of being falsified. Wetherick (1970) has criticised this conception of the task by arguing that it lacks representativeness with respect to how the scientist solves problems, in that it seems to depend on a

'trick'. Wason (1983), however, parries this blow by arguing that the hypothetico-deductive nature of science is based on the idea of falsifying the consequent of a conditional sentence, that there is a trick of some sort in all problem-solving experiments and that

> " ... it is through the discernment of nature's tricks that we make advances in knowledge, and in hindsight we may call them deceptively simple." (Wason, 1983, p.64)

In order to solve the task, a subject must appreciate the principle of falsification - and then work out which cards are able to falsify the statement, and which not. The most frequent selection patterns (around 80% of subjects) were found to be either P alone, or P and Q. Just which cards are seen as able to falsify will, of course, depend upon the interpretation of the conditional: for a statement of the form 'if P then Q', the falsifiers under both I-tables and D-tables are P and ~Q , whilst for E-tables, all four cards need to be turned over. For I-tables and D-tables, then, the typical selections reveal two characteristic errors: the failure to select ~Q (by most of the subjects) and the erroneous selection of Q (by about half of all subjects). For E-tables, the failure to select ~P and ~Q is erroneous. Whichever of these three interpretations subjects might make, then, the failure to select ~Q is erroneous.

There is, however, one important way in which I-tables and D-tables are distinguished. For D-tables, the only cards which might verify the conditional are the frequently selected P and Q cards. Wason thus concluded that this was the interpretation that most subjects were making and that the preferred mode of behaviour is to verify rather than to falsify - this he termed 'verification bias'.

It should be noted, however, that Wason (1966) used only AA rules (that is, rules in which both antecedent and consequent are affirmative) - and that for AA rules, Evans' (1972c) 'matching bias' would predict precisely this selection pattern (since P and Q are precisely the items named in the rule 'if P then Q'). Evans and Lynch (1973) (and later Manktelow and Evans, 1979, using both abstract and thematic materials), set out to distinguish between the 'verification bias' and 'matching bias' accounts. As has been seen, systematically permuting the presence of negative components allows us to compare the frequency of matching and mismatching cards whilst holding the logical case constant. The data from Evans and Lynch (1973) are summarised in Table 1.12, in which card selections are classified on the basis of 'true antecedent' (TA), 'false antecedent' (FA), 'true consequent' (TC) or 'false consequent' (FC). If the selection pattern were due to matching bias, then there should be a TA/FA preference for positive/negative antecedents and a TC/FC preference for positive/negative consequents.

conditional	logical case :	TA	FA	TC	FC
AA :	if P then Q	88	8	50	33
AN :	if P then not Q	92	4	8	58
NA :	if not P then Q	58	29	58	42
NN :	if not P then not Q	54	46	29	75
	overall:	73	22	36	52

Table 1.12 : Evans & Lynch (1973) : Percentage of selections by rule
(adapted from Table 2 from Evans & Lynch, 1973, p.394, in which frequencies were used with n=24)

As may be seen from Table 1.12, the results of this study were, indeed, in line with these predictions. Evans and Lynch (1973) claimed that these

findings refuted Wason's (1966) 'verification bias' account in favour of Evans' (1973a) 'matching bias' account, since only the latter was able to predict the selection patterns found in the presence of negative components.

Reich and Ruth (1982)[23] suggest, however, that in requiring 'verification bias' to manifest itself in a TC/FC preference, Evans & Lynch (1973) and Manktelow & Evans (1979) might have been excessively conservative, since there may be cases where verification behaviour may be masked by concurrent, but equally strong, falsification behaviour. This difficulty arises out of the observation that a number of the choices characteristic of verification, falsification or matching will overlap on certain of the rules: for example, on the AA rule, selection of TC (Q) is predicted for both matching and verifying behaviour; similarly, on the AN rule, selection of FC (Q) is predicted for both matching and falsifying behaviour.

On this basis, Reich and Ruth (1982) argue that an alternative method of analysis would be to consider just those selections which are exclusively predicted by just one of these three tendencies. This exclusivity occurs in just six out of cells out of the sixteen of Table 1.12: Table 1.13 summarises the exclusive predictions concerned. It will be noted from Table 1.13 that each tendency makes exactly two exclusive predictions, so that we might reasonably estimate the predominance of each simply by looking at the mean selections for the two cells concerned. If the Evans and Lynch (1973) data are recast in these terms, we find that the exclusive predictions of 'verification bias' average a selection rate of just 19%, whilst those of 'matching bias', and of the 'correct' falsification strategy, each average 38%.

---

23 This study is actually concerned with thematic, rather than abstract, materials. Nevertheless, the alternative method of analysis which they propose applies equally to abstract materials.

conditional	case	item	tendency
AA : if P then Q	FC	~Q	falsification
AN : if P then not Q	TC	~Q	verification
NA : if not P then Q	FA	P	matching
NA : if not P then Q	FC	~Q	falsifying
NN : if not P then not Q	FA	P	matching
NN : if not P then not Q	TC	~Q	verifying

Table 1.13 : "Exclusive" selection predictions by case and rule

(summarised from Reich & Ruth, 1982, p.401)

This appears to lend further weight to Evans and Lynch's (1973) conclusion that 'matching bias' is better able to account for the data than 'verification bias'. As Reich and Ruth (1982) admit, however, such an analysis is unable to make direct estimates of the response patterns predicted by each of these tendencies, since it (necessarily) ignores the larger part of the data.

Further support for the presence of 'matching bias' on the selection task - and for explanations of 'matching bias' in terms of 'relevance' (Evans, 1984) and 'availability' (Pollard, 1982) - comes from a study by Evans, Ball and Brooks (1987). The study was concerned to investigate in a direct fashion how selections might be influenced by subjects' perceived salience of particular cards. Evans, Ball and Brooks (1987) sought to demonstrate that such an 'attentional' bias was present by investigating the order in which subjects made decisions about the cards. In most earlier studies, subjects had been asked to say merely which cards should be turned over: in this study, however, what was of interest was the order of decisions (including rejections) rather than just of selections. Specifically, it was predicted that a card perceived to be salient would be either selected or rejected earlier than

it would be were it not perceived to be salient. In addition, it was predicted (in line with earlier studies) that matching cards would be selected more often than mismatching cards. These predictions were, indeed, borne out: firstly, there was a strong overall 'matching bias' effect; secondly, decisions about cards mentioned in the rule were made before those not so mentioned. Furthermore, in line with the argument that both these effects result from judgments of a card as 'salient', 'yes' responses (selections) were made significantly earlier overall than 'no' responses (rejections).

It has been argued above that the major importance of Wason's selection task to conditional reasoning research has been in the theoretical and experimental work which it has spawned, and the task will be referred to frequently throughout the remainder of this chapter. In particular, a considerable amount of research has been concerned with various effects of content and context when thematic, rather than abstract, materials are employed on the selection task. Before turning to the thematic reasoning literature, however, it is appropriate to take stock of the various determinants of behaviour so far identified, and to introduce three further theoretical approaches which have been proposed to account for them.

1.3 : Three More Theoretical Frameworks

In this section, three further theoretical approaches to the psychology of conditional reasoning are introduced. First of all, a general account of extra-logical determinants of behaviour is described which Evans (1987), in a critical review of contemporary approaches to reasoning, categorizes as the 'heuristics and biases' approach. Evans' (1984) 'heuristic and analytic' account of reasoning is then considered - this will be seen to be a 'hybrid' theoretical framework, inspired largely by the 'heuristics and biases' approach, and in which both logical and extra-logical determinants of behaviour might be accommodated. Finally, a radically different account of reasoning is described - the notion of 'pragmatic reasoning schemas' (Cheng and Holyoak, 1985; Cheng, Holyoak, Nisbett and Oliver, 1986).

It is clear from the studies reviewed so far that whatever system of logic (if any) people might in some sense 'possess', their behaviour in conditional reasoning experiments is (at least partially) determined by factors which are quite independent of the logical structure of the task or, indeed, of any formal relation which might obtain between antecedent and consequent. This is made quite clear by investigations where a combinatorial approach is used, such that the logical relationships are held constant when the data are pooled - so that any effects still remaining are then (necessarily) founded in extra-logical factors.

The operation of factors which make a task more difficult is simple to characterise: either they simply increase the likelihood that a subject will 'lose track' of what they are doing and so respond on the basis of a 'faulty' sequence of operations, or else the sequence of operations might be so long-

winded that the subject abandons it before its proper completion. Where, however, these extra-logical determinants of behaviour appear in the form of a 'bias' or 'heuristic', then there is an onus upon their proponents to justify the existence of such factors in terms of 'what they might be for' in situations other than experiments in the reasoning laboratory - it would clearly be preposterous to suggest that subjects 'possess' a set of response biases simply for use in reasoning experiments! Put another way, it might be argued that the biases and heuristics described so far - 'matching bias', 'verification bias' and 'negative conclusion bias' (a manifestation of a 'caution heuristic') - are no more than descriptive features of the data: in order to argue that they represent general determinants of behaviour, it is incumbent upon us to identify ways in which they might be adaptively useful in our everyday life.

1.3.1 : The Heuristics and Biases Approach

As a framework for conditional reasoning, the 'heuristics and biases' approach represents a relatively informal marriage of serendipitous discoveries within conditional reasoning research itself with a more theory-driven approach hailing from the field of intuitive statistical judgment.

The formal basis for the framework lies in the work of Tversky and Kahneman (1971, *et seq.*), who proposed a theory of non-rational intuitive statistical judgment in which responses are mediated by two main heuristics: 'representativeness' (Kahneman and Tversky, 1972; 1973; Tversky and Kahneman, 1971; 1974) and 'availability' (Tversky and Kahneman, 1973; 1974). The 'representativeness' heuristic is assumed to be

involved when people are called upon to assess the likelihood of a given sample set of statistical data, and involves a comparison of the salient characteristics (such as the mean) of the sample with the corresponding characteristics of the population from which it is drawn. Of more direct relevance to this consideration of behaviour on reasoning tasks, however, is the 'availability' heuristic which, in the field of statistical judgment, corresponds to the assumption that people make judgments on the basis of the information which they can retrieve from memory.

One of the earliest attempts to apply the 'heuristics and biases' approach to the field of deductive reasoning was that of Pollard (1982), who was particularly concerned with Tversky and Kahneman's (1973) 'availability heuristic'. Pollard (1982) suggests that two distinct sources of availability arises in reasoning studies. The first source of availability, which is relevant only to thematic tasks, is that derived from subjects' experience and which may involve previous exposure either to confirming or disconfirming instances of a statement, or else to associations between the constituent terms of a statement. The second source of availability, which is relevant to both abstract and thematic tasks, is that derived from salient characteristics of the stimulus presentation. Pollard argues that, in both cases, the availability of an item directly 'cues' a subject's response, so that the response is (non-rationally) a function of the cue rather than of a process of 'reasoning': furthermore, although this may lead to 'errors', it will nevertheless (incidentally) produce the 'correct' response some (and in some circumstances most, or even all) of the time.

As has been seen, contemporaneously with Kahneman and Tversky's work in intuitive statistical judgement, a number of extra-logical biases were being identified in the field of conditional reasoning. The first of these,

'matching bias', (Evans, 1973a) has been shown to have similar feature-matching characteristics to the 'atmosphere bias' of syllogistic reasoning (Woodworth and Sells, 1935) and to be compatible with accounts in terms of 'attentional' biases based on salience, such as 'availability' (Pollard, 1982) and 'relevance' (Evans, 1984). In a similar vein, 'negative conclusion bias' (Evans, 1977) might be accounted for in terms of a 'caution heuristic' or, within a hypothetico-deductive framework (see, for example, Popper, 1959) as a bias towards the least falsifiable statement.

Evans (1987) suggests that the general flavour of the 'heuristics and biases' approach is characterised by the assumption that people tend to rely on 'rules-of-thumb' or 'short-cuts' which, whilst effective in some situations, will lead to systematic errors and biases in others.

A rather less informal alliance of these ideas is to be found in Evans' (1984) 'heuristic and analytic' framework, which draws considerable inspiration from the work of Tversky and Kahneman, but which attempts to accommodate both logical and extra-logical determinants of behaviour. This framework is considered in the next section.

1.3.2 : Evans' Heuristic and Analytic Framework

Evans' (1984) proposed his 'heuristic and analytic framework' as a general theoretical framework for reasoning research, wherein he distinguishes these two different types of thought process, which he sees as operating in a strictly serial fashion - first heuristic, then analytic. 'Heuristic thought processes' are seen as being responsible for the selection of items for further

processing, such items coming not only from the information presented in the problem, but also from memory. As for Johnson-Laird's (1983a) mental models, then, we have a mechanism for the combination of the effects of a diverse range of factors. The outcome of the heuristic process is a judgment of 'relevance', such that items deemed 'irrelevant' will simply not be processed further: information deemed 'relevant', on the other hand, will go on to be processed by the analytic stage. Evans sees the heuristic stage as pre-attentive, unconscious, rapid and indescribable by the subject. It is concerned only with the selection of information: it is not concerned with how judgments are made but only with what judgments are about.

Whilst most authors would now agree that 'erroneous' performance does not necessarily imply that subjects are reasoning with some defective logic, there are those (for example, Pollard, 1979; Cohen, 1981) who still argue that a 'correct' response does imply that they are reasoning with 'correct' logic. Evans (1984) argues that in order to see whether subjects are reasoning logically, we must first ascertain beyond doubt what it is that they are reasoning about. Under such a view, then, rationality might be seen as residing in the analytic stage, but as amenable to test only when we can completely specify the heuristic stage.

The framework is compatible with a number of other theoretical approaches: Evans presents it as a development of the view that inferential behaviour should be seen as a form of problem solving (as espoused, for example, by Newell, 1980; Sternberg, 1982). Similarly, it is compatible with Einhorn and Hogarth's (1981) account of 'figure-ground effects' in decision-making, in that the analytic stage is concerned with the testing of the information generated by the heuristic stage.

How well, then, does the framework accommodate the phenomena described throughout this chapter? The phenomenon of 'matching bias' (Evans, 1972c) fits comfortably within this account, and is seen as resulting from a heuristic evaluation of certain items as 'irrelevant' and thus denied to the subsequent analytic stage of processing. In these terms, 'matching bias' might be seen as resulting from certain linguistic determinants of the topic of the sentence: Evans and Beck (1981), for example, have suggested that 'if' appears to direct attention towards the proposition which it modifies (the antecedent) and to lead to the supposition of its truth. Negatives, in a similar vein, are seen as directing attention to the propositions which they deny such that the negation of an item does not alter its topicality. It has already been noted that such an account is supported by Evans' (1983b) finding that in an evaluation task comparing explicit and implicit negation, subjects showed a significant reduction in matching bias as a consequence of explicit negation. This experiment also showed, however, that explicit negation produced an increase in logically correct classifications of TT and TF cases as 'true' and 'false' respectively, which is consistent with the framework in that the reduced elimination of items at the heuristic stage is seen as rendering them more likely to be processed analytically.

The data typical of the Wason Selection Task (Wason, 1966) are also consistent with the framework. On the abstract task, where there is an overall 'matching bias', and a bias towards the selection of TA over FA cases, Evans (1984) argues that selections are completely and utterly determined by the the heuristic stage, being based on linguistically-cued relevance judgments.

What, then, of the analytic stage of processing? Evans (1984) proposes that

whilst the function of this stage is to generate inferences, this is not generally carried out by the application of logical inference rules (of either a formal or an informal nature) to underlying propositional structures. Rather, he argues that on an abstract task, subjects make inferences in accordance with the normal linguistic usage of the syntactic form (which for conditional sentences is 'if P then Q' and, less frequently, 'P only if Q'). Evans argues that the ambiguity of these forms in an abstract setting is due to the fact that when they are used in natural language, there will, in general, be cues available from the context, and the content of the propositions, so as to determine the precise relationship between antecedent and consequent.

In essence, therefore, Evans (1984) argues that in verbal reasoning, analytic processes reflect the operation of linguistic and semantic systems rather than of an abstract logic. Performance on verbal reasoning tasks, then, reflects subjects' knowledge both of language and of particular contexts. Where, as in abstract tasks, the type of experience on which they must draw is not immediately apparent to them, then they must ask themselves in what sort of context such a sentence is normally used, and what types of inference are normally appropriate in such a context.

As will be seen when the thematic literature is reviewed, both the 'heuristics and biases' approach and Evans (1984) 'heuristic and analytic framework' are comfortably able to accommodate the way in which subjects' experience with the contexts and content of realistic reasoning scenarios tends to affect their behaviour on laboratory reasoning tasks. Before turning to the thematic literature, however, one final framework is introduced - Cheng and Holyoak's (1985) 'pragmatic reasoning schemas'.

1.3.3 : Cheng and Holyoak's 'Pragmatic Reasoning Schemas'

Lehman, Lempert and Nisbett (1988) claim to have identified[24] a number of 'pragmatic inference rules' - "naturally occurring inferential rules that people use to solve everyday-life problems" (*ibid*, p.432). In general, such rules might be characterised as follows:

> " *[They]* ... capture recurring regularities among problem goals and among event relationships that people encounter in everyday life. They are fully abstract in that they are not tied to any content domain (much like Piaget's formal operations) but they are not as independent of relationship types and problem goals as formal logical rules (which are included in Piaget's propositional operations) or the purely syntactic rule systems often studied by modern cognitive psychologists."
>
> (Lehman, Lempert and Nisbett, 1988, p.432)

A particular subtype of Lehman, Lempert and Nisbett's systems of 'pragmatic inference rules' are what Cheng and Holyoak (1985) term 'pragmatic reasoning schemas' which might, in turn, be characterised as follows:

> " A pragmatic reasoning schema consists of a set of generalized, context-sensitive rules which, unlike purely syntactic rules, are defined in terms of classes of goals (such as taking desirable actions or making predictions about possible future events) and relationships to these goals (such as cause and effect or precondition and allowable action). Although context-sensitive, the rules that comprise pragmatic schemas may extend beyond the scope of purely syntactic rules of logic, because they will serve to interpret "nonlogical" terms such as *cause* and *predict* as well as terms treated by formal logic, such as *if-then* and *only-if*."
>
> (Cheng and Holyoak, 1985, p.395)

---

[24] (together with their colleagues over the previous five years or so)

A number of studies have considered the way in which (inductive) pragmatic reasoning schemas might be involved in the prediction and explanation of events and in intuitive statistical inference (for example, Nisbett, Krantz, Jepson and Kunda, 1983; Jepson, Krantz and Nisbett, 1983; Fong, Krantz and Nisbett, 1986). In a different vein, there is a growing tradition of studies concerned with extent to which it is possible to train people in the use of pragmatic reasoning schemas (for example, Holland, Holyoak, Nisbett and Thagard, 1986; Cheng, Holyoak, Nisbett and Oliver, 1986; Nisbett, Fong, Lehman and Cheng, 1987; Lehman, Lempert and Nisbett, 1988).

Of particular interest for the present purposes are three specific types of (deductive) pragmatic inference schema. Firstly, 'permission' schemas (Cheng and Holyoak, 1985) are those concerned with situations wherein some precondition needs to be satisfied in order to permit some action to be carried out. Secondly, 'obligation' schemas (Cheng, Holyoak, Nisbett and Oliver, 1986) are those in which some obligation is incurred in virtue of the occurrence of some event: these first two sorts of schema are classified by Lehman, Lempert and Nisbett (1988) as 'contractual' schemas. Finally, 'causal' schemas (of the kind defined by Kelley, 1971; 1973) are argued to underly certain linguistic anomalies (Cheng and Holyoak, 1985) in addition to being directly involved in the solution of reasoning problems.

As will be seen in the review of the 'thematic' literature, a typical effect of subjects' prior experience with particular materials is that it will facilitate 'correct' performance on problems such as Wason's (1966) selection task. Cheng and Holyoak (1985) suggest that facilitation will result when such prior experience acts so as to evoke a particular type of schema which so happens to correspond to the material implication of formal logic: crucially,

however, they point out that not all types of schema will facilitate peformance, so that performance as evaluated by comparison with formal logic depends on what type of schema (if any) is evoked.

So far as abstract materials are concerned, however, Cheng and Holyoak (1985, p.396) argue that "An arbitrary rule, being unrelated to typical life experiences, will not reliably evoke any reasoning schemas". What this account has to say about behaviour on abstract laboratory tasks, therefore, is necessarily somewhat limited[25] - nevertheless, certain of the experimental conditions in Cheng and Holyoak (1975) and Cheng, Holyoak, Nisbett and Oliver (1986) did involve abstract materials.

Of what, then, does a pragmatic reasoning schema consist? The first type of schema described by Cheng and Holyoak (1985) is the 'permission' schema, in which taking a certain action requires that a certain precondition be satisfied. As has been seen, statements in propositional logic (such as, for example, 'if P then Q') may be put together into 'inference patterns' in which "P" and "Q" are context-free symbols, so that each may stand for any proposition without exception. Pragmatic reasoning schemas, in contrast, are context-sensitive, so that the components of the inference patterns are no longer context-free symbols but, rather, may be constrained to include concepts such as possibility, necessity, an action which is to be taken or a precondition which is to be satisfied. A given pragmatic reasoning schema is seen as being evoked in virtue of particular features of the problem in hand: in particular, its relevance to the problem domain and the structural similarity of the problem to the set of procedural rules that underlies the schema.

25 Their application to thematic reasoning tasks will be described in some detail in the relevant section.

Cheng and Holyoak (1985, p.397) summarise the "core" of the permission schema in four production rules. If we express a 'permission' type of rule in the general terms "If (some action) A is to be taken, then (some precondition) P must be satisfied", then the four rules underlying the permission schema are as follows:

Rule-1: If A is to be taken, then P must be satisfied.
Rule-2: If A is not to be taken, then P need not be satisfied.
Rule-3: If P is satisfied, then A may be taken.
Rule-4: If P is not satisfied, then A must not be taken.

This set of rules may be viewed as a set of procedural knowledge for application to the problem in hand - that is, to the combination of the rule and other information about the situation. In order to illustrate how the evocation of the permission schema might facilitate performance, it is appropriate to 'jump the gun' somewhat by the use of an example from a typical thematic study, Griggs and Cox (1982), which utilised a thematic selection task involving statements such as "If a person is drinking beer, then the person must be over 19 years of age". According to Cheng and Holyoak (1985), Rule-1 (which is phrased in very similar terms to the statement itself) has the same effect as Modus Ponens. Rule-2 says that when the action is not to be taken, it is no longer relevant whether the precondition is, or is not, satisfied: the DA inference is thus effectively blocked. Rule-3 says that when the precondition is satisfied, it is then permissible for the action to be taken (but the action is not dictated): the AC inference is thus effectively blocked by this rule. Finally, Rule-4 says that when the precondition is unsatisfied, it is not permissible for the action to be taken: this rule thus corresponds to Modus Tollens. It should be noted, however, that this correspondence should not be taken to imply that the

permission schema is *equivalent* to material implication in formal logic: firstly, the permission schema is context-sensitive, whereas formal logic is context-free; secondly, the permission schema includes deontic concepts such as 'must' and 'may' which are not permissible in the standard propositional logic.

To the extent that a given rule might evoke a permission schema, therefore, then 'correct' behaviour in accordance with a material implication interpretation will be facilitated. Whilst two of the selection task experiments in Cheng and Holyoak (1985) involved thematic materials, in which the ease of evocation of a permission schema was manipulated in terms of familiarity of the propositions and arbitrariness of their interrelationships, Experiment 2 involved totally abstract materials. In Experiment 2, the ease of evocation of a permission schema was manipulated in terms of how the task was described to the subject. In the 'card' condition, the task was presented as a standard abstract selection task involving statements such as "If the card has an 'A' on one side then it must have a '4' on the other". Alternatively, in the 'regulation' condition, the subject was told to suppose that they were an authority checking whether or not people were obeying certain regulations (of the general form "If one is to take action 'A', then one must first satisfy precondition 'P'") and then given four cards corresponding to 'has taken action 'A'', 'has not satisfied precondition 'P'', and so on. Stressing the precondition-action relationship in this way produced very considerable facilitation, despite the fact that the propositions involved were simply abstract symbols: some 61% solved the 'regulation' problem as against a typical 19% for the standard abstract 'card' problem. Cheng and Holyoak (1985) argue that these findings are evidence for the existence of an abstract permission schema.

Cheng, Holyoak, Nisbett and Oliver (1986) were concerned to investigate the effect of various types of training on the solution of the selection task. In Experiments 1 and 2 of that study, they found that neither training with examples of the selection problem nor training in the abstract principles of formal logic (even by way of a one-semester course in logic), could alone improve performance. A combination of the two methods of training did, however, produce a significant improvement: it was argued that logic training alone failed to produce an improvement due to subjects being unable to apply it to any concrete examples, whilst examples training alone failed due to subjects having no intuitive grasp of the rule which they were applying. The results of Experiments 1 and 2 also lent support to those of Cheng and Holyoak (1985), in that problems phrased so as to evoke a permission schema produced significantly more correct solutions than problems which were not so phrased[26].

Although in a different domain, Fong, Krantz and Nisbett (1986) had suggested that even purely abstract training in naturally occurring rule systems could be effective in encouraging people to use such systems. This, together with the findings of their Experiments 1 and 2, led Cheng, Holyoak, Nisbett and Oliver (1986) to propose that if people normally solve 'real world' problems by the use of pragmatic reasoning schemas, then training in the use of such schemas should facilitate performance on laboratory problems for which they were appropriate.

In Experiment 3 of their study, Cheng, Holyoak, Nisbett and Oliver (1986) investigated the effect of training subjects in the use of the second type of

26 Note that in this study, only certain of the thematic tasks were phrased so as to evoke 'permission' schemas, the abstract tasks being phrased in the standard way rather than as 'regulations'.

'contractual' schema - the 'obligation' schema - in which some obligation is incurred in virtue of the occurrence of some event. A thematic example of such a relationship might be a regulation of the form "If a student is a psychology major, then the student must take an introductory psychology course" (Cheng, Holyoak, Nisbett and Oliver, 1986, Appendix B, p.325), which may, more generally, be stated as "If (some event) E occurs, then (some action) O is obligatory". The "core" of the 'obligation schema' may be summarised in four production rules as before:

Rule-1: If E occurs, then O is obligatory.
Rule-2: If E does not occur, the obligation O does not arise.
Rule-3: If O is done ...
if E does occur, the obligation O is satisfied;
if E does not occur, the obligation O does not arise.
Rule-4: If O is not done, then E must not have occurred.

It should be clear that, like the rules for the permission schema, these rules should facilitate behaviour which is in accordance with a material implication interpretation: Rule-1 is equivalent to Modus Ponens, and Rule-4 to Modus Tollens; Rule-2 does not dictate whether or not the consequent should occur, and so blocks the DA inference; Rule-3 makes it clear that it is irrelevant whether or not its consequent (the original antecedent) should occur, and so blocks the AC inference.

Subjects were given either no training at all, training on the nature and use of obligation schemas or else 'contingency training', which was similar to that on obligation schemas except that the checking procedures were described in terms of the assessment of universal contingencies rather than of 'obligations'. Thematic materials were used in seven of the eight problems and, as before, some of the statements were designed to evoke an obligation schema, whilst others contained just arbitrary items. For the

'obligation' problems, 'obligation training' almost eradicated the errors altogether. For the 'arbitrary' problems, both types of training produced some improvement, but 'obligation training' more so than 'contingency training'. On two of the (intendedly) 'arbitrary' problems, a similar pattern of improvement occurred to that on the four 'obligation' problems: *post hoc* examination of the materials involved does, however, suggest that those two problems could reasonably be interpreted as representing rules of a 'regulation' nature for which the obligation schema would be applicable. Of the remaining two 'arbitrary' problems, one was a totally arbitrary statement of a categorical universal, whilst the other was a standard abstract selection task involving letters and numbers on cards. Whilst 'obligation training' did not improve performance any more than did 'contingency training' for these two problems, both types of training did produce some improvement.

Cheng, Holyoak, Nisbett and Oliver (1986) claim that their demonstration of the effectiveness of 'obligation training' in improving selection task performance lends support to the notion that people naturally use such pragmatic knowledge structures in reasoning. Clearly, this result is strongly consonant with Cheng and Holyoak's (1985) evidence for the effectiveness of an evoked permission schema in improving selection task performance.

This introduction to the 'pragmatic reasoning schemas' approach concludes by considering one final type of deductive schema - 'causal' schemas. The notion of causal schemas is not new, having first been proposed in the realm of social psychology by Kelley (1971; 1973), an attribution theorist concerned with their role in the attribution of causes to observed events. Kelley held that people have an understanding of the notions of necessity and sufficiency in causality and that they possess separate schemas for

checking evidence for hypotheses about causes which are both necessary and sufficient (as in material equivalence), sufficient but not necessary (as in material implication), necessary but not sufficient, or neither necessary nor sufficient.

This notion of separate types of causal schema (at least in terms of different sets of checking procedures) is supported by Cheng, Nisbett and Oliver (1987), who found that subjects appeared tacitly to assume a particular type of causality and then to follow the set of checking procedures appropriate to that type. Given that the material implication interpretation of the conditional might be precisely expressed in terms of necessity and sufficiency (in that the antecedent is sufficient but not necessary for the consequent), it follows that if a subject interprets the causality in a problem as being of this type, then their performance on the selection task will be correct.

Lehman, Lempert and Nisbett (1988) claim that people possess a range of distinct 'pragmatic reasoning schemas' for use in appropriate situations. It certainly seems reasonable that the schemas which they propose for use in inductive inference and intuitive statistical inference should be quite distinct from the (deductive) permission, obligation and causal schemas. Amongst the deductive schemas, it also seems reasonable that the 'contractual' schemas of permission and obligation should be distinct from causal schemas, in that the former are concerned, rather specifically, with the realm of social regulations.

In concluding this introduction to 'pragmatic reasoning schemas', it may be interesting to note one somewhat tangential finding of Cheng and Holyoak (1985) which relates to their emphasis upon reasoning being guided by

experientially-induced schematic structures which are evoked by pragmatic interpretation of the situation, rather than by the syntactic interpretation of sentences. The cross-linguistic Whorfian hypothesis - that our thought is shaped by the language we speak - has been argued to apply to conditional forms (see, for example, Bloom, 1981). Cheng and Holyoak's (1985) Experiment 1 (which will be described in the 'thematic' section) compared a group of U.S. students with a group of Hong Kong students. Whilst the choice of groups was based upon their respective cultural experience rather than the language which they spoke, it so happened that the materials were presented to the U.S. group in English, but to the Hong Kong group in Chinese. It turns out that the Chinese language has distinct forms for the two readings of 'if-then'[27], such that a Whorfian should predict more logical performance from a Chinese speaker, since the ambiguity of the English form simply does not arise. Cheng and Holyoak (1985) did not, however, find any difference whatever between the two groups where a straightforward selection task was employed: whilst this is only rather indirect evidence about the reality of pragmatic reasoning schemas, it does certainly lend support to the notion that reasoning behaviour is determined experientially rather than syntactically.

In this section, three theoretical frameworks have been introduced - the 'heuristics and biases' approach, Evans' (1984) 'heuristic and analytic' account and Cheng and Holyoak's (1985) notion of 'pragmatic reasoning schemas'. Each of these frameworks will be referred to from time to time, together with some of those described in earlier sections, throughout the next section, in which the 'thematic' literature is reviewed.

27 (*ruguo jui* for 'if' and *ruguo cai* for 'if-and-only-if')

1.4 : Research Using Thematic Materials

In this section, research using thematic materials is reviewed, and the often substantial effects of propositional content and the context in which the conditional statement is perceived to be uttered are described. The section begins by considering some of the motivations which underly the choice between abstract materials and thematic materials and some of the advantages and disadvantages of each are identified. In contrasting abstract and thematic materials, the broader issue of language as behaviour is addressed and its implications for tasks using each class of materials considered. Typical effects of the perceived context in which the conditional utterance is made are then described, following which the effects of subjects' beliefs about the truth of the rule *per se* are considered.

The use of thematic materials has often been found to facilitate performance on reasoning tasks: this is particularly true of the Wason Selection Task (Wason, 1966), thematic versions of which will be reviewed in some depth. Thematic facilitation on the selection task will, however, be seen to have been far from reliably replicated, and various possible explanations for the highly variable effects will be considered.

An understanding of the effects of thematic content is important not only to any explanation of 'real world' reasoning, but also to the evaluation of the theoretical standpoints encountered so far: the Piagetian standpoint, for example, predicts that adult formal operations should enable an analysis of a problem in a content-free fashion, so that any facilitation by the use of thematic materials would be difficult to explain. The immediate implications of thematic facilitation for the rationality debate are not clear

cut. The rationalists might claim support on the grounds that it arises out of improved understanding (and thus a more accurate interpretation). The non-rationalists, on the other hand, might argue that thematic content increases the influence of extra-logical factors.

1.4.1 : Abstract *versus* Thematic Materials

The primary motivation of the use of abstract materials is the avoidance of any possible effect of the semantic content of the propositions. By using propositions which are totally devoid of meaning it is held that what is observed must reflect a subject's understanding of the nature and meaning of conditionals *qua* conditionals. Tasks using abstract materials might, of course, be criticised on the grounds that they are rather too far removed from real life usage of the conditional. So substantial and varied will the effects of content be shown in this section to be, however, that the attractiveness of content-free materials is clear.

Such a view might, however, be argued to be somewhat nãive, since it will also be shown in this section that the perceived context in which an utterance is made will often be a substantial and systematic determinant of behaviour quite apart from the semantic content of the propositions. Critically underlying the use of abstract materials in the investigation of the essential nature of conditionals is the assumption that the utterance of the conditional statement is made *in vacuo* -that is, the statement is made outside of any linguistic context. It should be clear, however, that utterances in natural language are never issued *in vacuo*,but rather there will always exist a rich context within which the utterance might be interpreted, which

will involve such factors as the topic of the discourse, the social situation, the hearer's beliefs about the speaker's knowledge, intentions and beliefs (including the speaker's beliefs about the hearer's beliefs about the speaker's beliefs and so on).

Given that abstract tasks do attempt to make the conditional utterance *in vacuo,* there appear to be three possible ways in which such tasks might be fatally unrepresentative of real world usage of the conditional. The first potential pitfall arises where the attempt to avoid any perceivable context is successful: since reasoning outside of any particular context is something which subjects will never have experienced in their everyday life, any attempt to generalise the findings to real world reasoning will at best be dubious. Secondly, it might be that so alien to subjects will the absence of any context be that they will make an effort to infer a context in which to reason: since the experimenter will have expressly avoided giving any hint as to the context, it follows that the inferred context will be quite uncontrollable and, given that the perceived context will indeed be seen to affect behaviour, the results would then mean rather little. Thirdly, it might be argued that even in abstract experiments there is a context of sorts in that the statement is uttered in a reasoning experiment taking place in a psychological laboratory, such that the subject is at least able to set the statement within the context of an experimenter wishing to find out how people reason about logical relationships between abstract symbols: since such a context is far removed from any real life context, the generalisability of the results is once again brought into question.

The use of thematic materials is, however, far from being a guarantee that the effects of perceived context will be under the experimenter's control, since the context perceived by the subject will by no means necessarily be

that intended by the experimenter. This potential mismatch between what the speaker (or experimenter) intends to communicate in issuing a conditional utterance, and what the hearer (or subject) perceives to have been communicated, is considered further in the next section.

It is clear, then, that both abstract and thematic materials have both advantages and disadvantages. Abstract materials might appear to be better able to get at the way that subjects understand conditionals *qua* conditionals by avoiding the effects of propositional content - in so doing, however, they rather distance themselves from real world usage of the conditional. Thematic materials might appear rather more natural but the very existence of content effects brings into the question the generalisablity of the results of such experiments to other types of content. Both classes of material have potential problems with perceived context - abstract materials in so far as dealing with conditional statements in the absence of a context is outside the subject's experience and thematic materials in so far as the perceived context might not be that which was intended.

1.4.2 : Language as Behaviour - Speech Acts and Implicature

The idea that in speaking or writing words and sentences we may perform an act which goes beyond the propositional content of the utterance was first systematised by Austin (1962), who proposed the notion of 'speech acts'. For Austin, certain verbs (which he termed 'performative verbs') make the associated speech act explicit: for example, to say 'I request you to leave' makes explicit the fact that the sentence is one the utterance of which performs the act of requesting someone to do something - and it is quite clear

to the listener that this is what the speaker intends. The speech act 'request' will not in general, however, be made explicit in the utterance. For example, the imperative 'get out!' will convey the same information (that the speaker wishes the hearer to leave) without making that fact explicit. Similarly, the assertion 'I'm sick of the sight of you' will, in most circumstances, serve just as well: it should be clear, however, that there are some circumstances in which such an assertion would not have this effect, such as where the speaker was using it hyperbolically to communicate mild annoyance with the hearer (and, indeed, it may be that even the more direct 'Get out!' might be used, or perceived to be have been intended, in this way).

In using a performative verb, then, the act being performed in issuing the utterance is indisputable: there is little or no opportunity for the hearer to misinterpret the speaker's intentions. Where some less direct wording is used, however, the speaker's intentions will need to be interpreted by the hearer. Where, for example, the speaker intends in uttering 'I'm sick of the sight of you' to perform the act of requesting the hearer to leave, the successful performance of the act is dependent upon the hearer's interpretation of the broad context in which the utterance is made - which in turn will depend upon his beliefs about the speaker's general demeanour, how the speaker presently, and habitually, feels about the hearer, what has already passed between the hearer and the speaker in the current discourse, and so on. It is clear that once the meaning of an utterance becomes open to interpretation by the hearer, then the doors are flung open to error or failure of the intended act.

On this basis, Austin argues that in issuing an utterance, a speaker peforms three distinct types of act. The first of these he terms the 'locutionary act',

which is merely the making of sounds (or marks) which correspond to particular words in a language: it is clear that the successful performance of this type of act will in general be free from error, in that it is subject only to such relatively trivial requirements as the clarity of diction (or writing) of the speaker and the hearing (or eyesight) of the hearer, and to the hearer being familiar with the language in question and the words used. The second type of act he terms the 'illocutionary act', which is the act performed by the speaker in issuing the utterance. In the examples given, the illocutionary act intended (or, to use another of Austin's terms, the intended 'illocutionary force' of the utterance) is the request that the hearer leave: it is clear that the successful performance of this type of act is, however, prone to error in terms of how it will be interpreted, in that the hearer might misinterpret the illocutionary force as being a communication of mild annoyance rather than a genuine request to leave. The third type of act Austin terms the 'perlocutionary act', which is the effect actually brought about by issuing the utterance. In the example given, the utterance might (additionally) have the perlocutionary force of amusing, or insulting, the hearer: it is clear that this third type of act is quite beyond the control of the speaker, but, rather, is entirely dependent upon interpretation by the hearer.

The notion of speech acts was extended somewhat by Searle (1969), who set out certain necessary conditions for the successful performance of a speech act: for the example given, these would include such requirements as that the hearer was not deaf (or blind), that the hearer is able to leave, that the speaker believes that the hearer is able to leave and that the speaker wants the hearer to leave. A rather broader specification of the conditions under which discourse normally proceeds was proposed by Grice (1975), who held that linguistic communication is a co-operative enterprise in which the each

participant recognises a mutually accepted direction (which might be fixed *ab initio* or which might emerge and evolve as the exchange proceeds). The direction of the exchange might be fixed or rather vague: but at every stage there will be some possible conversational moves which would be excluded as unsuitable. Grice argues that there is a general principle (his 'Co-operative Principle') which governs all 'talk exchanges': "make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged". Adherence to the the Co-operative Principle is seen as being achieved by a series of 'maxims', the most important of which are the maxims of Quality ('say only that which you know to be true' and 'say only that for which you have sufficient evidence'), the maxims of Quantity (broadly, 'make your contribution neither more nor less informative than is required for the present purposes'), the maxim of Relation ('be relevant') and the maxims of Manner ('be perspicuous'). Grice comments that the first maxim of Quality ('say only that which you know to be true') should be seen as a prerequisite for the other maxims to operate: in a similar vein, Harnish (1976), has suggested a maxim of 'quantity-quality' ("make the strongest relevant claim justifiable by your evidence").

A crucial feature of Grice's argument is that the Co-operative Principle, and the maxims by which it is to be achieved, are known (and adhered to) by all speakers in a particular culture such that discourse might properly proceed on the assumption that all the participants are aware of them and will abide by them. It is clear, however, that participants in talk exchanges do frequently fail to observe these various maxims. Grice identifies four ways in which such failure might occur. Firstly, a speaker might quietly and unostentatiously violate a maxim, which will in general mislead the

hearer: for example, the maxim of Quality would be violated by telling a lie. Secondly, a speaker might openly and obviously opt out of a maxim: for example, he might opt out of the maxims of Quantity by saying 'I can say no more - my lips are sealed'. Thirdly, maxims might fail unavoidably due to a clash - for example, between a maxim of Quantity ('be as informative as is required') and a maxim of Quality ('do not say that for which you have insufficient evidence').

By far the most important type of failure in Grice's argument, however, is that where a speaker blatantly exploits the maxims so as to make what Grice terms a 'conversational implicature'. If it appears to the hearer that the speaker is not trying to mislead him, that he is not opting out and that there is no clash of maxims, then the hearer must reconcile what the speaker had said with his supposition that the speaker is indeed observing the Co-operative Principle and the maxims. Say, for example, that Henry had just said to Sid 'Hello Sid, how is Xavier getting on in his job?' and had received the reply 'Oh, quite well I think. He likes his colleagues and he hasn't been arrested yet'. On the face of it, Sid has failed to observe the maxims of Relation ('be relevant'), and thus might be regarded as having flouted those maxims. In these circumstances, Henry could simply assume that Sid was not observing the Co-operative Principle: Grice argues, however, that this is not what would typically happen. Rather Henry would proceed as follows: (1) Sid has said that Xavier hasn't been arrested yet - which appears to be irrelevant to my question - but (2) I have no reason to suppose that Sid is not observing the Co-operative Principle - (3) the only way that Sid can be observing it is if he believes Xavier to be dishonest - (4) Sid knows (and knows that I know that he knows) that I am capable of seeing that this supposition is necessary - (5) he has done nothing to stop me thinking that

he believes Xavier to be dishonest - so (6) he intends me to think, or at least is willing to let me think, that he believes Xavier to be dishonest. Sid has, then, *exploited* the maxims so as make a 'conversational implicature': he has implicated the additional piece of information that he believes Xavier to be dishonest.

Grice (1975) is by no means the only writer to have proposed a formulation of the laws of language use. For example, several principles had already been proposed by Ducrot (1972), the most important of which are the 'law of exhaustivity' (state the maximum of what you know or believe to be the case on the topic under discussion), the 'law of informativeness' (any statement of P that purports to be a source of information implies that the addressee is ignorant of P, or even that he expects that not-P) and the 'law of economy' (each particular quailification introduced in an affirmative statement should be informative).

In line with the basic concept of semantic informativeness (Bar-Hillel and Carnap, 1953), a definition of 'informativeness' for these purposes is proposed by Harnish (1976) and Atlas and Levinson (1981): 'P is more informative than Q if P entails Q, Q does not entail P, and P is neither logically true nor logically false)'. Atlas and Levinson (1981) further proposed two Maxims of Relativity ('do not say what you believe to be highly noncontroversial' and 'take what you hear to be lowly controversial') and one Principle of Informativeness ('the best interpretation of a sentence is the most informative proposition among the competing interpretations that is consistent with common ground').

It should be clear that there may from time to time be clashes between this Principle of Informativeness and Grice's Maxims of Quantity: Atlas and

Levinson (1981) suggested that in this case Grice's Maxims of Quantity should take precedence. Using a different formulation of the Principle of Informativeness, however, ('read as much into an utterance as is consistent with what you know about the world), Levinson (1983) suggested that there may be occasions when quantity should take second place to informativeness. Wilson and Sperber (1981) and Sperber and Wilson (1983) propose that the clash between quantity and informativeness is better disposed of by replacing all the Gricean maxims by a single "axiom of maximal relevance", whereby the speaker tries their best to produce the most relevant proposition.

Despite the differences in the details of these various formulations, it seems difficult to object to Politzer's (1986) assertion that we can

> " accept as an established fact that in a normal talk exchange, speaker and hearer are submitted to rules that determine the informativeness and the relevance of utterances ... *[and that]* ... Grice's maxims, whatever their deficiencies may be, represent the most coherent, and also the most influential, formulation of those rules." (Politzer, 1986, p.52, italics added)

Politzer (1986) distinguishes two kinds of limitation which are imposed upon the informativeness of utterences by the application of the various conversational principles described above. The first of these, which he terms the "principle of maximal[28] informativeness" concerns the lower limit of informativeness which an utterance should contain, and corresponds to Grice's first maxim of Quantity ('make your contribution as informative as is required')[29]. Politzer applies this principle to the quantitative aspects of

[28] It does, perhaps, seem rather perverse to call the principle governing the lower limit of informativeness the "principle of maximal informativeness" and that for the higher limit the "principle of minimal informativeness" - but these are the terms which Politzer uses.

utterances and describes how it relates to the typical phenomena from studies of disjunction and quantification.

Of more direct interest for our present purposes is Politzer's second principle - the "principle of minimal informativeness" - which concerns the upper limit of informativeness which an utterance should contain, and thus corresponds to Grice's second maxim of Quantity ('do not make your contribution more informative than is required'). This is of particular relevance here, since Politzer considers how this principle applies to inference schemata and then goes on to propose a developmental theory of the "conflicting development and cohabitation of language and logic" (Politzer, 1986, p.64) which he argues may be used to account for the observed discrepancy between subjects' behaviour in reasoning tasks and that dictated by formal logic.

Politzer's theory rests critically upon two observations. The first observation is that the laws of language use can lead to a conception of truth which is at odds with that of formal logic. He argues that in actual use, something which is relevant, but not asserted, is regarded as 'not the case'; in formal logic, on the other hand, something which has not been proved to be false is considered to be true (or, at least, of indeterminate truth). Language use, then, may be seen as being about the assertion of truth and so partitions the world into 'true' and 'not true', whereas formal logic is about falsification, and partitions the world into 'false' and 'not false'. On this basis, a statement which does not happen to be asserted by a speaker will thus be taken to be 'not true' *pragmatically,* yet 'not false' *logically.* This necessarily leads to the observation that the partitions made by each reading of an utterance may

[29] (and also to Ducrot's law of exhaustivity, to Atlas & Levinson's principle of informativenss and to Harnish's maxim of quality-quantity).

not always coincide - and that "language truth" is simply not the same thing as "logical truth".

The second observation, which is the basis for the developmental perspective of the theory, is that children appear to have acquired, and to adhere to, the laws of conversation from such an early age that the development of their logical system must necessarily take place in the presence of a well-developed system of pragmatics. For example, it is well-documented that young children use, and understand, pragmatic inferences (Ackerman, 1978a; 1979), indirect requests (Carrell, 1981), and directives (Ackerman, 1978b; Bock and Hornsby, 1981; Mitchell-Kernan and Kernan, 1977). Indeed, in her book on the acquisition of pragmatics, Bates (1976) proposes that children begin to show the ability to use conversational principles to construct indirect speech acts at the end of the preoperational period. On these grounds, Politzer argues that "the laws of pragmatics and communication are completely acquired before the beginning of adolescence" (Politzer, 1986, p.52).

On the basis of these two observations, Politzer (1986) argues that not only is the early development of pragmatics and logic contemporaneous but also, more importantly, that

> " the late logical development ... occurs against a background of already acquired principles of language use and communication that are incompatible with many rules of formal logic ... *[and thus that]* ... the development of logical thought tends toward a dual and internally contradictory system"
>
> (Politzer, 1986, p.65, italics added)

Politzer argues that the resulting conflict can only be resolved by the differentiation and coordination of these two (competing) subsystems: and it

necessarily follows that unless the logical system attains absolute dominance in the process of conflict resolution, we should expect to find non-logical behaviour in human reasoning. Moreover, to the extent that the coordination and differentiation is not completely achieved, we should expect to find individual differences in reasoning behaviour.

In those cases where pragmatics and logic do need to compete, Politzer proposes that the probability of each of them "taking control" will be determined by the situation and the context in which the utterance is made. On this basis, he suggests that when the overall situation leads to an insufficient difference between these probabilities, then behaviour is as likely to be determined by one subsystem as by the other - and so we should not be surprised by the high degree of within-subject inconsistency seen on many reasoning tasks.

Politzer (1986) also makes some useful observations and conjectures concerning situations in which it may actually be more correct to use one or other of the subsystems, such that people may need to be able to make a definite choice between the respective behaviours dictated by the pragmatic and logical systems. He suggests that there may be "domains of application of the individual's logical-linguistic equipment" such that, for example, one would not wish to use the same rules when testifying in court (when logical conventions might lead to deception) as when taking part in a reasoning experiment (when linguistic conventions might lead to fallacy). In a similar vein, Scribner (1977), in considering the effect of culture on studies of logical ability, suggested that when the format in which material is presented does not "belong to the subjects' repertoire", then they will tend to answer entirely on the basis of pragmatic laws, irrespective of their logical

development.

What should have started to become clear in this section is that in natural linguistic communication, the making of an utterance will tend to be a very much richer piece of behaviour than the mere assertion of the content of a proposition. The framing of an utterance has been seen to be subject to a relatively strict, shared, system of conversational principles. Moreover, its interpretation has been seen to require of the hearer considerable range of different types of knowledge: for example, knowledge of the conversational principles themselves, knowledge about the speaker's beliefs and intentions and, if Politzer's (1986) proposal of two conflicting subsystems is correct, knowledge about how to determine whether logical or pragmatic behaviour is the more appropriate in a given situation.

In addition to these rather particular types of knowledge, it will become clear in the next two sections that our knowledge and beliefs about the world, and the way that people typically behave in it, are also critically involved in our construction of an interpretation of the utterances we apprehend. The next section describes and reviews the particular effects of the perceived linguistic context in which an utterance is made: in the section which follows it, the effects of our personal beliefs about the truth of the rule *per se* are considered.

1.4.3 : Effects of Perceived Context

The context in which an argument is presented has also been shown to influence reasoning performance. In certain contexts, subjects tend to make a *pragmatic* inference - one that does not arise logically, but which rather is

a product of the context. Such phenomena have traditionally been the domain of the social psychologist (for example, Harris and Monaco, 1976; Harris, 1978, who were concerned with courtroom testimony), but have also been studied with respect to conditional reasoning. For example, Staudenmayer (1975), argued that certain contexts induce an equivalence interpretation of the logical structure of the situation, such that DA and AC follow logically.

Similarly, Geis and Zwicky (1971) suggested that certain contexts "invite" or "suggest" a DA or AC inference even though the statement is not explicitly phrased so as to express an equivalence relation (the only interpretation under which these two inferences are logically necessarily true). They argued that the conditional

(M)
if you mow the lawn
(G) then I'll give you $10

not only *"promises"* the truth of

M -> G

but also *"invites the inference of,* or *suggests"*

~M -> ~G

They argued that "in many cases, there is a quasi-regular association between the logical form of a sentence and the form of the inference it invites" and proposed the existence of a general principle of "Conditional Perfection" whereby "a sentence of the form X -> Y invites an inference of the form ~X -> ~Y" (*ibid.*, p.562). This principle of "Conditional Perfection" was argued to be operative in a range of linguistic contexts, including predictions, promises, threats and law-like statements. Geis and Zwicky further argued that "conditionals are understood to be perfected unless the hearer has reason to believe that the converse is false" - although it is not

immediately clear from this paper whether this is proposed to be the case in general or, rather, just in those contexts in which the principle is operative.

Fillenbaum (1975; 1976) argued that the DA inference in such contexts should not be seen as an error, but rather as being 'pragmatically reasonable'. Fillenbaum was concerned primarily with 'inducements' (such as conditional promises) and 'deterrents' (conditional threats and warnings). Although the speaker's purpose in uttering an inducement is to get the hearer to do something, whilst in a deterrent it is to get the hearer to refrain from doing something, the two types of utterance might be seen as similar in so far as they seek "to control the behavior of the addressee by signalling ... *[the speaker's]* ... intentions with regards to the actions of ... *[the hearer]*, ... indicating the consequences of such actions" (Fillenbaum, 1975, p.246, italics added).

One of the tasks employed by Fillenbaum (1975) was an inference endorsement task similar in some respects to those to be described for Experiments 5 and 8. Subjects were given conditional statements of the form 'if P then Q' in contexts which included promises, threats and contingent universals, and required to evaluate certain inferences which might follow therefrom: the DA inference was endorsed 84% of the time on average for promises and threats, as compared with only 67% of the time for contingent universals. In Fillenbaum (1976), the task was repeated for just promises and threats, and a precisely similar mean DA endorsement rate was found.

Whilst the inference endorsement task in Fillenbaum (1976) was limited to promises and threats, the range of inducement contexts was extended in other tasks to include warnings and 'bulls'[30]. These two contexts differ from

promises and threats in that the speaker is not now in control of the truth of the consequent, such that whilst the statement is still uttered so as to modify the hearer's behaviour, it is now a statement of the speaker's belief rather than of their contingent intentions. Whilst none of these further tasks was directly comparable with the tasks in the Experiments 5 to 8, the results did strongly suggest that promises and threats on the one hand, and warnings and bulls on the other, were interpreted in very similar ways. In addition, whilst warnings and bulls were not investigated directly in terms of the inferences which they were seen to support, Fillenbaum argues that on the basis of Grice's (1967) Co-operative Principle, the hearer might infer that the strength of association between antecedent and consequent is still sufficient to mean that "his action will affect in some substantial or significant degree the likelihood of obtaining those consequences" (Fillenbaum, 1976, p.247).

In a later account of the use of conditionals in inducements and deterrents, Fillenbaum (1986) refers to such utterances as "speech act conditionals" amounting to a request to the hearer and stresses that promises and threats are essentially tied to the perlocutionary effects upon the hearer, so that "the point of a promise is not merely to inform ... *[the hearer]* ... of the good consequences ... of some action, but, rather, to try to enforce that action by a (tacit) offer of these consequences" (Fillenbaum, 1986, p.179, italics added).

Fillenbaum (1986) offers a clear and plausible Gricean explanation of the tendency to commit the fallacy of Denial of the Antecedent (that is, to take the FT case as falsifying the conditional). The account is best described using Fillenbaum's notation whereby the consequent (Q) is signed to show the polarity and extent of the desirability of the consequent proposition from

30 Fillenbaum's 'bulls' correspond to the Tip context to be described for Experiments 5 to 8.

the hearer's point of view, so that Q+ is a moderately desirable consequent proposition, Q++ extremely desirable, Q- moderately undesirable and Q– extremely undesirable. In these terms, the Gricean analysis of the DA fallacy for a promise asks the question "Why, offered an inducement of the form 'If P then Q+' might the hearer believe that 'If not P then not Q+' follows?".

The analysis which follows deals only with an inducement, but the related analysis for a deterrent should be so obvious as not to require its explicit description. Fillenbaum argues that in so far as obtaining Q+ is contingent upon doing P, the inducement would lose all force if that contingency were eliminated, in as much as if Q+ were to come about whether or not the hearer did P, the conditionality of the inducement (which is the very thing which makes it an inducement rather than a mere offer of a good outcome) would be lost, such that the mention of P at all would appear to be superfluous or even irrelevant. In Gricean terms, such a statement would thus appear to be in violation of the first maxim of Quantity ('make your contribution as informative as is required') and/or the maxim of Relation ('be relevant'). In order to see the speaker as conforming to the Co-operative Principle, the hearer must find that P is relevant and non-superfluous - and the only way to do that is to make the Conversational Implicature that Q+ is contingent upon P. On these grounds, Fillenbaum argues, quite reasonably, that in inducement contexts, "the 'fallacious' invited inference is not only plausible, but *not* to make it would appear at best foolish, if not perverse" (Fillenbaum, 1986, p.183).

It should be noted that whilst recognition of the contingency (whether by a Conversational Implicature or otherwise) is necessary for the success of an

inducement or deterrent, it may not, however, be sufficent. The actual content of the propositions can also be shown to be important. An inducement or deterrent constitutes an attempt to get the hearer to do (or to refrain from doing) something *via* explicitly stated outcomes, or 'enforcers'. It turns out that the relative extremities of the propositions (that is, whether the enforcer is P+ or P++ when the outcome is Q- or Q–) may also determine the effectiveness of an inducement or deterrent. Fillenbaum (1977) asked subjects to judge various inducement and deterrent sentences in terms of whether they were ordinary or else somehow strange. The sentences were varied in terms not only of the extremity of the goodness (or badness) of the consequent outcome to the hearer (Q+, Q++, Q-, Q–) but also of the value (or cost) of the antecedent enforcer (P+, P++, P-, P–). In some sentences, the extremity of the consequent was similar to that of the antecedent, whilst in others either the antecedent or the consequent was far more extreme. Sentences in which the consequent was more extreme than the antecedent (for example, 'If P+ then Q–': "If you goof off any more then I'll fire you") were generally judged as normal (87% of the time for inducements, 80% for deterrents). Sentences in which the antecedent was more extreme than the consequent, on the other hand, (for example, 'If P++ then Q-': "If you save the child's life then I'll spit in your face") were generally judged as extraordinary (75% of the time for inducements, 60% for deterrents).

Fillenbaum (1977; 1986) interprets these findings as indicating that an inducement will be effective just in case the hearer is able to recognise the extent and polarity of the outcome (that is, whether it is Q+ or Q++) and the hearer wants Q badly enough to do P in exchange: in other words, P must be less costly, or at least no more costly, than Q is valuable. Similarly, a deterrent will be effective just in case the hearer wants to avoid Q badly enough to do P in exchange, where P is less valuable, or at least no more

valuable, than Q is costly.

Most of the literature concerned with effects of the perceived linguistic context of a conditional utterance have, like that cited above, been concerned with purposive conditionals such as threats, promises, tips and warnings. Another context to have received some attention, however, is the causal conditional, which is stated so as to inform the hearer that the occurrence of the antecedent will cause the consequent to come about.

One such study is that of O'Brien, Costa and Overton (1986), who were concerned to compare the evaluation of causal and conditional hypotheses in two different meaningful domains (mechanics and medicine) and a meaningless, arbitrary domain. The study used three different types of task, one of which (the 'diagnosis' task) is not of direct relevance here. The other two tasks involved presenting subjects with a conditional causal statement (for example, "If the thermostat is replaced, then the car will not overheat") either with, or without, a preceding 'setting' (in the example quoted, this was a two-sentence story about someone bringing a car into a garage with an overheating problem and Claudio the mechanic forming the hypothesis that it was overheating because the thermostat was defective). Subjects were given an outcome sentence (for example, "Claudio replaces the thermostat and the car no longer overheats"[31]) and asked to say whether it showed the hypothesis to be correct or incorrect or, alternatively, that it was not certain whether the hypothesis was correct.

Whilst the actual frequencies of each response in each situation are not

31 This corresponds to the TT situation - all four situations (TT, TF, FT & FF) were offered to each subject for each of four exemplars, but in just one of the three types of domain.

reported, these may be approximately calculated from the results presented. Where the setting was presented, around 48% of subjects overall considered the FT case to falsify the conditional (which is consistent with DA and AC inferences), although this occurred more markedly in the mechanical (around 64%) and arbitrary domains (around 55%) than in the medical domain (around 26%). Where no setting was presented, around 37% of subjects overall considered the FT case to falsify the conditional, and this again occurred more markedly in the mechanical (around 42%) and arbitrary domains (around 40%) than in the medical domain (around 29%).

O'Brien *et al* argue that the lower frequencies of falsification responses to the FT case in the medical domain were the result of subjects treating hypotheses in that domain as relatively less open to falsification, and cite in support the particular medical phenomenon of spontaneous remission, in which a patient may recover (Q) despite the non-administration of a hypothesised cure (not-P). That the general effect of providing a setting in which the conditional might be uttered was to increase the number of (non-logical) falsification responses might be interpreted in terms of subjects being better able to bring to bear their general understanding of the way in which the domain operates so that, presented with the setting for the example given above, the FT case (an apparent 'spontaneous remission' of the overheating problem) suggests that the mechanic was incorrect in formulating the hypothesis. The lack of such a setting, on the other hand, might make them less able to bring their knowledge of the *domain* to bear and thus more likely to rely upon their knowledge of *logic*.

Having described the effects of the perceived linguistic context in which an utterance is made, the effects of our personal beliefs about the truth of rule

and conclusion *per se* will now be considered.

1.4.4 : Effects of Belief

One of the earliest reasoning studies using thematic materials was that of Wilkins (1928), who was interested not only in thematic materials *per se,* but also in the effects of subjects' beliefs in the truth of the statements. She compared performance on classical syllogisms using materials which were thematic, abstract or unfamiliar (that is, nonsense or obscure) - and which were either believed or disbelieved. She found that performance with thematic materials was significantly better than that with abstract and unfamilar materials, and (marginally) significantly better with believed than with disbelieved materials. It thus appeared that thematic materials facilitated performance, but that such facilitation would be considerably reduced where there was a conflict between belief and logical structure - this phenomenon was dubbed the 'belief bias effect'.

A number of the rash of studies which followed in the wake of Wilkins were deficient on methodological grounds (for example, Janis and Frick, 1943; Morgan and Morton, 1944; Henle and Michael, 1956), whilst later methodologically sound studies failed to find strong, unambiguous effects (for example Kaufman and Goldstein, 1967; Frase, 1966; 1968). More recently, Evans, Barston and Pollard (1983) found substantial evidence for belief bias, together with an interaction between logic and belief, using thematic materials on a syllogistic reasoning task. In essence, they found that where a syllogism was invalid, but the conclusion was believable, subjects would generally endorse it. On the other hand, valid syllogisms

with unbelievable conclusions were only endorsed at an intermediate level. What this means in essence is that both logic and believability determined responses, but (perhaps not surprisingly) that the effect was far greater when both were in concert.

Oakhill and Johnson-Laird (1985) investigated the effect of belief on syllogistic reasoning by asking subjects to generate conclusions to categorical syllogisms. They were particularly concerned to investigate whether beliefs affected the process of reasoning *per se* or, rather, merely the process of evaluating conclusions offered by the experimenter. The theoretical basis for Oakhill and Johnson-Laird's study was Johnson-Laird's (1983a) 'Mental Models' view of reasoning (which was described in some detail earlier), in which reasoning consists of constructing mental models of the premises and then searching for models which are counterexamples to putative conclusions. Under such an account, a conclusion which a subject believes to be true might affect the reasoning process by biasing them against such a search for counterxamples, whilst a conclusion which they disbelieve might cause them to search especially assiduously. Some support is offered to such an account by Pollard's (1982) notion that the 'availability' heuristic of the statistical judgement literature (Tversky and Kahnemann, 1973) may apply in deductive reasoning also, such that subjects will be more readily endorse a conclusion which is easily available from general knowledge. In syllogistic reasoning, Oakhill and Johnson-Laird suggest that the available conclusions are those which are initially formulated on the basis of the mental model of the premises. Where such a conclusion accords with a subject's belief, then it may be accepted without any further attempt to validate it; where, on the other hand, it is contrary to a subject's belief, then they may be more likely to attempt to refute it by searching for a counterexample.

Oakhill and Johnson-Laird suggest, however, that there is an alternative way in which belief might act, which involves no search whatever: if a putative conclusion is disbelieved, then a subject may modify it to one in which they believe or else say that there is no valid conclusion. For determinate syllogisms (that is, where the premises offered supported a valid conclusion), it was found that subjects were more likely to produce that (correct) conclusion if it was believable than if it was unbelievable, although only where the truth of such conclusions rested upon a matter of definition as opposed to a matter of empirical fact. For indeterminate syllogisms, it was found that subjects were far more likely to produce the correct response ('no valid conclusion') where the conclusion *suggested* by the premises was definitionally, rather than empirically, false. In summary, then, subjects tended not to produce conclusions which were definitionally false: in as much as this was true irrespective of whether illicit conversion would alter the valid conclusion, it was concluded that the belief has its effect not in the distortion of the interpretations of the premises but, rather, "in the process of making the inference and formulating and evaluating putative conclusions" (Oakhill and Johnson-Laird, 1985, p.566) - that is, directly upon the process of inference itself.

Although most reasoning studies concerned with the effects of belief have used categorical syllogisms, one or two conditional reasoning studies have either investigated belief effects directly, or else refer to them in discussion. Van Duyne (1976), from an informal analysis of subjects' justifications for responses on the Wason Selection Task (Wason, 1966), reported that subjects were more likely to respond in accordance with logic when they believed the rule presented to be 'sometimes true' rather than 'always true', and

suggests that this might be the result of a motivation to fulfil their expectations, such that more effort is made to falsify a rule about the *actual* truth of which they are less certain.

In evaluating these findings, however, Pollard and Evans (1981) argued that the use of *post hoc* justifications was methodologically unsound in this study, given the considerable evidence (for example, Evans and Wason, 1976; and, from a wider range of cognitive and social studies, Nisbett and Wilson, 1977) that such justifications may often be *post hoc* rationalisations rather than reflections of genuine introspection. On these grounds, Pollard and Evans (1981) reanalysed Van Duyne's data on the basis of actual response frequencies and found no such effect. Pollard and Evans (1981) pointed out, however, that Van Duyne's procedure had not tested the effect of belief as strongly as it might have done, since the conditional sentences used were generated by the subjects themselves, such that the difference in the strength of subjects' belief in the two types of sentence (that is, 'sometimes true' or 'always true') might easily have been so small as to mask any effect.

With this in mind, Pollard and Evans (1981) sought to extend Van Duyne's (1976) study using sentences generated by subjects as being not only 'always true' and 'usually true', but also 'always false' and 'usually false', such that the effect of necessity/contingency was manipulated by the use of 'always'/'usually' and the effect of truth by 'true'/'false'. In an analysis of response selections on a standard selection task, falsity was found to have a strong facilitation effect in terms of increased selection of not-Q items, but no effect upon the frequency of selection of any other item. In addition, falsity was found to increase significantly the likelihood that subjects would

correctly justify their selections of both logically correct items (P and not-Q), but not their rejection of the logically incorrect items (not-P and Q). For neither measure was there any significant effect of contingency.

Pollard and Evans propose that these results might be explained in terms of an association between the antecedent and the consequent. In 'believed true' sentences, P and Q are positively associated which, in the absence of any stronger determinant of response, may produce a non-logical bias towards the (logically incorrect) selection of P and Q. In 'believed false' sentences, on the other hand, P and Q are negatively associated, while P and not-Q are positively associated which, again in the absence of any stronger determinant of response, may produce a non-logical bias towards the (logically correct) selection of P and not-Q.

Whilst this study does provide some evidence that a subject's belief in the truth of a conditional rule might render them less likely to behave in accordance with logic, it is only concerned with selection frequencies and thus is unable to say anything about the effects of belief upon the tendency to make logically fallacious conditional inferences.

With the exception of Pollard and Evans (1981), the literature is devoid of investigations of the effects of belief on conditional inference - and that study was limited to response frequencies on the Wason Selection Task. Nevertheless, this study, when taken together with the substantial body of work reported on syllogistic reasoning, makes it quite reasonable to suppose that subjects' beliefs might affect their behaviour in other, more traditional, types of conditional reasoning task. Indeed, it might be argued that if Johnson-Laird (1983a) is correct in his suggestion that we reason by the formulation and manipulation of mental models, then the very fact that

beliefs can be reliably shown to affect syllogistic reasoning means necessarily that they will affect conditional reasoning also, since both types of reasoning share the mechanism argued to be critically affected by belief - that is, the search for counterexamples to a putative conclusion.

In describing the particular effects of context and belief, it will have become clear from a number of the studies reported that thematic materials often tend somehow to facilitate performance on reasoning tasks, by leading more often to behaviour which is in accordance with that dictated by formal logic. Such effects of thematic facilitation are particularly widely reported, however, on the Wason Selection Task (Wason, 1966), thematic versions of which will now be reviewed in some depth. Thematic facilitation on the selection task will, however, be seen to have been far from reliably replicated, and various possible explanations for the highly variable effects will be considered.

1.4.5 : Thematic Facilitation in the Wason Selection Task

A very substantial proportion of investigations into thematic facilitation have been concerned with, or at least spawned by, thematic versions of the Wason Selection Task (Wason, 1966), which has already been described and discussed in some detail with respect to its use with abstract materials.

The earliest study to find thematic facilitation on this task was Wason and Shapiro (1971), using materials in which towns were paired with modes of transport[32]. An attempt by Johnson-Laird, Legrenzi and Legrenzi (1972) to

[32] For example, "Every time I to Manchester I travel by car".

replicate that study, using statements about sealed and unsealed envelopes and the values of postage stamps, was entirely successful, as was that of Van Duyne (1974), using universities and fields of study. The effect thus appeared to be established. These early findings were soon, however, to come under fire. The earliest reported failure to replicate was that of Manktelow and Evans (1979), using types of drink and types of meal, and a rash of further failures were to follow.

Materials similar to the the 'towns and transport' materials of Wason and Shapiro (1971) failed to produce the effect in studies by Brown, Keats, Cox and Seggie (1980), Griggs and Cox (1982) and Yachanin and Tweney (1982). Materials similar to the the 'two-tier postal system' of the replication by Johnson-Laird, Legrenzi and Legrenzi (1972) failed to produce the effect for Griggs and Cox (1982) whilst, using materials similar to the 'universities and fields of study' materials of the Van Duyne (1974) replication, Yachanin and Tweney (1982) were similarly unsuccessful.

These failures to replicate by no means knocked the effect out of court, however, as a number of successful replications have been reported, particularly where the materials are at least broadly analagous to the 'two-tier postal system' of Johnson-Laird, Legrenzi and Legrenzi (1972). Examples of such replications are Griggs and Cox (1982) using the Florida drinking-age laws, D'Andrade (described in Rumelhart, 1979; 1980) using the requirement in Sears stores for a signed receipt for purchases of over a certain sum, and Golding (1981) using a two-tier postal system.

Given this mass of conflicting evidence, it is clear that an explanation of the effect in terms of the use of thematic materials *per se* must be far too simplistic. Griggs and Cox (1982) offer an explanation in terms of 'memory-

cueing'. Specifically, performance is facilitated when subjects can recall (1) past experience with the content of the problem, (2) the relationship expressed and (3) a counterexample - that is, some situation which would represent the TF case. So far as their drinking-age problem was concerned, all three of these criteria were met. The requirements that their subjects be familiar with the particular relationship (that between age and legal drinking) was confirmed by way of a questionnaire - which also confirmed that their subjects were *not* familiar with the two-tier postal system of their failed attempt to replicate Johnson-Laird, Legrenzi and Legrenzi (1972).

Johnson-Laird, Legrenzi and Legrenzi's (Italian) subjects were, however, familiar with such a rule, and Griggs and Cox's memory-cueing hypothesis was lent further support by Golding's (1981) postal task, in which she used two groups according to age: those in the 'age>45' group (who were able to recall such a system in Britain) were successful on the task, but those in the younger group were not. Where the familiarity criterion has been satisfied in the examples considered so far, it has been satisfied directly. Griggs (1983) does, however, make an interesting observation about the success of D'Andrade's Sears problem (described in Rumelhart, 1979; 1980), as replicated by Mandler (1980), in so far as none of the subjects in these studies did, in fact, have direct experience of the requirement in Sears stores that receipts for purchases over a certain sum be endorsed by the manager's signature. What they did have, however, was experience with *analogous* materials and relationships, such that Griggs (1983) was led to suggest that the facilitation might arise not only through memory-cueing, but also through 'reasoning-by-analogy' (and thus that experience with the specific content is not required).

Griggs (1983) explains the failure of Yachanin and Tweney (1982) to replicate the results of Van Duyne's (1974) 'university and fields of study' problem as being due to their not having employed rules where the correct selections represented well-known counterexamples, whereas Van Duyne's correct selections consisted of academically bizarre[33] combinations of subject and university.

Yachanin and Tweney (1982) have proposed that formal reasoning (such as that required by the abstract version of the selection task) imposes additional cognitive demands, so motivating subjects to use heuristics so as to reduce the cognitive load. Griggs (1983) suggests that matching bias might be seen as a short-circuiting heuristic in that it involves the use of systematic responses in place of actual reasoning operations: he believes, in fact, that all cases of facilitation might be seen as effects of the use of short-circuiting strategies (and cites in support the lack of any significant transfer of ability from abstract to thematic versions of the task). Griggs is, here, in line with, for example, Manktelow and Evans (1979), Pollard (1979) and Manktelow (1979) in holding that in interpretating the results, it is necessary to ask when the tasks are, in fact, *reasoning* tasks and when, rather, they are *memory* tasks.

In this vein, Griggs argues that some of the disparities within the results of the thematic studies reported above might be the result of qualitative differences in the perceived nature of the task (whether consciously so or not). Doherty (reported by Griggs, 1983), for example, argues that in D'Andrade's Sears problem, Griggs and Cox's (1982) drinking-age problem and Johnson-Laird, Legrenzi and Legrenzi's (1972) two-tier postage problem, subjects reason *from* a rule rather than *about* a rule (as in the original

33 (And clearly so to the subjects).

abstract selection task), in so far as the rules in question are 'procedural' rules (that is, rules which are to be followed) rather than rules which are simply true or false (as in the abstract task). Indeed, a number of commentators (for example, Griggs and Cox ,1983; Yachanin, 1986) have suggested that virtually all the studies where thematic facilitation was found used materials which had to do with laws and regulations. When such materials are employed, Doherty argues that the task confronting the subject is now qualitatively different in that the subject now seeks only to confirm whether each of the four instances obeys some procedural rule, rather than to determine the truth status of some putative universal generalisation. In essence, then, the task appears to have changed from checking the truth status of a rule to checking conformity with a rule, and Griggs (1983) argues that these seem to be psychologically distinct operations.

Griggs (1983) suggests that additional alterations to the nature of the task might result from the wording of the instructions given to subjects, and that in general the instructions in the thematic tasks reported might have biased subjects towards falsification (thus facilitating correct performance). He identifies three distinct types of instruction: subjects might be asked (1) to select items which determine whether the rule is *violated* (as in Johnson-Laird, Legrenzi and Legrenzi, 1972; Van Duyne, 1974; Cox and Griggs, 1982; Griggs and Cox; 1982), (2) to say whether the items *break* the rule (as in Golding, 1981) or (3) to determine whether or not the rule was *followed* (as in D'Andrade, reported in Rumelhart, 1979; 1980).

Whilst arguing that directing the subject towards falsification in this way might be a necessary condition for substantial facilitation, Griggs does, however, suggest that it is not a sufficient condition for falsification. Indeed,

Griggs (1984) and Yachanin (1986), asking subjects to check for violations of the rules where the materials were otherwise arbitrary, have demonstrated that merely directing subjects to check for violations is not in itself sufficient to lead to facilitation.

In a similar vein to those writers who have suggested that the use of familiar materials might turn a reasoning task into a memory task (for example, Manktelow and Evans, 1979; Pollard, 1979; Manktelow, 1979; Griggs, 1983), Politzer (1986) proposes that the use of familar/thematic content might lead the subject to by-pass a logical rule that is somehow blocked by pragmatic principles. He argues that facilitation might come about by familar content allowing the subject to subsume the situation under a 'superrule' which they have already experienced and which happens to control their response logically.

Rumelhart (1980) offers an explanation of the effects which, like Griggs' (1983) proposals, is based on subjects' specific experience of the materials, but which is rather more specific about how such experience might lead to facilitation. He proposes that problems are encoded into 'schemata', so that

> " ... once we can 'understand' the situation by encoding it in terms of a relatively rich set of of schemata, the conceptual constraints of the schemata can be brought into play and the problem readily solved. It is as if the schema already contains all of the mechanisms ordinarily required in the use of the schemata. Thus, understanding the problem and solving it are essentially the same thing ..." Rumelhart (1980)

The notion of schemata is not, of course, new, having been introduced by Bartlett (1932) in order to explain certain phenomena of memory for stories. Essentially, schemata might be seen as packets of information stored in

memory which represent general knowledge about objects, situations, events or actions and which are created and maintained dynamically (see, for example, Cohen, 1986).

Mandler (1980) proposes that thinking is controlled by schemata, which are constructed around specific content areas and able to be "flexibly applied to varying content within a given area" (Mandler, 1980, p.27). Mandler holds that the facilitation seen in D'Andrade's Sears problem is evidence that reasoning knowledge is context-dependent and that performance on the selection task is the result only of "knowledge of what matters in a familiar setting" (Mandler, 1980, p.32). In a similar vein, Rumelhart (1979) had to say of the D'Andrade study that "this is exactly the result we would expect if our knowledge of reasoning is embedded in task-specific procedures rather than in general rules of inference".

The role of schema-based world knowledge in conditional reasoning was investigated directly by Clement and Falmagne (1986), who manipulated both relatedness (in terms of how readily the premises could be related to each other within a common schema) and the imageability of the content. Whilst neither variable produced a main effect, it was found that subjects performed significantly better (in terms of recognising DA and AC as indeterminate rather than necessarily true) when the premises were both schematically related and highly imageable. Clement and Falmagne argued that such facilitation arose

> " ... because (1) relatedness permitted access of knowledge schemata and schema-related elaborative processing, allowing generation of counterexamples to invalid inferences, and (2) imaginal representations supported elaboration as well as processing of problems in working memory." (*ibid.*, p.306).

A rather more specific schema-based account of facilitation effects on the selection task is offered by Cheng and Holyoak (1985) and Cheng, Holyoak, Nisbett and Oliver (1986) in terms of their notion of 'pragmatic reasoning schemas', which have already been described in some detail both generally and with respect to their application in abstract reasoning tasks. For the pragmatic reasoning schema camp, the role of prior experience is seen as being in the induction and evocation of certain types of schema. Given that only some types of schema will lead to responses which accord with formal logic, it follows that the logicality of a subject's behaviour will thus depend upon exactly what type of schema (if any) is evoked by the interaction of the problem with their prior experience. As was seen earlier, Cheng *et al* propose that permission and obligation schemas both consist of a set of four rules which, between them, provide for Modus Ponens (MP) and Modus Tollens (MT) inferences, yet effectively block the fallacious Denial of the Antecedent (DA) and Affirmation of the Consequent (AC) inferences.

It was also seen earlier that logical behaviour may be facilitated by the causal schema, but only when the particular version of the causal schema evoked is one which is concerned with checking hypotheses for causes which are sufficient, but not necessary, for the outcome. Evidence presented by the pragmatic reasoning schema theorists has already been presented in support of the reality of such schemas, in as much as training in the use of particular types of schema can be shown to facilitate logical performance (for example, Cheng and Holyoak, 1985; Cheng, Holyoak, Nisbett and Oliver, 1986; Fong, Krantz and Nisbett, 1986; Cheng, Nisbett and Oliver, 1987; Lehman, Lempert and Nisbett, 1988).

Cheng and Holyoak (1985) argue that most of the thematic studies which have shown a thematic facilitation effect have used materials which would

fit a permission schema - that is, one in which taking a particular action requires the satisfaction of a particular precondition[34]. On these grounds, they proposed that if a 'permission schema' could somehow be evoked by the experimental conditions, then logical performance should be facilitated. Experiment 1 of their study used one group of students from the U.S. and another from Hong Kong. Each group was presented with two types of selection task problem: one problem concerned a two-tier postal system, whilst the other used a supposed law by which immigrants (but not transit visitors) to the Philippines were required to have been inoculated against cholera. Whilst neither group had direct experience of such an immigration law, just the Hong Kong group had recent, direct experience of a two-tier postal system. Cheng and Holyoak argued that if the permission schema could be evoked, then both groups should perform well on both tasks; without such evocation, however, just the Hong Kong group should do well on just the postal problem.

The method used to evoke the permission schema was to present a detailed rationale for the rule in advance of the rule itself, such that the rule was quite clearly about an action and its precondition. The results upheld the predictions impressively strongly, the vast majority of subjects in both groups performing in accordance with logic on both tasks where a rationale was given, but just the Hong Kong group performing well on just the postal problem in the absence of a rationale.

Cheng and Holyoak (1985) argue that their results cannot be accounted for by an explanation in terms merely of specific experience, since neither

[34] This is clearly consonant with the observation by Griggs and Cox (1983) and Yachanin (1985) that most of these 'successful' tasks tended to have used materials which had to do with laws and regulations.

group had experience of such an immigration law. What facilitated performance in this case was the provision of additional information which made explicit the fact that the rule was about a regulation in which the antecedent and consequent were related to each other as action and precondition respectively. The point is made more emphatically still by the observation that the rationale enabled the U.S. group to do just as well as the Hong Kong group on the postal problem despite their having no experience whatever of such a system.

Whilst offering considerable corroborative evidence in support of the reality of pragmatic reasoning schemas, those theorists do not deny the possible existence and use of alternative knowledge structures in reasoning, not only between, but also within subjects. Indeed, Cheng, Holyoak, Nisbett and Oliver (1986) have proposed a set of relations between logical rules, specific experience and pragmatic reasoning schemas. As to the actual mechanism of facilitation, Cheng and Holyoak (1985) point out that permission schemas, for example, consist of procedural rules for checking whether a rule has been violated: whilst Griggs (1984) and Yachanin (1986) have shown that directing subjects towards checking for violations is not in itself sufficient for facilitation, the very fact that pragmatic reasoning schemas are context-sensitive[35] may be taken to suggest that the permission schema will only be evoked (and the subject thus led towards checking for violations) precisely when the rule expresses an action-precondition relationship.

It should be noted, however, that not all evocations of 'contractual' schemas will necessarily result in facilitation. For example, some obligations are pragmatically biconditional. Cheng and Holyoak (1985, p.410) give the

[35] (in that they will be evoked on the basis of domain relavance and structural similarity of the problem to the schema).

example "If a child has reached age six, then he or she must enter school", in which the consequent clearly does imply the antecedent pragmatically, and argue that in such situations, pragmatic considerations may result in nonlogical behaviour irrespective of the fact that the subject may check for violations.

Whilst pragmatic reasoning schemas do appear to be able to account well for behaviour on the selection task, it is clear that they cannot account for the whole of such behaviour. For example, Cheng and Holyoak admit that they cannot account for the observations of Wason and Green (1984) that subjects perform better when presented with an array of choices which is restricted to the consequent (Q and not-Q) cases only, nor when the rule is concerned with features of a unitary object (as in "If the figure on the card is a triangle then it has been coloured red") as opposed to features of disjoint objects.

The fact that subjects perform better on the Reduced Array Selection Task (RAST: see, for example, Johnson-Laird and Wason, 1970; Wason and Green, 1984), and with rules about unitary objects, is generally thought likely to be due to the reduction in cognitive load imposed by such versions of the task. Indeed, it is for this reason that the RAST has been employed in a number of developmental studies of conditional reasoning. An example of a developmental study using RAST which is of particular relevance here is that of Girotto, Light and Colbourn (1988), in which rules were phrased either as straightforward quantified universals[36] (such as 'All the buzzing

[36] The decision to use the quantified form was, presumably, in order to make the rules sound as natural as possible. There is strongly suggestive evidence, however, that the two forms (quantified and conditional) need not be differentiated in developmental studies. For example, neither Johnson-Laird and Wason (1970) using the RAST, nor Kuhn (1977) using a syllogistic task, found any differences in children's reasoning performance between these two forms.

bees are inside') or else so as to evoke a permission schema (such as 'All the buzzing bees must stay inside in the evening'). When the straightforward rule was given, only 11% of solutions were correct (which, for the RAST, is just not-Q), whilst the 'permission' rule yielded a remarkable 70% of correct solutions. Quite apart from the fact that this lends strong support to the general notion that a suitable pragmatic reasoning schema may be evoked by the nature of the materials so as to facilitate performance (in line with Cheng and Holyoak, 1985), this result was obtained using nine- and ten-year-old children, who are at their "pre-formal" stage of development and thus would not be expected to be able to solve the task using the formal operations of logic (that solution of this task requires formal operation reasoning is well-established - see, for example, Beth and Piaget, 1966; Kuhn, 1977; O'Brien and Overton, 1980; 1982).

This section has described various accounts put forward to explain thematic facilitation. Whilst the Wason selection task has been used as a vehicle for their presentation, it should be clear that none of the various mechanisms proposed to account for the effects of thematic materials is in any way necessarily restricted in application to that task.

This completes the review of the literature on reasoning research using thematic materials. In the final section which follows, the various theoretical approaches identified throughout this chapter will be summarised and contrasted so as to consider just how well each is able to account for the panoply of established effects over the whole range of conditional reasoning tasks, both abstract and thematic.

1.5 : The Major Approaches Revisited

This review has described a wide of range of studies of human behaviour in conditional reasoning. It has been made clear that whilst performance may accord with standard logic some of the time, this is by no means always the case. Where performance does diverge from logic, however, it often does so in a far from random fashion: rather, the 'errors' that people make tend very often to be systematic and predictable. A number of theoretical approaches of conditional reasoning have been described which attempt to explain the mechanisms that underlie what people do when presented with a reasoning task. Some of these approaches have been seen to be concerned mainly with accounting for systematic diversions from logical behaviour, whilst others were concerned with the mechanisms determining behaviour which is in accord with formal logic.

Byrne and Johnson-Laird (1990) identify three classes of theory about the processes involved in deductive reasoning. The first class of theory, exemplified by Braine (1978), rests upon the assumption that the mind contains formal rules of inference. The second class of theory, exemplified by Cheng and Holyoak's (1985) pragmatic reasoning schemas, proposes mental rules of inference with specific content and in a particular form. Finally, the third class of theory, exemplified by Johnson-Laird's (1983a) mental models approach, characterises inference as a search for alternative semantic interpretations. Byrne and Johnson-Laird (1990) do admit the possibility that each theory might might account for how some people reason some of the time but argue that such a state of affairs is neither parsimonious nor open to empirical falsification. Unwholesome as this position may be, however, it seems likely that a single, unifying theory of

deductive inference will continue to elude us.

In this final section, each of the major theoretical approaches will be revisited briefly not only so as to make clear the differences between them, but also in an attempt made to identify some common ground and to consider what sort of flavour a unified account of conditional reasoning might be likely to have. Since detailed expositions of, and references to, the various approaches appear earlier in this chapter, these will largely be taken as read for the purposes of this section.

First and foremost, we have seen that the suggestion that people possess some 'competence' system of formal logic for use in reasoning (as implied by all but the later work of Piaget, for example) is in difficulty on a number of grounds. As Wason (1977) has pointed out, people simply do not apply the formal combinatorial analysis required by such a system. More importantly, the general competence implied by possession of a formal logical system is at odds with the fact that reasoning performance tends to be highly task-specific, context-specific and content-specific. Perhaps the most damning indictment of the suggestion that people possess a formal logical system is the difficulty of accounting for how they acquire it, given that adults do not display consistently logical behaviour from which a child night acquire the system by observation.

Nevertheless, that people do sometimes behave in accord with logic must be accounted for. One approach which overcomes many of the objections to a formal logical system is that people possess some sort of natural logical system, which leads to similar behaviour to that dictated by formal logic but without the need for knowledge of formal proof procedures. Whilst such

approaches offer a plausible account of how people get many reasoning problems right, and others wrong, they are generally unable to make predictions about behaviour outside the operations which they embody explicitly. Moreover, they find themselves in some difficulty in accounting for the effects of problem content and for the systematic nature of many reasoning errors. In addition, they come up against a similar problem to that of the formal logic approaches in terms of accounting for how such systems are acquired.

Johnson-Laird's (1983a) mental models approach overcomes many these difficulties by proposing that reasoning consists in the construction and manipulation of integrated abstract models of the problem (such that the effects of problem content and context are automatically taken into account) and then carrying out a search for counterexamples to putative conclusions. The problem of acquisition is not entirely resolved, however, since even though the procedures involved are seen as compositional, Johnson-Laird is obliged to admit that their construction ultimately relies upon "an innate armamentarium of data and procedures". Perhaps a more serious problem for this account, however, as Evans (1989) points out, is that it fails to explain why reasoning performance on abstract tasks tends to be so very poor, since construction and manipulation of the models involved for purely abstract materials ought to be especially easy.

Not only does the mental models account fail to predict that abstract tasks will be so difficult, but it also fails to explain the highly systematic nature of many of the nonlogical behaviours typical on such tasks. Rather more successful in accounting for such patterns of behaviour have been the various approaches in terms of heuristics and biases, which stress the systematic effects of totally extra-logical factors such as the perceived

salience of particular aspects of the task, a general tendency towards caution, a preference for conclusions which accord with one's belief and the use of 'rules of thumb' and 'short cuts' to reduce the cognitive effort required in solving a problem. Such accounts are, indeed, able to predict many of the typical response patterns on reasoning tasks, whilst being based on determinants of behaviour which have a high face validity - that is, which have highly plausible reasons to exist as general problem-solving strategies.

Much of the success of such accounts tends, however, to have been in the realm of abstract materials, and the predictions which they make often do not travel well to thematic problems. This is not to say, however, that there is anything incorrect about such accounts: rather, it is more likely to be the case that the additional, richer determinants of behaviour which exist in thematic reasoning situations are simply stronger in their effect. Many of the heuristics and biases seen to affect behaviour on abstract tasks may, perhaps, be seen as primitive strategies which will operate in the absence of any sounder basis upon which to behave.

When thematic materials are used, such that the task more closely approximates the sort of reasoning which people undertake in their everyday life, it has been seen that peoples' knowledge and experience about such things as the problem domain, socially shared rules of discourse and the world in general tend to have a significant effect upon reasoning behaviour - and that such effects can often result in behaviours which are more in accord with formal logic.

A number of approaches to the explanation of such effects have been described. Whilst apparently somewhat disparate in nature, what such

accounts have in common is that various aspects of the task have the effect of cueing stored knowledge which is thus made available for use in solving the problem. Whilst the range of particular types of knowledge representation implicated by these various approaches has been considerable, most of the accounts might be characterised as being essentially schema-based. What differs amongst the various accounts is the level of abstraction of the knowledge contained in such schemas. At one end of the scale, Johnson-Laird's mental models account involves schemas which are highly abstract: at the other end of the scale, theorists such as Rumelhart (1979) and Mandler (1980) talk of schemas which are highly domain-specific. Cheng and Holyoak's (1985) pragmatic reasoning schemas are at an intermediate level of abstraction, such that they contain rich knowledge about how certain rather particular sorts of event might be related to each other, but are not tied to any particular content domain.

Schema-based accounts might appear to be rather a different kettle of fish from those which have to do more directly with peoples' memories of specific experiences, or those which emphasise the use of heuristics and biases. These various approaches do, however, turn out to have a number of features in common. For example, Wason (1983) has argued that schema-based accounts are at least analogous to the accounts by Griggs in terms of reasoning-by-analogy, and by Pollard in terms of 'scenarios' and 'availability', whilst Griggs (1983) points to a more specific similarity in arguing that memory-cueing, reasoning-by-analogy and schema-cueing can all be seen as reasoning strategies which 'short-circuit' a 'logical' reasoning mechanism.

None of the various accounts described throughout this chapter has been

proposed in the absence of reasonable and plausible arguments for its reality - and most have gained considerable support from empirical test of their specific predictions. The important point is that none of the accounts is only able - nor indeed claims to be able - to make predictions about behaviour across the whole range of reasoning tasks and problem domains. Perhaps unsurprisingly, those accounts which claim to be broadest in application are also those which are least specific about the finer details of the mechanisms involved, whilst those which make the most specific predictions are applicable only in a relatively restricted range of situations.

It may be that a unified account of human conditional reasoning is an unreasonable thing for which to strive, particularly in view of the observation by a number of writers that the solution of many of the typical tasks may not actually involve reasoning at all. It is clear, however, that any such account would have to take into account not only the operation of a very large number of disparate determinants of behaviour, but also to specify the mechanism for the resolution of the competition between conflicting determinants in any given situation.

Chapter Two

Experiments 1 to 4:
The 'Fuzzy' Experiments

2.1 : Introduction to Experiments 1 to 4

This Chapter presents Experiments 1 to 4. The essential notion underlying these Experiments was that people interpret conditional statements of the form "if P then Q" or "X only if Y" in a 'fuzzy' way - that is, as if these respective statements were actually something more along the lines of "if P then *as near as makes no odds* Q" and *"generally speaking* X only if Y".

The way in which the term 'fuzzy' will be used in the context of this thesis was set out in the previous chapter, but it may be useful to summarise its use here. Where a statement about a given set of instances is said to be 'fuzzily true' with respect to that set, then this denotes that the set contains some certain (small) proportion of (logically falsifying) TF cases and that a substantial proportion of the remaining items are (logically verifying) TT cases. Saying that a statement is 'fuzzily false' with respect to such a set, on the other hand, is intended to convey that the set contains some certain (small) proportion of (logically verifying) TT cases and that a substantial proportion of the remaining items are (logically falsifying) TF cases.

As was stated earlier, the ascription of the term 'fuzzy' is assymmetrical with respect to formal logic. Where a statement is 'fuzzily true' with respect to a set of instances (or a set is to be constructed such that a given statement about it is shown to be true), then even a single TF case in the set renders the statement *logically* false. Where, on the other hand, a statement is 'fuzzily false' with respect to a set of instances (or a set is to be constructed such that a given statement about it is shown to be false), then the inclusion of any number of TT cases does not in any way make a logically false statement logically true, since so long as at least one TF case is present, the statement is still logically false.

2.2 : Experiment 1

Experiment 1 was an evaluation task in which subjects were shown computer-generated arrays of 240 coloured symbols and were required to indicate (by moving an arrow on a scale running from "ABSOLUTELY TRUE" through "CAN'T TELL" to "ABSOLUTELY FALSE") the extent to which a given statement about the display was true or false.

The arrays of symbols were constructed so as to be fuzzy to varying extents - that is, so that the statement about the array was either absolutely true or false, or else fuzzily true or false. This manipulation was achieved by employing five different types of array with respect to the respective proportions of verifying examples (TT cases) and counterexamples (TF cases). Whilst the number of FT and FF cases were each held constant at 60 of each, the number of TT and TF cases were varied as follows:

'absolutely true'	TT	**120**	: **0**	TF
'fuzzily true'	TT	**116**	: **4**	TF
(intermediate)	TT	**60**	: **60**	TF
'fuzzily false'	TT	**4**	: **116**	TF
'absolutely false'	TT	**0**	: **120**	TF

It was hypothesised as follows:

(1) fuzzily true statements (TT:TF 116:4) would be rated as true

(2) fuzzily true statements (TT:TF 116:4) would be rated as less true than absolutely true statements (TT:TF 120:0)

(3) all other statements (TT:TF 60:60, 4:116, 0:120) would be rated false

(4) fuzzily false statements (TT:TF 4:116) would be rated as less false than absolutely false statements (TT:TF 0:120)

It is clear that Experiment 1 represented something of a departure from the traditional methodology, in so far as the arrays and statements presented - and the responses permitted - involved notions of fuzzy truth.

2.2.1 : Design and Methodology

Subjects:

24 first-year BSc (Hons) Psychology undergraduates at Plymouth Polytechnic (18 female, 6 male), mean age 20.0 years.

Materials:

The experiment was conducted on a BBC-B microcomputer fitted with a Computer Concepts Graphics ROM. The graphical display was presented on a Microvitec Cub 653 high resolution monitor in BBC BASIC Mode 1. Subjects were presented with an array of 240 coloured symbols (12 up-down x 20 left-right) above which was a statement about the colour and shape of the symbols in the array. At the top of the screen there was presented a scale marked with three labels - "ABSOLUTELY FALSE", "CAN'T TELL" and "ABSOLUTELY TRUE". Subjects pressed keys on the computer keyboard to move an arrow along the scale, the arrow jumping between 21 marked graduations. A screen dump of a typical screen is presented in Appendix 1a.

The statements were of the general syntactic form 'if P then Q' (IT) or 'P only if Q' (OI) and conformed to the following specific conventions:

(IT): IF THE SYMBOLS ARE (NOT) <colour/shape>
THEN THEY ARE (NOT) <shape/colour>

(OI): THE SYMBOLS ARE (NOT) <colour/shape>
ONLY IF THEY ARE (NOT) <shape/colour>"

The polarity (that is, the presence or absence of "NOT") of the antecedent and consequent propositions was manipulated systematically such that for each of the syntactic forms IT and OI, subjects received all four possible combinations of affirmative ('A') and negated ('N') antecedents and consequents. There were thus eight syntax x polarity combinations, which might be denoted by ITAA (which signifies IT syntax with both the antecedent and the consequent affirmed), ITAN, ITNA, ITNN, OIAA, OIAN, OINA and OINN.

Nine different symbols were used, comprising all possible combinations of three colours (yellow, blue and red) and three shapes (circular, square and triangular). Half the subjects were given a shape at the antecedent of each rule and a colour at the consequent throughout: the remaining subjects were given the reverse. The actual colour/shape named in each proposition of the statement was randomly determined for each trial.

The number of each of the nine symbols to be displayed was determined on the basis of its logical truth case with respect to the antecedent and consequent propositions in the statement (and their respective polarity). Of the 240 symbols in the array, there were always 60 symbols representing the FT case(s) and 60 the FF case(s). The ratio of TT to TF symbols was systematically varied between 120:0, 116:4, 60:60, 4:116 and 0:120, such that subjects received each of these five TT:TF proportions exactly once for each of the eight possible syntax x polarity combinations, making 40 trials in all.

As has been stated above, the number of discrete types of symbol encompassed by the various truth cases was variously one, two and four: where more than one symbol could be used to represent a truth case, each such symbol was used in equal number. The absolute position of each symbol in the 20 x 12 array was randomly determined for each trial, whilst the order in which the 40 syntax x polarity x proportion combinations were presented was randomly determined for each subject.

Procedure:

Subjects were briefed using text presented on the computer screen, the only verbal instructions being a reinforcement of the request that they should not commence the experiment until they were absolutely certain that they understood what they were required to do (subjects were asked to call the experimenter for assistance if they were still in difficulty after seeing the instructions twice: no subject needed to do so, however). The instructions were followed by an opportunity for familiarisation with the use of the computer keyboard in moving the arrow on the scale and in going on to the next task. The instructions given are presented in full in Appendix 1b, but that part of them concerned with the task itself is reproduced below:

> "... You will be presented with an array of 240 coloured symbols above which will be shown a statement about the shape and colour of the symbols in the array. At the very top of the screen will be shown a graduated scale ranging from "ABSOLUTELY FALSE" through "CAN'T TELL" to "ABSOLUTELY TRUE", and you will be required to indicate by moving an arrow on the scale how true you feel the statement to be with respect to the array of symbols: note that you should only use the extreme ends of the scale when you feel that the statement is ABSOLUTELY true or false ..."

Following the 40 trials, subjects were debriefed by text presented on the computer screen, as presented in Appendix 1c. No formal timing was undertaken, but all subjects took between 20 and 30 minutes from starting to read the instructions to the completion of the experiment.

2.2.2 : Results

The dependent variable is an ordinal encoding of the position to which subjects moved the arrow on the 21-point rating scale. For the purposes of analysis, the position is encoded from -10 to +10, such that -10 represents "ABSOLUTELY FALSE", 0 "CAN'T TELL" and +10 "ABSOLUTELY TRUE". The advantage of encoding the arrow position in this way is that all negative values represent ratings as "FALSE" whilst all positive ratings represent ratings as "TRUE".

Whilst all of the different analyses to be reported below were carried out on these real values from -10 to +10 (that is, taking into account both the direction and the extent of the rating), certain additional analyses were undertaken in terms of the absolute values of the ratings from 0 to 10 (that is, taking into account just the extent of the rating, but not the direction). The purpose of these absolute analyses was to provide a measure of the determinacy of responses. That the determinacy of response might otherwise be masked may be demonstrated by supposing that half of all responses to a given item were to be "ABSOLUTELY TRUE" (+10), whilst the other half were to be "ABSOLUTELY FALSE" (-10). The mean real response for that item would thus be 0 (which represents the totally indeterminate "CAN'T TELL" response): the absolute values, on the other hand, would be 10 for both types of response, so that the mean value (10)

would now reflect the absolute determinacy of the individual responses.

Since there was no significant effect of the concepts at the antecedent and consequent positions (*ie* which was the shape and which the colour), and accordingly the two groups have been collapsed, which permits an N of 24 throughout.

[a] Real Ratings (-10 ... +10):

The mean real ratings are shown in Table 2.1.

TT:TF ratio syntactic form	120:0 absolutely true	116:4 fuzzily true	60:60	4:116 fuzzily false	0:120 absolutely false	overall
IT syntax						
AA	+8.21	-1.29	-6.79	-8.83	-10.00	-3.74
AN	+9.17	-2.00	-7.29	-8.71	-9.92	-3.75
NA	+6.17	-1.17	-7.63	-7.46	-9.04	-3.83
NN	+7.08	-0.13	-8.04	-9.25	-7.83	-3.63
IT overall	+7.66	-1.15	-7.44	-8.56	-9.20	-3.74
OI syntax						
AA	+7.33	-1.71	-7.04	-9.12	-9.08	-3.92
AN	+8.00	-4.04	-8.25	-9.33	-9.08	-4.54
NA	+3.79	-3.67	-7.79	-7.79	-6.38	-4.37
NN	+4.38	-3.42	-8.54	-8.04	-6.38	-4.40
OI overall	+5.88	-3.21	-7.91	-8.57	-7.73	-4.31
IT+OI syntax						
AA	+7.77	-1.50	-6.92	-8.98	-9.54	-3.83
AN	+8.58	-3.02	-7.77	-9.02	-9.50	-4.15
NA	+4.98	-2.42	-7.71	-7.63	-7.71	-4.10
NN	+5.73	-1.77	-8.29	-8.65	-7.10	-4.02
IT+OI overall	+6.77	-2.18	-7.67	-8.57	-8.46	-4.02

Table 2.1 : Experiment 1: mean real ratings: -10 = ABSOLUTELY FALSE
0 = CAN'T TELL
+10 = ABSOLUTELY TRUE

An analysis of variance was carried out using the GENSTAT statistical package and the following significant effects were identified:

(i) main effect of TT:TF ratio ($F=104.35$; $df=4,88$; $p<0.001$) such that the 120:0 TT:TF ratio alone produced ratings which were 'true' on average, whilst 116:4 produced ratings which were less false than any of the other proportions.

A Scheffé multiple comparison test showed the difference between the mean real ratings to be highly significant ($p<0.001$) between 120:0 and 116:4, and between 116:4 and 60:60, but there were no significant differences between 60:60, 4:116 and 0:120. The multiple comparison tests are summarised in Table 2.2.

TT:TF ratio	120:0 absolutely true	116:4 fuzzily true	60:60	4:116 fuzzily false	0:120 absolutely false
mean rating	+6.77	-2.18	-7.67	-8.57	-8.46
	I_____I	I_____I	I_____I	I_____I	
	p<0.001	p<0.001	n.s.	n.s.	

Table 2.2 : Experiment 1:
Scheffé comparisons of mean real ratings x TT:TF ratio

(ii) syntax x ratio interaction ($F=3.249$; $df=4,88$; $p=0.016$) such that for the IT syntax the 120:0 TT:TF ratio was rated more 'true' (+7.66 *vs* +5.88), and the 116:4 less 'false' (-1.15 *vs* -3.21) than for the OI

syntax, whilst the 0:120 was rated more 'false' (-9.20 *vs* -7.73): in essence, then, the IT syntax made the 'true' ratings more true and the 'false' ratings more false. The interaction is illustrated graphically in Figure 2.1.

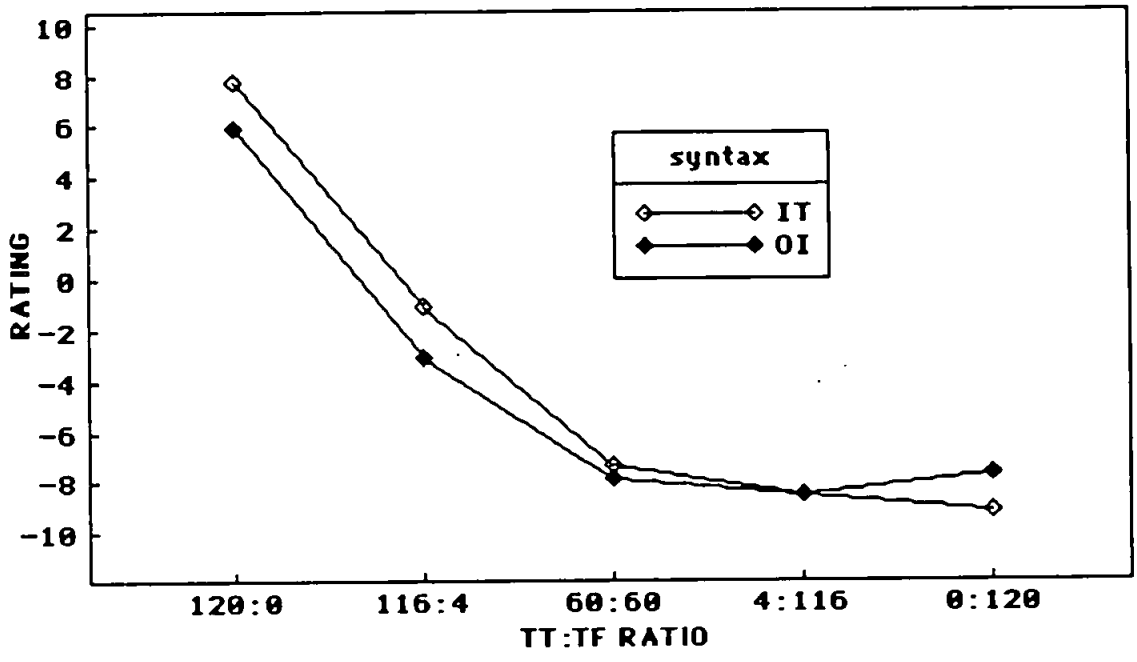


Figure 2.1 : Experiment 1: mean real ratings x TT:TF ratio x IT/OI syntax

(iii) antecedent polarity x ratio interaction ($F=2.96$; $df=12, 264$; $p<0.001$) such that for (AA and AN) statements with an affirmative antecedent, the 120:0 TT:TF proportion was rated more 'true' than it was for (NA and NN) statements with a negative antecedent (AA +7.77, AN +8.58 *vs* NA +4.98, NN +5.73), whilst 0:120 was rated more 'false' (AA -9.54, AN -9.50 *vs* NA -7.71, NN -7.10): in essence, then, affirmative antecedents (AA and AN) make the 'true' ratings more true and the 'false' ratings more false. This interaction is illustrated graphically in Figure 2.2:

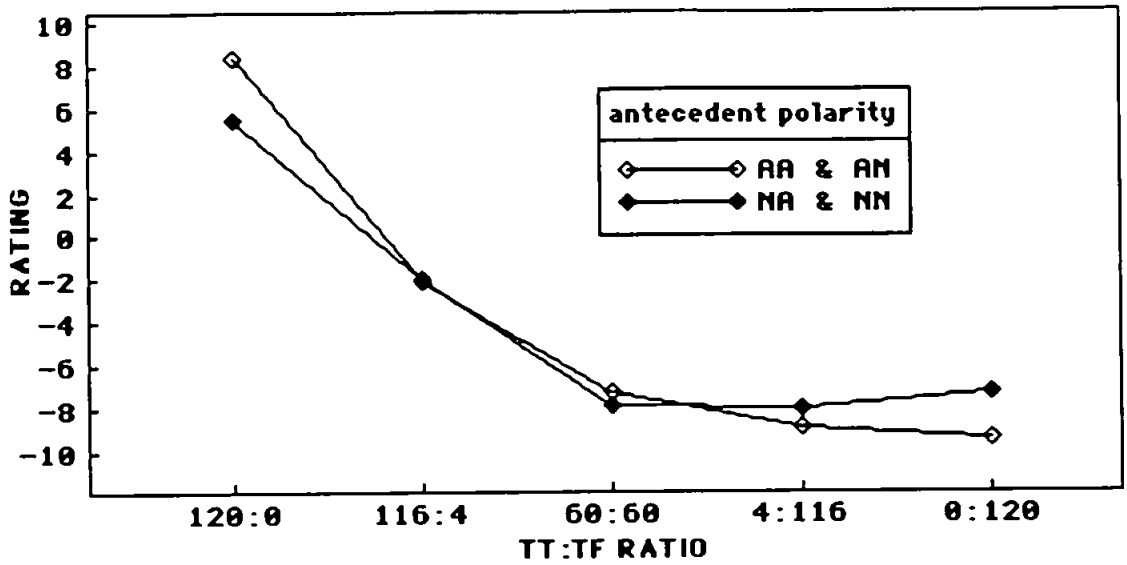


Figure 2.2 : Experiment 1: mean real ratings x TT:TF ratio x antecedent polarity

[b] Absolute Ratings (0 ... 10):

The mean absolute ratings are shown in Table 2.3.

TT:TF ratio syntactic form	120:0 absolutely true	116:4 fuzzily true	60:60	4:116 fuzzily false	0:120 absolutely false	overall
IT syntax						
AA	9.79	8.46	8.04	9.25	10.00	9.11
AN	9.50	8.17	7.54	9.38	9.92	8.90
NA	9.50	8.50	7.96	8.29	9.88	8.83
NN	9.58	7.96	8.04	9.25	9.50	8.87
IT overall	9.59	8.27	7.90	9.04	9.82	8.93
OI syntax						
AA	9.25	8.13	7.63	9.13	9.92	8.81
AN	8.83	8.79	8.25	9.33	9.92	9.03
NA	8.88	7.67	8.13	8.63	8.88	8.43
NN	8.13	7.83	8.54	8.54	8.88	8.38
OI overall	8.77	8.10	8.14	8.91	9.40	8.66
IT+OI syntax						
AA	9.52	8.29	7.83	9.19	9.96	8.96
AN	9.17	8.48	7.90	9.35	9.92	8.96
NA	9.19	8.08	8.04	8.46	9.38	8.63
NN	8.85	7.90	8.29	8.90	9.19	8.63
IT+OI overall	9.18	8.19	8.02	8.97	9.61	8.79

Table 2.3 : Experiment 1: mean absolute ratings:
10 = ABSOLUTELY TRUE/FALSE
0 = CAN'T TELL

An analysis of variance identified just one significant effect: a main effect of TT:TF ratio ($F=9.468$; $df=4,88$; $p<0.001$) such that the ratings were consistently more extreme the more disparate the TT and TF counts. A Scheffé multiple comparison test showed the difference between the mean ratings to be (just) significant ($p<0.05$) between 120:0 and 116:4, significant ($p<0.01$) between 120:0 and 60:60 and highly significant ($p<0.001$) between 60:60 and 0:120. These findings are summarised in Table 2.4.

TT:TF ratio	120:0 absolutely true	116:4 fuzzily true	60:60	4:116 fuzzily false	0:120 absolutely false
mean rating	9.18	8.19	8.02	8.97	9.61
		p<0.05	n.s.	n.s.	n.s
			p<0.01	p<0.001	

Table 2.4 : Experiment 1:
Scheffé comparisons of mean absolute ratings x TT:TF ratio

[c] Individual Differences:

Subjects differed in their tendency to rate the statements at intermediate points on the scale in that four of the 24 subjects rated the statements as absolutely true (+10) or absolutely false (-10) throughout (and thus might be seen as working with a strictly 2-valued logic), and a further two subjects only diverged from this pattern by the inclusion of 'pure' 'CAN'T TELL' (0)

ratings (and thus might be seen as working with a strictly 3-valued logic). Six of the 24 subjects (25%) were thus resistant to the notion of intermediate degrees of truth, although this 'non-fuzzy' pattern of responding was slightly more prevalent on IT than on OI syntax, two further subjects being 'non-fuzzy' on IT syntax only.

[d] Noticeability of the Counterexamples:

As has been seen, the ratings on the 'fuzzily true' 116:4 TT:TF arrays are significantly different from those on the 'absolutely true' 120:0 arrays and the variously 'false' 60:60, 4:116 and 0:120 arrays. This 'half way house' position might, of course, have been a product merely of the TF falsifiers being noticed on some trials but not on others. To show a genuine difference in the responses evoked by these 'fuzzily true' displays, therefore, it must be demonstrated that there was a qualitative difference in response on at least the majority of subjects over at least the majority of the eight syntax x polarity combinations (IT/OI x AA/AN/NA/NN).

If the four falsifying symbols were noticed by a subject, then this would be indicated by a lower ('more false'/'less true') rating on the 116:4 array for a given syntax x polarity combination than on the corresponding 120:0 array. The extent to which subjects demonstrated this pattern of response is shown at Table 2.5, from which it might be seen that the responses of virtually all subjects were influenced by the presence of the four TF cases in at least six of the eight (IT/OI x AA/AN/NA/NN) trials concerned. This table distinguishes the six 'non-fuzzy' subjects identified above (ie those six who responded -10, 0 or +10 over all the IT and OI statements: note that a 'non-fuzzy' subject is only counted as having noticed the falsifiers where 120:0 was rated +10 and 116:4 0 rated or -10, or 120:0 was rated 0 and 116:4 rated -

10).

It is thus clear that by far the majority of subjects were influenced by the presence of the four TF falsifiers by far the majority of the time, and thus that the significant difference between the 'fuzzily true' 116:4 TT:TF displays and all the other displays cannot be attributed to subjects simply noticing the falsifiers on some occasions but not on others.

N of trials	ALL subjects (N=24)	FUZZY subjects (N=18)	NON-FUZZY subjects (N=6)
8 of 8	5 /24	4 /18	1 /6
7 of 8	9 /24	6 /18	3 /6
6 of 8	7 /24	7 /18	-
≥6 of 8	21 /24	17 /18	4 /6

Table 2.5 : Experiment 1: number of responses 'reliably' influenced by the presence of the 4 x TF cases in the 8 x 116:4 TT:TF trials (*ie* on at least 6 of 8 trials)

[e] Matching Status:

In order to faciltate comparison with the findings of experiments to be reported later, it may be useful at this point to identify any possible effects of the matching status of items present in the array in terms of Evans' (1972c) 'matching bias', which was described in the last chapter. The predictions of the matching bias account are necessarily rather less specific for evaluation tasks than they are for construction tasks - or, indeed for any task where

observable choices are involved, such as Wason's (1966) selection task. Nevertheless, to the extent that responses might be subject to matching bias in an evaluation task then it is, if nothing else, clear that those responses should be more determinate. In Experiment 1, matching bias would cause subjects to attend particularly strongly to the item which had both the colour and shape named in the statement (the double match), and to attend least of all to items which had neither attribute (the double mismatch).

The predictions which might thus be made for AN and NA rules involve the 4 TF items in the fuzzily true 116:4 TT:TF arrays. On AN rules, where the TF case is the double match, then since it should be at its most likely to be attended to, the 'false' response to the 116:4 TT:TF array should be at its most determinate. By contrast, on NA rules, when TF is the double mismatch, it should be at its least likely to be attended to, such that the (erroneous) 'true' response to that array should be at its most determinate.

For the fuzzily true 116:4 TT:TF array, then, matching bias would appear predict that the determinacy of responses should be at its greatest for AN and NA rules. These predictions about determinacy (in terms of the absolute ratings) lead to corresponding predictions about the real ratings: since the determinacy on AN rules is with respect to a 'false' response, whilst that for NA rules relates to a 'true' response, we should expect the lowest (most 'false' overall) real rating for this array with AN rules, and the highest (*ie* least negative) real rating with NA rules. Examination of the absolute data in Table 2.2 appears to lends no support whatever to these predictions. The predictions concerning the real ratings are partially supported by the data in Table 2.1, however, in so far as AN rules do produce markedly more 'false' ratings overall than the other three forms: the prediction for NA rules is not,

however, supported.

The corresponding predictions for AA and NN rules involve the 4 TT items in the fuzzily false 4:116 TT:TF arrays and depend upon the extent to which subjects are less ready to give a 'false' response to that array than to the 0:120 TT:TF array. On AA rules, where TT is the double match, then since it should be at its most likely to be attended to, the 'false' response to the 4:116 TT:TF array should be at its least determinate, since the fuzziness of the array will thus be made more apparent. Finally, on NN rules, where TT is the double mismatch, then since it should be at its least likely to be attended to, the 'false' response to the 4:116 TT:TF array should be at its most determinate.

For the fuzzily false 4:116 TT:TF array, then, matching bias would appear predict that the determinacy of responses should be at its greatest for NN rules, yet at its lowest for AA rules. These predictions about determinacy (in terms of the absolute ratings) lead to corresponding predictions about the real ratings: we should expect the lowest (most 'false' overall) real rating for this array with NN rules, yet the highest (*ie* least negative) real rating with AA rules. Examination of the data in Tables 2.1 and 2.2 does, however, appear to lends no support whatever to any of these predictions.

2.2.3 : Discussion

This experiment was designed to investigate the notion of degrees of truth. It was hypothesised that subjects would rate fuzzily true statements as less true than absolutely true statements, but that fuzzily true statements would still be rated as true rather than false. Similarly, it was hypothesised that subjects would rate all false and fuzzily false statements as false, but that they would be rated as more false the less fuzzy their falsity.

Looking firstly at the absolutely and fuzzily true statements, it was only the absolutely true statements (those true of arrays with no logically falsifying TF cases whatever) that were rated as true overall - whilst the fuzzily true statements (those with 116 TT cases and just 4 TF cases) were indeed rated as very much less false than the more clearly false statements, they were still rated as false overall. Although the prediction that these fuzzily true statements would still be rated as true was not supported, the fact that it was rated in a very significantly different way from any of the other variously false statements might be taken as a demonstration that a small proportion of counterexamples will be tolerated to the extent, at least, that a contingent universal statement about such an array is not seen as having been rendered absolutely false by their presence. Such an assertion would, however, be unfounded if it could be shown that the intermediate rating of the fuzzily true statements arose out of subjects failing to notice the TF falsifiers on just some of the trials. It is clear from Table 2.5, however, that this is not the case, a clear majority of subjects rating that statement as less true than the absolutely true statement on at least six of the eight trials concerned.

The hypotheses relating to the absolutely false and fuzzily false statements fared worse still in that there were no significant differences whatever in the extent to which these were rated as false. There does appear, then, to be an assymetry between a fuzzily true situation and a fuzzily false situation. This is not, perhaps, especially surprising. As Kyburg (1983) argues, universal generalisations in our normal life do tend to permit occasional counterxamples - and are still useful to as generalisations despite the fact that they are sometimes incorrect. It is, however, difficult to imagine a situation in which a universal generalisation would be of use to us if it were to be incorrect most of the time - and, indeed, a putative generalisation will not have to be observed to be incorrect very often before it is modified or discarded. Universal generalisations in our everyday experience, then, might sometimes be fuzzily true, but they will tend not to be retained if less true than that. If this argument is correct, then we should not be surprised that statements of all three degrees of falsity were rejected equally strongly.

Before moving on from this consideration of the main effects of fuzziness, it may be timely to admit to one possible *caveat* upon the interpretation of these results. In their briefing, subjects were told that they should indicate "how true" they felt the statement to be on a "graduated scale". To give such an instruction may well imply to the subjects that there do exist varying degrees of truth (at least for the purposes of this experiment) and that there will be something about the arrays to be presented that will mean that the experimenter is hoping for a response that falls somewhere on the scale. This push towards an intermediate level of response is further invited by the additional instruction that they should only use the extreme ends of the scale when they felt that the statement was absolutely true or false. To the extent that this might have been a demand characteristic of the experiment, however, it is clear that it would have been far from universally effective

since, as we have seen, six of the 24 subjects resisted the intermediate positions throughout, responding only 'absolutely true', 'absolutely false' or 'can't tell' throughout.

In addition to considering the main effect of fuzziness, it is necessary to consider the interaction of fuzziness with both syntactic form and polarity. The syntactic effect was characterised as the IT form resulting in the 'true' ratings being more true and the 'false' ratings more false. In so far as there was no significant effect of syntax upon the absolute ratings, it is clear that this increase in the overall real rating must come not from greater determinacy but, rather, from greater consistency between subjects (*ie* a greater proportion of subjects giving a similar response). It is clear from Figure 2.1 that the relevant differences between IT and OI syntax are to be found in just three types of array: the absolutely true 120:0, the fuzzily true 116:4 and the absolutely false 0:120. In the case of the absolutely true and absolutely false arrays, the question of fuzziness simply does not arise. There is thus no reason whatever to propose that these two differences reflect any more than the well-established fact that (for whatever reason) people simply do make more errors in reasoning with OI syntax than with IT syntax. In the case of these two non-fuzzy arrays, an increase in reasoning errors on the OI form might be expected to manifest itself in an increased failure to endorse demonstrably true statements and to reject demonstrably false statements - and this is precisely what the data reveal.

One such potential error is that of conversion. Conversion is an error arising out of the illicit interchanging of the propositions in a conditional statement, such that 'if P then Q' becomes (or is treated as if it had become) 'if Q then P', whilst 'X only if Y' becomes 'Y only if X'. It should be clear that

by converting the conditional, the (original) TF case is no longer that in which the antecedent is true and the consequent false, since the antecedent and consequent have now changed places. Rather, TF is now the case in which the consequent is true and the antecedent false - in other words, it has become the FT case (and, correspondingly, the original FT case has, by conversion, become the TF case). The presence in an array of what was originally the TF case will thus no longer be taken to falsify the array, but rather to be logically *irrelevant*. Similarly, the presence of what was originally the FT case is no longer irrelevant, but will now be taken to falsify the array.

Evans (1975) reports a truth table evaluation task in which IT and OI syntax were compared. For IT syntax, the TF case was rated as false 81% of the time, whilst the FT case was rated as false only 29% of the time. For OI syntax, on the other hand, the two cases were rated as false roughly as often, TF 58% of the time and FT 57%. Given that the overall level of rating as false is almost identical for each syntax (a mean of 55% for IT and 57.5% for OI), it is clear that to the extent that FT is rated more often as false with OI syntax, TF is correspondingly rated *less* often as false, which is consistent with some of the subjects (some of the time) having converted the conditional in the OI syntax.

An explanation in terms of task difficulty with OI syntax (*ie* an increased tendency to make reasoning errors) might also be offered for the fuzzily true 116:4 arrays, since although the arrays are now fuzzy, subjects' responses to them are still dependent upon their ability to reason with conditional statements in one syntactic form or the other. For example, to the extent that subjects make errors on the OI form through the illicit

conversion of the premises (from 'P only if Q' to 'if Q then P'), then in the fuzzily true array, the logical falsifier will change from the (fuzzily present) TF case to the (massively present) FT case, and thus the overall real rating might be expected to be lower for the OI form than for the IT form.

If the IT/OI differences are to be dismissed simply in terms of increased task difficulty on the OI form, then we must consider why no such difference is to be found on the intermediate 60:60 array nor the fuzzily false 4:116 array. In the intermediate 60:60 array, all four truth cases are present in equal number: thus, whatever a subject might take the falsifier to be, it will be there in considerable force, and thus no IT/OI difference would be expected. On the fuzzily false 4:116 array, we are concerned not with falsification but, rather, with verification. Since the critical item is now the TT case, which is neither unchanged by any conversion of the premises, nor undergoes any change in truth status dependent upon the truth table with which a subject might reason, there is, here again, no reason to expect any IT/OI difference.

An explanation in terms of task difficulty seems equally attractive for the interaction between fuzziness and the polarity of the antecedent, which was characterised as affirmative antecedents (AA and AN) resulting in the 'true' ratings being more true and the 'false' ratings more false. It is clear from Figure 2.2 that the relevant differences between affirmative and negative antecedents syntax are to be found in just the absolutely true 120:0 and the absolutely false 0:120 arrays. If it might be demonstrated that the presence of negatives in the NA and NN forms is a source of additional difficulty, then the arguments above concerning the IT/OI difference on these two non-fuzzy arrays might reasonably be applied here *mutatis mutandis*. It turns out that the source of such differential difficulty is particularly easy to identify.

Where the antecedent is affirmed, then it refers to just one shape (or colour): since the array consists of all combinations of three shapes and three colours, however, a negated antecedent now refers to two separate shapes (or colours), such that a search for items satisfying the antecedent now involves twice as many discrete types of symbol and will thus clearly be more difficult, time-consuming and prone to error.

2.3 : Experiment 2

Experiment 2 was a construction task in which subjects were required to fill a computer-presented grid with coloured symbols so as to show a given statement, should it be made about the grid, to be either true or false. Conceived as a complementary experiment to Experiment 1, it was hypothesised that subjects would incorporate into the arrays a certain proportion of counterexamples so as to make the statements not absolutely true or false, but rather 'fuzzily' true or false.

Whilst Experiments 1 and 2 involve two quite distinct types of operation, they were intended to be complementary: in so far it might be argued that these operations involve the same underlying mechanisms of inference and understanding, a convergence of the results of the two experiments would lend weight to the validity of each.

Experiment 2 was highly prospective in nature, it being hypothesised merely that subjects would choose to include in their arrays a certain degree of fuzziness - that is, a certain (small) number of counterexamples. The specific hypotheses were thus as follows:

(1) in the verification condition (where statements were to be shown to be true) the arrays would include a (small) number of logically falsifying TF cases

(2) in the falsification condition (where statements were to be shown to be false) the arrays would include a (small) number of logically verifying TT cases

2.3.1 : Design and Methodology

Subjects:

24 first-year BSc (Hons) Psychology undergraduates at Plymouth Polytechnic (17 female, 7 male), mean age 20.6 years.

Materials:

The experiment was conducted on a microcomputer as specified for Experiment 1. Subjects were presented with a blank 6 x 6 grid of 36 squares above which was shown a statement about the colour and shape of symbols. Subjects were required to press keys on the computer keyboard to place symbols into the blank grid so as to show the statement to be either true or false, as indicated to the right of the grid. As the mechanics of placing and replacing symbols were necessarily somewhat elaborate, detailed instructions were displayed at all times. A screen dump of a typical screen is presented in Appendix 2a.

The statements were systematically varied over the eight syntactic forms presented in the description of Experiment 1 *(q.v.)*, that is, ITAA, ITAN, ITNA, ITNN, OIAA, OIAN, OINA and OINN. Each syntax x polarity combination was presented once so as to be shown to be false, and once so as to be shown to be true, and accordingly there were 16 trials in all, the order of presentation of which was randomly determined for each subject.

The same nine symbols were used as those described for Experiment 1, that is, all possible combinations of three colours (yellow, blue and red) and three shapes (circular, square and triangular). Half the subjects were given a

shape at the antecedent of each rule and a colour at the consequent throughout: the remaining subjects were given the reverse. The actual colour/shape in each proposition of the statement was randomly determined for each trial.

Procedure:

Subjects were briefed using text presented on the computer screen, the only verbal instructions being a reinforcement of the request that they should not commence the experiment until they were absolutely certain that they understood what they were required to do (subjects were asked to call the experimenter for assistance if they were still in difficulty after seeing the instructions twice: no subject needed to do so, however). The instructions were followed by an opportunity for familiarisation with the use of the computer keys in moving the cursor around the grid, selecting symbols to be placed into the grid at the cursor and going on to the next task. The instructions given are presented in full in Appendix 2b, but that part of them concerned with the task itself is reproduced below:

> " ... You will be presented with an empty 6 x 6 grid of 36 squares. Above the grid will be shown a statement concerning the shape and colour of symbols which are to be put into the grid, and to the right of the grid will be an instruction to show the statement to be TRUE or FALSE. What you have to do is to fill the grid with symbols so that the statement given at the top of the screen is either TRUE or FALSE with respect to the appearance of the grid...
> ...The fact that there are nine symbols available does not mean that they all must necessarily be included in the grid. You may use any of the symbols as often or as little as you wish ..."

Following the 16 trials, subjects were debriefed by text presented on the

computer screen, as presented in Appendix 2c. No formal timing was undertaken, but all subjects took between 30 and 50 minutes from starting to read the instructions to the completion of the experiment.

2.3.2 : Results

The dependent variables are the number of symbols selected relating to each truth case, and thus four separate analyses have been carried out (for TT, TF, FT and FF) within each condition (that is, whether the statement was to be shown to be true or false), making eight analyses in all. No overall metric of behaviour has been devised as it is considered that any such metric would be totally arbitrary and thus of at least dubious validity. There was no significant effect of the concepts at the antecedent and consequent positions (that is, which was the shape and which the colour), with the minor exception of its participation in a (just) significant 3-way interaction in one of the eight analyses, and accordingly the two groups have been collapsed, which permits an N of 24 throughout.

[a] Selections so as to show the Statements to be TRUE:

The mean numbers of symbols selected in the verification condition are shown at Table 2.6. The purpose of showing the means of the PQ and ~P~Q diagonals will be made clear in a later section of this chapter, in which the question of matching status is addressed.

truth case / polarity:	AA	AN	NA	NN	overall
IT syntax					
TT	20.37	19.42	27.67	22.46	22.48
TF	1.46	1.50	-	3.50	1.61
FT	3.08	5.17	1.54	3.25	3.26
FF	11.08	9.92	6.79	6.79	8.65
IT diagonals	9.68 ~P~Q				PQ 7.55
OI syntax					
TT	19.54	21.46	16.13	18.75	18.97
TF	1.46	1.71	7.17	2.92	3.31
FT	5.08	7.75	4.71	6.58	6.03
FF	9.92	5.08	8.00	7.75	7.69
OI diagonals	10.90 ~P~Q				PQ 8.43
IT + OI combined					
TT	19.96	20.44	21.90	20.60	20.72
TF	1.46	1.60	3.58	3.21	2.46
FT	4.08	6.46	3.13	4.92	4.65
FF	10.50	7.50	7.40	7.27	8.17
IT + OI diagonals	10.29 ~P~Q				PQ 7.99

Table 2.6 : Experiment 2: mean number of symbols selected so as to show the statements to be TRUE

An analysis of variance was carried out upon the number of symbols in each truth case and the following significant effects were identified:

[1] TT case:

(i) main effect of syntax ($F=6.202$; $df=1,22$; $p=0.021$) such that more TT symbols were selected on the IT statements overall (mean 22.48) than on the OI statements overall (mean 18.97)

(ii) syntax x polarity interaction ($F=3.751$; $df=3,66$; $p=0.015$) such that for IT syntax less TT selections were made on statements with affirmative antecedents than on those with negative antecedents (AA+AN 39.79 *vs* NA+NN 50.13), whilst for OI syntax, the reverse was the case (AA+AN 41.00 *vs* NA+NN 34.88)

[2] TF case:

syntax x polarity interaction ($F=3.989$; $df=3,66$; $p=0.011$) such that

(a) TF selections for IT syntax were zero on the NA polarity whilst for OI syntax the NA polarity produced the highest TF count of all (7.17)

(b) if selections are pooled on the basis of affirmative and negative antecedents, then for IT syntax there was little difference (AA+AN 2.96 *vs* NA+NN 3.50) whilst for OI syntax, there were considerably more selections for statements with negative antecedents (AA+AN 3.17 *vs* NA+NN 10.09)

(c) if selections are pooled on the basis of affirmative and negative consequents, then for IT syntax there were more selections for negative than for affirmative consequents (AA+NA 1.46 *vs* AN+NN 5.00), whilst for OI syntax, the reverse was the case (AA+NA 8.63 *vs* AN+NN 4.63)

[3] FT case:

(i) (weak) main effect of syntax ($F=4.451$; $df=1,22$; $p=0.046$) such that less FT symbols were selected on the IT statements overall (mean 3.26) than on the OI statements overall (mean 6.03)

(ii) (weak) concept order x syntax x polarity interaction ($F=2.778$; $df=3,66$; $p=0.048$) - exactly what is the nature of this 3-way interaction is, however, far from being immediately apparent.

[4] FF case: no significant effects

[b] Selections so as to show the Statements to be FALSE:

The mean numbers of symbols selected in the falsification condition are shown at Table 2.7.

truth case / polarity:	AA	AN	NA	NN	overall
IT syntax					
TT	2.71	3.75	2.04	5.83	3.58
TF	25.04	20.33	19.17	24.00	22.14
FT	3.79	5.13	7.04	2.67	4.66
FF	4.46	6.79	7.75	3.50	5.63
IT diagonals 8.65	~P~Q				PQ 8.40
OI syntax					
TT	5.21	2.62	5.21	4.29	4.33
TF	18.96	14.33	14.50	5.88	13.42
FT	5.37	9.92	10.00	7.12	8.10
FF	6.46	9.13	6.29	18.71	10.15
OI diagonals 8.79	~P~Q				PQ 12.06
IT + OI combined					
TT	3.96	3.19	3.63	5.06	3.96
TF	22.00	17.33	16.83	14.94	17.78
FT	4.58	7.52	8.52	4.90	6.38
FF	5.46	7.96	7.02	11.10	7.89
IT + OI diagonals 8.72	~P~Q				PQ 10.23

Table 2.7 : Experiment 2: mean number of symbols selected so as to show the statements to be FALSE

An analysis of variance was carried out upon the number of symbols in each truth case and the following significant effects were identified:

[1] TT case: no significant effects

[2] TF case:

(i) main effect of syntax ($F=4.085$; $df=1,22$; $p<0.001$) such that more TF selections were made for IT syntax (mean 22.14) than for OI syntax (mean 13.42)

(ii) syntax x polarity interaction ($F=4.085$; $df=3,66$; $p=0.010$) such that for IT syntax there was little difference in the number of TF selections between those for statements with affirmative antecedents and those with negative antecedents (AA+AN 45.37 *vs* NA+NN 43.17), whilst for OI syntax, considerably more TF selections were made where the antecedent was affirmative (AA+AN 33.29 *vs* NA+NN 20.38).

[3] FT case:

main effect of syntax ($F=6.327$; $df=1,22$; $p=0.020$) such that less FT selections were made for IT syntax (mean 4.66) than for OI syntax (mean 8.10)

[4] FF case:

(i) main effect of syntax ($F=23.803$; $df=1,22$; $p<0.001$) such that less FF selections were made for IT syntax (mean 5.63) than

for OI syntax (mean 10.15)

(ii) syntax x polarity interaction ($F=7.345$; $df=3,66$; $p<0.001$) such that

(a) if selections are pooled on the basis of affirmative and negative antecedents, then for IT syntax there was no difference whatever (AA+AN 11.25 *vs* NA+NN 11.25) whilst for OI syntax, there were considerably more FF selections for statements with negative antecedents (AA+AN 15.59 *vs* NA+NN 25.00)

(b) if selections are pooled on the basis of affirmative and negative consequents, then for IT syntax there was little difference (AA+NA 12.21 *vs* AN+NN 10.29), whilst for OI syntax, there were considerably more FF selections for statements with negative consequents (AA+NA 12.75 *vs*AN+NN 27.84).

[c] Inclusion of Counterexamples:

Table 2.8 presents an analysis of the number of 'fuzzy' arrays constructed in each condition and for each syntactic form. For IT syntax, only 7 subjects (29%) ever included a falsifying TF item in the verification condition, and even then only in one of the four relevant arrays. In the falsifying condition, 12 subjects (50%) never included a verifying TT item, whilst of the remaining 12, 7 only did so on just one of the four arrays.

For OI syntax, Table 2.8 suggests that rather more fuzzy arrays were constructed, 14 subjects (58%) including TF items in the verification condition and 15 subjects (63%) including TT items in the falsifying

condition. Here again, though, by far the majority of 'fuzzy' subjects showed themselves to be so on just one of the four relevant arrays in each case.

	VERIFICATION condition N of 24 subjects with ≥1 TF case		FALSIFICATION condition N of 24 subjects with ≥1 TT case	
N of fuzzy arrays	IT	OI	IT	OI
1	7	8	7	8
2	-	3	2	4
3	-	3	1	1
4	-	-	2	2
fuzzy subjects	7	14	12	15
non-fuzzy subjects	17	10	12	9

Table 2.8 : Experiment 2: fuzzy arrays constructed & 'fuzzy subjects' x condition x syntactic form

[d] Matching Status:

In order to consider the extent to which the matching status of items might have influenced their selection, it is helpful to re-present the data shown in Tables 2.6 and 2.7 above so as to make explicit the matching status of each truth case for each syntactic form. Given a shape (colour) at the antecedent and a colour (shape) at the consequent, it is clear that a given symbol can fall into one of four categories in terms of the colour and shape it possesses: it can match on the attribute at the antecedent only, the attribute at the consequent only, neither or both.

	polarity:	AA	AN	NA	NN	overall
IT syntax						
double match	pq	20.37	1.50	1.54	6.79	7.55
antecedent only	p~q	1.46	19.42	6.79	3.25	7.73
consequent only	~pq	3.08	9.92	27.67	3.50	11.04
double mismatch	~p~q	11.08	5.17	-	22.46	9.68
OI syntax						
double match	pq	19.54	1.71	4.71	7.75	8.43
antecedent only	p~q	1.46	21.46	8.00	6.58	9.38
consequent only	~pq	5.08	5.08	16.13	2.92	7.30
double mismatch	~p~q	18.75	7.17	7.75	9.92	10.90
IT + OI combined						
double match	pq	19.96	1.60	3.13	7.27	7.99
antecedent only	p~q	1.46	20.44	7.40	4.92	8.55
consequent only	~pq	4.08	7.50	21.90	3.21	9.17
double mismatch	~p~q	20.60	3.58	6.46	10.50	10.29

Table 2.9 : Experiment 2: mean number of matching and mismatching items selected so as to show the statements to be TRUE

explanatory note: for the statement *'If the symbols are circular then they are yellow'*, the double match is a yellow circle, the antecedent-only match any non-yellow circle, the consequent-only match any yellow non-circle and the double match any non-yellow non-circle

Table 2.9 shows the number of symbols selected corresponding to each of these categories of matching status across each syntactic form in the verification condition (note that the double match is the PQ diagonal shown in Tables 2.6 and 2.7, and the double mismatch the ~P~Q diagonal). It is clear from this table that in the verification condition - where the statements are to be shown to be true - the symbols constituting the double match are selected less often overall than those with any other matching status. To the

extent that matching status appears to have any effect at all upon selections, then, it is in a small overall preference for items which constitute the double mismatch - that is, items neither of the attributes of which is mentioned in the statement. There is accordingly no evidence whatever of matching bias where the statements are to be verified.

	polarity:	AA	AN	NA	NN	overall
IT syntax						
double match	pq	2.71	20.33	7.04	3.50	8.40
antecedent only	p~q	25.04	3.75	7.75	2.67	9.80
consequent only	~pq	3.79	6.79	2.04	24.00	9.15
double mismatch	~p~q	4.46	5.13	19.17	5.83	8.65
OI syntax						
double match	pq	5.21	14.33	10.00	18.71	12.06
antecedent only	p~q	18.96	2.62	6.29	7.12	8.75
consequent only	~pq	5.37	9.13	5.21	5.88	6.40
double mismatch	~p~q	4.29	14.50	9.92	6.46	8.79
IT + OI combined						
double match	pq	3.96	17.33	8.52	11.10	10.23
antecedent only	p~q	22.00	3.19	7.02	4.90	9.28
consequent only	~pq	4.58	7.96	3.63	14.94	7.78
double mismatch	~p~q	5.06	16.83	7.52	5.46	8.72

Table 2.10 : Experiment 2: mean number of matching and mismatching items selected so as to show the statements to be FALSE

explanatory note: for the statement *'If the symbols are circular then they are yellow'*, the double match is a yellow circle, the antecedent-only match any non-yellow circle, the consequent-only match any yellow non-circle and the double match any non-yellow non-circle

Table 2.10 shows a similar analysis of selections in the falsification condition. It is clear from this table that in the falsification condition also,

matching status appears to have very little effect upon selections. With IT syntax, mean overall selections of the double match and the double mismatch are almost identical. With OI syntax, however, there is a small tendency for the double match to be preferred over items with any other matching status.

(1) fewer TF selections with negative antecedents (NA+NN vs AA+AN)	
IT : yes (just)	43.17 < 45.37
OI : yes	20.38 < 35.29
(2) more TF selections with negative consequents (AN+NN vs AA+NA)	
IT : yes (just)	44.33 < 44.21
OI : yes	20.21 < 33.46
(3) most TF selections on AN	
IT : no	NA 19.17 < AN 20.33 < NN 24.00 < AA 25.04
OI : no	NN 5.88 < AN 14.33 < NA 14.50 < AA 18.96
(4) most FT on NA	
IT : yes	NN 2.67 < AA 3.79 < AN 5.31 < NA 7.04
OI : yes (just)	AA 5.37 < NN 7.12 < AN 9.92 < NA 10.00
(5) most FF on NN	
IT : no	NN 3.50 < AA 4.46 < AN 6.79 < NA 7.75
OI : yes	NA 6.29 < AA 6.46 < AN 9.13 < NN 18.71

Table 2.11 : Experiment 2: conformity of selection patterns with those predicted by matching bias *(per Evans, 1972c)* [falsification condition only]

Since matching bias was originally identified on the falsification version of a construction task (Evans, 1972c), it is pertinent to look more closely at performance in this condition. In particular, Evans' (1972c) matching bias account makes five specific predictions about selections on such tasks. The

predictions are set out in Table 2.11, together with the relevant summary data from the present experiment.

In the falsification condition, then, IT syntax produced (somewhat weak) support for three of the five predictions, whilst strongly rejecting the remaining two. OI syntax, on the other hand, strongly supported three of the predictions.

[e] Ranges of Discrete Items Available

As has already been explained, the use of all three combinations of three shapes and three colours means that there are unequal numbers of discrete items available across the four types of matching status - there is always exactly one item constituting the double match, two each the antecedent-only and consequent-only matches, and four the double mismatches.

Whilst this observation will not be discussed further for Experiment 2, it will be useful to the overall discussion of Experiments 1 to 4 to present in Tables 2.12 and 2.13 additional analyses of the data from the present experiment in terms of the range of discrete symbols available. Since there are two types of 'half' match - antecedent-only and consequent-only - the figures shown for the 'half' match in Tables 2.12 and 2.13 are the mean of those two types.

truth case	matching status: N available	double match: 1	'half' match: 2	double mismatch: 4
IT syntax				
TT		20.37	23.54	22.46
TF		1.50	2.48	-
FT		1.54	3.16	5.17
FF		6.79	8.35	11.08
IT mean		7.55	9.38	9.68
OI syntax				
TT		19.54	18.79	18.75
TF		1.71	2.19	7.17
FT		4.71	5.83	7.75
FF		7.75	6.54	9.92
OI mean		8.43	8.34	10.90
IT + OI combined				
TT		19.96	21.17	20.60
TF		1.60	2.33	3.58
FT		3.13	4.50	6.46
FF		7.27	7.45	10.50
IT + OI mean		7.99	8.86	10.29

Table 2.12 : Experiment 2: mean number of symbols selected so as to show the statements to be TRUE
x N of discrete symbols available

truth case	matching status: N available	double match: 1	'half' match: 2	double mismatch: 4
IT syntax				
TT		2.71	2.89	5.83
TF		20.33	24.52	19.17
FT		7.04	3.23	5.13
FF		3.50	7.27	4.46
IT mean		8.39	9.48	8.65
OI syntax				
TT		5.21	3.91	4.29
TF		14.33	12.42	14.50
FT		10.00	6.24	9.92
FF		18.71	7.71	6.46
OI mean		12.06	7.57	8.79
IT + OI combined				
TT		3.96	3.41	5.06
TF		17.33	18.47	16.83
FT		8.52	4.74	7.52
FF		11.10	7.49	5.46
IT + OI mean		10.23	8.53	8.72

Table 2.13 : Experiment 2: mean number of symbols selected so as to show the statements to be FALSE x N of discrete symbols available

It is clear from Table 2.12 that in the verification condition, as the range of discrete items available increases, so too does the frequency of selection, such that the more different items there are available for a given matching

status, the more will be selected overall. The effect is, however, relatively weak, in that whilst the difference in availability doubles between each successive type (from 1 to 2, and from 2 to 4), the differences in mean selection rates is of a far smaller order - and, moroever, is highly inconsistent across syntactic forms and truth case. It appears from Table 2.13 that in the falsification condition, however, the range of discrete symbols available appears to have no effect whatever upon the corresponding frequencies of selection.

2.3.3 : Discussion

As mentioned earlier, this experiment was largely prospective in nature. In line with the notion of fuzziness which underlay the conception of both this experiment and Experiment 1, it was hypothesised that subjects would construct their arrays so as to be not absolutely true or false, but rather fuzzily true (by including a certain proportion of logically falsifying TF cases) or fuzzily false (by including a certain proportion of logically verifying TT cases). It will be recalled, however, that the nature of this fuzziness is assymmetrical in terms of formal logic, in that something fuzzily true is logically false, whilst something fuzzily false is logically just as false as something absolutely false.

Looking first at the verification condition, where the statement was to be shown to be true with respect to the array, it does on the face of it seem as if subjects have constructed fuzzily true arrays, in that arrays contained a mean of 2.46 TF items (around 7% of the 36 items in the array). Any analysis of the use of counterexamples will, however, inevitably be clouded by the presence of errors, and the fact that a considerable number of subjects who included a counterexample at all only did so in one or two arrays would seem to suggest strongly that these are the results of errors rather than of the intentional inclusion of counterexamples so as to make the arrays 'fuzzy'.

As seen earlier, one well-established source of error is that of illicit conversion of the propositions in the conditional rule. As was seen in the discussion of Experiment 1, the effect of conversion is that what is, in reality, the TF case will now be perceived as the FT case, whilst the original FT case

will be perceived as the TF case. The consequence of conversion in a construction task, therefore, is that to include what was originally the TF case is no longer to include a logical falsifier, but rather a logically irrelevant case: similarly, to include what was originally the FT case is now to include a logical falsifier.

Table 2.6 shows that in the verification condition, TF selections were more than twice as frequent with OI syntax than with IT syntax. Examination of the raw data reveals, however, that in the vast majority of cases, there were no FT items whatever in those arrays where TF items were included, which is consistent with the TF and FT cases having been interchanged in terms of their logical status. Further support for this explanation comes from particular examination of TF selections on the AN polarity: whilst this is by far the major source of the higher overall TF selections on OI statements, the number of FT cases selected for AN are lower than for any other polarity.

The fact that as TF selections increase in the verification condition FT selections go down is consistent with subjects having a general strategy of avoiding falsifiers, but treating the FT case as the falsifier rather than TF as a consequence of conversion. This avoidance is not necessarily at odds with the notion of the intentional incorporation of fuzziness, since it was only ever hypothesised that a certain (small) proportion of counterexamples would be included. Nevertheless, the selection rates for counterexamples are so low - and, worse, so inconsistent - in both conditions, that it seems very unlikely that these items are included as a consequence of an intention to incorporate fuzziness. Rather, the inclusions of counterexamples seem likely to be the results of errors, both systematic (as in conversion of the conditional) and more haphazard (as in misreading the statement - for example, by failing to notice the word "not").

What *does* turn out to be extremely interesting about these data, however, has nothing whatever to do with notions of fuzziness. This experiment is in most ways very similar indeed to that of Evans (1972c), in which the phenomenon of 'matching bias' was first reported. What makes the comparison particularly important is the fact that although matching bias has been reliably reported countless times over the years, the studies involved appear without exception[37] to have used evaluation tasks, whereas the original (Evans, 1972c) experiment - like the present experiment - was a construction task.

Any comparison of the data from Evans (1972c) with those from the present experiment is necessarily restricted to IT syntax (since that is all that Evans used) and to the falsification condition, since in the Evans study there was a ceiling effect in the verification condition in that subjects selected nothing but TT items. As has been seen, the IT data from the present experiment are very substantially at odds with Evans (1972c) in that matching status appears to have had very little effect whatever upon selections, the five specific predictions of matching bias receiving only scant support, as might be seen from Table 2.11.

In as much as matching bias does not appear ever to have been replicated on a construction task, it might be tempting to hold this finding up as a failure to replicate which casts serious doubt upon the reliability of the data from Evans (1972c). Given, however, that matching bias does seem to be an extremely robust phenomenon on evaluation tasks, and on the selection

[37] Matching bias has also been reported on the selection task (Wason, 1966), which might be argued to be similar to a construction task in that subjects are asked actively to select items rather than merely to evaluate states of affairs.

task, it might be more sensible to resist that temptation and to consider, rather, the extent to which various differences between these two experiments might have contributed to the disparity between the results.

The first difference is that the Evans experiment was conducted manually in the presence of the experimenter, whilst the present experiment was conducted on a microcomputer in the absence of the experimenter. It seems reasonable to suppose that the social effects of the presence of the experimenter in Evans (1972c) might have caused subjects to make more errors - including those resulting from matching bias. Manktelow and Evans (1979) do, however, report no effect whatever of the presence of the experimenter on the manifestation of matching bias in Wason's selection task. Given the fact that the selection task is so very difficult for subjects, however, it is, perhaps, unsurprising that they continued to make errors when the experimenter was absent.

The two studies also appear to have differed in terms of task difficulty, in as much as the subject's task in the present experiment was rather more closely-defined and thus potentially less demanding - and, perhaps, correspondingly less stressful. Evans (1972c) required subjects to produce verifying or falsifying items one at a time until the subject felt that all possible solutions had been exhausted: the task in the present experiment, on the other hand, was the provision of exactly one array consisting of exactly 36 symbols.

Two further differences, also related to task difficulty, lie in the relative complexity of the materials. Matching bias might be defined as a tendency to prefer, or to attend to, items the attributes of which are explicitly

mentioned in the conditional statement. In Evans (1972c), the conditional statement was about two different coloured symbols and their spatial relationship, so that a typical statement might have been

If there is a red triangle on the left
then there is a green square on the right

In the present experiment, the statements were simply about the attributes of a single type of symbol, as in

If the symbols are yellow
then they are circular

A double match for Evans would be a particular pair of symbols spatially oriented in a particular way (in the example given, a red triangle on the left and a green square on the right). In the present experiment, however, the double match is a rather more simple entity - a single symbol, the yellow circle. Wason and Green (1984) reported that reasoning performance was rather better when the rules referred to unitary objects: in as much as subjects' errors on reasoning tasks might be expected to include those resulting from behaviour determined by matching bias, this implies that we should expect matching bias to manifest itself less often when the entities referred to by the rule are simplified in this way.

A further difference in the relative complexity of the materials arises out of a disparity in the range of items from which subjects could make their selections. Evans offered his subjects 16 items, being one each of every possible combination of four colours and four shapes. In the present experiment, on the other hand, subjects were offered just 9 items, being symbols made up of every possible combination of three colours and three shapes.

Comparison of the items available, however, is not a simple matter of 16 *vs* 9. It will be recalled that Evans' statements had the entire specification of one symbol at the antecedent and of a second at the consequent. The 16 items available, then, comprised one way of making the double match (the pair of symbols mentioned in the statement), 14 ways of making the antecedent-only match (the antecedent symbol paired with a symbol other than the consequent symbol), 14 ways of making the consequent-only match in a similar fashion and 91 ways of making the double mismatch (every possible pair of symbols excepting either of those mentioned in the statement). It is not clear how Evans dealt with a pair of symbols which matched those named in the statement but which were in the wrong spatial order, but clearly if spatial order were to be taken into account then the double mismatch could, it seems, be constructed in a further 120 ways! Even this analysis turns out to be something of an oversimplification, since there are clearly two qualitatitively different ways in which a given symbol can mismatch that named in either proposition of the statement: it can be dissimilar in both shape and colour, or it can be dissimilar in just either one of those attributes.

The position in the present experiment is rather different and, thankfully, rather more simple to characterise. Since the statements now have a simple attribute at each of the antecedent and the consequent, the nine symbols now comprise one double match, two antecedent-only matches, two consequent-only matches and four double mismatches.

It seems reasonable to suppose that the considerable difference in the numbers of ways of constructing 'half' matches and double mismatches

might also have played a part in the disparity between the studies. It is clear that this is inextricably linked with the difference in response requirements, where Evans required his subjects to carry on constructing arrays until there were no more that fitted the specifications. It seems somewhat implausible, however, that any subject would have gone meticulously through all 211 possible arrays where the NA rule is to be falsified - that is, where the double mismatch constitutes the TF case! Evans (1972c) is silent about how many arrays were constructed in each case, the analysis being based simply on the initial construction. The difference between the two experiments is clearly more complex than this, however, since Evans' subjects could - if they so wished - stop after one response.

A number of the differences between Evans (1972c) and the present experiment thus appear to be in terms of task difficulty as a result, for example, of a difference in cognitive load. If a task is more difficult, then subjects might be expected to make more errors: to the extent that those errors are systematic, then specific determinants of behaviour such as matching bias might be expected to become more prevalent. Similarly, whilst the difference in terms of the presence of the experimenter is not directly to do with task difficulty, it is, nevertheless, a potential cause of an increased error rate. In summary, then, the failure to find any sign of matching bias in the present experiment might be no more than a reflection of the fact that the task is relatively easier, better-defined and less stressful to perform than that of Evans (1972c).

There is, however, a crucial logical difference between the two studies which has nothing whatever to do with task difficulty. In Evans (1972c), each case chosen by a subject had *in itself* to constitute a logically verifying or falsifying instance whereas, in the present study, the truth status had to

apply merely to the completed array of 36 items. Since the results of Experiment 1 show clearly that subjects (correctly) consider an array to be false when it contains only one or two logically false items, it is, perhaps, unreasonable to expect matching bias to affect the selection of logically true or irrelevant items, since such items may be seen as little more than 'fillers' in the array. This is not, however, to denigrate the importance of such items or to suggest that subjects included them carelessly: if the array constructed does reflect a subject's 'mental model' of a world in which the rule is true (or false) then such 'fillers' might be taken to reflect the deliberate inclusion of items to which the rule simply does not apply.

In considering just how 'matching bias' might come about, Evans (1984) proposes that items matching those mentioned in the rule are judged as more *relevant* to the situation than those which mismatch. If subjects do indeed deliberately set out to include *irrelevant* items in their arrays, then we should not, after all, be surprised to find that such arrays contain a preponderance of mismatching items - since such items are precisely those which the operation of matching bias fails to signify as relevant.

This difference in the status of mismatching items in arrays (as in this experiment) as opposed to individual instances (as in Evans 1972c) does not, however, relieve the matching bias account of its specific predictions about the selection of the critical falsifying items. The matching bias account predicts that matching TF falsifiers should be more reliably avoided in the verification condition, and more reliably included in the falsification condition. In terms of selection patterns, this is to say that when the TF case constitutes the double-match (on the AN rule), then it should be selected less often than for any other rule form in the verification task and more often

than for any other rule form in the falsification task. These predictions are not, however, borne out by the data: for the verification task, Table 2.6 shows TF selection on the AN rule to be far from at its lowest; conversely, for the falsification task, Table 2.7 shows TF selection on the AN rule to be far from at its highest.

2.4 : Experiment 3

Experiment 3 was designed as a construction task similar to Experiment 2 in all respects save that the presentation of the experiment was now entirely manual, it was conducted in the presence of the experimenter and only IT syntax was used. The experiment was designed to investigate the possibility that the lack of any effect of matching status upon selections in Experiment 2 was due, at least in part, to the fact that that experiment was conducted on a microcomputer in the absence of the experimenter, whereas the construction task of Evans (1972c), where matching bias was first identified (and, indeed, those evaluation task studies reported in the literature as having data consistent with matching bias) had all been conducted manually in the presence of the experimenter.

It was hypothesised that the data would be more consistent with the literature to the extent that the matching status of items would influence their selection. As has been seen, this leads to five specific predictions as identified by Evans (1972c):

(1) there will be fewer TF selections on rules with negative antecedents (NA and NN)

(2) there will be more TF selections on rules with negative consequents (AN and NN)

(3) TF selections will be made most often on AN rules

(4) FT selections will be made most often on NA rules

(5) FF selections will be made most often on NN rules

2.4.1 : Design and Methodology

Subjects:

24 first-year BSc (Hons) Psychology undergraduates at Plymouth Polytechnic (18 female, 6 male), mean age 19.1 years.

Materials:

Coloured plastic symbols were to be placed into a 6 x 6 grid marked on a plastic baseboard. The symbols consisted of circles, squares and equilateral triangles of diameter/side 20mm and fashioned from 3mm coloured perspex in primary shades of yellow, blue and red: the symbols thus comprised the same nine discrete symbols used in Experiments 1 and 2. The symbols were contained in open bins of black plastic connected together in the left-to-right order Cy Cb Cr Sy Sb Sr Ty Tb Tr, where C/S/T refers to circle/square/triangle and y/b/r refers to yellow/blue/red. There were 39 of each type of symbol available in the bins, such that the entire grid could, if so required, be filled with a single type of symbol.

The baseboard was made from 6mm smoked grey perspex and was 150mm square. A 6 x 6 grid was formed centrally upon the baseboard by glueing 1mm x 0.5mm off-white plastic strip at intervals of 22.5mm in each direction, leaving a surround of approximately 6mm. The baseboard was designed so as to be a good visual match for the grid presented on the computer screen in Experiment 2 and it is considered that this requirement was well met. It should be noted, however, that for ease of manipulation it was felt that the grid should be somewhat larger than it was in the computer-presented version: the actual grid in this manual version

measured 138mm square as compared with approximately 110mm high x 105mm wide on the computer screen in Experiment 2 (this was as close to square as the distortion of the screen would allow).

The experiment was conducted in a sound-proofed cubicle, the subject being seated across a table from the experimenter. The grid was placed in front of the subject with the bins containing the symbols behind it. Subjects were shown task sheets of folded paper on each of which was printed a statement about the colour and shape of symbols and the instruction "FILL THE GRID TO MAKE THE STATEMENT TRUE" (or "FALSE"). Each task sheet measured 140mm x 215mm and the printing thereon was carried out on an Epson FX-80 printer using the Printmaster ROM to design the typeface of the statement to be as close as possible to BBC BASIC MODE 1 text (as used in the computer-presented Experiment 2). An example of a typical task sheet is presented at Appendix 3.

The statements presented were systematically varied over the four polarities of IT syntax described for Experiments 1 and 2, that is, ITAA, ITAN, ITNA and ITNN. Each polarity combination was presented once so as to be shown to be false, and once so as to be shown to be true, and accordingly there were 8 trials in all, the order of presentation of which was randomly determined for each subject. Half the subjects were given a shape at the antecedent of each rule and a colour at the consequent throughout: the remaining subjects were given the reverse. The actual colour/shape in each proposition of the statement was randomly determined for each trial.

Procedure:

Subjects were seated at the table, introduced to the experimental materials and briefed as follows:

1) they would be shown a sheet of paper on which would appear a statement about the (colour/shape) and (shape/colour) of symbols - subjects were then shown the first task sheet briefly by way of illustration

2) they would be required to fill the grid with symbols such that were the statement to be made about the grid then it would be either true or false, depending on the instruction appearing below the statement

3) there were four 'true' trials and four 'false' trials and these would be presented in a random order

4) they could use as many or as few different types of symbol as they wished in any sort of combination - in case they wished to use just one type, there were enough of each symbol to do so

5) there was no single 'correct' answer, but rather an almost infinite number of ways in which subjects could respond consistent with the instructions - the task should not in any way be seen as a test of ability or intelligence

6) they could change their mind as often as they wished during a trial and could take as long as they wanted - no timing of any sort was being undertaken

7) at the end of each trial, the experimenter would code the selections made by writing numbers on a chart on the reverse of the task sheet - subjects were shown how these codes were arrived at to reassure them that they merely referred to which symbols were present and in no way were marks of correctness of response

8) after the grid had been coded, the experimenter would ask the subject to help him to replace the symbols used into the bins ready for the next trial

No formal timing was undertaken, but subjects all took between 20 and 35 minutes to complete the experiment.

2.4.2 : Results

As for Experiment 2, the dependent variables are the number of symbols selected relating to each truth case, and thus four separate analyses have been carried out (for TT, TF, FT and FF) within each condition (whether the statement was to be shown to be true or false), making eight analyses in all, no overall metric of behaviour having been devised. There was no significant effect of the concepts at the antecedent and consequent positions (that is, which was the shape and which the colour), and accordingly the two groups have been collapsed, which permits an N of 24 throughout.

[a] Selections so as to show the Statements to be TRUE:

The mean numbers of symbols selected in the verification condition are shown in Table 2.14.

truth case / polarity:	AA	AN	NA	NN	overall
[IT syntax only]					
TT	18.83	23.88	25.92	25.92	23.64
TF	1.13	-	1.17	1.58	0.97
FT	2.33	6.29	0.50	2.17	2.82
FF	13.71	5.83	8.42	6.33	8.57
diagonals	11.77	~P~Q			PQ 6.42

Table 2.14 : Experiment 3: mean number of symbols selected so as to show the statements to be TRUE

An analysis of variance upon the number of symbols in each truth case identified the following significant effects:

[1] TT case:

main effect of polarity ($F=4.707$; $df=3$; $p=0.005$) such that mean TT selections were considerably lower on AA than on the other polarities (AA 18.83 < AN 23.88 < NA 25.92 = NN 25.92)

[2] TF case: no significant effects

[3] FT case:

main effect of polarity ($F=9.605$; $df=3$; $p<0.001$) such that mean FT selections were considerably lower on NA and considerably higher on AN than on the other polarities
(NA 0.50 < NN 2.17 < AA 2.33 < AN 6.29)

[4] FF case:

main effect of polarity ($F=8.948$; $df=3$; $p<0.001$) such that mean FF selections were considerably higher on AA than on the other polarities (AN 5.83 < NN 6.33 < ITNA 8.42 < AA 13.71)

[b] Selections so as to show the Statements to be FALSE

The mean numbers of symbols selected in the falsification condition are shown at Table 2.15.

polarity: / truth case	AA	AN	NA	NN	overall
[IT syntax only]					
TT	0.29	3.83	3.42	8.50	4.01
TF	20.75	15.46	22.25	17.79	19.06
FT	6.29	12.63	4.08	3.29	6.57
FF	8.67	4.08	6.25	6.42	6.35
diagonals	13.01	~P~Q			PQ 6.56

Table 2.15 : Experiment 3: mean number of symbols selected so as to show the statements to be FALSE

An analysis of variance was upon the number of symbols in each truth case identified the following significant effects:

[1] TT case:

main effect of polarity ($F=3.485$; $df=3$; $p=0.021$) such that mean TT selections are considerably lower on AA and considerably higher on NN than on the other polarities (AA 0.29 < NA 3.42 < AN 3.83 < NN 8.50)

[2] TF case: no significant effects

[3] FT case:

main effect of polarity ($F=8.004$; $df=3$; $p<0.001$) such that FT selections are considerably higher on AN than on the other polarities (NN 3.29 < NA 4.08 < AA 6.29 < AN 12.63)

[4] FF case: no significant effects

[c] Matching Bias

As for Experiment 2, Tables 2.14 and 2.15 are re-presented as Tables 2.16 and 2.17 so as to make explicit the matching status of the symbols selected corresponding to each truth case. Table 2.16 shows this analysis for the verification condition.

	polarity:	AA	AN	NA	NN	overall
[IT syntax only]						
double match	pq	18.83	-	0.50	6.33	6.42
antecedent only	p~q	1.13	23.88	8.42	2.17	8.90
consequent only	~pq	2.33	5.83	25.92	1.58	8.92
double mismatch	~p~q	13.71	6.29	1.17	25.92	11.77

Table 2.16 : Experiment 3 : mean number of matching and mismatching items selected so as to show the statements to be TRUE

explanatory note: for the statement *'If the symbols are circular then they are yellow'*, the double match is a yellow circle, the antecedent-only match any non-yellow circle, the consequent-only match any yellow non-circle and the double match any non-yellow non-circle

Table 2.17 shows the corresponding analysis for the falsification condition.

	polarity:	AA	AN	NA	NN	overall
[IT syntax only]						
double match	pq	0.29	15.46	4.08	6.42	6.56
antecedent only	p~q	20.75	3.83	6.25	3.29	8.53
consequent only	~pq	6.29	4.08	3.42	17.79	7.90
double mismatch	~p~q	8.67	12.63	22.25	8.50	13.01

Table 2.17 : Experiment 3 : mean number of matching and mismatching items selected so as to show the statements to be FALSE

explanatory note: for the statement *'If the symbols are circular then they are yellow'*, the double match is a yellow circle, the antecedent-only match any non-yellow circle, the consequent-only match any yellow non-circle and the double match any non-yellow non-circle

From even a cursory examination of Tables 2.16 and 2.17, it is clear that there is even less evidence of matching bias in this experiment than there was in Experiment 2. Whereas in Experiment 2, however, the data were merely somewhat inconclusive in terms of the effect of matching status, in the present experiment there appears to be a strong preference (in both the verification and falsification conditions) for the selection of double mismatching items - that is, those the attributes of which match neither the shape nor the colour named in the statement. If matching has any effect at all, then, it is by way of *avoidance* rather than selection.

As for Experiment 2, an analysis is presented at Table 2.18 of the extent to which the data on the falsification version of the task support the five specific predictions made by an account of behaviour in terms of matching bias.

(1) fewer TF selections with negative antecedents (NA+NN vs AA+AN)	
[IT] : yes (just)	43.17 < 45.37
(2) more TF selections with negative consequents (AN+NN vs AA+NA)	
[IT] : no	16.63 < 21.50
(3) most TF selections on AN	
[IT] : no	AN 15.46 < NN 17.79 < AA 20.75 < NA 22.25
(4) most FT on NA	
[IT] : no	NN 3.29 < NA 4.08 < AA 6.298 < AN 12.63
(5) most FF on NN	
[IT] : no	AN 4.08 < NA 6.25 < NN 6.42 < AA 8.67

Table 2.18 : Experiment 3: conformity of selection patterns with those predicted by matching bias *(per Evans, 1972c)*
[falsification condition only & IT syntax only]

It is clear from the foregoing that there is no evidence of matching bias in the overall data from this experiment. It was suggested in the discussion of Experiment 2, however, that the selection of mismatching instances might reflect the deliberate inclusion of items perceived to be irrelevant. Under this view, it was seen that the matching bias account might only be expected to make predictions about the critical falsifying cases. Specifically, this means that when the TF case constitutes the double-match (on the AN rule), then it should be selected less often than for any other rule form in the verification task and more often than for any other rule form in the falsification task. Whilst the verification task prediction is indeed borne out (Table 2.14 shows a zero TF selection on just the AN rule), this is not the case for the falsification task, for which Table 2.15 shows TF selection on the

AN rule to be at its lowest rather than at its highest.

[d] Ranges of Discrete Items Available

Given the strong preference for selection of the double mismatching items over the double matching items, there would appear to be a greater possibility in this experiment than in Experiment 2 that the selections were influenced by the range of discrete symbols available. Table 2.19 shows for the verification condition the mean number of symbols selected on the basis of the range available. As before, the figures given for the 'half' match are the means of the selection rates for the antecedent-only and consequent-only matches.

	matching status	double match	'half' match	double mismatch
truth case	N available	1	2	4
[IT syntax only]				
TT		18.83	24.90	25.92
TF		-	1.35	1.17
FT		0.50	2.25	6.29
FF		6.33	7.12	13.71
mean		6.42	8.90	11.77

Table 2.19 : Experiment 3: mean number of symbols selected so as to show the statements to be TRUE x N of discrete symbols available

Table 2.19 suggests that the number of symbols available appears to have had a considerable overall effect upon the selection rates for each truth status. An

analysis of variance identified the following significant effects:

[1] TT selections:

main effect of N available (F=5.942; df=2; p=0.005) such that the more discrete symbols available, the more selected

[2] TF selections: no significant effect

[3] FT selections:

main effect as above (F=12.344; df=2; p<0.001)

[4] FF selections:

main effect as above (F=10.434; df=2; p<0.001)

Table 2.20 presents a similar analysis for the falsification condition.

truth case	matching status / N available	double match 1	'half' match 2	double mismatch 4
[IT syntax only]				
TT		0.29	3.62	8.50
TF		15.46	19.27	22.25
FT		4.08	4.79	12.63
FF		6.42	5.17	8.67
mean		6.56	8.21	13.01

Table 2.20 : Experiment 3: mean number of symbols selected so as to show the statements to be FALSE x N of discrete symbols available

Again, it is clear that the number of symbols available appears to have had a considerable overall effect upon the selections made. An analysis of variance identified the following significant effects:

[1] TT selections:

main effect of N available ($F=5.787$; $df=2$; $p=0.006$) such that the more discrete symbols available, the more selected

[2] TF selections: no significant effects

[3] FT selections:

main effect as above ($F=8.276$; $df=2$; $p<0.001$)

[4] FF selections:

no significant effects

2.4.3 : Discussion

In Experiment 2 there was very little evidence that selections were being influenced by matching bias, despite that experiment being very similar to that of Evans (1972c). It was suggested that this might, at least in part, be due to Experiment 2 having (novelly) been presented on a microcomputer in the absence of the experimenter. In this present experiment, it was hypothesised that by reverting to manual presentation in the presence of the experimenter, subjects' selections would be far more strongly consistent with an explanation in terms of matching bias.

Taking the verification data first, Table 2.16 might be directly compared with the IT section of Table 2.9 for Experiment 2. It is clear from this this comparison that far from the data merely showing no clear sign of matching bias, they now indicate a very strong tendency to prefer items constituting the double mismatch over all the other types of matching status - and, indeed, to select least of all the double match predicted by matching bias to be the most popular. The falsification data of Table 2.17 show an even stronger preference for the double mismatch, these items now being selected twice as often as items constituting the double match. Far from resulting in a greater tendency to behave in way consistent with matching bias, then, the change of mode of presentation has produced a strong effect in the opposite direction. As was suggested earlier, however, this preference for mismatching items might be explained in terms of the deliberate inclusion of items perceived to be irrelevant - which, under the matching bias account, are precisely those which fail to match the shape and colour mentioned in the rule.

It will be recalled from the analysis of Experiment 2 that it was suspected in that experiment that the unequal numbers of items available corresponding to each matching status might have had an effect on the selection rates, but that this suspicion proved to be unfounded: whilst in the verification condition, the selections did show a trend in the same direction as the number of discrete symbols available, this was extremely weak; in the falsification condition there was, if anything, a trend in the reverse direction.

A similar analysis of the data from the present experiment does reveal an apparent main effect of the number of discrete symbols available upon

selections of at least some of the truth cases - TT, FT and FF for the verification task, TT and FT for the falsification task. This is not, however, to say that the number of symbols has *caused* the corresponding selection rates since, as has already been explained, the number of discrete symbols available is confounded with the matching status - that is, there is no reason to suppose that it might not be the matching status *per se* that is responsible for the selection patterns. It must, however, be conceded that this is relatively unlikely, since such a preference for mismatching items (or avoidance of matching items) would be in direct opposition not only with the findings of the construction task of Evans (1972c) but also, rather more seriously, with the established body of supporting evaluation studies.

It seems likely, on balance, that the discrete number of symbols available was genuinely, and directly, a major determinant of responses. Interestingly enough, there is a way in which this observation this might be related to the matching bias account such that we should not be particularly surprised that matching bias does not appear to have manifested itself. Pollard (1982), in considering the applicability to the study of conditional reasoning of Tversky and Kahneman's (1973) 'availability heuristic'[38], argued that matching bias might be a special case of availability bias such that the items whose attributes match those named in a rule are more available or somehow more easily 'come to mind'. It is clear that in the present experiment, the items which were made more numerously (physically and explicitly) available are precisely those which would would be more (psychologically) 'available' in Pollard's terms. In short, if matching bias does, indeed, arise out of availability, then the 'mismatching bias' apparent in the present experiment is only to be expected.

---

[38] (applied by them in the field of intuitive statistical judgment).

2.5 : Experiment 4

In order to remove the confound between the number of discrete symbols available and the matching status of each symbol, Experiment 4 was designed as a computer-presented construction task similar to Experiment 2 in all respects save that subjects were now offered only four symbols from which to make their selections.

As has been explained earlier, given three colours and three shapes, and a statement mentioning exactly one colour and exactly one shape, then there is exactly one symbol which characterises both the colour and the shape (the double match), two which characterise only that concept appearing as the antecedent of the statement (the antecedent-only matches), two which characterise only that concept appearing as the consequent of the statement (the consequent-only matches) and four which characterise neither the colour nor the shape (the double mismatches).

The four different symbols were made available to subjects in each trial comprised the unique double match, one each at random of the two antecedent-only and the two consequent-only matches and one at random of the four double mismatches. Thus whether the symbols were to be considered in terms of their logical case or in terms of their matching status, there was now exactly one of each available throughout.

The primary hypothesis for this experiment is that the choice of just four symbols will result in the re-emergence of matching bias - the five specific matching bias hypotheses of Experiment 3 will thus stand here also:

(1) there will be fewer TF selections on rules with negative antecedents (NA and NN)

(2) there will be more TF selections on rules with negative consequents (AN and NN)

(3) TF selections will be made most often on AN rules

(4) FT selections will be made most often on NA rules

(5) FF selections will be made most often on NN rules

In addition, whilst the prospective hypotheses of Experiment 2 (that subjects would include a small proportion of counterexamples) were not supported in that experiment, they were retained in the present experiment for the purposes of replication:

(6) in the verification condition (where statements were to be shown to be true) the arrays would include a (small) number of logically falsifying TF cases

(7) in the falsification condition (where statements were to be shown to be false) the arrays would include a (small) number of logically verifying TT cases

2.5.2 : Design and Methodology

Subjects:

20 first-year BSc (Hons) Psychology undergraduates at Plymouth Polytechnic (16 female, 4 male), mean age 20.2 years.

Materials:

The experiment was conducted on a microcomputer as specified for Experiment 1. Responses were collected via a purpose-built keypad as illustrated in Appendix 4a. As for Experiment 2, subjects were presented with a blank 6 x 6 grid of 36 squares above which was shown a statement about the colour and shape of symbols, and were required to press keys on the keypad to place symbols into the blank grid so as to show the statement to be either true or false, as indicated to the right of the grid. As the mechanics of placing and replacing symbols were necessarily somewhat elaborate, detailed instructions were displayed at all times. A screen dump of a typical screen is presented in Appendix 4b.

As for Experiments 1 to 3, the statements were systematically varied over the eight syntactic forms described for those earlier experiments - that is, ITAA, ITAN, ITNA, ITNN, OIAA, OIAN, OINA and OINN. Each syntax x polarity combination was presented once so as to be shown to be false, and once so as to be shown to be true, and accordingly there were 16 trials in all, the order of presentation of which was randomly determined for each subject.

Four different symbols were available to subjects in each trial from which to select those which were to appear in the grid. The four symbols available were a particular subset of the nine possible combinations of three colours (yellow, blue and red) and three shapes (circular, square and triangular), selected on the basis set out above. Half the subjects were given a colour at the antecedent of each rule and a shape at the consequent throughout: the remaining subjects were given the reverse. The actual colour/shape in each proposition of the statement was randomly determined for each trial.

Procedure:

Subjects were briefed using text presented on the computer screen, the only verbal instructions being a reinforcement of the request that they should not commence the experiment until they were absolutely certain that they understood what they were required to do (subjects were asked to call the experimenter for assistance if they were still in difficulty after seeing the instructions twice: no subject needed to do so, however). The instructions were followed by an opportunity for familiarisation with the use of the keypad in moving the cursor around the grid, selecting symbols to be placed into the grid at the cursor and going on to the next task. The instructions given are presented in full in Appendix 4c, but that part of them concerned with the task itself is reproduced below:

> " ...You will be presented with an empty 6 x 6 grid of 36 squares. Above the grid will be shown a statement concerning the shape and colour of symbols which are to be put into the grid, and to the right of the grid will be an instruction to show the statement to be TRUE or FALSE. What you have to do is to fill the grid with symbols so that the statement given at the top of the screen is either TRUE or FALSE with respect to the appearance of the grid
> ...
> ... On each task there will be four different symbols available
> ... The fact that there are four symbols available does not mean that they all must necessarily be included in the grid. You may use any of the symbols as often or as little as you wish ..."

Following the 16 trials, subjects were debriefed by text presented on the computer screen, as presented in Appendix 4d. No formal timing was undertaken, but all subjects took between 30 and 50 minutes from starting to read the instructions to the completion of the experiment.

2.5.2 : Results

As for Experiments 2 and 3, the dependent variables are the number of symbols selected relating to each truth case, and thus four separate analyses have been carried out (one each for TT, TF, FT and FF) within each condition (whether the statement was to be shown to be true or false), making eight analyses in all, no overall metric of behaviour having been devised. There was no significant effect of the concepts at the antecedent and consequent positions (that is, which was the shape and which the colour), and accordingly the two groups have been collapsed, which permits an N of 20 throughout.

[a] Selections so as to show the Statements to be TRUE

The mean numbers of symbols selected in the verification condition are shown at Table 2.21.

truth case	polarity:	AA	AN	NA	NN	overall
IT syntax						
TT		21.20	23.90	17.90	19.90	20.72
TF		1.50	-	3.30	3.55	2.09
FT		3.60	4.85	3.90	4.85	4.30
FF		9.70	7.25	10.90	7.70	8.89
IT diagonals	9.44	~P~Q				PQ 8.20
OI syntax						
TT		25.30	24.35	22.35	22.25	23.56
TF		0.60	1.80	2.95	1.25	1.65
FT		2.65	4.60	3.30	7.25	4.45
FF		7.45	5.25	7.40	5.25	6.34
OI diagonals	9.31	~P~Q				PQ 8.91
IT + OI combined						
TT		23.25	24.13	20.13	21.08	22.14
TF		1.05	0.90	3.13	2.40	1.87
FT		3.13	4.73	3.60	6.05	4.38
FF		8.58	6.25	9.15	6.48	7.61
IT + OI diagonals	9.38	~P~Q				PQ 8.56

Table 2.21 : Experiment 4: mean number of symbols selected so as to show the statements to be TRUE

An analysis of variance identified just one significant effect: on the FF case, there was a main effect of syntax (F=5.271; df=1; p=0.034) such that FF selections were higher on the IT statements than on the OI statements (8.89 *vs* 6.34).

[b] Selections so as to show the Statements to be FALSE

The mean numbers of symbols selected in the falsification condition are shown at Table 2.22. An analysis of variance identified significant effects on the selection rates for TF and FT cases only as follows:

[1] TF case:

main effect of polarity (F=4.794; df=3; p=0.005) such that mean TF selections were rather higher on statements with positive antecedents than on those with negative antecedents
(AA 20.98+AN 18.08 *vs* NA 11.95+NN 15.23)

[2] FT case:

main effect of polarity (F=3.002; df=3; p=0.038) such that mean FT selections were considerably higher on NA than on the other polarities (AN 4.90 < NN 5.27 < AA 5.60 < NA 10.83)

truth case	polarity:	AA	AN	NA	NN	overall
IT syntax						
TT		1.65	2.80	3.20	2.70	2.59
TF		21.15	19.20	13.40	17.65	17.85
FT		4.90	5.65	8.85	2.55	5.49
FF		8.30	8.35	10.55	13.10	10.08
IT diagonals	7.51	~P~Q			PQ	10.70
OI syntax						
TT		3.20	4.70	1.65	2.95	3.13
TF		20.80	16.95	10.50	12.80	15.26
FT		6.30	4.15	12.80	8.00	7.81
FF		5.70	10.20	11.05	12.25	9.80
OI diagonals	5.82	~P~Q			PQ	11.30
IT + OI combined						
TT		2.42	3.75	2.42	2.82	2.86
TF		20.98	18.08	11.95	15.23	16.56
FT		5.60	4.90	10.83	5.27	6.65
FF		7.00	9.27	10.80	12.67	9.94
IT + OI diagonals	6.67	~P~Q			PQ	11.00

Table 2.22 : Experiment 4: mean number of symbols selected so as to show the statements to be FALSE

[c] Matching Data

As for Experiments 2 and 3, Tables 2.21 and 2.22 are re-presented below as

Tables 2.23 and 2.24 so as to make explicit the matching status of the symbols selected corresponding to each truth case. It will be recalled that unlike those two earlier experiments, there is now just one item relating to each matching status.

	polarity:	AA	AN	NA	NN	overall
IT syntax						
double match	~pq	21.20	-	3.90	7.70	8.20
antecedent only	p~q	1.50	23.90	10.90	4.85	10.29
consequent only	~pq	3.60	7.25	17.90	3.55	8.07
double mismatch	~p~q	9.70	4.85	3.30	19.90	9.44
OI syntax						
double match	pq	25.30	1.80	3.30	5.25	8.91
antecedent only	p~q	0.60	24.35	7.40	7.25	9.90
consequent only	~pq	2.65	5.25	22.35	1.25	7.88
double mismatch	~p~q	7.45	4.60	2.95	22.25	9.31
IT + OI combined						
double match	pq	23.25	0.90	3.60	6.48	8.56
antecedent only	p~q	1.05	24.13	9.15	6.05	10.09
consequent only	~pq	3.13	6.25	20.13	2.40	7.98
double mismatch	~p~q	8.58	4.73	3.13	21.08	9.38

Table 2.23 : Experiment 4: mean number of matching and mismatching items selected so as to show the statements to be TRUE

explanatory note: for the statement *'If the symbols are circular then they are yellow'*, the double match is a yellow circle, the antecedent-only match any non-yellow circle, the consequent-only match any yellow non-circle and the double match any non-yellow non-circle

It is clear from Table 2.23 that in the verification condition of Experiment 4 there is no evidence whatever of matching bias in the overall data in that there is virtually no difference at all in the selection rates across the four

types of matching status - and to the extent that any difference does exist, the double mismatch is (very marginally) preferred to the double match. The prediction of the matching bias account that when the TF case constitutes the double-match (on the AN rule), then it should be selected less often than for any other rule form is, however, supported in that Table 2.14 shows TF selection to be at its lowest overall on the AN rule.

Table 2.24 presents a similar analysis for the falsification condition:

	polarity:	AA	AN	NA	NN	overall
IT syntax						
double match	pq	1.65	19.20	8.85	13.10	10.70
antecedent only	p~q	21.15	2.80	10.55	2.55	9.26
consequent only	~pq	4.90	8.35	3.20	17.65	8.53
double mismatch	~p~q	8.30	5.65	13.40	2.70	7.51
OI syntax						
double match	pq	3.20	16.95	12.80	12.25	11.30
antecedent only	p~q	20.80	4.70	11.05	8.00	11.14
consequent only	~pq	6.30	10.20	1.65	12.80	7.74
double mismatch	~p~q	5.70	4.15	10.50	2.95	5.82
IT + OI combined						
double match	pq	2.42	18.08	10.83	12.67	11.00
antecedent only	p~q	20.98	3.75	10.80	5.27	10.20
consequent only	~pq	5.60	9.27	2.42	15.23	8.13
double mismatch	~p~q	7.00	4.90	11.95	2.82	6.67

Table 2.24 : Experiment 4: mean number of matching and mismatching items selected so as to show the statements to be FALSE

explanatory note: for the statement *'If the symbols are circular then they are yellow'*, the double match is a yellow circle, the antecedent-only match any non-yellow circle, the consequent-only match any yellow non-circle and the double match any non-yellow non-circle

There is some indication from the data in Table 2.24 that in the falsification condition subjects might have been influenced by the matching status of the items from which they were able to make their selections. Whilst none of the mean selection rates diverges substantially from the mean of 9.00 per cell, the double match and the double mismatch do diverge in a complementary fashion, such that to the extent to which the double match is preferred, this is apparently at the expense of the double mismatch.

Table 2.25 considers the extent to which the five traditional predictions of matching bias (Evans, 1972c) are supported by these data. It is clear from this table that four of the five predictions of matching bias are indeed borne out by the data from the falsification condition of this experiment, for both IT and OI syntax. It should be noted, however, that the one prediction which is not supported is that about TF selection on the AN form.

prediction 1:	
fewer TF selections with negative antecedents (NA+NN *vs* AA+AN):	
IT : yes	30.05 < 40.35
OI : yes	23.30 < 37.75
prediction 2:	
more TF selections with negative consequents (AN+NN *vs* AA+NA):	
IT : yes	36.85 > 34.55
OI : yes	39.75 > 31.30
prediction 3:	
most TF selections on AN:	
IT : no	NA 13.40 < NN 17.65 < AN 19.20 < AA 21.15
OI : no	NA 10.50 < NN 12.80 < AN 16.95 < AA 20.80
prediction 4:	
most FT on NA:	
IT : yes	NN 2.55 < AA 4.90 < AN 5.65 < NA 8.85
OI : yes	AN 4.15 < AA 6.30 < NN 8.00 < NA 12.80
prediction 5:	
most FF on NN:	
IT : yes	AA 8.30 < AN 8.35 < NA 10.55 < NN 13.10
OI : yes	AA 5.70 < AN 10.20 < NA 11.05 < NN 12.25

Table 2.25 : Experiment 4: conformity of selection patterns with those predicted by matching bias
(per Evans, 1972c)
[falsification condition only]

2.5.3 : Discussion

This experiment was designed to remove the confound of the number of discrete symbols available with the matching status of the symbols, such confound having been seen as a possible reason for there being no evidence of matching bias on Experiments 2 and 3. To the extent that the data from the present experiment might be interpreted as exhibiting matching bias, then this would suggest that the lack of such an effect on Experiment 2 - and the reverse, 'mismatching bias' on Experiment 3 - were merely artifacts of the materials made available to subjects from which to make their choices.

Looking first at the verification data, it is clear that standardising the availability of items has had no effect whatever - the selection rates whether looked at from the point of view of truth case or of matching status are virtually indistinguishable from those in Experiment 2, of which the present experiment is an exact replication in all respects save the number of discrete items available. It will be recalled, however, that in the verification condition of Evans (1972c), there was an absolute ceiling effect in that all subjects chose the TT case 100% of the time, so that matching status failed to exert any influence, selections being governed wholly by factors leading to conformity with logic.

In Experiment 2 and the present experiment, however, there is no such ceiling effect, so that any influence of matching status might be expected to reveal itself. It appears from the data overall, however, that matching status has no effect whatever, none of the mean selection rates over the four different types of matching status diverging more than a trivial amount from the overall mean of 9 per cell. If, however, as suggested earlier,

mismatching cases were to have been deliberately selected so as to reflect a state of the world in which the rule is not always applicable, then matching bias should only be expected to reveal itself in the avoidance of falsifying TF items on the NA rule: and this is, indeed, apparent from Table 2.14.

In the falsification condition, however, matching status does appear to have had an influence on selection rates even on the overall data, in that for both IT and OI syntax, subjects select the double match most often overall and the double mismatch least often overall, four of Evans' (1972c) five predictions being supported. This is, perhaps, embarrassing for the explanation offered earlier for the absence of an overall effect of matching bias on Experiments 2 and 3. If the preponderance of mismatching items in those experiments resulted wholly from the deliberate selection of items perceived to be irrelevant, then a reduction in the range of such items should have had little or no effect upon the probability of their selection.

The well-established effect of matching status on the falsification task appears, then, to have re-emerged as a consequence of the removal of the confound of matching status with the range of items available. This clearly begs the question of how these two factors might be related. In Experiment 2, neither the range available nor the matching status appeared to have had any great effect on selections, whilst in Experiment 3, where the range available did appear to have had some effect, the effect of matching status was, if anything, to render items whose attributes matched those named in the rule *less* likely to be selected.

At first sight, this series of results might be taken to indicate that matching bias and the physical or visual availability of items are two distinct

determinants of behaviour which compete in the determination of subjects' responses. There is, however, no necessary reason to propose two separate determinants of behaviour if we adopt Pollard's (1982) suggestion that 'matching bias' is a special case of a more general 'availability bias'. 'Availability bias' is concerned with the *psychological* availability of items which is, in turn, determined by the whole panoply of features of the materials and the task with which the subject is presented. Under such an account, it is quite clear that the naming of items of attributes in a rule would be one factor in determining what was psychologically available, whilst the *physical* availability of items (the range of discrete symbols explicitly presented) would be yet another. What would have been competing in these experiments would, in these terms, have been not two separate biases, but rather two distinct types of feature of what was presented to the subject. Furthermore, what they would have been competing for would have been not to determine the subject's behaviour directly but, rather, to determine the relative extent to which various items were *psychologically* available.

It is clear, however, that psychological availability in these terms is far from the only determinant of selection in this experiment. Even with the removal of the confound of physical availability with matching status, the TT case in the verification condition accounts for just 62% of all selections, whilst the TF case accounts for just 46% of all selections in the falsification condition. This suggests that the requirement to produce an array rather than individual instances might indeed have led some subjects to construct an array so as to reflect the hypothetical nature of the conditional - that is, an array to which the rule is only relevant part of the time.

2.6 : General Discussion of Experiments 1 to 4

This general discussion begins by considering the extent to which the findings of the complementary investigations of Experiments 1 and 2 can provide support for the general notion that the abstract conditional might be understood as involving 'fuzziness'. As will have become clear from the individual discussions above, however, to the extent that fuzziness was apparent in those experiments, there is some reason to suppose that it might largely have resulted from erroneous (or, at least, non-logical) reasoning arising out of task difficulty.

The apparent influence of various extra-logical determinants upon responses in all four Experiments will thus be addressed and consideration given to the extent to which the phenomena observed might be accommodated within the various theoretical frameworks identified and described in Chapter 1.

The essential notion which underlay the first two experiments was that people ordinarily interpret conditional statements in a 'fuzzy' way: that is, in such a way that they will tolerate - or even expect - a certain (small) proportion of observations that are logically inconsistent with the statements. Experiment 1, an evaluation task, offered a limited amount of support for this contention. Whilst 'fuzzily true' statements were rated as 'false' overall, they were far less often so rated than statements about arrays which contained more substantial proportions of falsifiers - and there was very strong evidence that this effect was not a result of the four falsifiers in the fuzzily true arrays simply failing to be noticed.

Experiment 1 also suggested that people differ in how they interpret the conditional relation, such that 6 of the 24 subjects behaved strictly in accordance with a strict 2-valued or 3-valued logic throughout. WIthin-subject behaviour was, however, highly consistent, in so far as 17 of the remaining 18 subjects gave a differentially 'fuzzy' response to the fuzzily true array at least 75% of the time.

To the extent that responses did reflect an underlying understanding of the conditional as fuzzy, however, this was only so where the fuzziness arose out of the presence of a very small number of falsifiers in an otherwise true array: once the number of falsifiers was substantial, variations in the TT:TF ratio had no effect. One possible reason for this assymmetry was identified as the fact that, as Kyburg (1983) argues, universal generalisations in our everyday life to seem to permit occasional counterexamples and are little the less useful for the fact that they are thereby rendered logically false; once a statement is false more than occasionally, however, then it ceases to be a useful rule by which to act.

It must, however, be acknowledged that at least some of the apparent fuzziness might have resulted from an (inescapable) demand effect in Experiment 1 in that it may have been implied by the instructions and mechanics of the task that there would be trials upon which the appropriate response would be not 'absolutely true' or 'absolutely false' but, rather, somewhere in between - and not necessarily the strictly indeterminate 'can't tell'. That most subjects did appear to be happy to entertain the notion of degrees of truth for the purposes of this experiment does not necessarily imply, therefore, that they would do so in their everyday reasoning.

Further support for the general notion that people ordinarily interpret conditional statements in a 'fuzzy' way was provided by the construction tasks of Experiment 2 to 4 in which, for both verification and falsification versions of the task, subjects did include a certain (very small) number of counterexamples. It is clear, however, that some of these counterexamples might have been included through errors, both systematic and random: indeed, it is difficult to conceive of any way in which a construction task could be designed such that occasional inclusions of logically erroneous items could be guaranteed to be the deliberate incorporation of fuzziness rather than a low rate of error.

In the conception of this series of experiments, the term 'fuzzy' was taken to denote that subjects would tolerate, or intentionally include, a small proportion of counterexamples to a conditional rule: in these terms, these four experiments appear to have provided some support for this notion. Furthermore, to the extent that fuzziness has to do with *relevance* rather than truth, then the results of Experiments 2 to 4 support a fuzzy interpretation of the conditional in that subjects did include a substantial proportion of FT and FF items[39], about which the statement was irrelevant (or which were irrelevant to the truth of the statement). Whilst the inclusion of these items (over and above those needed merely to 'solve' the task) might have resulted from the requirement that the entire array be filled, most subjects nevertheless consistently chose to employ FT and FF items in addition to the more relevant TT items (for the verification task) or TF items (for the falsification task).

[39] 19% of selections overall were FT or FF cases:

	FT	FF
verification	12%	22%
falsification	18%	23%

Whilst the results of these four experiments do appear to reflect a fuzzy interpretation of the conditional, it is clear that subjects' behaviour might, in part, have resulted from various extra-logical determinants of behaviour which are independent of a fuzzy understanding of the conditional. It is thus necessary to consider the range of non-logical phenomena displayed over these four experiments as a whole and to consider the extent to which they might be accommodated within the various theoretical frameworks identified and described in Chapter 1.

Experiment 1 revealed a consistent effect of syntactic form, such that with statements with the IT syntax, absolutely true statements were rated more consistently as 'true', variously false statements as 'false' and fuzzily true statements more often as 'false'. If we disregard the notion of fuzziness and deal only with logical correctness, then this might alternatively be expressed as the fact that less errors were found on IT syntax than on OI syntax. Furthermore, it has been demonstrated above that a large proportion of such errors might be explained through a single type of error which is well-established as being more prevalent with OI syntax: that of illicit conversion from 'P only if Q' to 'if Q then P'.

IT/OI differences were also to be found in the construction tasks of Experiments 2 and 4[40]. The effects were particularly pronounced in Experiment 2, in which more errors involving the critical TT and TF items occurred with OI syntax than IT syntax - especially where negatives were involved. Here again, the involvement of illicit conversion was strongly implied by the tendency towards the mutual exclusivity of FT and TF cases, which become logically interchanged as a consequence of conversion. Furthermore, to the extent that matching bias was identified at all on

40 (IT/OI differences do not arise in Experiment 3, since only IT syntax was used).

Experiment 2, this was only on OI syntax.

The effects of negativity were also seen to permeate the four experiments. In Experiment 1, the polarity of the antecedent proposition (*ie* whether affirmed or negated) had similar effects upon the ratings to the IT/OI difference characterised above: here again, if we disregard the notion of fuzziness and deal only with logical correctness, then this might alternatively be expressed as the fact that less errors were found with AA and AN rules than with NA and NN rules. Quite apart from the well-established fact that the presence of negatives leads to an increased error rate, a further source of difficulty arises in Experiments 1 to 3, in that it is always the case that twice as many discrete types of symbol possess the negation of an attribute than its affirmation. The effect of negatives was not restricted to the evaluation task of Experiment 1: in Experiment 2, whilst there was no main effect of negativity, there was a significant increase in logical errors on the verification task when negative antecedents were combined with OI syntax.

The consideration of extra-logical determinants of behaviour on these experiments thus far has dealt only with those determinants which are destructive - that is, which cause logical errors. One further extra-logical determinant, which will cause errors only sometimes, is matching bias (Evans, 1972c), which was addressed at length in the various individual analyses and discussions in view of the surprise generated by its non-appearance on Experiment 2: indeed, it will be recalled that the failure to find matching bias on that Experiment was the primary motivation for Experiment 3 and that, in turn, for Experiment 4.

Possible reasons for the failure to find matching bias on Experiment 3 and with IT syntax on Experiment 2 have been considered at length above. Of particular importance to this present line of discussion is the relationship of matching bias to the logical status of items whose selection it appears to influence. Since on none of the experiments was matching bias found on the verification task, we may restrict this discussion to the falsification task. On none of the tasks where matching bias was found overall did it appear to have had any noticeable effect upon selections of the (logically necessary) falsifying TF case nor the verifying (logically unnecessary, but relevant) TT case. Rather, the whole of the observed overall matching bias effect appears to come from the fact that selections of the (logically irrelevant) FT and FF cases are by far at their highest when they constitute the double match (on NA rules for the FT case and NN rules for the FF case).

It is also relevant to the present discussion to consider the relationship between matching bias and task difficulty. In Experiment 2, matching bias was only found with OI syntax: on Experiment 4, although matching bias is now to be seen on IT syntax also, by far the greater effect is with OI syntax. In considering the effect of matching bias on tasks made more difficult by the presence of negatives, we are, of course, in some difficulty given the fact that matching status and truth status are perfectly counterbalanced by the permutation of the presence of negative components in the rules. It will be noted, however, that the only selections to be substantially influenced by matching are those occurring upon rules with negative antecedents (*ie* FT on NA and FF on NN).

As was seen earlier, it is well-established that people find OI syntax more difficult to reason with than IT syntax and that errors also increase with the use of negative components. Given that matching bias is at its strongest

when OI syntax is used and when the antecedent proposition is negated, it would appear that its influence is enabled (or, at least enhanced) by the difficulty of the task.

It was suggested earlier that the operation of factors which make a task more difficult (for example, the lack of any content or context, the syntactic form or the presence of negatives) might be characterised as follows: either they simply increase the likelihood that a subject will 'lose track' of what they are doing and so respond on the basis of a 'faulty' sequence of operations, or else the sequence of operations might be so long-winded that the subject abandons it before its proper completion. A complementary and more informal characterisation arising out of the general flavour of the findings in these experiments might be that task difficulty somehow 'opens the door' to various extra-logical determinants of response.

This characterisation is quite compatible with a number of the theoretical approaches described earlier. Johnson-Laird (1983a; 1986), in specifying his mental models approach to conditional reasoning, proposes that the function of the antecedent is to establish a context in terms of a state of affairs in which the consequent can be evaluated. He argues, however, that most conditionals lack clear-cut truth conditions, in that their antecedents, and the beliefs they trigger, place insufficient constraints upon the possible set of antecedent models, errors then arising out of limitations of the working memory system, in so far as it is not possible to hold and manipulate too large a set of models. It seems clear that an abstract reasoning task, with neither content nor context about which the subject would have any beliefs or knowledge whatever, might almost be guaranteed to contain too little information for the proper constraint of the set of possible antecedent

models, such that reasoning on such tasks might be expected to be highly error-prone.

The suggestion that task difficulty 'opens the door' to extra-logical factors sits less comfortably with the 'heuristics and biases' approach, however. The heuristics and biases approach, which stresses the operation of various heuristic short-cuts such as Tversky and Kahneman's (1973) 'availability bias' generally has it that people do not reason at all but, rather, that their responses are determined by extra-logical factors which may, sometimes, lead to logically correct behaviour. Whilst it is attractive to suppose that such biases may be the primary, or sole, determinants of response where the task is especially difficult, there seems to be no necessary reason to suggest that people do not actually do something which we might wish to call 'reasoning' on rather easier problems. This is not, however, necessarily to suggest that people in any way possess a formal logical system or that they use logical inference rules: rather, as Evans (1984) proposes in his 'heuristic and analytic framework', such 'logical' reasoning might be done on the basis of the normal linguistic usage of the syntactic form.

The proposed characterisation is not otherwise entirely consonant with Evans' (1984) 'heuristic and analytic framework', however. Evans argues for a strictly serial operaton of the two components, the heuristic system selecting items upon which the analytic system may then operate, such selection being on the basis of a relevance judgment. In abstract tasks, errors are seen as arising out of the insufficency of the information from content and context upon which to make the correct selection of items at the heuristic stage. That some sort of selection needs to take place before the analytic system can start to operate at all is unobjectionable. What is more

open to question is the notion that the heuristic system has no subsequent part to play once the analytic system has begun to operate. There appears to be no necessary reason, for example, why the analytic system should not be able to fall back upon the heuristic system when a problem turns out to be too difficult to solve in a reasonable time. Similarly, where a problem is ambiguous (and thus has multiple analytic solutions) there is no reason why the heuristic system should not be called upon to select from amongst those solutions. Even this suggests a serial (albeit oscillating or cascading) relationship between the two systems: there is, however, no reason to suppose that both systems should not be able to operate in parallel.

Common to Evans' (1984) heuristic and analytic approach and to Johnson-Laird's (1983a) mental models approach is the notion that the lack of information in abstract tasks will somehow handicap the 'normal' mechanism of everyday reasoning - and this is entirely compatible with the present suggestion that task difficulty somehow 'opens the door' to extra-logical factors. Rather more direct support is, however, to be found in Cheng and Holyoak's 'pragmatic reasoning schemas' approach, central to which is the assumption that logically correct reasoning depends upon the elicitation of a particular kind of pragmatic reasoning schema which happens to lead to behaviour which accords with formal logic[41]. The pragmatic reasoning schema account states quite explicitly, however, that an arbitrary rule will not reliably evoke any particular kind schema - or even any schema at all. If pragmatic reasoning schemas are, indeed, how we do our normal everyday reasoning, then this system may be brought to its knees by an abstract reasoning task, such that responses will then necessarily need to be determined by extra-logical factors.

[41] (for example, the permission or obligation schema, or certain kinds of causal schema).

This concludes the consideration of Experiments 1 to 4. It has been demonstrated that people will tolerate, or even expect, some degree of 'fuzziness' in the use of the conditional, lending support to the notion that people understand this connective in a 'fuzzy' way. This position is in no way weakened by the fact that few subjects intentionally included counterexamples in the arrays which they constructed, since it reasonable to suppose that they would strive in such a task to produce a 'correct' solution.

It has been proposed that typical patterns of response on conditional reasoning tasks might be characterised as task difficulty 'opening the door' to extra-logical determinants of behaviour, and this has been shown to be largely compatible with the major theoretical approaches described in Chapter 1. One source of task difficulty which is peculiar to (or, at least, necessarily vastly more prevalent in) abstract tasks is the paucity of information arising out of the content of the propositions and the context in which the conditional statement is made. In Chapter 3, a further four experiments will be described which investigate the effects of removing (or, at least, alleviating) this source of task difficulty by the use of thematic materials and the explicit provision of a context in which the statement is made.

Chapter Three

Experiments 5 to 8:
The 'Contexts' Experiments

3.1 : Introduction to Experiments 5 to 8

This Chapter presents Experiments 5 to 8. These represent a considerable departure from Experiments 1 to 4 which, by their use of abstract materials, were quite deliberately devoid of any linguistic context. These later experiments, on the other hand, used thematic materials and were designed so as to investigate reasoning behaviour in a variety of 'real life' contexts. As was seen in Chapter 1, there is evidence that in certain linguistic contexts - notably those in which the utterance of the conditional statement constitutes the issue of a promise or threat (see, for example, Geis and Zwicky, 1971; Fillenbaum, 1975; 1976) - it seems that certain logically erroneous interpretations might be seen as quite proper on pragmatic grounds. Experiments 5 to 8 constitute a systematic investigation of the effects of various linguistic contexts, in sharp contrast to studies like Experiments 1 to 4, which go to some pains to avoid such effects.

Common to all four experiments to be reported in this chapter is the manipulation of the linguistic context in which the conditional statement is uttered. In the next section, a coarse-grained taxonomy of natural language uses of the truth-functional double-affirmative conditional of the form 'if P then Q' will be offered, which categorizes and describes the set of eight linguistic contexts used in these four experiments.

A brief descriptive overview of the experiments is then presented so as to make explicit the relationships between them. Each individual experiment is then reported in detail, following which the series as a whole is discussed, and issues arising therefrom addressed.

3.1.1 : A Taxonomy of Usage

It is not by any means suggested that the taxonomy which follows is exhaustive - nor that the categories offered are finely-grained enough to capture all of the subtleties of usage of even just the truth-functional double-affirmative conditional in 'if P then Q', quite apart from conditionals of other polarities and syntactic forms and non-truth-functional conditionals. Rather, it is offered as a set of classifications which is small enough to be workable, yet sufficiently factorial in nature to serve as a basis for useful prospective investigations.

Eight distinct natural language uses were identified for the truth-functional double-affirmative conditional in the form 'if P then Q'. Seven of these (Temporal, Causal, Promise, Threat, Tip, Warning and Universal) were used throughout the series of four experiments, whilst the eighth (Intentional) - and the factorial analysis which its inclusion affords - was only identified only after Experiments 5 and 6 had been conducted, and was thus used in Experiments 7 and 8 only. The eight contexts will described in detail below, but it is useful first to set out various dimensions upon which they might be distinguished:

[a] temporality

With the exception of the Universal context, it will be seen that every one of the contexts to be described is 'temporal' in the sense that the antecedent is always temporally distinct from (and in general precedes) the consequent.

[b] : causality

With the exception of the Temporal and Universal contexts, it will be seen that in every one of the contexts to be described the antecedent and consequent are *causally* connected, such that the antecedent causes (or at least is capable of causing) the consequent.

[c] : speaker's intention

In the Temporal, Causal, Intentional and Universal contexts, the speaker's intention in uttering the conditional statement is merely to communicate to the hearer some item of information. In the Promise, Threat, Tip and Warning contexts, on the other hand, the intention is to modify the hearer's behaviour.

[d] : speaker's control over the consequent

In the Intentional, Promise and Threat contexts, the speaker has control over the truth of the consequent (and thus, over the truth of the conditional), whereas in the remaining contexts the speaker has no such control.

[e] : hearer's control over the antecedent

In the Promise, Threat, Tip and Warning contexts, the hearer has control over the truth of the antecedent, whereas in the remaining contexts the hearer might merely *observe* whether the antecedent is true or false.

[f] : subjective quality of the consequent

In the Promise and Tip contexts, the consequent is always a subjectively Good Thing for the hearer, whilst in the Threat and Warning contexts it is always a subjectively Bad Thing. The remaining contexts are not reliably distinguished on this dimension, but the consequent will typically be comparatively neutral to the hearer.

The eight contexts identified might be specified as follows:

1 Temporal

eg: " If the next southbound train is for Ealing
then the one after it will be for Wimbledon "

Whilst in all the other contexts (apart from the Universal) the antecedent and consequent are related temporally, the Temporal context is distinguished from the rest in terms of there being no causal link whatever between the antecedent and the consequent (other than that they might both be effects of a common cause): the temporal connection is the *only* link that subsists between them.

This context is characterised by there existing some natural or artificial formal time order within which the two propositions are related to each other.

2 Causal

eg: " If the lorry is heavier than the legal limit
then the alarm bell will ring "

In common with the Intentional context, the Causal context is distinguished from the Promise, Threat, Tip and Warning contexts in two ways. Firstly, the speaker in uttering the statement intends merely to communicate some item of information to the hearer. Secondly, the hearer has no control over the truth of the antecedent, but rather might merely observe whether it is true or false.

The Causal and the Intentional contexts might be distinguished in that in the Intentional context the speaker has control over the truth of the consequent, whilst in the Causal context the speaker has no such control.

3 **Intentional**

eg: " If it's fine over the weekend
then I'll go fishing "

This is similar to the Causal context except in so far as the speaker does now have control over the truth of the consequent. In addition, the causal link is somewhat different, in that in the Causal context the antecedent directly causes the consequent, whilst in the Intentional context it merely enables the speaker to carry out the stated intention. In the Intentional context, then, the speaker informs the hearer of his future intention (the consequent) given that some precondition (the antecedent) should turn out to be satisfied.

4 **Promise**

eg: " If you wash the car
then I'll let you borrow it tonight "

As stated above, in the Promise, Threat, Tip and Warning contexts the utterance of the statement is motivated by the speaker's desire to modify the behaviour of the hearer. In addition, it is the case that the hearer always has control over the truth of the antecedent. These four contexts might be further distinguished, however, on two dimensions - the control of the speaker over the truth of the consequent (and thus over the truth of the conditional) and whether the consequent is a subjectively Good Thing or Bad Thing for the hearer.

In the Promise and Threat contexts, the speaker has direct control over the truth of the consequent, whilst in the Tip and Warning contexts, the speaker does not have any such control. The Promise and Threat contexts might be distinguished in terms of the quality of the consequent - in the Promise context this is always a subjectively Good Thing for the hearer.

5 **Threat**

eg " If you wear jeans to work again tomorrow
then I'll fire you "

This context is precisely similar to the Promise context except that the consequent is now a subjectively Bad Thing for the hearer.

6 **Tip**

eg: " If you stand by the pillar
then you'll be served immediately "

As stated above, the Tip and Warning contexts might be distinguished from the Promise and Threat contexts on the basis

of the speaker's control: in the Promise and Threat contexts the speaker is in direct control of the truth of the consequent, whilst in the Threat and Warning contexts the speaker has no such control.

The Tip and Warning contexts might be distinguished in terms of the quality of the consequent - in the Tip context this is always a subjectively Good Thing for the hearer.

7 **Warning**

eg: " If you pull his tail again
then he'll bite you "

This context is precisely similar to the Tip context except that the consequent is now a subjectively Bad Thing for the hearer.

8 **Universal**

eg " If the student is doing Economics
then he is a socialist "

This context differs from all the other contexts in that the antecedent and consequent are not temporally related in any way - causally or otherwise. The Universal context is characterised by its being a conditional expression of a contingent universal: in this case, the contingent universal "All students doing Economics are socialists".

It may be helpful to later discussion to summarise graphically the status of the various contexts on the two 'control' dimensions and in terms of the subjective quality of the consequent so as to make their factorial relationship explicit. Figure 3.1 depicts such an analysis:

			H's control over antecedent	
		(quality)	YES	NO
S's control over consequent	YES	*(good)*	Promise	Intentional
		(bad)	Threat	
	NO	*(good)*	Tip	Temporal Causal Universal
		(bad)	Warning	

Figure 3.1 : The eight contexts in terms of:

Speaker's control over the truth of the consequent
Hearer's control over the truth of the antecedent
Subjective quality of the consequent

The eight contexts might be distinguished upon one further dimension: the subjective non-contingent likelihood of the consequent - that is, how likely it might seem to be that the consequent will become true independently of the antecedent becoming true. It is stressed, however, that whereas the defining characteristics of the foregoing six dimensions were all strictly adhered to in the construction of the materials, this further dimension is merely described (entirely intuitively) in terms of where upon it the various contexts might *typically* fall in everyday usage.

[g] : subjective non-contingent likelihood of the consequent

It seems unobjectionable to propose that where the consequent appears unlikely to become true in virtue of any event other than the antecedent becoming true, then an equivalence interpretation might be expected, since

the FT case would be expected to be false. Where, by contrast, there appear to be a range of alternative plausible ways in which the consequent might come to be true quite independently of the antecedent coming to be true, then an implication interpretation might be expected, since the FT case would no longer mean that the statement was false (since we can readily accept that the consequent has become true even though the antecedent has not become true).

As was seen in Chapter 1, the utterance of a particular statement in a particular context can constitute a communicative act which is intended to, and does, convey rather more than just the propositional content of the statement. Furthermore, given certain mutually shared conventions of linguistic exchange, such as Grice's (1975) Co-operative Principle with its related Maxims, the speaker in making such an utterance assumes that the hearer is capable of working out this additional information and that they will make an effort to do so. In addition, it is clear that in order to convey just the propositional content of a conditional statement of the form 'if P then Q', there will typically be a number of equally suitable alternative linguistic forms available. The fact that 'if P then Q' is selected might itself add to the information in the utterance: where, for example, such a statement is an expression of the fact that the set of all P-things is subsumed by the set of all Q-things, then the selection of a conditional in preference to a categorical universal such as 'all Ps are Qs' might be seen as communicating that the relationship is not absolute, or is uncertain or fuzzy.

Looking first at the Temporal context, it might be argued that a conditional statement such as 'if the next train is for Wimbledon then the one after it will be for Ealing' would be selected by the speaker in order to

communicate that Wimbledon and Ealing trains are reliably temporally paired in this way: since natural language exchanges are typically required to be relevant as well as reliable and truthful, however, the selection of this particular linguistic form also implies that the hearer is particularly likely to be able to observe the relationship in the immediate future. It also seems clear that this linguistic form would not be selected unless the speaker wished to communicate their belief that neither train was likely to be paired with any other. In terms of the subjective non-contingent likelihood of the consequent, then, it is suggested that the consequent will not typically be seen as likely to become true in the Temporal context other than by virtue of the antecedent becoming true. In other words, it is suggested that the antecedent will typically be seen as (or, at least, treated *as if* it were) necessary and sufficient for the consequent - and a tendency towards an equivalance interpretation might be expected.

In the Causal context, it might be argued that a conditional statement such as 'if the lorry is over the weight limit then the alarm bell will ring' conveys the fact that the alarm bell is there precisely for the purpose of determining whether the lorry is over the weight limit. Should there be other events which might cause the bell to ring, then it is argued that a co-operative speaker would point explicitly to this fact in their utterance, saying instead, for example, 'the bell is set off, *for example,* by ...' or *'one of the ways* in which the bell is set off is ...'. By selecting this particular linguistic form from amongst the various alternatives, it may be argued that the speaker intends to imply that no other similar events will cause the bell to ring - or, at least, that no other similar events are sufficiently likely to occur for it to be relevant to mention them. It is thus suggested that the consequent will not typically be seen as likely to become true in the Causal context other than

by virtue of the antecedent becoming true.

In the Intentional context, it might be argued that the utterance 'if it's fine over the weekend then I'll go fishing' is intended to convey that the speaker intends to perform some future action but that the ability to perform that action is conditional upon the occurrence of the antecedent. Had the speaker intended go fishing at the weekend come hell or high water, then 'I'm going fishing at the weekend' would have sufficed: that is, the specific reference to the weather would only be relevant were it being identified as a precondition. On this basis, it is suggested that in the Intentional context, the consequent will not typically be seen as likely to become true other than by virtue of the antecedent becoming true.

In the Promise context, it is quite clear that it would be ridiculous (or at least singularly unhelpful) to select a conditional statement such as 'if you wash the car then I'll let you borrow it tonight' unless the speaker wished to convey that washing the car was (effectively, at least) the only way in which the loan would be permitted. Since the speaker's purpose in uttering the statement is to induce the hearer to wash the car, the omission of other relevant antecedents from the statement would actually reduce the chances of achieving the goal. The hearer is thus strongly invited to infer that no such relevant alternative antecedents exist.

A similar situation obtains with the Threat context. Given the conditional statement 'if you wear jeans to work again tomorrow then I'll fire you' it is clear that the consequent is somewhat unlikely to occur other than by virtue of the antecedent occurring. Should there be a range of relevant (that is, likely to occur) antecedents which might result in the consequent, then a co-operative speaker would make these explicit: whilst alternative

antecedents must exist (the hearer kneeing the speaker in the groin, for example) the fact that none is mentioned implies that the speaker considers none of them to be sufficiently likely to occur to be worth referring to. In both the Promise and Threat contexts, therefore, it is suggested that the consequent will not typically be seen as likely to become true other than by virtue of the antecedent becoming true.

The position in the Tip and Warning contexts is, however, somewhat different. In the Tip context, a conditional statement such as 'if you buy Acme Industries stock then you'll become a rich man' is clearly a statement of the speaker's belief about one relevant way in which the hearer might become rich. Were it the case that the speaker believed the purchase of Acme Industries stock to be the only way in which the hearer might become rich, then it seems intuitively more reasonable to suppose that they would have selected a linguistic form which made this rather more obvious to the hearer, such as 'the *only* way in which you will become a rich man is ...'.

Similarly, a conditional statement uttered in the Warning context such as 'if you pull the dog's tail again then he'll bite you' is clearly a statement of the speaker's belief about one relevant way in which the hearer might get himself bitten: in addition, it communicates that the dog is disposed to bite people. Were it the case that the speaker believed that pulling the dog's tail was the only action which was likely to result in the hearer being bitten, then they might be expected to have selected a linguistic form which made this rather more obvious to the hearer, such as 'the dog *only ever* bites people when ...'. Unlike the contexts considered so far, then, it is suggested that in the Tip and Warning contexts the consequent will typically be seen as

having the potential to become true anyway, independently of the truth of the antecedent. In other words, it is suggested that the antecedent will typically be seen as (or, at least, treated *as if* it were) sufficient *but not absolutely necessary* for the consequent.

As has already been argued, the selection of a conditional statement in the Universal context (in preference to a categorical universal) does seem strongly to imply that there exists some uncertainty about the co-occurrence of the antecedent and the consequent, such that the speaker implies in selecting a conditional that there are plausible alternative ways in which the consequent might become true other than by virtue of the antecedent becoming true - that is, that the relationship is not absolute, or is uncertain or fuzzy. This seems to suggest that here again, the antecedent will typically be seen as (or, at least, treated *as if* it were) sufficient *but not necessary* for the consequent.

This intuitive analysis of the subjective non-contingent likelihood of the consequent in each of the eight contexts is summarised in Figure 3.2:

Given *'If P then Q':* Q subjectively likely to become true by virtue of something other than P?	
typically MORE so	typically LESS so
	Temporal
	Causal
Tip	Promise
Warning	Threat
Universal	Intentional

Figure 3.2 : The eight contexts in terms of the subjective non-contingent likelihood of the consequent

3.1.2 : Overview of the Four Experiments

It is relevant here to comment on how certain of these contexts relate to the established literature. It is clear that by far the larger part of the literature is concerned with what is termed in this series of experiments the Universal - the use of 'if P then Q' as a paraphrase of 'all P are Q'. Much of the literature concerns investigations where the conditional is used quite clearly in this way (for example, Evans, 1977: 'if the letter is A then the number is 7' is a clear paraphrase of the categorical universal 'all A-pairs are 7-pairs'). The addition of thematic content to the materials does not alter the situation (for example, Wason & Shapiro, 1971: 'if I go to Manchester then I travel by car' is a paraphrase of 'all Manchester-trips are car-trips'); neither does making the relationship between the propositions one with which the subject is

familiar (for example, Johnson-Laird, Legrenzi & Legrenzi, 1972: 'if a letter is sealed then it has a 50 lire stamp on it' is a paraphrase of 'all sealed-letters are 50-lire-letters'). There are exceptions to the (almost) universal Universal - such as Fillenbaum's (1975) work on inducements, which was described in Chapter 1 - but these are few and far between, and tend to have been held up as *anomalies*. It might be argued, however, that it is the Universal which should more properly be considered as anomalous, in so far as its use with abstract materials in a laboratory setting will tend to be far removed from subjects' experience of everyday usage of the conditional.

An even more important observation about the established literature of conditional reasoning is that almost without exception the conditional sentences offered to subjects have been presented *in vacuo* - that is, without any attempt to establish explicitly the linguistic context in which the interpretation should be made[42,43]. As has been seen, a widely accepted notion about discourse is that it is a co-operative enterprise with a mutually shared and quite complex underlying rule system (for example, Grice, 1975), and that many utterances constitute acts of communication which go far beyond the literal meaning of the words uttered (for example, Austin, 1962; Searle, 1969). Certainly, in any natural language situation, the hearer will always have a more or less accurate understanding of the purpose of the speaker in making a statement and of what is and is not likely to be entailed by the speaker choosing to phrase a particular assertion in a particular way.

On this basis, it seems clear that the failure of traditional laboratory

[42] Explicit attempts to establish the context have, however, been made in studies of *disjunctive* reasoning (for example Newstead, Griggs & Chrostowski, 1984).

[43] The most substantial exception to this general position is, of course, the work of Fillenbaum (for example, 1975; 1976; 1977; 1986) on inducements.

reasoning experiments to establish the *context* in which the statements are uttered necessarily must mean that the *purpose* of the utterance is left similarly indeterminate. This is not, however, to say that the statements will necessarily be interpreted as if no purpose existed: since part of the interpretation of everyday conditionals necessarily includes a consideration of the speaker's purpose, it seems reasonable to suppose that subjects will attempt to *infer* a purpose as part of the process of interpreting laboratory conditionals also. Where (as almost always) the conditional statement is being used as a paraphrase of a categorical universal, it is hard to imagine a subject inferring any other purpose than something like 'this is uttered so that the experimenter can find out how I reason with statements about classes of things'.

It is perfectly clear that the conditional is used in natural communication in many other ways than to talk about classes of things - and, indeed, that this use might be a relatively rare one. It is important to realise, however, that even when it is used in this way, the purpose of the utterance is still somewhat removed from that of its counterpart in the literature. When a speaker utters a Universal statement in natural language, it is surely with at least some expectation that the hearer will be able (and willing, and interested ...) to put it to the test in the reasonably near future. The underlying purpose may be of many different flavours: for example, for the speaker to demonstrate expertise along the lines of 'put it to the test and see how right I am', or to assist the hearer along the lines of 'here's a handy way of finding out if P is Q'. Whatever the purpose, however, it is clear that never will it be that suggested above to be implied when a Universal conditional is uttered *in vacuo* in a laboratory reasoning experiment.

It is argued, then, that by forcing subjects to reason with disembodied statements devoid of natural 'purpose', the studies in the literature have failed to address (with any ecological validity) the way that people interpret conditionals in their everyday life. It must, however, be conceded that there are occasions when the purpose is sufficiently clear from the content of the statement for there to be no need for any separate steps to be taken to establish the context explicitly. Certainly, this will generally be true of inducements - Promises, Threats, Tips and Warnings - where the very subject matter of the statement will usually be sufficient to establish that the speaker's purpose in making the utterance is to modify the hearer's behaviour in some way. Additionally, it will (in general) within these four contexts be clear from the content whether the statement is one about contingent future action by and under the control of the speaker (Promises and Threats) or about the speaker's *belief* about some contingent future event outside his control (Tips and Warnings).

As has been seen, Fillenbaum (for example, 1975; 1976) did carry out a certain amount of investigation into reasoning behaviour with conditional statements classifiable as inducements, although this was mainly concerned with Promises and Threats, Tips being addressed only cursorily and Warnings (as defined for this experiment) not at all. In so far as it is comparable with this experiment, Fillenbaum's work suggested two specific predictions: firstly, that Promises, Threats, Tips and Warnings would be far more likely to give rise to an equivalence reading than to an implication reading (and probably more so than the other four contexts); secondly, that Promises and Threats on the one hand, and Tips and Warnings on the other, would give rise to similar response patterns to each other such that they formed 'natural pairs'.

Experiment 5, which used seven of the eight contexts identified earlier[44], involved two tasks - a truth table evaluation task and an inference endorsement task - in the hope that a convergence of the results from the two experiments would lend weight to the validity of each. As the experiment was entirely prospective in nature, the inferences presented for endorsement comprised everything which could follow from the affirmed or negated antecedent - that is, the 'classical' MP (P->Q) and DA (~P->~Q) together with (for the purposes of control) P->~Q and ~P->Q. Whilst the omission of the two 'classical' consequent-to-antecedent inferences, AC (Q->P) and MT (~Q->~P) meant that the underlying truth tables could not be inferred directly, it was considered that the endorsement rates for MP (P->Q) and AC (Q->P) might at least be able to provide support for the differential tendency towards the equivalence or implication interpretation as identified by the truth table evaluation task.

Experiment 6 involved only a truth table evaluation task. In addition to seeking replication of the results of Experiment 5 using a fresh set of materials, the subjective non-contingent likelihood of the consequent was explicitly manipulated in this experiment.

Following Experiments 5 and 6, it was realised that the contexts employed included none in which the speaker had control over the truth of the consequent, but the hearer had no control over the truth of the antecedent. As is clear from Figure 3.1, the inclusion of such a context permits a factorial consideration of the results: the Intentional context, which fulfils these criteria, was accordingly added to the taxonomy.

[44] Experiments 5 and 6 used the Temporal, Causal, Promise, Threat, Tip, Warning and Universal contexts, but not the Intentional context, which was not identified until after those two experiments had been conducted.

By incorporating the Intentional context, Experiment 7, provided an opportunity for a factorial comparison of the data on the dimensions of speaker's and hearer's control. In addition, a memory task was added to the procedure so as to ensure that subjects would pay close attention to the fine details of the manipulation, rather than choosing to ignore the text and to process the conditional statements *in vacuo*.

Experiment 8 used an inference endorsement task similar to that which formed part of Experiment 5, except that the four inferences now offered to subjects were the four 'classical' inferences, (MP, AC, DA and MT), the endorsement pattern of which permits the underlying truth tables to be directly inferred. By using the same materials as were used for Experiment 5 it was, in addition, hoped that Experiment 8 might provide a direct replication of the response patterns on the two inferences which were offered for endorsement in both experiments (MP and DA). More importantly, it was hoped that the endorsement patterns over all four inferences would reveal a distribution of inferred truth tables similar to that of the truth table evaluation tasks of Experiments 5 to 7.

3.2 : Experiment 5

The first in this series of four experiments, Experiment 5 was almost entirely prospective, such predictions as were made being relatively broad and based entirely upon an established literature. The experiment constituted a systematic investigation into the effects of seven particular linguistic contexts (Temporal, Causal, Promise, Threat, Tip, Warning and Universal) upon the interpretation of truth-functional double-affirmative conditional statements of the form 'if P then Q'. Those predictions that were made all concerned the four contexts which might be classified as 'inducements': the Promise, Threat, Tip and Warning contexts. Following Fillenbaum (1975; 1976), it was hypothesised as follows:

H1 the Promise, Threat, Tip and Warning contexts would be more likely to give rise to an equivalence interpretation than to an implication interpretation

H2 the proportion of such interpretations would be higher than in any of the other contexts (Temporal, Causal and Universal)

H3 the response patterns in the Promise and Threat contexts would be extremely similar to each other, as would those in the Tip and Warning contexts, such that these contexts would form respective 'natural pairs'

These hypotheses may be restated in operational terms as follows:

H1 *truth table task:* more E-tables than D-tables on Promise, Threat, Tip & Warning
inference task: DA will be generally endorsed on Promise, Threat, Tip & Warning

H2 *truth table task:* higher E-table:D-table ratio on these contexts than on Temporal, Causal and Universal
inference task: more DA endorsements on Promise, Threat, Tip & Warning than on Temporal, Causal & Universal

H3 *truth table task:* similar E-table:D-table ratios on Tip/Warning and Promise/Threat respectively
inference task: similar DA endorsement rate on Tip/Warning and Promise/Threat respectively

3.2.1 : Design and Methodology

Subjects:

20 1st-year BSc (Hons) Psychology undergraduates at Plymouth Polytechnic (17 female, 3 male) mean age 21.2 years.

Materials:

Subjects performed two separate tasks - a truth table evaluation task and an inference endorsement task. Half the subjects did the truth table task first and half the inference task: there was in every case a delay of between six and eight days between performance of the tasks, which was intended to minimise any possible effects of memory of the materials, since each subject received identical texts in the two tasks.

The experiment was conducted on a BBC-B microcomputer fitted with a Solidisk 32K Sideways RAM which was used to hold the textual materials. Responses were collected by way of purpose-built keypads as illustrated at Appendix [5a] for the truth table evaluation task and Appendix [5b] for the inference endorsement task.

In each trial, there was presented on the computer screen a short passage of

text comprising two sentences. The second sentence in every case ended with a quoted statement reported to have been said by one person to another. The statement was in every case a conditional of the form "If <proposition-1> then <proposition2>". The propositions were in every case phrases of the form <noun-phrase + verb-phrase>, where the verb-phrase was neither negated (as in '... did not press the button') nor of an essentially negative nature (as in '.. failed to press the button'). In traditional conditional reasoning terms, then, the conditional statement was a double-affirmative ('AA').

Each text belonged to one of seven of the contexts described earlier. In order to establish to which linguistic context each of the texts 'belonged', five candidate texts for each context were prepared by the experimenter and offered for rating to two expert judges (both acknowledged experts in the field of linguistic effects in human conditional reasoning). Only when 100% agreement was obtained was a text assigned to a particular category for use in the experiment. The instructions given to the expert judges for these purposes are shown in Appendix [5c].

Corresponding to each double-affirmative conditional statement were four pairs of sentences comprising all possible combinations of the unnegated and explicitly negated first and second propositions (and so if the propositions are denoted by 'P' and 'Q' respectively, the four pairs of outcome sentences correspond to PQ, P~Q, ~PQ and ~P~Q. These sentence pairs thus represented all possible states of affairs which could obtain following the utterance of the statement, and were either paired (to form an 'outcome pair' for the truth table evaluation task) or presented as the second premise and conclusion of an inference (for the inference endorsement task): details of the presentation are given in the relevant sections for the two tasks below.

So as to make the tasks as natural as possible, the outcome sentences were presented in a conversational style so that, for example, the proposition "... the student is doing Economics ..." gave rise to the affirmed outcome "The student was doing Economics" and the negated outcome "The student wasn't doing Economics". The 28 texts used, and the four pairs of outcome sentences associated with each, are presented in full in Appendix [5d].

[a] : Truth Table Evaluation Task

Presentation of Materials

On the truth table evaluation task, subjects indicated whether given pairs of sentences representing possible outcomes supported, contradicted or told them nothing about the quoted conditional statement. Four such outcome pairs were presented in each trial, each of which consisted of either the affirmed or negated outcome sentence relating to the first proposition of the conditional statement followed by ";" then, on a fresh line, either the affirmed or negated outcome sentence relating to the second proposition.

The top-to-bottom order in which the four combinations of outcome pairs were presented was determined at random in each trial, as was the order in which the 28 texts (7 contexts x 4 exemplars) were presented. A typical text, and its attendant outcome pairs, might thus have been presented as shown in Figure 3.3.

Sandy, a staunch Everton supporter, was discussing with his father whether it was safe to wear his team's colours when travelling to the forthcoming game with Arsenal. Sandy's father thought not and told him

> **"If you wear Everton's colours to the match**
> **then you'll be beaten up on the train."**

Sandy did wear Everton's colours to the match;
he was beaten up on the train.

Sandy did wear Everton's colours to the match;
he wasn't beaten up on the train.

Sandy didn't wear Everton's colours to the match;
he was beaten up on the train.

Sandy didn't wear Everton's colours to the match;
he wasn't beaten up on the train.

Figure 3.3 : **Experiment 5:**
Truth table evaluation task:
A typical text and the related outcome pairs
(shown here in the arbitrary order PQ, P~Q, ~PQ, ~P~Q)

The screen also contained instructions regarding the rating of each item - a typical screen for the truth table evaluation task is shown in Appendix [5e]. Subjects pressed buttons on the keypad labelled "SUPPORTS the Statement", "CONTRADICTS the Statement" and "TELLS US NOTHING ABOUT the Statement" so as to rate each of the four outcome pairs and their response (which they could alter at any time during that particular trial) was

recorded on the screen next to the pair in question.

Procedure

Subjects were briefed using text presented on the computer screen, the only verbal instructions being a reinforcement of the request that they should not start the experiment until they were absolutely certain that they understood what they were required to do (subjects were asked to call the experimenter for assistance if they were still in difficulty after seeing the instructions twice: no subject need to do so, however). The instructions given are presented in full in Appendix [5f], but that part of them concerned with the task itself is reproduced below:

> " ... You will notice that the keypad in front of you has twelve grey keys set out in four rows and three columns. Each of the four rows corresponds to one of the states of affairs shown on the screen: the top row refers to the first state of affairs, the next to the second and so on. The three columns are labelled so as to correspond to the three possible responses. If you think, for example, that the first state of affairs SUPPORTS the statement, you should press the key in the top row of the column labelled SUPPORTS. The response which you have made will appear on the screen next to the state of affairs concerned. Do not worry if you press the wrong key or wish to change your mind - simply press the key which you meant to press and the response on the screen will change accordingly.
>
> When you have indicated what you think about each state of affairs, press CONTINUE to go on to the next trial. You cannot go on to the next trial until you have responded to ALL FOUR states of affairs. If you try to do so then the computer will beep and a question mark will flash against the remaining items to remind you ..."

Following the experiment, subjects were debriefed by text presented on the computer screen, as presented in Appendix [5g]. No formal timing was

undertaken, but subjects all took between 20 and 30 minutes to complete the task.

[b] : Inference Endorsement Task

Presentation of Materials

On the inference endorsement task, subjects indicated whether a given inference followed from the quoted statement. Just one of the possible inferences was presented in each of the 112 trials (7 contexts x 4 exemplars x 4 inferences). The inference to be evaluated in each case consisted of either the affirmed or negated outcome sentence relating to the first proposition of the conditional statement followed by "Therefore," followed by either the affirmed or the negated outcome sentence relating to the second proposition. Figure 3.4 shows by way of example the way in which the text given in Figure 3.3 would have been presented for endorsement with its associated DA (~P->~Q) inference.

The screen also contained instructions regarding the rating of each item: a typical screen for the inference endorsement task is shown in Appendix [5h]. Subjects indicated their response by pressing one of the two large buttons on the keypad marked "FOLLOWS" and "DOESN'T FOLLOW" respectively. As before, their response was indicated on the screen and they were free to change it before going on to the next trial.

Sandy, a staunch Everton supporter, was discussing with his father whether it was safe to wear his team's colours when travelling to the forthcoming game with Arsenal. Sandy's father thought not and told him

> **" If you wear Everton's colours to the match**
> **then you'll be beaten up on the train."**

Sandy didn't wear Everton's colours to the match

Therefore,
he wasn't beaten up on the train

Figure 3.4 : Experiment 5: Inference endorsement task:
A typical text and the related DA inference

The order of presentation of the 116 trials was determined at random with the restriction that all 28 texts were presented in some order with some inference type in a consecutive block before any of them could be presented again (although the inference type associated with a given text in each of its four successive exposures was randomly ordered): this was to ensure that two conflicting inferences from the same exemplar would be unlikely to occur consecutively.

Procedure

Subjects were briefed using text presented on the computer screen, the only verbal instructions being a reinforcement of the request that they should not

start the experiment until they were absolutely certain that they understood what they were required to do (subjects were asked to call the experimenter for assistance if they were still in difficulty after seeing the instructions twice: no subject need to do so, however). The instructions given are presented in full in Appendix [5i], but that part of them concerned with the task itself is reproduced below:

> " ... You are required to indicate, using the keypad provided, whether the conclusion FOLLOWS or DOESN'T FOLLOW from the quoted statement and the state of affairs described beneath it. The response which you make will appear on the screen. Do not worry if you press the wrong key or wish to change your mind - simply press the other key and the response on the screen will change accordingly. When you have made your response, press CONTINUE to go on to the next trial. You cannot go on to the next trial until you have made a response. If you try to do so then the computer will beep and a question mark will flash on the screen to remind you ..."

Following the experiment, subjects were debriefed by text presented on the computer screen, as presented in Appendix [5g]. No formal timing was undertaken, but subjects all took between 30 and 50 minutes to complete the task.

3.2.2 : Results

[a] : Truth Table Evaluation Task

The dependent variable on the truth table evaluation task is the decision indicated by a subject via the keypad. The responses available on the keypad were "SUPPORTS the Statement", "CONTRADICTS the Statement" and "TELLS US NOTHING ABOUT the Statement", which will be referred to as

'T', 'F' and '?' respectively.

Context		Tem	Cau	Pro	Thr	Tip	War	Uni	All
outcome [case]									
P & Q	T	100	100	100	100	100	100	99	100
	F	-	-	-	-	-	-	-	-
[TT]	?	-	-	-	-	-	-	1	0
P &~Q	T	-	-	-	-	-	-	-	-
	F	99	99	99	98	98	96	98	98
[TF]	?	1	1	1	3	3	4	3	2
~P & Q	T	-	-	-	-	-	1	-	0
	F	79	48	56	50	26	30	36	46
[FT]	?	21	53	44	50	74	69	64	54
~P &~Q	T	43	60	68	70	51	58	38	55
	F	1	-	2	-	-	1	-	1
[FF]	?	56	40	30	30	49	41	63	44

Table 3.1 : Experiment 5: Truth table evaluation task:
Summary of of raw responses (%)
T = "SUPPORTS the Statement"
F = "CONTRADICTS the Statement"
? = "TELLS US NOTHING ABOUT the Statement"

[N per context = 80 = 20 subjects x 4 exemplars]
[Tem=Temporal, Cau=Causal, Pro=Promise, Thr=Threat,
Tip =Tip, War=Warning, Uni=Universal]
(failures to sum to 100% are due to rounding)

Table 3.1 presents the summary response data, showing the raw frequency of 'T', 'F' and '?' responses over the seven contexts. As may be seen from this

table, there is a massive ceiling effect on responses to the TT and TF cases which are, respectively, seen as supporting and contradicting the statements by virtually all subjects on virtually every trial. The FT and FF cases show very substantial variability in responses, although there are still responses which are reliably avoided in each case: the FT case is virtually never seen as supporting the statement, whilst the FF case is virtually never seen as contradicting the statement.

On the FT case, the mean responses across all contexts are roughly evenly distributed between 'F' and '?'. A 1-way analysis of variance of the frequency of 'F' responses does, however, show a highly significant effect of context ($F = 12.216$; $df = 6$; $p < 0.0001$). Pairwise t-test comparisons of each two adjacently ranked contexts reveal just one significant difference: more 'F' responses to FT on Temporal (79%) than on its nearest neighbour, Promise (56%) ($t = 3.672$; $p < 0.001$; 1-tailed). The rank-ordered frequencies do, however, reveal two distinct clusters in the remaining six contexts, such that there are significantly less 'F' responses to FT on Tip+Warning+Universal pooled (31%) than on Causal+Threat+Promise pooled (51%) ($t = 3.812$; $p < 0.001$; 1-tailed).

The roughly even overall distribution between 'T' and '?' on the FF case is similarly misleading. A 1-way analysis of variance of the frequency of 'T' responses shows a highly significant effect of context ($F = 5.866$; $df = 6$; $p < 0.0001$), although no single pairwise t-test comparison of two adjacently ranked contexts reaches significance. Three distinct clusters are apparent in the rank-ordered contexts: Universal+Temporal (40%), Tip+Warning+Causal (56%) and Promise+Threat (69%). The pairwise difference between Universal+Temporal pooled and Tip+Warning+Causal pooled is significant ($t = 4.458$; $p < 0.001$; 1-tailed), but not that between

Tip+Warning+Causal and Promise+Threat.

That the only effects of context are between 'F' and '?' on the FT case and between 'T' and '?' on the FF case is not, of course, surprising, in that it is only these responses to these two cases which distinguish between the two truth tables typically taken to underly performance on conditional reasoning tasks - E-tables (reflecting a material equivalence interpretation) and D-tables (Wason's "defective" truth table) - as is clear from Figure 3.3. (note that whilst the "logical" I-table is also traditionally associated with such tasks it is very rarely produced[45], and is normally considered more as a standard of logical performance than as a typical pattern of behaviour).

given : *"if P then Q"*

truth case	material implication I-table	material equivalence E-table	defective D-table	X-table	Y-table
TT	T	T	T	T	T
TF	F	F	F	F	F
FT	T	F	?	F	?
FF	T	T	?	?	T

Figure 3.3 : Truth tables:
I-table (material implication)
E-table (material equivalence)
D-table (Wason's "defective implication" table)
X-table
Y-table

Table 3.2 combines the raw response data so as to show the truth tables

[45] This experiment is no exception: as may be seen from Table 3.2, the I-table accounts for only one response pattern out of 560.

which might be inferred to be underlying the responses. As might be seen from Table 3.2, a further two response patterns occurred sufficiently frequently on this experiment to suggest that they might genuinely reflect alternative underlying interpretations of the logical structure of the situation rather than just being random fluctuations in the data. So as not to pre-empt any future interpretations of these two novel truth tables, they will, for the time being, be referred to by the quite arbitrary labels 'X-tables' and 'Y-tables'. Figure 3.3 sets out the response patterns giving rise to each of these five truth tables.

			Response				
Truth case:	TT	T	T	T	T	T	
	TF	F	F	F	F	F	
	FT	T	F	?	F	?	
	FF	T	T	?	?	T	
Truth table:		I	E	D	X	Y	none
Context:	Temporal	-	41	19	38	1	1
	Causal	-	34	27	12	25	1
	Promise	-	39	16	14	27	4
	Threat	-	43	21	6	27	3
	Tip	-	21	43	4	29	3
	Warning	1	25	36	4	30	4
	Universal	-	20	46	12	18	4
	Total %	-	32	30	13	22	3

Table 3.2 : Experiment 5: Truth table evaluation task:
Inferred truth tables x context (%)
[N = 80 per context = 20 subjects x 4 exemplars]
(failures to sum to 100% are due to rounding)

The predominant truth tables are E-tables (material equivalence) and D-

tables (Wason's 1966 "defective implication" truth table). It is interesting to note, however, that whilst predominant overall, these tables account between them for only 62% of the data. A finding such as this might merely reflect massive inconsistencies in the data, such that the remaining 38% of responses simply cannot be classified when combined over all four truth cases. This is clearly not the case here, however, as the remaining data turn out to be almost exclusively made up of two particular response patterns, which have been specified earlier as X-tables and Y-tables. Far from there being massive inconsistencies in the data, then, it is clear that subjects did in fact respond in an extremely consistent fashion, only 3% of the data failing to conform to one of these four truth tables.

Table 3.2 reveals certain substantial effects of context. If we treat as classifiable response patterns just the (traditional) E-tables and D-tables, then it appears that conditional statements in the Temporal, Promise and Threat contexts are far more frequently seen as representing an equivalence relation than a ("defective") implication relation. In the Tip, Warning and Universal contexts, however, the reverse is the case: it appears that conditional statements in these contexts are far more frequently seen as representing a ("defective") implication relation than an equivalence relation. On the basis of this comparison, the Causal context is the 'odd man out', responses being roughly evenly distributed between the two.

When one considers in addition the two novel truth tables - X-tables and Y-tables - however, the effect of context is seen to be rather more complex. Most particularly, X-tables occur very much more often in the Temporal context than in any other. Furthermore, this preponderance of X-tables appears to be at the expense of Y-tables, in that only one Y-table appears on this context: further support for a mutually exclusive relationship between X-tables and Y-

tables is provided by the observation that the sum of the frequencies of the two tables does not vary substantially over the seven contexts.

The analysis presented in Table 3.2 is for all 560 responses - that is, it is collapsed across subjects. Table 3.3 presents a similar analysis in terms of the modal response patterns of individual subjects, using as a criterion for a particular classification within a given context that a subject should show a similar response pattern on at least three of the four exemplars in that context. In addition, Table 3.3 shows the number of subjects modally consistent across all seven contexts, the criterion now being that they should be thus classified on at least six of the seven contexts.

			Response				
Truth case:	TT	T	T	T	T	T	
	TF	F	F	F	F	F	
	FT	T	F	?	F	?	
	FF	T	T	?	?	T	
Truth table:		I	E	D	X	Y	(none)
Context:	Temporal	-	35	10	30	-	25
	Causal	-	30	25	-	25	20
	Promise	-	25	10	5	20	40
	Threat	-	35	20	-	20	25
	Tip	-	20	35	-	20	25
	Warning	-	20	30	-	30	20
	Universal	-	20	40	-	5	35
Within contexts:							
≥3/4 overall		-	27	24	5	17	27
Across contexts:							
≥3/4 on ≥6/7		-	20	15	-	5	n/a

Table 3.3 : Experiment 5:
Truth table evaluation task:
Inferred truth tables x context:
% of subjects modally consistent *within* contexts (≥3/4)
and *across* contexts (≥6/7)

[N per context = 20 subjects]

(failures to sum to 100% are due to rounding)

[b] : Inference Endorsement Task

The dependent variable on the inference endorsement task is the endorsement or rejection of the four inferences across the seven contexts. The raw data are summarised in Table 3.4 which, in order to facilitate later

comparison with the results of Experiment 8, also shows the column rank for each cell for the MP and DA inferences (7=highest, 1=lowest).

Inference:		[MP] (P->Q) % (rank)	[DA] (~P->~Q) % (rank)	(P->~Q) %	(~P->Q) %
Context:	Temporal	96 (4.5)	78 (4)	-	4
	Causal	98 (6)	91 (7)	-	3
	Promise	96 (4.5)	81 (5)	3	6
	Threat	100 (7)	84 (6)	3	6
	Tip	86 (2)	65 (2)	5	4
	Warning	87 (3)	73 (3)	8	8
	Universal	81 (1)	49 (1)	-	3
	overall %	92	74	3	5

Table 3.4 : Experiment 5:
Inferences endorsed x context (%)
[N per context = 80 = 20 S x 4 exemplars]

Looking firstly at the MP inference, a 1-way ANOVA showed a highly significant effect of context ($F = 4.315$; $df = 6$; $p = 0.0006$). Consideration of the ranked pairs of means revealed two distinct clusters, such that MP endorsements on Temporal+Causal+Promise+Threat (98% overall) are significantly greater than those on Tip+Warning+Universal (85% overall) ($t = 2.643$; $p = 0.008$; 1-tailed).

For the DA inference, a 1-way ANOVA showed a highly significant effect of context ($F = 8.492$; $df = 6$; $p < 0.0001$). Consideration of the ranked pairs of means revealed two distinct clusters, such that DA endorsements on Temporal+Causal+Promise+Threat (84% overall) are significantly greater

than those on Tip+Warning+Universal (62% overall) ($t = 3.754$; $df = 19$; $p < 0.001$; 1-tailed).

Additionally, it will be noted that the column ranks are extremely similar for both inferences - that is, there appears to be a consistent overall effect of context upon the acceptability of these two inferences. Comparisons of the endorsement rates for MP and DA revealed significant Spearman's rank correlations in only three of the seven contexts, however: Temporal ($rho = 0.454$; $p = 0.044$), Threat ($rho = 1.0$; $p < 0.001$) and Tip ($rho = 0.479$; $p = 0.033$).

Since the P->~Q and ~P->Q inferences are intuitively nonsensical as inferences (and were included only for the purposes of control) it is not surprising to note that they are rejected almost universally. As was mentioned earlier, the omission of the other two 'classical' inferences, MT (~Q->~P) and AC (Q->P) means that it is not possible to use these raw data to infer the underlying truth tables directly. Nevertheless, the endorsement rates for MP (P->Q) and DA (~P->~Q) might at least be able to offer support for the differential tendency towards the equivalence or implication interpretation as identified by the truth table evaluation task, in so far as MP should be endorsed under either interpretation, whilst DA should only be endorsed under an equivalence interpretation.

If the comparison with the truth table evaluation task is restricted to the traditional D-tables and E-tables, then we should expect that the endorsement of DA should be higher on the Temporal, Promise and Threat contexts (which produced around twice as many E-tables as D-tables on the truth table task) than on the Tip, Warning and Universal contexts (for which the reverse was the case). This is, indeed, borne out by the present data in that

DA endorsement on Temporal+Promise+Threat pooled (81%) is significantly greater than that on Tip+Warning+Universal pooled (62%) ($t = 3.234$; $p = 0.002$; 1-tailed). In that the Causal context fell somewhere between these two clusters of contexts on the truth table task, it might be expected to produce an intermediate level of DA endorsement: this expectation is not, however, borne out by the present data, which show DA endorsement in the Causal context (91%) to be higher even than that on the Temporal, Promise and Threat contexts (81% pooled).

[c] : Order Effects

Half the subjects did the truth table evaluation task first and half the inference endorsement task first, with an intervening period of between six and eight days between the tasks so as to minimise the possibility of carry-over effects. It is thus important to compare the data from each set of subjects so as to establish that the intervening period did indeed eliminate any such effects.

Table 3.5 shows a comparison of the raw responses to the truth table evaluation task in terms of task order: given the ceiling effects of 'T' responses to TT and 'F' responses to TF seen in Table 3.1, this table presents only the FT and FF cases. There is an apparent overall effect of task order on the truth table evaluation task which may be characterised in terms of the determinacy of response: where the truth table task was done first, determinate 'F' responses to FT were in the minority (38%), whereas where that task was done second, they were in the majority (55%) - and a similar observation might be made about 'T' responses to FF, which rose from 47% when the truth table task was done first to 64% when second. Whilst this

appears to be a general effect across all seven contexts, when the individual contexts are compared, however, none of the differences turns out to be statistically significant.

Context		Tem	Cau	Pro	Thr	Tip	War	Uni	All
Truth case									
task order									
FT : ~PQ									
interence	T	-	-	-	-	-	-	-	-
task	F	88	55	68	58	33	38	45	55
first	?	13	45	33	43	68	63	55	45
truth-table	T	-	-	-	-	-	3	-	0
task	F	70	40	45	43	20	23	28	38
first	?	30	60	55	58	80	75	73	61
FF : ~P~Q									
interence	T	50	70	88	80	58	63	40	64
task	F	-	-	-	-	-	3	-	0
first	?	50	30	13	20	43	35	60	36
truth-table	T	35	50	50	63	45	53	35	47
task	F	3	-	5	-	-	-	-	1
first	?	63	50	45	38	55	48	65	52

Table 3.6 : Experiment 5:
Truth table evaluation task:
Raw responses to FT and FF x task order (%):
T = "SUPPORTS the Statement"
F = "CONTRADICTS the Statement"
? = "TELLS US NOTHING ABOUT the Statement"

[N per context = 80 = 20 S x 4 exemplars]

[Tem=Temporal, Cau=Causal, Pro=Promise, Thr=Threat, Tip =Tip, War=Warning, Uni=Universal]

(failures to sum to 100% are due to rounding)

Table 3.6 shows a comparison of the endorsement rates on the inference task in terms of task order: given the very small level of endorsement of the P->~Q and ~P-> Q inferences seen in Table 3.4, this table presents only the MP and DA inferences. As may be seen from Table 3.6, when the inference task was done first, the MP inference (which is logically correct irrespective of the underlying interpretation) is only endorsed 85% of the time, whereas when it was done second, this increases to a ceiling of 99%. Analysis of the individual contexts, however, reveals only two significant effects: on the Tip context (73% vs 100%) ($t = 2.4$; $p = 0.027$; 2-tailed) and on the Universal context (63% vs 100%) ($t = 3.308$; $p = 0.004$; 2-tailed). The effect upon DA inferences was smaller still (78% when the inference task was done first vs 71% when second). Analysis of the individual contexts, however, failed to find a single significant effect.

Context	Tem	Cau	Pro	Thr	Tip	War	Uni	All
Inference *task order*								
MP : P->Q								
inference->truth-table	93	96	96	100	73	78	63	85
truth-table->inference	100	100	98	100	100	98	100	99
DA : ~P->~Q								
inference->truth-table	83	98	90	93	63	78	43	78
truth-table->inference	73	85	73	75	68	68	55	71

Table 3.6 : Experiment 5:
Inference endorsement task:
MP and DA inferences endorsed x context x task order (%)
[N per context = 80 = 10 S x 4 exemplars x 2 task orders]
[Tem=Temporal, Cau=Causal, Pro=Promise, Thr=Threat,
Tip =Tip, War=Warning, Uni=Universal]

The general absence of order effects in this experiment is not, of course, especially surprising, since the delay of six to eight days was intended to secure just that.

3.2.3 : Discussion

[a] : Truth Table Evaluation Task

Whilst intended to be a broadly prospective experiment, three specific predictions were made about responses in the truth table evaluation task in those four contexts which might be classified as 'inducements' (the Promise, Threat, Tip and Warning contexts).

The first prediction, that the Promise, Threat, Tip and Warning contexts would be more likely to give rise to an equivalence reading than an implication reading, was only partially supported. The prediction was strongly supported on the Promise and Threat contexts - and it is, of course with these contexts that the majority of the literature on context effects is concerned. It is clear from Table 3.2 that E-tables do predominate strongly in these two contexts, accounting for some 41% of the response patterns overall, as compared with for D-tables (19% overall). The Tip and Warning contexts, on the other hand, gave rise to a very different pattern of responses from that which had been predicted, having the second and third lowest frequencies of E-tables (23% overall), the D-table of "defective" implication predominating in each case (40% overall).

The second prediction, that each of these four inducement contexts would

give rise to a higher level of equivalence interpretations than the other three contexts, was supported only somewhat weakly. The prediction was strongly supported on the Promise and Threat contexts, where there were considerably more E-tables (41% overall) than in the Universal and Causal contexts (27% overall), although no more than in the Temporal context (41% overall). The position in the Tip and Warning contexts, however, is quite the opposite of that predicted, in that there were very considerably less E-tables in these two contexts (23% overall) than in all except the Universal context (20%).

The third prediction relating to the four inducement contexts was that the response patterns in the Promise and Threat contexts would be extremely similar, as would those in the Tip and Warning contexts, so that these contexts would form 'natural pairs'. With the minor exception of the Promise context giving rise to rather more X-tables than its partner (the level of X-tables on both these contexts being very low anyway), this prediction was strongly upheld for both pairs of contexts: Promise and Threat both produced a very high level of E-tables and a very low level of D-tables, whilst Tip and Warning both produced the reverse.

The tendency towards D-tables on the Tip and Warning contexts is somewhat at odds with Fillenbaum (1975; 1976), whose findings and arguments strongly implied that *all* inducements should produce a tendency towards an equivalence reading - although, as pointed out earlier, Fillenbaum addressed Tips only somewhat cursorily, and Warnings not at all.

The 'mirror-image' of the response patterns on these contexts does, however, appear to lend support to the intuitive analysis presented earlier in terms of

the subjective non-contingent likelihood of the consequent. Where the consequent appears unlikely to become true in virtue of any event other than the antecedent becoming true (as in Promises and Threats), then a strong equivalence interpretation might indeed be expected, since the FT case would be expected to be false. Where, by contrast, there appear to be a range of alternative plausible ways in which the consequent might come to be true quite independently of the antecedent coming to be true (as in Tips and Warnings), then an equivalence interpretation may no longer be so appropriate in that the FT case would no longer necessarily mean that the statement was false (since we can more readily accept that the consequent has become true even though the antecedent has not become true).

The analysis in terms of the subjective non-contingent likelihood of the consequent appears to receive support from the results of the present experiment on the 'non-inducement' contexts also: the Temporal and Causal contexts both gave rise to a high frequency of (equivalence) E-tables and a low frequency of (implication) D-tables, whilst the reverse was the case for the Universal context.

One rather surprising feature of the results of this experiment is the fact that the 'traditional' E-tables and D-tables, whilst predominating, only account between them for some 62% of the data. As has been seen, this cannot be accounted for in terms of massively inconsistent reponses, since the remaining 32% of responses turn out to be made up almost entirely of two particular patterns: if these two patterns might be accorded the status of 'genuine' truth tables[46], then a trifling 3% of the response patterns remain

[46] (with, for the moment, the deliberately uninformative labels 'X-tables' and 'Y-tables')

unclassified.

When these two novel tables are taken into account, the effects of context can be seen to be somewhat richer than appears from a comparison merely of E-tables and D-tables. For example, whereas on that simpler analysis the Temporal context was indistinguishable from the Promise and Threat contexts, it might now be seen to differ considerably in that, unlike any other context, there is a very high level indeed of X-tables (indeed, these are almost as frequent as E-tables): furthermore, there are virtually no Y-tables whatever in the Temporal context, whereas these account for between 18% and 30% of the data in each of the other contexts (26% overall).

It is, of course, possible that the fact that these two response patterns make up virtually all the otherwise unclassified response patterns is nothing more than a freak result: several additional features of the data would appear to argue against this, however. Firstly, there is a strong suggestion of mutual exclusivity between the two patterns, in that where the frequency of X-tables is high (in the Temporal context and, to a lesser extent, in the Universal context), the frequency of Y-tables is low. Secondly, it is inconceivable that the fact that there are almost as many X-tables on the Temporal context as there are E-tables (and considerably more than there are D-tables) would have arisen purely by random inconsistencies in the data. Thirdly, the fact that there are comparatively few X-tables in contexts other than the Temporal context lends considerable weight to their being a genuine reflection of the underlying interpretation in that context. Fourthly, the analysis of the number of subjects modally consistent in their responses within contexts reveals that in all except the Universal context around a quarter of subjects produced either X-tables or Y-tables in at least three of the four exemplars within those contexts.

In the light of these observations, there would appear to be very strong evidence indeed that X-tables and Y-tables do have some 'reality'. Their occurrence does, of course, require to be replicated on further experiments (and with different materials) before they can be seen as genuinely reflecting an underlying interpretation. Furthermore, some account needs to be given as to how such interpretations might arise - and of the circumstances in which they might be useful interpretations to hold.

Some consideration also needs to be given to the very novelty of X-tables and Y-tables: if they genuinely underlie some subjects' interpretations of the conditional, then why have they not been reported in the literature? One study which does report the frequency of various non-traditional response patterns on a truth table task is Johnson-Laird and Tagart (1969), who derived subjects' truth tables for various sentential forms of conditional relations (such as 'if P then Q', 'not P if not Q' and so on), using abstract statements about letters and numbers. Despite the inclusion of virtually all possible response patterns in which the antecedent was true, however, not a single X-table was reported for any of the linguistic forms employed. The Y-table was reported, but accounted for just two of 24 responses - and those were to 'not P if not Q' rather than to 'if P then Q'. Evans (personal communication) argues strongly that the X-table is frequently to be seen (if in small measure) amongst the raw data of conditional reasoning tasks: the problem is, however, that this response pattern has not been accorded the status of a genuine truth table, so that its occurrence would generally have been hidden in the catch-all 'unclassified' category of reported response patterns.

It seems reasonable to suppose that had either of these response patterns

occurred at the substantial level seen in the present experiment, then they would have been reported and discussed in the literature. If they are to be accorded the status of genuine response patterns, therefore, it must be shown that this experiment differs from those reported in the literature in a sufficently gross way that qualitatively different response patterns should not be surprising. Fortunately, such a difference exists: unlike the studies typical of the literature[47] - which, as has been argued, have used the conditional as an expression of a contingent universal - the present experiment involved the explicit manipulation of linguistic context. As may be seen from the typology of contexts presented earlier, the relations between the antecedent and consequent in six of the seven contexts used are qualitatively quite different from the materials traditionally employed - and so it is not, perhaps, too surprising that qualitatively different response patterns have emerged. Whilst the appearance of X-tables and Y-tables in the Universal context is rather at odds with the traditonal literature, it should be noted that the combined frequency of these tables is lower than in any of the other contexts. Furthermore, it is clear from Table 3.3 that only one subject out of twenty consistently produced Y-tables in the Universal context, no subject consistently producing X-tables: this is in stark contrast to all the other contexts, in which between 20% and 30% of subjects consistently produced either X-tables or Y-tables.

From the discussion presented so far, it seems reasonable to conclude that X-tables and Y-tables might be genuine reflections of subjects' underlying interpretations of the relationship between antecedent and consequent in these various linguistic contexts. A consideration of how such interpretations might arise - and what they 'mean' - will be deferred until rather later in this chapter. It may be useful at this point, however, to

[47] (with the exception of Fillenbaum, 1975; 1976 *et seq.*)

consider how they are related to the established truth tables. Evans (1982), in considering the effect of 'defective' truth tables generally on the selection task, refers to the table in which just the FF case is seen as irrelevant (the X-table) as reflecting 'defective equivalence': this is contrasted with the D-table of 'defective implication' (in which both the FT and the FF case are seen as irrelevant). It is appears from Figure 3.3 that the label 'defective equivalence' for the X-table is unobjectionable if a 'defective' table is to be so termed in virtue of its reflecting an incomplete representation of a 'complete' (fully determinate) table upon which it is based: the only fully determinate table upon which the X-table can be based is the E-table of material equivalence, since it is at odds with the I-table of material implication in the FT case (which is 'F' in the X-table but 'T' in the I-table).

What appears not to be so unobjectionable, however, is the use of the term 'defective implication' for the D-table, since the cases about which that table is determinate are consonant with both the E-table and the I-table. It is certainly tempting to consider the D-table to be 'defective implication', since otherwise the I-table of formal logic must be accepted as playing no significant part in determining peoples' behaviour on reasoning tasks. It is argued, however, that there is actually no more reason to suppose that the D-table is a defective I-table than that it is a defective E-table. The appearance of the Y-table is not, of course, able to resolve this issue, since it is indeterminate about the critical FT case and thus, like the D-table, is consonant with both the I-table and the E-table.

[b] : Inference Endorsement Task

The three specific predictions made about performance in the inference endorsement task received similarly mixed support. The first prediction, that DA would be generally endorsed on the four 'inducement' contexts was upheld strongly for Promise and Threat (83% overall), but less so for Tip and Warning (69% overall). It is interesting to note, however, that the rate of MP endorsment on Tip and Warning was also rather lower (87% overall) than the universal endorsement which might have been expected. This may be taken suggest that there is something about these contexts which somehow makes the logical task harder to carry out: this suggestion is considered more fully later in the discussion.

The second prediction, that there would be more DA endorsements on Promise, Threat, Tip and Warning than on Temporal, Causal and Universal, received only relatively weak support. Whilst the DA endorsement rate was indeed higher on Promise and Threat (83% overall) than on Temporal (78%) or Universal (49%), it was actually rather lower than on the Causal context (91%). Rather more violence was done to this prediction by the Tip and Warning contexts, which produced a lower DA endorsement rate (69% overall) than any context except Universal (49%).

The third prediction, that Promise/Threat and Tip/Warning would produce similar DA endorsement rates, was upheld: the two respective pairs are adjacently ranked and there is no significant difference between them.

Before discussing the combined picture which emerges from the two separate tasks, the effect of context upon the endorsement rates for the two

traditional inferences in this task - MP and DA - will be considered further. MP inferences (which are logically and, it has been argued, *trivially*[48], true under the four truth tables) are traditionally almost universally endorsed[49] for double-affirmative rules in 'if...then' - and indeed this is the case in four of the seven contexts employed in the present experiment.

It is interesting to note that the three contexts in which MP was not universally endorsed were precisely those identified in Figure 3.2 as being contexts in which (in everyday usage) there seemed typically to be a number of ways in which the consequent might be expected to become true other than by the antecedent becoming true - that is, in which there is a relatively higher subjective non-contingent likelihood of the consequent.

Even more interesting to note, however, is that the lowest MP endorsement rate of all was to be found on the Universal context, in which MP was rejected 19% of the time. The reason that this is particularly surprising is that it is this context which is closest in nature to the abstract materials employed in the majority of the studies in the literature, in which almost universal endorsement is typically reported. Just why this might have occurred in the present experiment is unclear, although it is possible that the inclusion of the P->~Q and ~P->Q inferences (which have been argued to be intuitively nonsensical) might have led subjects to view the task as somewhat bizarre in the Universal context and thus to be uncertain about what was required of them. In the other six contexts, subjects would have

48 (although only for *truth-functional* conditionals - Johnson-Laird (for example, 1983a) has argued that MP may be far from trivial for non-truth-functional conditionals such as 'if you want a cup of tea, then there's one made')

49 (for example, Taplin, 1971: 92%; Taplin and Staudenmayer, 1973: 99%; Evans, 1977: 100%)

been able to rely to an extent upon their experience of what normally follows and does not follow in those particular contexts outside the reasoning laboratory: it has been argued, however, that the use of a conditional to communicate a categorical universal relationship is somewhat unnatural, so that subjects might not in this context have been able to draw so helpfully upon their experience, and given a forced choice between "FOLLOWS" and "DOESN'T FOLLOW", it might thus be expected that there would be rather more rejections.

It appears that such an explanation in terms of subjects' comparative inexperience with conditionally-phrased universals is wholly consonant with the pragmatic reasoning schemas account (Cheng and Holyoak, 1985; Cheng, Holyoak, Nisbett and Oliver, 1986), which identifies a facilitation in reasoning performance when the subject matter suggests a conditional realationship between antecedent and consequent of a sort with which subjects will typically be used to dealing - such as permissions and obligations. If peoples' conditional reasoning is, indeed, dependent upon such over-learned schemas, then the fact that universals will not typically be expressed in 'if...then' would, presumably, mean that the scenarios presented in the Universal context should fail to elicit a useful schema - and so constitute a much harder task altogether (and one which is qualitatively rather different from that posed by the other six contexts).

Unlike the MP inference, the DA inference should only be endorsed by subjects who interpret the relation between antecedent and consequent as one of material equivalence. That it is endorsed rather less in the Tip, Warning and Universal contexts is, therefore, consonant with the fact that it was in these contexts in the truth table evaluation task that the lowest

frequency of E-tables was found.

[c] : Overall Discussion of Experiment 5

This experiment employed two distinct types of task in the hope that a convergence of the results might lend weight to the validity of each. The major findings of the truth table task may be summarised as follows:

1 the Temporal, Causal, Promise and Threat contexts produced a greater tendency to an equivalence interpretation (39% E-tables overall) than did the other contexts (22% E-tables overall)
2 the Promise/Threat and Tip/Warning contexts produced very similar results, so forming 'natural pairs'
3 two novel response patterns were observed sufficiently often to be considered as if they genuinely reflected subjects' underlying interpretations: X-tables were particularly strongly present in the Temporal context (38%), whereas Y-tables were consistently present in all contexts except Temporal - but rather less so in the Universal context (18% as compared with 28% for the other five contexts)

The major findings of the inference task may be summarised as follows:

1 the Temporal, Causal, Promise and Threat contexts produced almost universal endorsement of the MP inference (98% overall), the other contexts less so (85% overall)
2 the Temporal, Causal, Promise and Threat contexts produced a high level of endorsement of the DA inference (84% overall), the Threat and Warning contexts a rather lower rate (69% overall) and the Universal context an especially low rate (49%)
3 the Promise/Threat and Tip/Warning contexts produced very similar results, so forming 'natural pairs'

It is clear that the results from the two tasks do converge very considerably.

Firstly, the high level of DA endorsement in the Temporal, Causal, Promise and Threat contexts is consistent with the high level of E-tables in those contexts. Secondly, the similarity of the findings on the Promise/Threat and Tip/Warning contexts provides strong support for the notion that these contexts should be seen as constituting 'natural pairs'.

Before the comparison of the two tasks can be extended to the 'novel' X-tables and Y-tables, it is necessary to consider what inferences those two tables might be expected to support. Whilst the AC (Q->P) and MT (~Q->~P) inferences were not used in this experiment, it may be useful to the discussion of later experiments to include those inferences in this consideration. It was stated in Chapter 1 that for the purposes of this thesis an inference would be taken to be logically and necessarily valid *if and only if* the counterexample to its conclusion is expressly forbidden: that is, a truth table need not expressly permit the conclusion itself. As was seen in that chapter, all the traditional truth tables have the TT case as 'T' and the TF case as 'F', so that MP and MT are always valid. The validity of the DA and AC inferences depends critically upon the FT case, which is 'F' for the I-table (so that DA and AC are invalid), 'T' for the E-table (so that DA and AC are valid) and '?' for the D-table (so that DA and AC may be argued to be indeterminate). X-tables have the FT case as 'F', so expressly forbidding the DA and AC inferences. Y-tables, on the other hand, have the FT case as '?', so that DA and AC are indeterminate. The relationship between the five truth tables and the four classical inferences is summarised in Figure 3.4. This figure also ascribes Evans' (1982) label 'defective equivalence' to the X-table and distinguishes the D-table and Y-table by labelling the former 'wholly defective'[50] and the latter 'partially defective'.

[50] (note that it was argued earlier that there is no necessary reason to suppose that the D-table is a 'defective' version of the I-table as is implied by the

	material implication I-table	material equivalence E-table	wholly defective D-table	defective equivalence X-table	partially defective Y-table
truth case:					
TT	T	T	T	T	T
TF	F	F	F	F	F
FT	T	F	?	F	?
FF	T	T	?	?	T
inference:					
P->Q MP	√	√	√	√	√
~Q->~P MT	√	√	√	√	√
~P->~Q DA	x	√	?	√	?
Q->P AC	x	√	?	√	?

Figure 3.4 : Truth tables and the inferences which they support:
I-table (material implication)
E-table (material equivalence)
D-table ('wholly defective')
X-table ('defective equivalence')
Y-table ('partially defective')
[√ = necessarily valid, ? = indeterminate, x = necessarily invalid]

It is clear from Figure 3.4 that the inferences supported by the X-table are precisely those supported by the E-table - and that the Y-table is similarly allied to the D-table. This is, perhaps irksome to any proponent of the reality of X-tables and Y-tables, in so far as they are not distinguished from the traditional tables in terms of the inferences which they support. Nevertheless, this is an important observation in that it may well indicate that subjects' tendency to an equivalence interpretation might often have

label 'defective implication', since it is just as consonant with the E-table as it is with the I-table).

been underestimated in the literature in as much as X-tables might have been relegated to the catch-all 'unclassified' category and thus not taken further into account.

With this in mind, it is relevant to reconsider the results from this experiment with X-tables being lumped together with E-tables to produce a composite indicator of an equivalence interpretation. On the truth table task, the Temporal, Causal, Promise and Threat contexts continue to show a far stronger tendency to equivalence (57% E-tables/X-tables overall) than do the Tip, Warning and Universal contexts (29% E-tables/X-tables overall). In so far as these four contexts continue to demonstrate a higher tendency to equivalence, these results remain consistent with those on the inference task.

There is an indication, however, that to view X-tables as reflecting the same sort of interpretation as E-tables might be somewhat over-simplistic, since the very large proportion of X-tables in the Temporal context would, on such a basis, suggest that the Temporal context is interpreted rather more often as equivalence (79% E-tables/X-tables) than its erstwhile bedfellows, the Causal, Promise and Threat contexts (49% E-tables/X-tables overall).

Any more detailed discussion of the nature of X-tables and Y-tables might be somewhat premature in the absence of a replication with different subjects and/or different materials, and thus will be deferred until later in this chapter. One final observation might be made before moving on to Experiment 6, however. It was observed earlier that X-tables and Y-tables appear to be mutually exclusive, in so far as the proportion of these two tables *combined* varies very little over the seven contexts, as is clear from Table 3.2. This appears to lend support to the notion that X-tables are somehow allied

with E-tables, and Y-tables with D-tables, since it is clear from Table 3.2 that E-tables and D-tables exist in a similar exclusive relationship - which is unsurprising in as much as E-tables are taken to reflect an equivalence interpretation and D-tables a (necessarily logically exclusive) implication interpretation.

3.3 : Experiment 6

Experiment 6 employed a truth table evaluation task similar to that used in Experiment 5, although with a fresh set of materials. In addition to manipulating the context in which the conditional was reported as having been uttered, a further manipulation was introduced in terms of the subjective non-contingent likelihood of the consequent - that is, how likely it seemed to be that the consequent would become true other than by virtue of the antecedent becoming true.

It was hypothesised that when the consequent appeared comparatively unlikely to become true other than in virtue of the antecedent becoming true[51] there would be more responses in line with equivalence, and that this effect would be independent of context. Since it would have been premature to speculate about the precise nature of X-tables and Y-tables, no specific prediction was made about the effect on the frequency of these tables. It was, however, predicted that despite the fresh set of materials, the patterns of truth tables emerging in Experiment 5 would be broadly replicated in this experiment although, clearly, if the manipulation of the subjective non-contingent likelihood of the consequent should turn out to be effective, then these patterns might be somewhat altered. These general hypotheses might be summarised as follows:

H1 where the scenario indicated that the consequent was relatively less likely to become true irrespective of whether the antecedent became true then this would lead to a greater tendency to an equivalence interpretation than when the reverse was indicated

H2 the relative proportions of the four truth tables overall would

[51] (that is, when there was a low subjective non-contingent likelihood of the consequent)

be broadly similar to that found on Experiment 5

These hypotheses may be restated in operational terms as follows:

H1	there would be more E-tables when the non-contingent likelihood of the consequent was indicated to be low
H2a	there would be more E-tables than D-tables in the Temporal, Causal, Promise and Threat contexts overall
H2b	there would be less E-tables than D-tables in the Tip, Warning and Universal contexts overall
H2c	response patterns in the Promise/Threat and Tip/Warning contexts would be similar so forming 'natural pairs'
H2d	there would be more X-tables in the Temporal context overall than in any other context
H2e	the combined proportion of X-tables and Y-tables would be relatively constant over all seven contexts

3.3.1 : Design and Methodology

Subjects:

24 social science undergraduates at Plymouth Polytechnic (12 female, 12 male), mean age 20.9 years.

Materials:

The experiment was conducted on a BBC-B microcomputer fitted with a Solidisk 32K Sideways RAM which was used to hold the textual materials. Responses were collected by way of the purpose-built keypad illustrated at Appendix [5a].

In each trial, there was presented on the computer screen a short passage of text which in every case ended with a quoted statement constituting a double-affirmative conditional statement, constructed as specified for Experiment 5. Each text belonged to one of the seven contexts used in Experiment 5, although the particular materials used were prepared afresh for the present Experiment. In preparing these fresh materials, the specifications for each context set out earlier were followed carefully, although there was one minor departure in the Causal context. In three of the four exemplars within the Causal context, the antecedent was of the form "If I <perform some action>": that is, it referred to a possible future act by the speaker. One of the exemplars in question, for example, had the speaker demonstrating the operation of a safety system, and involved the following conditional statement:

> **" If I jam the coolant valve open**
> **then the warning siren will sound."**

The mechanism of the connection between the antecedent and the consequent remained unchanged, in that there was still no human intervention involved once the antecedent had become true. What this alteration does mean, however, is that the analysis in terms of control has now changed for the three exemplars in question, since the speaker is now in control of the antecedent, whereas in the remaining Causal exemplars - and all those used in Experiment 5 - neither the speaker nor the hearer had control over the truth of the antecedent. In considering the effects of control, the question of interest with respect to the hearer's control over the truth of the antecedent is whether or not the hearer can choose to test the conditional statement: this much is unchanged, however, as the hearer can still only observe whether or not the antecedent comes to be true. The question of interest with respect to the speaker's control over the truth of the consequent

is whether the speaker can thereby control the truth of the whole conditional: this is also unchanged, since the speaker still has no direct control over what happens once the performance of the action renders the antecedent true. On these grounds, it might be argued that although an analysis in terms of the control of the parties over the various propositions is now altered for three of the Causal exemplars, it is not altered in any material way.

Each context had a total of four 'stories' associated with it. There were two separate versions of each story, corresponding to a dimension of subjective non-contingent likelihood of the consequent. One version (the 'lo') was designed to evoke a situation wherein it was subjectively unlikely that the consequent would become true other than by the antecedent becoming true. The other version (the 'hi') was designed to evoke a situation wherein it was subjectively likely that the consequent would become true anyway, even should the antecedent not become true. The double-affirmative conditional statement with which the passage ended was the same for each of the two versions. There was thus a total of 56 texts (7 contexts x 4 exemplars x 2 versions).

Candidates for each of the 56 texts were prepared by the experimenter, offered for rating to two expert judges and modified where necessary to obtain 100% agreement on the context and version to which a text 'belonged'. Clearly, there could be no objective criterion upon which a text could unquestionably be assigned the label 'lo' or 'hi' as defined above. The criterion for this judgement was thus that the two versions of a particular 'story' be quite clearly distinct on this dimension.

Corresponding to each double-affirmative conditional statement were four pairs of sentences comprising all possible permutations of the unnegated and explicitly negated first and second propositions. These sentence pairs represented all possible states of affairs which could obtain following the utterance of the statement, and were paired so as to form an 'outcome pair'. As for Experiment 5, the outcome sentences were presented in a conversational style so as to make the task as natural as possible. The 56 texts used, and the four pairs of outcome sentences associated with each, are presented in full in Appendix [6].

Of the four exemplars in each context, each subject received two lo versions and two hi versions. In order to ensure that each version was presented an equal number of times across all subjects, subjects were 'yoked'. This meant that for each odd-numbered subject the decision about which version of each story would be presented was determined at random, subject only to the restriction that two of the four exemplars in each context would be in the hi version and two in the lo. The immediately following even-numbered subject received the opposite version of each story to that received by the preceding odd-numbered subject. The order in which the 28 stories were presented was determined at random for each subject independently.

Subjects indicated whether given pairs of sentences representing possible outcomes supported, contradicted or told them nothing about the quoted conditional statement. Four such outcome pairs were presented in each trial, each of which consisted of either the affirmed or negated outcome sentence relating to the first proposition of the conditional statement followed by ";" and either the affirmed or negated outcome sentence relating to the second proposition: the four pairs thus corresponded to PQ, P~Q, ~PQ and ~P~Q. The top-to-bottom order in which the four outcome pairs were presented was

determined at random in each trial. Figure 3.5 shows the lo version of a typical text, and its attendant outcome pairs (the example given in that Figure is the lo version of one of the Threat exemplars with the outcome pairs in the order PQ, ~P~Q, ~PQ, P~Q) - the alternative text used in the corresponding hi version is shown in Figure 3.6.

Little Billy's mother strongly disapproved of parents smacking their children for doing something wrong. On one particular day, when his mother was feeling particularly fragile after a heavy evening out, Billy seemed to be in an especially naughty mood. When he turned the radio up full blast, his mother yelled at him

" If you turn that radio up again
then I shall smack you."

Billy did turn the radio up again;
his mother did smack him.

Billy didn't turn the radio up again;
his mother didn't smack him.

Billy didn't turn the radio up again;
his mother did smack him.

Billy did turn the radio up again;
his mother didn't smack him.

Figure 3.5 : Experiment 6:
Truth table evaluation task:
A typical 'lo' text and the related outcome pairs

Little Billy's mother strongly believed that children should be smacked for the merest naughtiness. On one particular day, when his mother was feeling particularly fragile after a heavy evening out, Billy seemed to be in an especially naughty mood. When he turned the radio up full blast, his mother yelled at him ...

Figure 3.6 : Experiment 6:
Truth table evaluation task:
The 'hi' version of the text shown in Figure 3.5

The screen also contained instructions regarding the rating of each item as illustrated for Experiment 5 in Appendix [5e]. Subjects pressed buttons on the keypad labelled "SUPPORTS the Statement", "CONTRADICTS the Statement" and "TELLS US NOTHING ABOUT the Statement" so as to rate each of the four outcome pairs, and their response (which they could alter at any time during that trial) was recorded on the screen next to the pair in question.

Procedure:

Subjects were briefed using text presented on the computer screen, the only verbal instructions being a reinforcement of the request that they should not start the experiment until they were absolutely certain that they understood what they were required to do (subjects were asked to call the experimenter for assistance if they were still in difficulty after seeing the instructions twice: no subject needed to do so, however). The instructions given are identical to those given for the truth table evaluation task of Experiment 5,

which are shown in full in Appendix [5f]. Following the experiment, subjects were debriefed by text presented on the computer screen, as shown in Appendix [5g]. No formal timing was undertaken, but subjects all took between 30 and 50 minutes to complete the task.

3.3.2 : Results

The dependent variable is the decision indicated by a subject via the keypad. As before, the responses available on the keypad were "SUPPORTS the Statement", "CONTRADICTS the Statement" and "TELLS US NOTHING ABOUT the Statement", which will be referred to as 'T', 'F' and '?' respectively. Table 3.7 presents the summary response data, showing the raw frequency of 'T', 'F' and '?' responses over the seven contexts, with the lo and hi versions collapsed to facilitate comparison with the relevant data from from Experiment 5 (as were shown in Table 3.1).

The overall patterns of raw responses are very similar to those found in Experiment 5. There is a massive ceiling effect on responses to the TT and TF cases which are, respectively, seen as supporting and contradicting the statements by virtually all subjects on virtually every trial. The FT and FF cases again show very substantial variability in responses, although one particular response is reliably avoided in each case - the FT case is almost never seen as supporting the statement, whilst the FF case is almost never seen as contradicting the statement.

Context		Tem	Cau	Pro	Thr	Tip	War	Uni	All
outcome [case]									
P & Q	T	99	99	100	99	100	96	100	99
	F	1	1	-	1	-	2	-	1
[TT]	?	-	-	-	-	-	2	-	0
P &~Q	T	1	-	3	1	-	1	1	1
	F	97	100	96	93	99	98	95	97
[TF]	?	2	-	1	6	1	1	4	2
~P & Q	T	-	1	-	2	-	2	-	1
	F	50	40	67	51	36	29	46	46
[FT]	?	50	59	33	47	64	69	54	54
~P &~Q	T	31	41	57	58	41	43	40	44
	F	2	1	2	1	2	1	1	1
[FF]	?	67	58	41	41	57	56	59	54

Table 3.7 : Experiment 6:
Summary of raw responses [lo+hi pooled] (%):
T = "SUPPORTS the Statement"
F = "CONTRADICTS the Statement"
? = "TELLS US NOTHING ABOUT the Statement"
[N per context = 96 = 24 subjects x hi/lo x 2 exemplars]
[Tem=Temporal, Cau=Causal, Pro=Promise, Thr=Threat,
Tip =Tip, War=Warning, Uni=Universal]
(failures to sum to 100% are due to rounding)

Table 3.8 re-presents the raw responses to the FT and FF cases shown in Table 3.7, further analysed to show the hi and lo versions of the task (given the massive ceiling effects in the TT and TF cases, there is nothing to be gained by analysing these cases in a similar way).

Context		Tem		Cau		Pro		Thr		Tip		War		Uni		All	
outcome [case]		lo	hi	lo	hi	lo	hi	lo	hi	lo	hi	lo	hi	lo	hi	lo	hi
~P & Q	T	-	-	-	2	-	-	4	-	-	-	2	2	-	-	1	1
	F	58	42	48	31	65	69	50	52	35	38	31	27	48	44	48	43
[FT]	?	42	58	52	67	35	31	46	48	65	63	67	71	52	56	51	56
~P &~Q	T	33	30	48	33	58	56	56	60	42	40	40	46	40	40	45	43
	F	-	4	2	-	2	2	2	-	-	4	-	2	-	2	1	2
[FF]	?	67	67	50	67	40	42	42	40	58	56	60	52	60	58	54	54

Table 3.8 : **Experiment 6:**
Summary of raw responses by version (lo/hi) (%):
T = "SUPPORTS the Statement"
F = "CONTRADICTS the Statement"
? = "TELLS US NOTHING ABOUT the Statement"

[N per context = 48 = 24 subjects x 2 exemplars]

[Tem=Temporal, Cau=Causal, Pro=Promise, Thr=Threat,
Tip =Tip, War=Warning, Uni=Universal]

(failures to sum to 100% are due to rounding)

Comparisons of the responses to the hi and lo versions in each context revealed just three significant effects of this manipulation. In the Temporal context, there were significantly more 'F' responses to FT in the lo version (58%) than in the hi (42%) ($t = 2.145$; $p = 0.043$; 2-tailed); in the Causal context, there were significantly more 'F' responses to FT in the lo version (48%) than in the hi (31%) ($t = 2.326$; $p = 0.029$; 2-tailed) and significantly more 'T' responses to FF in the lo version (48%) than in the hi (33%) ($t = 2.29$; $p = 0.032$; 2-tailed). Essentially, this indicates that the manipulation had a significant effect upon responses in just two contexts out of seven (and only

slightly so in the Temporal context): in these circumstances, the remainder of the analysis will carried out upon the pooled data from both the lo and hi versions of the task (as shown in Table 3.7).

On the FT case, the mean responses across all contexts are roughly evenly distributed between 'F' and '?'. A 1-way analysis of variance of the frequency of 'F' responses does, however, show a highly significant effect of context ($F = 8.651$; $df = 6$; $p < 0.0001$). Pairwise t-test comparisons of each two adjacently ranked contexts reveal two significant differences. There are significantly more 'F' responses to FT on Promise (67%) than on its nearest neighbour, Threat (51%) ($t = 2.46$; $p = 0.011$; 1-tailed); and there are (just) significantly more 'F' responses to FT on Tip (36%) than on its nearest neighbour, Warning (29%) ($t = 1.772$; $p = 0.045$; 1-tailed). The rank-ordered frequencies reveal four clusters in the seven contexts, such that there are significantly less 'F' responses to FT on Warning (29%) than on Tip+Causal pooled (38% overall) ($t = 5.658$; $p < 0.001$; 1-tailed); significantly less 'F' responses to FT on Tip+Causal pooled (38% overall) than on Universal+Temporal+Threat pooled (49% overall) ($t = 6.613$; $p < 0.001$; 1-tailed); and significantly less 'F' responses to FT on Universal+Temporal+Threat pooled (49% overall) than on Promise (67%) ($t = 6.665$; $p < 0.001$; 1-tailed).

On the FF case, the mean responses across all contexts are roughly evenly distributed between 'T' and '?'. A 1-way analysis of variance of the frequency of 'T' responses shows a highly significant effect of context ($F = 5.456$; $df = 6$; $p < 0.0001$). Whilst pairwise t-test comparisons of each two adjacently ranked contexts fail to reveal a single significant difference, the rank-ordered frequencies do reveal three distinct clusters in the seven contexts, such that there are significantly less 'T' responses to FF on

Temporal (31%) than on Universal+Causal+Tip+Warning pooled (41% overall) ($t = 5.555$; $p < 0.001$; 1-tailed) and significantly less 'T' responses to FF on Universal+Causal+Tip+Warning pooled (41% overall) than on Promise+Threat pooled (58% overall) ($t = 2.457$; $p = 0.011$; 1-tailed).

Table 3.9 presents the inferrable truth tables, with the lo and hi versions collapsed to facilitate comparison with the relevant data from Experiment 5 (as were shown in Table 3.2).

			Response				
Truth case:	TT	T	T	T	T	T	
	TF	F	F	F	F	F	
	FT	T	F	?	F	?	
	FF	T	T	?	?	T	
Truth table:		I	E	D	X	Y	none
Context:	Temporal	-	28	44	19	3	6
	Causal	-	26	45	11	15	3
	Promise	-	49	27	13	5	6
	Threat	-	40	32	4	15	9
	Tip	-	27	50	6	14	3
	Warning	-	23	51	3	17	6
	Universal	-	28	44	15	8	5
	Total %	-	32	42	10	11	6

Table 3.9 : Experiment 6:
Truth table evaluation task:
Inferred truth tables x context [lo+hi pooled] (%)
[N = 96 per context = 24 subjects x 4 exemplars]
(failures to sum to 100% are due to rounding)

Table 3.10 re-presents the truth table analysis shown in Table 3.9, further

analysed to show the hi and lo versions of the task.

			Response									
Truth case:	**TT**	**T**	**T**		**T**		**T**		**T**			
	TF	**F**	**F**		**F**		**F**		**F**			
	FT	**T**	**F**		**?**		**F**		**?**			
	FF	**T**	**T**		**?**		**?**		**T**			
Truth table:		**I**	**E**		**D**		**X**		**Y**		**none**	
			lo	**hi**	**lo**	**hi**	**lo**	**hi**	**lo**	**hi**	**lo**	**hi**
Context:	**Temporal**	-	33	23	38	50	25	13	-	6	4	8
	Causal	-	31	21	35	54	15	8	17	13	2	4
	Promise	-	48	50	27	27	13	13	6	4	6	6
	Threat	-	40	40	33	31	2	6	13	17	13	6
	Tip	-	29	25	52	48	4	8	13	15	2	4
	Warning	-	21	25	52	50	4	2	13	21	10	2
	Universal	-	29	27	42	46	17	13	8	8	4	6
	Total lo	-	33		40		11		10		6	
	Total hi	-		30		44		9		12		5

Table 3.10 : **Experiment 6:**
Truth table evaluation task:
Inferred truth tables x context x version (lo/hi) (%)
[N = 48 per context = 24 subjects x 2 exemplars]
(failures to sum to 100% are due to rounding)

It is clear from Table 3.10 that the lo/hi manipulation produced no substantial overall effect in any of the seven contexts. In these circumstances, further analysis of the data will thus be based upon the collapsed data (lo+hi), as presented in summary in Table 3.9.

As for Experiment 5, the predominant truth tables are E-tables and D-tables. Whilst the overall proportion of E-tables (32%) is identical to that in Experiment 5, the overall proportion of D-tables is, however, substantially increased (from 30% in Experiment 5 to 42% in the present experiment).

These predominant tables account between them for 74% of the data: of the remaining 26% of response patterns, only 21% are accounted for by X-tables and Y-tables, in contrast to Experiment 5 in which they accounted for 35% of the data: nevertheless, these four tables continue between them to account for practically all the data, only 6% of response patterns remaining unclassified. Since the proportion of E-tables is unchanged, it is clear that the reduction in X-tables and Y-tables is reflected directly in the increase in D-tables.

Table 3.9 reveals certain substantial effects of context. Looking first just at E-tables and D-tables, it is only in the Promise and Threat contexts that there is a preference for E-tables (although that in the Threat context is rather weaker). In each of the other five contexts there is a substantial preference for D-tables. These results differ from those of Experiment 5 in two major respects: firstly, there is now a somewhat looser correspondence between Promise and Threat, which produced very similar 'E'/'D' preferences in Experiment 5; secondly, whereas no strong preference was shown in the Causal context in Experiment 5, there is now a fairly substantial 'D'/'E' preference.

When X-tables and Y-tables are considered in addition, further differences between the two experiments are revealed. Whilst it is still the case that more X-tables occur in the Temporal context than in any other context, these now account for a rather smaller proportion of all responses in that context (19% as compared with 30% in Experiment 5). A similar observation might be made about the frequency of Y-tables on the remaining six contexts: in Experiment 5 this was of a similar order to that of E-tables and D-tables, whereas in the present experiment E-tables and D-tables predominate

throughout. One important feature of the data does remain, however: there does still appear to be an element of mutual exclusivity between X-tables and Y-tables, higher frequencies of X-tables going with lower frequencies of Y-tables, and *vice versa*.

The analysis presented in Table 3.9 is for all 672 responses - that is, it is collapsed across subjects. Table 3.11 presents an analysis in terms of the modal response patterns of individual subjects, using as a criterion for a particular classification within a given context that a subject should show a similar response pattern on at least three of the four exemplars in that context. In addition, Table 3.11 shows the proportion of subjects modally consistent across all seven contexts, the criterion now being that they should be thus classified on at least six of the seven contexts.

The main observation to be made upon a comparison the analysis in Table 3.11 with that in Table 3.3 for Experiment 5 is the substantial reduction in the number of subjects consistently producing X-tables and Y-tables - and, as would be expected in view of the analyses presented earlier, this is reflected almost exclusively in the increase in modally consistent D-table responses.

			Response				
Truth case:	TT	T	T	T	T	T	
	TF	F	F	F	F	F	
	FT	T	F	?	F	?	
	FF	T	T	?	?	T	
Truth table:		I	E	D	X	Y	(none)
Context:	Temporal	-	25	29	4	-	42
	Causal	-	13	42	4	8	33
	Promise	-	42	26	4	-	29
	Threat	-	29	25	-	13	33
	Tip	-	17	42	-	4	38
	Warning	-	17	54	-	8	21
	Universal	-	17	25	-	-	58
Within contexts: ≥3/4 overall		-	23	35	2	5	36
Across contexts: ≥3/4 on ≥6/7		-	24	25	-	-	n/a

Table 3.11 : Experiment 6:
Truth table evaluation task:
Inferred truth tables x context:
% of subjects modally consistent *within* contexts (≥3/4)
and *across* contexts (≥6/7)

[N per context = 24 subjects]

(failures to sum to 100% are due to rounding)

3.3.3 : Discussion

It was somewhat surprising to find that the manipulation of the subjective non-contingent likelihood of the consequent produced no significant effect upon the response patterns either overall or within any but the Causal

context, the frequencies of the various truth tables being almost identical between the lo and hi conditions. As might be seen by reference to the materials presented in Appendix [6a], the manipulation was relatively gross, such that subjects were left in little doubt about that likelihood: indeed, in designing the materials, it had been feared that so gross was the manipulation that there was a danger of a very strong demand effect, in that it might become quite clear to subjects that they were being pushed towards, or away from, an equivalence interpretation.

One possible explanation for the lack of effect was that subjects might have failed to notice the details in the texts with which the manipulation was effected. There are several possible ways in which this might have come about. In constructing the materials, it became apparent that certain contexts 'belonged' very strongly to either the lo or hi condition, such that to use a conditional statement in whichever of the conditions was 'unnatural' seemed anomalous: this was particularly true in the case of the Promise and Threat contexts, where the conditional seemed to belong naturally to the lo condition - that is, the condition in which it seems subjectively unlikely that there exist any alternative antecedents. The effect of this was to make the construction of the texts extremely difficult if the statement was to sound natural in the context of the preceding text, and in many cases naturalness was only achieved by rather careful - and often somewhat longwinded - wording of the texts. As a result, many of the texts were considerably longer than their counterparts in Experiment 5, such that subjects might be excused for skipping through them somewhat cursorily, thus perhaps missing the critical elements of the manipulation. Another related potential reason for missing the details is that the overall 'flavour' of the texts (that is, whether they were about a Threat, a Temporal relationship or whatever) was clear

from even a cursory reading of the text: once subjects had established the sort of situation in which the conditional was uttered then it may be that they would not be inclined to read further, and thus would respond on the basis of the perceived context rather than the specific details of the text.

Support for the notion that subjects did not read the texts thoroughly was provided by comments made by subjects following the experiment, several volunteering without any enquiry on the part of the experimenter that they did not bother to read the texts, but rather responded on the basis of the conditional statement alone. It is true that in several of the contexts (notably the Promise and Threat contexts) the 'flavour' of the context might easily be divined from the statement alone: it certainly cannot be the case for the majority of subjects that they ignored the texts, however, given the strong overall effects of context over all seven contexts[52].

The manipulation of the subjective non-contingent likelihood of the consequent in this experiment was directly inspired by consideration of the apparent 'natural pairing' of the Threat/Promise and Tip/Warning contexts. The failure of the present data to support this explanation means that it is incumbent upon us to consider whether there is any alternative essential feature of these four contexts such that the observed pairing and opposition might be expected. As was seen in the introduction to this chapter, another potential explanation is the fact that whilst in the Promise and Threat

[52] In Experiment 7, as will be seen, an attempt was made to make it more likely that subjects would read the details of the manipulation by warning them that they would receive a memory test following the truth table evaluation task. In addition, by framing the questions such that they asked incisively about just those details by which the subjective non-contingent likelihood of the consequent had been manipulated, an independent measure of how carefully subjects had read that particular aspect of the texts was provided.

contexts the speaker has control over the truth of the consequent (and thus over the truth of the conditional), they have no such control in the Tip and Warning contexts.

The explanation falls down, however, when it is realised that the speaker does not have control over the truth of the consequent in the other three contexts either: whilst in Experiment 6 those other three contexts were indeed indistinguishable from the Tip and Warning contexts in terms of the frequencies of E-tables, they were most certainly not so in Experiment 5, where the frequency of E-tables in the Temporal and Causal contexts, at least, bore a far closer resemblance to those in the Promise and Threat contexts. It is necessary, therefore, to find some additional (but preferably related) feature which will serve to distinguish the four 'inducement' contexts from the other three. As was seen earlier in this chapter, a further distinction might indeed be made on the basis of control - but this time in terms of the hearer's control over the truth of the antecedent: in the four 'inducement' contexts, the hearer has control over the truth of the antecedent, by performing or desisting from some action, whereas they have no such control in the other three contexts.

It is clear, however, that the present seven contexts do not provide an opportunity for a factorial analysis of the effect of control (and thus of any effect of interaction between the control invested in each party), since there is amongst the seven no context in which the speaker has control over the truth of the consequent, but the hearer does not have control over the truth of the antecedent. This observation led, quite serendipitously, to the realisation that one important natural language usage of 'if P then Q' was missing from the range of contexts: that in which the speaker expresses a future intention which is contingent upon the truth of an antecedent over which neither the

speaker nor the hearer has any control, as in 'if it's fine over the weekend then I shall go fishing'. As will be seen, this context was added to the range of contexts investigated in Experiment 7, so permitting such a factorial analysis.

In addition to investigating the effect of manipulating the subjective non-contingent likelihood of the consequent, it was hoped that this experiment would provide a replication of the various context effects seen in Experiment 5 - and various specific comparative hypotheses were set out in the Introduction to this experiment.

Hypothesis 2a (that there would be more E-tables than D-tables in the Temporal, Causal, Promise and Threat contexts overall) was only partially supported. There was, indeed, an E/D preference in the Promise (E 49%; D 27%) and Threat (E 40%; D 32%) contexts, but the reverse was the case in the Temporal (E 28%; D 44%) and Causal (E 26%; D 45%) contexts. All of these observations are, of course, consistent with the general swing away from X-tables and Y-tables towards D-tables: in other words, with the general reduction in determinacy of responses to the FT and FF cases.

Hypothesis 2b (that there would be less E-tables than D-tables in the Tip, Warning and Universal contexts overall) was strongly supported (Tip: E 27%; D 50% - Warning: E 23%; D 51% - Universal: E 28%; D 44%), as might be expected given the general swing towards D-tables.

Hypothesis 2c (that response patterns in the Promise/Threat and Tip/Warning contexts would be similar so forming 'natural pairs') was reasonably well supported in terms of the frequency of truth tables but not so

in terms of raw responses: as has been seen, the only two significant differences between 'F' responses to FT in adjacently ranked contexts were between Promise (67%) and Threat (51%) and between Tip (36%) and Warning (29%). It is clear from Table 3.9 that for Promise and Threat this difference is reflected in the relative proportions of X-tables and Y-tables (which differ just on the FT case: 'F' for X-tables and '?' for Y-tables) and in the rather higher E/D preference in the Promise context. When the truth table patterns are considered, however, then the position is rather more encouraging than the analysis of raw responses might suggest: the frequency of E-tables in the Promise and Threat contexts is very considerably larger than in any other context and that of D-tables very considerably smaller. A similar - and complementary - situation obtains in the Tip and Warning contexts, in which the frequencies of E-tables in each are amongst the lowest and those of D-tables by far the highest. Despite the significant differences in raw responses, therefore, it is clear that the overall characters of the Promise/Threat and Tip/Warning contexts are still such as to suggest 'natural pairs'.

Hypothesis 2d (that there would be more X-tables in the Temporal context overall than in any other context) was strongly supported, but the effect was not quite so strong in the present experiment (in which 26% of all X-tables occurred in the Temporal context) as in Experiment 5 (where the corresponding figure was 42%): again, this is clearly the result of a general swing away from X-tables and Y-tables towards D-tables.

Finally, Hypothesis 2d (that the combined proportion of X-tables and Y-tables would be relatively constant over all seven contexts) received fairly strong support (this figure varying between 18% and 26% across the contexts).

The continued appearance of X-tables and Y-tables lends considerable weight to the contention that these tables are genuine reflections of the underlying interpretation: whilst the overall frequencies of these tables were somewhat diminished, their relative frequencies across the contexts were very similar indeed, and the apparent mutual exclusivity between them was still quite evident.

In the discussion of Experiment 5, it was it was argued that the D-table is by no means necessarily a defective version of the I-table, since it is equally concordant with the E-table. More specifically, it was suggested that the X-table might be a defective version of the E-table, whilst the Y-table might be seen as a somewhat *less* defective version of the D-table. The overall swing in the present experiment from X-tables and Y-tables towards D-tables does, however, bring this characterization into question, since the reduction in X-tables should, on this basis, have been reflected in an increase in E-tables rather than D-tables - but the frequency of E-tables remains unchanged at 32%.

That the frequencies of E-tables are unchanged by the reduction in X-tables and Y-tables leads strongly to the suggestion that D-tables, X-tables and Y-tables are all defective versions of the same truth table, X-tables and Y-tables merely being rather less defective - or more *determinate*. It is argued that the inclusion of indeterminate values in these three tables indicates a 'defective' table only in as much as it involves a degree of uncertainty about the truth values which obtain when the antecedent is false: and that X-tables and Y-tables simply reflect rather less uncertainty than D-tables. If X-tables, Y-tables and D-tables are indeed simply more or less uncertain (or more or less defective) versions of the same 'complete' truth table then, as was seen

earlier, that table must be the E-table, since the X-table is at odds with the I-table on the FT case.

In summary, then, it is argued that the apparent relationship between D-tables on the one hand, and X-tables and Y-tables on the other, indicates that *all three* are more or less defective E-tables. Further weight is added to such a contention by the observation that out of the 1236 truth tables inferred in Experiments 5 and 6 combined, only one was the I-table of the logically correct material implication: if D-tables were indeed defective versions of the I-table, then it is surely inconceivable that in only one response pattern out of 1236 would the 'defectiveness' be overcome.

3.4 : Experiment 7

Experiment 7 was a repetition of Experiment 6 with the addition of materials corresponding to the eighth context - Intentional. Following the truth table evaluation task, subjects were tested on a memory task, the prior warning about which was hoped to encourage subjects to read the texts more carefully than it was suspected that they might been read in Experiments 5 and 6. In addition, the memory task was to provide an independent measure of how carefully subjects had read those particular details in the texts which were critical to the manipulation of the subjective non-contingent likelihood of the consequent.

It was hypothesised that such a more careful reading of the texts would mean that the manipulation of the subjective non-contingent likelihood of the consequent would now produce the effect originally predicted: that the lo condition would produce a greater tendency to an equivalence interpretation. Additionally, it was anticipated that should this manipulation not produce the predicted effect, then performance on the memory task would be likely to be relatively poor.

As has been seen, the introduction of the eighth context - Intentional - permits a factorial analysis on the basis of hearer's and speaker's control. It was predicted that there would be a main effect of the speaker's control over the truth of the consequent such that the response patterns would be more in line with an equivalence interpretation where the speaker did have such control. No prediction was made about a main effect of the hearer's control over the truth of the antecedent, but it was anticipated that there would be an interaction such that there would be considerably more equivalence

interpretations where both parties had control. In addition, it was predicted that those patterns of truth tables which were common to Experiments 5 and 6 would be broadly replicated in this experiment. These general hypotheses might be summarised as follows:

H1 where the scenario indicated that the consequent was relatively less likely to become true irrespective of whether the antecedent became true then this would lead to a greater tendency to an equivalence interpretation than when the reverse was indicated

H2 *if H1 was not supported,* performance on the memory task would be relatively poor

H3 in the Promise, Threat and Intentional contexts (in which the speaker has control over the truth of the consequent) there would be a greater tendency to an equivalence interpretation than in the other five contexts

H4 there would be an interaction between speaker's and hearer's control such that there would be a considerably greater tendency to an equivalence interpretation in the Promise and Threat contexts (where both parties have their respective types of control)

H5 the relative proportions of the four truth tables overall would be broadly similar to that found on Experiments 5 and 6 - particularly to Experiment 6, which employed identical materials (except for the Intentional context)

These hypotheses may be restated in operational terms as follows:

H1 there would be more E-tables when the non-contingent likelihood of the consequent was indicated to be low

H2 *if H1 was not supported,* performance on the memory task would be relatively poor

H3 there would be more E-tables in the Promise, Threat and Intentional contexts overall than in the other five contexts overall

H4 Promise and Threat contexts would produce the highest overall proportion of E-tables and the Temporal, Causal and Universal contexts the lowest, with intermediate levels in the Intentional, Tip and Warning contexts

H5a there would be more E-tables than D-tables in the Promise and Threat contexts overall

H5b there would be less E-tables than D-tables in the Tip, Warning and Universal contexts overall

H5c response patterns in the Promise/Threat and Tip/Warning contexts would be similar so forming 'natural pairs'

H5d there would be more X-tables in the Temporal context overall than in any other context

H5e the combined proportion of X-tables and Y-tables would be relatively constant over all eight contexts

H5f to the extent that the combined overall level of X-tables and Y-tables might differ from that on Experiments 5 and 6, then this would be reflected in a complementary change in the overall level of D-tables

3.4.1 : Design and Methodology

Subjects:

24 first-year BSc (Hons) Psychology undergraduates at Plymouth Polytechnic (18 female, 6 male), mean age 19.5 years.

Truth Table Task Methodology:

The truth table task was carried out in precisely the same fashion, and with the same apparatus, as Experiment 6. The materials were identical to those described for Experiment 6, with the addition of the Intentional context specified earlier. There were thus now 32 trials in total, using materials drawn from a total of 64 texts (8 contexts x 4 exemplars x 2 versions). As before, candidates for the four exemplars of the new Intentional context were prepared by the experimenter, offered for rating to two expert judges and modified where necessary to obtain 100% agreement that they belonged to a

given version (lo or hi) of that context. The additional four Intentional texts used, and the four pairs of outcome sentences associated with each, are presented in full in Appendix [7a].

Memory Task Materials:

Following the truth table evaluation task, subjects were given a memory task comprising one question about each of the 32 texts. The questions were asked in the same order as that in which the texts had originally been presented to that subject. In every case, the question related to that portion of the passage that determined whether it was the lo or hi version. Subjects were forced to choose between two possible answers, each of which was (as close as grammatically possible) a verbatim extract from the respective lo or hi version of the relevant passage. Each question/answer set was preceded by a short 'reminder' phrase to assist the subject in bringing the relevant passage to mind. The question/answer set for the example texts given in Figures 3.5 and 3.6 (and the relevant 'reminder') are shown in Figure 3.7.

Whether the lo answer was shown above the hi answer on the screen, or *vice versa,* was determined at random in each case. Subjects pressed keys on the keypad to select one of the two answers. In order to permit changes of mind and to obviate errors resulting from subjects keeping their fingers on the keys, the item selected was highlighted on the screen and subjects had to press a separate key to go on to the next trial. Summary instructions as to how to use the keypad remained on the screen throughout the memory task: a typical screen is shown at Appendix [7b].

the reminder:

[. . . Little Billy turning up the radio . . .]

the question:

What was Little Billy's mother's attitude towards parents smacking their children?

lo answer:

She strongly disapproved

hi answer:

She strongly believed that they should be smacked for the merest naughtiness

Figure 3.7 : Experiment 7:
Memory task:
The 'reminder', the question and the 'lo' and 'hi' answers relating to the texts shown in Figures 3.5 and 3.6
(the italicised items were not displayed to subjects)

Memory Task Procedure:

Following the reasoning task, subjects were presented with instructions regarding the memory task. The instructions given are presented in full in Appendix [7c], but that part of them concerned with the task itself is reproduced below:

> "... two possible answers, A and B. Any key in the top row will select answer A, and any key in the bottom row will select answer B. Don't worry if you press the wrong key by mistake - your choice will be noted on the screen and you can change your mind simply by pressing another key. When you are happy with your selection, then you should press CONTINUE to go on to the next trial ..."

Following the experiment, subjects were debriefed by text presented on the computer screen, as presented in Appendix [5g]. No formal timing was undertaken, but subjects all took between 40 and 60 minutes to complete the two tasks.

3.4.2 : Results

[a] : Truth Table Evaluation Task

As before, the dependent variable is the decision indicated by a subject via the keypad as detailed for Experiments 5 and 6. Table 3.12 presents the summary response data, showing the raw frequency of 'T', 'F' and '?' responses over the eight contexts, with the lo and hi versions collapsed.

The overall patterns of raw responses are similar to those found in Experiments 5 and 6 in as much as there is a massive ceiling effect on responses to the TT and TF cases which are, respectively, seen as supporting and contradicting the statements by virtually all subjects on virtually every trial. The FT and FF cases again show very substantial variability in responses, although one particular response is reliably avoided in each case - the FT case is almost never seen as supporting the statement, whilst the FF case is almost never seen as contradicting the statement.

Unlike Experiments 5 and 6, however, there are considerably less indeterminate '?' responses to the FT and FF cases: whereas these accounted for roughly half the FT and FF responses in those two experiments, they

only constitute about a third of those responses in the present experiment, two thirds of responses to the FT and FF cases now being the determinate 'F' and 'T' respectively.

Context		Tem	Cau	Pro	Thr	Tip	War	Uni	Int	All
outcome [case]										
P & Q	T	99	98	100	100	97	100	99	99	99
	F	-	-	-	-	1	-	1	1	0
[TT]	?	1	2	-	-	2	-	-	-	1
P &~Q	T	-	1	-	-	-	2	-	-	0
	F	99	97	95	98	96	96	94	98	97
[TF]	?	1	2	5	2	4	2	6	2	3
~P & Q	T	-	-	-	-	-	1	-	-	0
	F	72	67	85	69	63	55	68	67	68
[FT]	?	28	33	15	31	38	44	32	33	32
~P &~Q	T	39	54	82	77	54	68	56	79	64
	F	4	3	5	1	2	5	2	2	3
[FF]	?	57	43	13	22	44	27	42	19	33

Table 3.12 : Experiment 7:
Summary of raw responses [lo+hi pooled] (%):
T = "SUPPORTS the Statement"
F = "CONTRADICTS the Statement"
? = "TELLS US NOTHING ABOUT the Statement"
[N per context = 96 = 24 subjects x hi/lo x 2 exemplars]
[Tem=Temporal, Cau=Causal, Pro=Promise, Thr=Threat,
Tip =Tip, War=Warning, Uni=Universal, Int=Intentional]
(failures to sum to 100% are due to rounding)

Table 3.13 re-presents the raw responses to the FT and FF cases shown in

Table 3.12, further analysed to show the hi and lo versions of the task (given the massive ceiling effects in the TT and TF cases, there is nothing to be gained by analysing these cases in a similar way).

Context		Tem		Cau		Pro		Thr		Tip		War		Uni		Int		All	
outcome [case]		lo	hi	lo	hi	lo	hi	lo	hi	lo	hi	lo	hi	lo	hi	lo	hi	lo	hi
~P & Q	T	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	0	-
	F	83	60	77	56	90	81	71	67	65	60	52	58	71	65	69	65	72	64
[FT]	?	17	40	23	44	10	19	29	33	35	40	46	42	29	35	31	35	28	36
~P &~Q	T	42	35	58	50	81	83	75	79	48	60	69	67	60	52	79	79	64	63
	F	8	-	2	4	4	6	2	-	4	-	4	6	2	2	-	4	3	3
[FF]	?	50	65	40	46	15	10	23	21	48	40	27	27	38	46	21	17	33	34

Table 3.13 : **Experiment 7:**
Summary of raw responses by version (lo/hi) (%):
T = "SUPPORTS the Statement"
F = "CONTRADICTS the Statement"
? = "TELLS US NOTHING ABOUT the Statement"

[N per context = 48 = 24 subjects x 2 exemplars]

[Tem=Temporal, Cau=Causal, Pro=Promise, Thr=Threat,
Tip =Tip, War=Warning, Uni=Universal, Int=Intentional]

(failures to sum to 100% are due to rounding)

Comparisons of the responses to the hi and lo versions in each context failed to reveal any significant effect of this manipulation upon responses to the FF case; in the FT case, significant differences were found between the lo and hi conditions in just two of the eight contexts. In the Temporal context, there were significantly more 'F' responses to FT in the lo version (40%) than in the hi (29%) ($t = 3.412$; $p = 0.002$; 2-tailed), whilst in the Causal context, there

were significantly more 'F' responses to FT in the lo version (37%) than in the hi (27%) ($t = 2.46$; $p = 0.022$; 2-tailed).

Both these differences replicate those found in Experiment 6. As in Experiment 6, however, the overall picture is one in which the manipulation had a significant effect upon responses in just two contexts: in these circumstances, the remainder of the analysis will carried out upon the pooled data from both the lo and hi versions of the task (as shown in Table 3.12).

On the FT case, the mean responses across all contexts are roughly two thirds 'F' and one third '?'. A 1-way analysis of variance of the frequency of 'F' responses does, however, show a highly significant effect of context ($F = 4.107$; $df = 7$; $p = 0.0004$). Whilst pairwise t-test comparisons of each two adjacently ranked contexts fails to reveal a single significant difference, three clusters are apparent over the eight contexts, such that there are significantly less 'F' responses to FT on Warning (55%) than on Tip+Causal+Intentional+Universal+Threat+Temporal pooled (68% overall) ($t = 13.29$; $p < 0.001$; 1-tailed) and significantly less 'F' responses to FT on Tip+Causal+Intentional+Universal+Threat+Temporal pooled (68% overall) than on Promise (85% overall) ($t = 10.722$; $p < 0.001$; 1-tailed).

In comparing these data with those from Experiments 5 and 6, the major observation to be made is that the 'natural pairing' of responses to the FT case in the Promise and Threat contexts found in Experiment 5 (but not Experiment 6), has again failed to be replicated: indeed, 'F' responses to the FT case in the Threat context (69%) are significantly different from those in the Promise context (85%) ($t = 2.996$; $p = 0.003$; 2-tailed).

On the FF case, the mean responses across all contexts are roughly two thirds 'T' and one third '?'. A 1-way analysis of variance of the frequency of 'T' responses shows a highly significant effect of context ($F = 12.646$; $df = 7$; $p < 0.0001$). Pairwise t-test comparisons of each two adjacently ranked contexts reveal three significant differences. There are significantly more 'T' responses to FF on Causal (52%) than on its nearest neighbour, Temporal (37%) ($t = 2.394$; $p = 0.012$; 1-tailed); there are significantly more 'T' responses to FF on Warning (65%) than on its nearest neighbour, Universal (54%) ($t = 1.905$; $p = 0.034$; 1-tailed); and there are significantly more 'T' responses to FF on Threat (74%) than on its nearest neighbour, Warning (65%) ($t = 2.099$; $p = 0.023$; 1-tailed). The rank-ordered frequencies reveal four clusters in the eight contexts, such that there are significantly less 'T' responses to FF on Temporal (37%) than on Causal+Tip+Universal pooled (55% overall) ($t = 7.933$; $p < 0.001$; 1-tailed), significantly less 'T' responses to FF on Causal+Tip+Universal pooled (55% overall) than on Warning (65%) ($t = 5.454$; $p < 0.001$; 1-tailed) and significantly less 'T' responses to FF on Warning (65%) than on Threat+Intentional+Promise pooled (79% overall) ($t = 14.117$; $p = 0.011$; 1-tailed).

In comparing these data with those from Experiments 5 and 6, the 'natural pairing' of responses to the FT case in the four 'inducement' contexts has again failed to be replicated, but now in terms of the Tip and Warning contexts: 'T' responses to the FF case in the Tip context (52%) are significantly different from those in the Warning context (65%) ($t = 2.122$; $p = 0.022$; 2-tailed).

Table 3.14 presents the inferrable truth tables, with the lo and hi versions collapsed to facilitate comparison with the relevant data from Experiments 5

and 6. It will be noted that just one out of 768 patterns accorded with the (logically correct) I-table of a material implication interpretation.

			Response				
Truth case:	TT	T	T	T	T	T	
	TF	F	F	F	F	F	
	FT	T	F	?	F	?	
	FF	T	T	?	?	T	
Truth table:		I	E	D	X	Y	none
Context:	Temporal	-	30	21	35	7	6
	Causal	-	43	19	22	11	5
	Promise	-	69	3	9	8	10
	Threat	-	54	9	13	22	2
	Tip	-	39	25	18	10	8
	Warning	1	38	15	13	25	9
	Universal	-	40	21	21	10	8
	Intentional	-	56	13	6	20	5
	Total %	0	46	16	17	14	7

Table 3.14 : Experiment 7:
Truth table evaluation task:
Inferred truth tables x context [lo+hi pooled] (%)
[N = 96 per context = 24 subjects x 4 exemplars]
(failures to sum to 100% are due to rounding)

Table 3.15 re-presents the truth table analysis shown in Table 3.14, further analysed to show the hi and lo versions of the task. It is clear from Table 3.15 that the lo/hi manipulation produced no substantial overall effect in any of the seven contexts. In these circumstances, further analysis of the data will thus be based upon the collapsed data (lo+hi), as presented in summary in Table 3.14.

		Response									
Truth case:	TT	T		T		T		T			
	TF	F		F		F		F			
	FT	T		F		?		F		?	
	FF	T		T		?		?		T	
Truth table:		I		E		D		X		Y	
		lo	hi	lo	hi	lo	hi	lo	hi	lo	hi
Context:	Temporal	-	-	33	27	10	31	38	33	6	8
	Causal	-	-	50	35	13	25	25	19	8	15
	Promise	-	-	71	67	4	2	10	8	4	13
	Threat	-	-	54	54	10	8	13	13	19	25
	Tip	-	-	40	38	25	25	21	15	6	15
	Warning	2	-	40	35	17	13	10	15	23	27
	Universal	-	-	40	40	17	25	21	21	10	10
	Intentional	-	-	63	50	15	10	6	6	17	23
	Total lo	0		49		14		18		12	
	Total hi		-		43		17		16		17

	none	
	lo	hi
Context: Temporal	13	-
Causal	4	6
Promise	10	10
Threat	4	-
Tip	8	8
Warning	8	10
Universal	13	4
Intentional	-	10
Total lo	8	
Total hi		6

Table 3.15: Experiment 7:
Truth table evaluation task:
Inferred truth tables x context x version (lo/hi) (%)
[N = 48 per context = 24 subjects x 2 exemplars]
(failures to sum to 100% are due to rounding)

Looking at the data overall, the predominant truth table by far is now the E-table, accounting for 46% of the responses, unlike Experiment 5 (where E-tables and D-tables occurred roughly equally often) and Experiment 6 (where D-tables predominated, although not so markedly as E-tables do here). In addition, there is a small increase in the proportion of Y-tables, and a substantial increase in the proportion of X-tables (from 10% in Experiment 6 to 17% in the present experiment). Between them, X-tables and Y-tables account for 31% of the responses, which compares closely with their 35% in Experiment 5.

Table 3.14 reveals certain substantial effects of context. Whilst there is a preference for E-tables over D-tables overall, this is virtually absolute in the Promise and Threat contexts, where there are almost no D-tables at all. The Tip and Warning contexts, on the other hand, have almost the lowest proportions of E-tables, whilst Tip has by far the highest proportion of D-tables.

The one context in which E-tables do not predominate is the Temporal context, where the X-table is the most frequently occurring response pattern. The suggestion of mutual exclusivity found in Experiments 5 and 6 between X-tables and Y-tables is apparent again in the present experiment, in that a higher frequency of X-tables goes with a lower frequency of Y-tables (and *vice versa*) - there is, however, one substantial exception to this pattern: in the Promise context, the particularly high proportion of E-tables appears to be reflected in a very low frequency of both X-tables and Y-tables.

The analysis presented in Table 3.14 is for all 768 responses - that is, it is collapsed across subjects. Table 3.16 presents an analysis in terms of the modal response patterns of individual subjects, using as a criterion for a particular classification within a given context that a subject should show a similar response pattern on at least three of the four exemplars in that context. In addition, Table 3.16 shows the proportion of subjects modally consistent across all eight contexts, the criterion now being that they should be thus classified on at least seven of the eight contexts: it will be noted that the addition of an eighth context makes this a somewhat more demanding criterion.

			Response				
Truth case:	TT	T	T	T	T	T	
	TF	F	F	F	F	F	
	FT	T	F	?	F	?	
	FF	T	T	?	?	T	
Truth table:		I	E	D	X	Y	(none)
Context:	Temporal	-	17	13	21	-	50
	Causal	-	33	4	8	-	54
	Promise	-	71	-	8	4	17
	Threat	-	46	8	8	8	29
	Tip	-	29	13	13	-	46
	Warning	-	29	13	-	8	50
	Universal	-	29	17	13	4	38
	Intentional	-	46	4	4	8	38
Within contexts: ≥3/4 overall		-	38	9	9	4	40
Across contexts: ≥3/4 on ≥7/8		-	21	-	-	-	n/a

Table 3.16 : Experiment 7:
Truth table evaluation task:
Inferred truth tables x context:
% of subjects modally consistent *within* contexts (≥3/4)
and *across* contexts (≥6/7)

[N per context = 24 subjects]

(failures to sum to 100% are due to rounding)

Given the very strong overall preference for E-tables in this experiment, it is not surprising to note from Table 3.16 that this is by far the strongest modally consistent response pattern across contexts. Another major difference in the present experiment is the far larger number of modally consistent X-tables than was found on Experiments 5 and 6: in those earlier experiments, X-tables were found to be by far the least frequent modally

consistent response pattern, whereas in the present experiment they occur consistently more often even than D-tables.

It will be recalled that the main purpose in including the Intentional context was to permit a factorial analysis of the data in terms of the speaker's control over the truth of the consequent and the hearer's control over the truth of the antecedent. Whilst it is clear even from a subjective visual examination of Table 3.14 that there are substantial qualitative differences between the various contexts, it is not a simple matter to subject this to a factorial analysis, since what is of interest is the various admixtures of tables inferred for each context.

It has been argued, however, that D-tables, X-tables and Y-tables are all more or less 'defective' versions of the (complete) E-table. It will be argued in the general discussion of these four experiments that the E-table is the 'naturally canonical' table for 'if P then Q' in everyday usage and that the particular effects of the various contexts result in the introduction of two distinct types of uncertainty so as to give the FT and/or the FF truth cases an indeterminate truth value. It will be helpful to that later discussion to redescribe the data in terms of the proportion of classifiable response patterns in a given context which remain in the 'naturally canonical' form of E-tables: that is, the proportion of E-tables out of the total number of E-tables, D-tables, X-tables and Y-tables. In addition, this proportion will provide a single 'metric of equivalence' which will facilitate the desired factorial analysis. The I-table of material implication is excluded from the total number of classifiable tables on the grounds that there was only one such pattern out of 768 in the present experiment (and only two out of 2000 responses in Experiments 5, 6 and 7 combined), which strongly suggests that this is not a

genuinely intended response pattern at all. Table 3.17 sets out this metric of equivalence - which will be referred to hereafter as the 'E-statistic' - for the eight contexts, together with the rank order of each (1=lowest, 8=highest).

Metric of equivalence:		$\frac{E}{E+D+X+Y}$					
		E-statistic (rank)				E-statistic (rank)	
				ranked			
Context:	Temporal	.323	(1)	Context:	Temporal	.323	(1)
	Causal	.462	(5)		Tip	.413	(2)
	Promise	.764	(8)		Warning	.434	(3)
	Threat	.545	(6)		Universal	.441	(4)
	Tip	.413	(2)		Causal	.462	(5)
	Warning	.434	(3)		Threat	.545	(6)
	Universal	.441	(4)		Intentional	.590	(7)
	Intentional	.590	(7)		Promise	.764	(8)

Table 3.17 : Experiment 7: Metric of equivalence (E / (E+D+X+Y)) x context

It is clear from Table 3.17 that there appear to be substantial effects of context in such an analysis. A 1-way analysis of variance revealed a highly significant effect of context on this metric ($F = 9.078$; $df = 7$; $p < 0.001$). Pairwise comparison of all adjacently ranked pairs of means revealed just one significant difference: the E-statistic was significantly higher in the Promise context (0.764) than in its immediate neighbour, Intentional (0.590) ($t = 3.331$; $p = 0.002$; 1-tailed). It is clear that the hypothesised pairing of the Promise and Threat contexts fails to appear here also, the E-statistic in the Promise context (0.764) being significantly different from that in the Threat context (0.545) ($t = 3.516$; $p = 0.002$; 2-tailed).

In order to carry out the desired factorial analysis, it is necessary to combine the statistics in Table 3.17 so as to show the mean metrics of equivalence for the one or more contexts falling into each category on the basis of the hearer's control over the truth of the antecedent and the speaker's control over the truth of the consequent. This collapsed analysis is presented in Table 3.18.

Metric of equivalence:		$\frac{E}{E+D+X+Y}$		
Speaker's control over the truth of the consequent?		**NO**	**YES**	***(Hearer's main effect)***
Hearer's control over the truth of the antecedent?	**NO**	**(Temporal) (Causal) (Universal) .409**	**(Intentional) .590**	**(.499)**
	YES	**.424 (Tip) (Warning)**	**.655 (Promise) (Threat)**	**(.539)**
(Speaker's main effect)		**(.416)**	**(.622)**	**(.519)**

Table 3.18 : Experiment 7: Metric of equivalence (E / (E+D+X+Y)) x context

Table 3.18 shows a clear main effect of the speaker's control over the truth of the consequent. Where the speaker has such control, then the equivalence metric is considerable higher: a 2-way ANOVA shows this difference to be significant ($F = 25.518$; $df = 1$; $p < 0.0001$). There is no significant main effect

of the hearer's control over the truth of the antecedent and whilst there appears to be some sign of a modest interaction, in that the speaker's control has rather more of an effect where the hearer also has control (such that the metric is higher - and thus the interpretation is more in line with 'pure' equivalence - where both parties have control), this effect is not significant. Whilst this factorial analysis cannot be carried out in its entirety for Experiments 5 and 6 (in the absence of the Intentional context), partial analyses for those experiments will be presented in the discussion of the present experiment for the purposes of comparison.

[b] : Memory Task

The dependent variable in the memory task is the number of errors made in choosing between the two alternative answers offered to each question: it will be recalled that one question was asked relating to each of the 32 texts with which a subject had been presented on the truth table evaluation task. Table 3.19 shows the distribution of errors over the two versions and eight contexts, together with the proportion of errors relating to each context and to each truth table.

Two points should be borne in mind when reading this table. Firstly, the number of questions asked relating to each cell is, of course, the number of truth tables produced for that cell as shown in Table 3.14, and it thus varies very considerably - the only directly comparable figures are thus the total number of errors for a given context x version combination, and the proportion of errors for a given truth table over all eight contexts. Secondly, it should be noted that the zero symbol "0" denotes that no subject who produced a truth table in a given cell made an error upon the corresponding

question: the dash symbol "-", on the other hand, denotes that no subject produced a truth table in that cell.

		Response														
Truth case:	TT	T		T		T		T		T						
	TF	F		F		F		F		F						
	FT	T		F		?		F		?						
	FF	T		T		?		?		T						
Truth table:		I		E		D		X		Y		none		overall % error		
		lo	hi	lo	hi	lo	hi	lo	hi	lo	hi	lo	hi	lo	hi	tot
Context:	Temporal	-	-	0	2	0	0	0	5	0	1	0	-	0%	17%	8%
	Causal	-	-	2	4	2	3	2	2	1	0	0	0	15%	19%	17%
	Promise	-	-	2	12	0	0	0	2	0	2	0	1	4%	35%	20%
	Threat	-	-	5	1	1	1	1	0	0	0	0	-	15%	4%	9%
	Tip	-	-	0	2	1	2	1	0	0	2	1	0	6%	13%	9%
	Warning	0	-	6	2	0	0	1	1	0	0	0	0	15%	6%	10%
	Universal	-	-	1	11	1	5	2	8	0	1	0	1	8%	54%	31%
	Intentional	-	-	2	3	0	1	0	1	2	0	-	1	8%	13%	10%
	Total lo	0		18		5		7		3		1				
	Total hi		-		37		12		19		6		3			
	Total lo+hi	0		55		17		26		9		4				
	% errors lo	0%		10%		9%		10%		7%		3%		9%		
	% errors hi		-		22%		18%		31%		9%		13%		20%	
	% errors lo+hi	0%		16%		14%		20%		8%		8%				14%

Table 3.19 : Experiment 7:
Memory task errors x version x context:
"0" = truth tables produced, but no memory task errors
"-" = not applicable (no truth tables produced)
[N per context = 96 = 24 S x 2 versions x 2 exemplars]

Two overall features of the data are immediately apparent from Table 3.19. Firstly, there is a very low level of error overall, 86% of questions being

answered correctly; secondly, there is an extremely large effect of the lo/hi manipulation, errors on the lo versions being far lower overall than on the hi versions.

A 2-way ANOVA revealed highly significant effects of context ($F = 5.978$; $df = 7$; $p < 0.0001$) and of version (lo/hi) ($F = 23.328$; $df = 1$; $p = 0.0001$) and a highly significant interaction effect ($F = 7.108$; $df = 7$; $p < 0.0001$). The error data summarised in Table 3.19 are illustrated graphically in Figure 3.8.

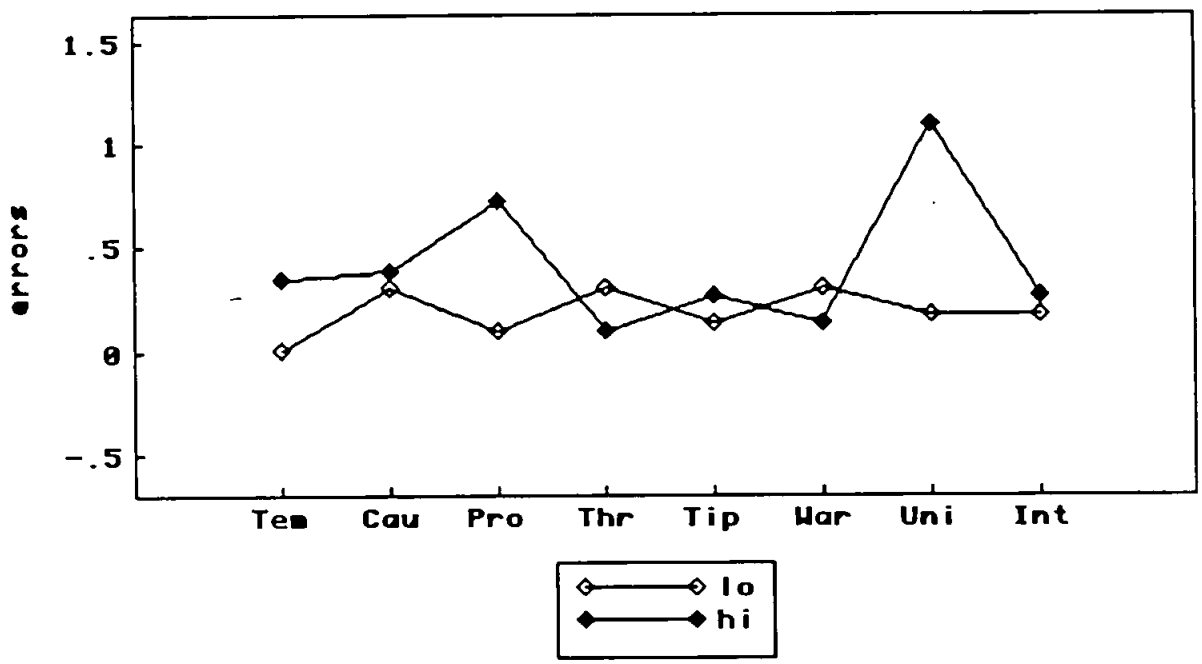


Figure 3.8 : Experiment 7:
Memory task errors x version x context
[max per point = 2 = 2 exemplars]

Pairwise comparisons of the lo and version reveal significant differences in four of the eight contexts: there were significantly more errors in the hi version of the Temporal (lo 0%, hi 17%) ($t = 2.563$; $p = 0.178$; 2-tailed), Promise (lo 4%, hi 35%) ($t = 3.978$; $p = 0.001$; 2-tailed) and Universal (lo 8%,

hi 54%) ($t = 4.836$; $p < 0.001$; 2-tailed) contexts, and significantly less errors in the hi version of the Threat context (lo 15%, hi 4%) ($t = 2.46$; $p = 0.022$; 2-tailed). When the total errors in each context are ranked, pairwise comparisons between each adjacently ranked pair of contexts reveals just one significant difference in that there were more errors overall in the Universal context (31% overall) than in its immediate neighbour the Promise context (20% overall). In addition, the hypothesised pairing of Promise and Threat contexts failed to appear in that the overall level of error in the Promise context (20% overall) was significantly different from that in the Threat context (9% overall) ($t = 2.198$; $p = 0.038$; 2-tailed) - and, as has been seen, whilst each of these contexts showed a significant lo/hi difference, this was in a different direction in each case.

It is also clear from Figure 3.8 that the interaction arises out of the fact that the higher overall level of error on the hi versions arises almost exclusively from the Temporal, Promise and Universal contexts, whilst the Threat context shows a strong effect in the opposite direction.

3.4.3 : Discussion

Once again, the manipulation of the subjective non-contingent likelihood of the consequent has failed to produce the hypothesised effect. Following subjects' informal comments after Experiment 6, it was hypothesised that should this effect not appear, then performance on the memory task would be relatively poor. It is clear, however, that the lack of a lo/hi effect cannot, after all, be attributed to subjects having failed to read the particular details in the texts by which the manipulation was effected, since the results of the

memory task show that subjects must have read these details rather carefully, given the very low level of error (14% overall). This overall level of accuracy on the memory task is even more remarkable when the demands of the memory task *qua* memory task are taken into account, in that subjects had to remember relatively fine details about 32 texts some of which would have been presented as long as 50 minutes before the test: additionally, no clue was given to subjects about what aspects of the texts they were required to memorise.

Where the manipulation of the subjective non-contingent likelihood of the consequent did have an effect, however, was on the memory task, in that errors in the lo condition were a paltry 9% overall (and never rose above 15%), whilst those in the hi condition were consistently higher (20% overall). Whilst there is indeed a main effect of version overall, it is clear that the higher level of error on the hi version arises almost exclusively from just two contexts, which account between them for 56% of the hi version errors (and 44% of all errors in the task).

On the basis of these data, it is proposed that the larger part of the errors result from the forced choice between the two alternative answers to each question, which would have led to subjects being obliged to guess when they could not recall an answer. Given the low level of error on the lo versions, it is clear that the majority of such guesses must have been in terms of the answer relating to the lo version of the text, so that the low level of error in the lo condition reflects lo guesses as well as genuine lo recollections: on this basis, it is proposed that the higher level of error in the hi condition would have resulted from a failure to recollect followed by a lo guess.

If this interpretation is correct, then the fact that the effect is seen to reside

almost exclusively in just two contexts suggests very strongly indeed that the lo versions of the texts in these conditions are in some sense more acceptable, or more plausible. This is consonant with the comment in the discussion of Experiment 6 that in certain contexts it was particularly difficult to construct hi versions of the texts which still sounded 'natural'. This would appear, then, to lend weight to the notion that the use of the conditional in particular contexts is not only less natural than the use of other linguistic expressions, but also that its selection in preference to the alternative expressions implies something about the relationship between antecedent and consequent that goes far beyond the propositional content of the utterance. In this connection, it is interesting to note that by far the highest level of errors arises on the Universal context in which, it was argued, the use of the conditional form is at its least natural. Furthermore, in as much as the high level of hi errors in that context presumably arose through lo guesses, this provides considerable support for the suggestion that the use of 'if P then Q' in that context leads subjects very strongly to interpret the relationship as one of equivalence.

In addition to investigating the effect of manipulating the subjective non-contingent likelihood of the consequent, specific predictions were made about the effect of speaker's and hearer's control. Hypothesis H3 predicted that in the Promise, Threat and Intentional contexts (in which the speaker has control over the truth of the consequent) there would be a greater tendency to an equivalence interpretation than in the other six contexts: this prediction is strongly supported in that these contexts did indeed have the three highest levels of E-tables (60% overall, compared with 38% overall for the other five contexts).

Hypothesis H4 predicted that there would be an interaction between speaker's and hearer's control such that there would be a considerably greater tendency to an equivalence interpretation in the Promise and Threat contexts (where both parties have their respective types of control). Whilst this interaction did not reach significance, it is clear from Table 3.18 that this cell is, as predicted, by far the highest. Together with the strong support for Hypothesis H3, there would thus appear to be considerable evidence for the importance of the perceived control of the parties to the utterance over the truth of its propositions.

Whilst the absence of the Intentional context in Experiments 5 and 6 precludes the factorial analysis available for the present experiment, it is relevant to compare at least the main effects of control in those earlier experiments with the present findings. Table 3.20 sets out the E-statistics for the seven contexts of Experiment 5.

A 1-way analysis of variance revealed a highly significant effect of context on this metric ($F = 4.899$; $df = 6$; $p = 0.0002$). Pairwise comparisons failed, however, to reveal any significant differences between adjacently ranked pairs of means. Additionally, it should be noted that neither the Tip/Warning contexts nor the Promise/Threat contexts differed significantly from each other, so supporting the hypothesised natural pairing.

Metric of equivalence:		E-statistic (rank)			E / (E + D + X + Y)	E-statistic (rank)	
					ranked		
Context:	Temporal	.417	(6)	Context:	Universal	.200	(1)
	Causal	.342	(4)		Tip	.225	(2)
	Promise	.400	(5)		Warning	.250	(3)
	Threat	.425	(7)		Causal	.342	(4)
	Tip	.225	(2)		Promise	.400	(5)
	Warning	.250	(3)		Temporal	.417	(6)
	Universal	.200	(1)		Threat	.425	(7)
					Intentional	n/a	n/a

Table 3.20 : Experiment 5:
Metric of equivalence (E + E+D+X+Y) x context

Table 3.21 summarises the main effects of control in Experiment 5. Whilst a proper factorial analysis is not possible in the absence of the Intentional context, an indication of the main effects of speaker's and hearer's control might be obtained by individual pairwise comparisons of all three cells. The 'neither' cell of Temporal+Causal+Universal (0.319) is (just) significantly lower than the 'both' cell of Promise+Threat (0.412) ($t = 2.224$; $p = 0.038$; 2-tailed) but not from the 'hearer only' cell of Tip + Warning (0.237). Taken together, these findings suggest that there is no significant main effect of the hearer's control over the truth of the antecedent. That the 'both' cell (0.412) is significantly greater than both the 'neither' cell (0.319) and the 'hearer only' cell (0.237) ($t = 3.573$; $p = 0.002$; 2-tailed) suggests, however, a strong main effect of the speaker's control over the truth of the consequent. Both these findings are in line with those of the factorial analysis of Experiment 7.

Metric of equivalence:		$\frac{E}{E+D+X+Y}$		
Speaker's control over the truth of the consequent?		NO	YES	*(Hearer's main effect)*
		(Temporal) (Causal) (Universal)		
Hearer's control over the truth of the antecedent?	NO	.319	n/a	(.319)
	YES	.237 (Tip) (Warning)	.412 (Promise) (Threat)	(.325)
(Speaker's main effect)		(.278)	(.412)	

Table 3.21 : Experiment 5:
Metric of equivalence (E / (E+D+X+Y)) x context

Table 3.22 sets out the E-statistics for the seven contexts of Experiment 6. A 1-way analysis of variance revealed a highly significant effect of context on this metric ($F = 5.681$; $df = 6$; $p < 0.0001$). Pairwise comparisons revealed just one significant difference between adjacently ranked pairs of means in that the E-statistic was significantly lower in the Tip context (0.292) than in the Threat context (0.462) ($t = 2.914$; $p = 0.004$; 1-tailed). As for Experiment 5, it should be noted that neither the Tip/Warning contexts nor the Promise/Threat contexts differed significantly from each other, so supporting the hypothesised natural pairing.

Metric of equivalence:				$\frac{E}{E+D+X+Y}$			
		E-statistic (rank)				E-statistic (rank)	
				ranked			
Context:	Temporal	.288	(3)	Context:	Warning	.253	(1)
	Causal	.267	(2)		Causal	.267	(2)
	Promise	.507	(7)		Temporal	.288	(3)
	Threat	.462	(6)		Universal	.288	(4)
	Tip	.292	(5)		Tip	.292	(5)
	Warning	.253	(1)		Threat	.462	(6)
	Universal	.288	(4)		Promise	.507	(7)
					Intentional	n/a	n/a

Table 3.22 : Experiment 6:
Metric of equivalence (E / (E+D+X+Y)) x context

Table 3.23 summarises the main effects of control in Experiment 6. The 'neither' cell of Temporal+Causal+Universal (0.281) is very significantly lower than the 'both' cell of Promise+Threat (0.484) ($t = 4.257$; $p < 0.001$; 2-tailed), but marginally (but not significantly) higher than the 'hearer only' cell of Tip+Warning (0.273). Given the disparity between these two comparisons, it is not possible to draw any conclusions about the overall main effect of the hearer's control over the truth of the antecedent. These findings are, however, at least consistent with the interaction discussed above, in so far as the hearer's control was only seen to exert any effect at all when the speaker was also in control, and that such effect was relatively dramatic. As for Experiment 5, that the 'both' cell (0.484) is significantly greater than both the 'neither' cell (0.281) and the 'hearer only' cell (0.273) ($t = 4.177$; $p < 0.001$; 2-tailed), suggests strongly that there is a main effect of the speaker's control over the truth of the consequent.

Metric of equivalence:		E / (E + D + X + Y)		
Speaker's control over the truth of the consequent?		**NO**	**YES**	***(Hearer's main effect)***
		(Tempóral) **(Causal)** **(Universal)**		
Hearer's control over the truth of the antecedent?	**NO**	**.281**	**n/a**	**(.281)**
	YES	**.273** **(Tip)** **(Warning)**	**.484** **(Promise)** **(Threat)**	**(.379)**
(Speaker's main effect)		**(.277)**	**(.484)**	

Table 3.23 : Experiment 6:
Metric of equivalence (E / (E+D+X+Y)) x context

Taken together, these *post hoc* analyses of Experiments 5 and 6 lend considerable weight to the findings of Experiment 7, in that the speaker's control over the truth of the consequent appears to increase the tendency towards equivalence, whilst no such effect results from the hearer's control over the truth of the antecedent. Additonally, further weight is added to the notion that the two types of control will interact, such that the effect of the speaker's control will be by far at its strongest when both parties are in control.

In addition to the major hypotheses already discussed, it was hoped that this

experiment would provide a replication of the various context effects seen in Experiments 5 and 6 - and various specific comparative hypotheses were set out in the Introduction to this experiment. Hypothesis 5a (that there would be more E-tables than D-tables in the Promise and Threat contexts overall) was particularly strongly supported, these contexts producing 56% E-tables overall as compared with just 6% D-tables overall.

Hypothesis 5b (that there would be less E-tables than D-tables in the Tip, Warning and Universal contexts overall) was not, however, supported, each producing more E-tables than D-tables (39% E-tables overall, 20% D-tables overall) which is, perhaps, unsurprising given the strong overall swing in this experiment from D-tables towards E-tables.

Hypothesis 5c predicted that the 'natural pairing' of the Promise/Threat and Tip/Warning contexts observed on Experiment 5 (although not on Experiment 6) would re-emerge in the present experiment. There is no support whatever for this hypothesis: indeed, there is significant evidence to refute it. The analysis of raw responses showed that 'F' responses to FT were significantly lower in the Warning context, and significantly higher in the Promise context, than in all other contexts; similarly, 'T' responses to FF were significantly lower in the Tip context than in the Warning context. Furthermore, the E-statistic for the Promise is significantly higher than that for the Threat context and whilst the Promise context is strongly implicated in the predominance of errors in the hi condition of the memory task, the Threat context actually produces significantly less errors in the hi condition. Whilst the Tip and Warning contexts do not differ markedly in terms of the truth tables which they produce, and the Promise and Threat contexts are similar at least in that they attract an especially high level of E-

tables, it would appear on balance that these contexts are not paired in as strong and natural a fashion as was originally proposed.

The remaining hypotheses were concerned with X-tables and Y-tables. Hypothesis 5d (that there would be more X-tables in the Temporal context overall than in any other context) was strongly supported which, taken together with the evidence from Experiment 5 (with different materials), provides considerable support for the notion that the X-table is a genuine reflection of at least one third of all subjects' modal interpretation in this context. Just how the X-table might be especially related to this particular context will be considered later in the general discussion of these four experiments.

Hypothesis 5e (that the combined proportion of X-tables and Y-tables would be relatively constant over all eight contexts) was strongly supported across all but two of the contexts. In the Temporal context, the particularly high level of X-tables meant that any such effect would have been swamped, although the level of Y-tables is indeed lower than in any other context; and in the Promise context, the predicted effect has been virtually eliminated by the particularly high level of E-tables, which meant that very few responses remained to be otherwise classified.

Finally, Hypothesis 5f predicted that to the extent that the combined overall level of X-tables and Y-tables might differ from that in Experiments 5 and 6, then this would be reflected in a complementary change in the overall level of D-tables. Whilst the level of D-tables on the present experiment is very much lower than in Experiment 5, it is clear that this is reflected almost entirely in the increase in the level of E-tables, since the levels of X-tables

and Y-tables are virtually identical in the two experiments. It is clear from comparison with Experiment 6 (in which there were rather fewer X-tables and Y-tables than in either of the other experiments, but rather more D-tables) that the reduction in D-tables in the present experiment is only partially accounted for by the increase in E-tables, the remaining difference being accounted for by the increase in X-tables and Y-tables to their previous level. These comparisons appear to lend strong support to the notion that all these four tables are actually more or less defective versions of the same thing (the E-table): that is, that the D-table is a 'wholly defective' version of the E-table, while X-tables and Y-tables are merely rather less defective versions.

If this is the case, however, then it is necessary to account for this swing towards more complete truth tables between Experiment 6 and the present experiment, since (with the exception of the new Intentional context) precisely the same materials were employed - so that the only possible sources of this difference are the different subject pool and the introduction of the memory task. Experiments 5 and 6 used not only different subject pools, but also different materials: given that the 32% proportion of E-tables remained stable despite these differences, it seems more reasonable to suppose that the larger part of this difference must have arisen not from the use of different subjects but, rather, from the use of the memory task.

If D-tables, X-tables and Y-tables are, indeed, all more or less defective E-tables, then the present swing to E-tables might be explained in terms of more careful reading of the details of the text producing a less uncertain (and thus less defective) interpretation. On this basis, the 'canonical' interpretation of the conditional might thus be seen as the E-table irrespective of context, but to the extent that a context gives rise to *uncertainty*,

this will result in either or both the FT and FF cases being assigned the indeterminate value '?'. Where both cases are assigned the value '?', the D-table results: where it just the FT case then the Y-table results, and where just the FF case then the X-table results.

The observation that there are reliable and substantial context effects on the relative frequencies of X-tables and Y-tables suggests, moreover, that there must be certain features intrinsic to particular contexts which lead to more or less uncertainty about each particular truth case. A consideration of just what these features might be might thus be able to throw some light upon the very essence of these four types of truth table. Given the general importance of such a discussion to this whole series of experiments, it will be deferred until after the report of Experiment 8, which follows.

3.5 : Experiment 8

Experiments 6 and 7, and part of Experiment 5, employed a truth table evaluation task in which subjects had to indicate the extent to which given subsequent states of affairs supported, contradicted or told them nothing about given conditional statements. In considering the results from those experiments, it was found that the inferrable truth tables (derived from the patterns of response over all four possible states of affairs) were almost exclusively drawn from just four possible types - two 'traditional' tables, E-tables and D-tables, and two novel tables, X-tables and Y-tables.

As has been seen, to the extent that a subject 'possesses' a particular truth table for a given conditional statement, this should be reflected not only in the interpretation of a subsequent state of affairs in terms of its support for the statement, but also in the evaluation of a given inference on the basis of the statement.

Experiment 8 required subjects to evaluate the four 'classical' inferences, (MP, MT, DA and AC), over the range of contexts. Figure 3.4 (in the discussion of Experiment 5) set out clearly the relationship between the four truth tables and the four inferences, using the definitions of 'valid', 'invalid' and 'indeterminate' which were declared for the purposes of this thesis in Chapter 1. Figure 3.4 summarised the argument that in terms of the inferences supported, E-tables and X-tables were indistinguishable (both having all four inferences as valid); D-tables and Y-tables were argued to be similarly indistinguishable (both having MP and MT as valid, but DA and AC as indeterminate)[53].

[53] It was further argued that only the I-table of formal logic explicitly rejects

It was hypothesised that the results would provide corroborative support for the overall patterns of truth tables[54] found in Experiments 5, 6 and 7. Furthermore, since the materials were identical to those used in Experiment 5[55], it was hypothesised that the findings of the inference endorsement task in that experiment would be broadly replicated. These general hypotheses might be summarised as follows:

H1 a high level of DA and AC endorsement would be found in those contexts in which a high combined level of E-tables and X-tables was seen in the truth table tasks of Experiments 5, 6 and 7

H2 the endorsement patterns in the Promise and Threat contexts would be extremely similar to each other, as would those in the Tip and Warning contexts, such that these contexts would form respective 'natural pairs'

H3 the inference endorsement patterns would be similar to those found in the inference task of Experiment 5

These hypotheses may be restated in operational terms as follows:

H1a there would be a higher level of DA endorsement in the Temporal, Promise, Threat and Intentional contexts than in any of the other four contexts

H1b there would be a higher level of AC endorsement in the Temporal, Promise, Threat and Intentional contexts than in any of the other four contexts

H2a there would be similar DA endorsement rates on Tip/Warning and Promise/Threat respectively

H2b there would be similar AC endorsement rates on Tip/Warning and Promise/Threat respectively

---

the DA and AC inferences as invalid.

54 The inclusion of the two consequent-to-antecedent inferences, AC and MT, permits the underlying truth tables to be inferred directly. This was not the case in Experiment 5, which used only the two antecedent-to-consequent inferences, MP and DA, together with the (nonsensical) P->~Q and ~P->~Q.

55 (with the addition of the materials for the Intentional context from Experiment 7)

H3a the Temporal, Causal, Promise and Threat contexts would produce almost universal endorsement of the MP inference, the other contexts less so

H3b the Temporal, Causal, Promise and Threat contexts would produce a high level of endorsement of the DA inference, the Tip and Warning contexts a rather lower rate and the Universal context an especially low rate

3.5.1 : Design and Methodology

Subjects:

80 social science undergraduates at Plymouth Polytechnic (60 female, 20 male, mean age 21.5 years) participated in the experiment without extrinsic reward.

Materials:

Subjects were presented with a printed booklet consisting of one page for each of 32 trials, together with briefing and debriefing pages. The page for each trial consisted of a short passage of text which in every case ended with a quoted statement reported to have been said by one person to another. The statement was a conditional constructed in the manner described for Experiments 5, 6 and 7 so as to be an 'AA' double-affirmative of the form 'if P then Q'.

The 32 trials consisted of four texts belonging to each of the eight contexts previously described (Temporal, Causal, Promise, Threat, Tip, Warning, Universal and Intentional). Beneath the text and quoted statement were

shown two sentences which represented possible states of affairs and which were separated by the word "Therefore".

So as to make the task as natural as possible, the premises of the inference were, as before, expressed in a conversational style. Subjects had to indicate (by ticking a box) whether, on the basis of the passage and state of affairs shown, the conclusion followed or did not follow. The 32 texts used and the wording of the inferences associated with each are set out in full in Appendix [8a], but a typical task would be that shown in Figure 3.4 for the inference endorsement task of Experiment 5.

Since there was a possibility that endorsement/rejection of a given inference within a particular text might affect subsequent behaviour on a different inference for that text, it was decided that each subject would receive each text only once. For any given context, then, a given subject was presented with each of the four exemplars in that context exactly once. Similarly, within a given context a given subject was presented with each of the four inference types exactly once. In order to ensure that all the exemplars in every context were presented an equal number of times, each successive set of four subjects was 'yoked' such that by the time all four had been run, each of the four exemplars within a given context would have been allocated each of the four inferences exactly once.

This was achieved by allocating a latin square to each of the eight contexts within each successive set of four subjects. For any given top row of a 4 x 4 latin square - that is, some permutation of {1 2 3 4}) - there are exactly 6 possible ways of transforming it into the next three rows: since there are 24 (*i.e.* 4 !) possible permutations of {1 2 3 4}, there are 144 different 4 x 4 latin

squares. The latin square applied to each context for each set of four subjects was randomly selected from this possible set.

Procedure:

The frontispiece of the booklet contained briefing instructions, and whilst these are presented in full in Appendix [8b], that part of them concerned with the task itself is reproduced below:

> " ... 32 items to each of which you have to respond by ticking a box. Each item consists of a short passage of text which ends with a statement made by someone who is mentioned in the passage.
>
> Beneath the passage will be shown a sentence describing a state of affairs and a conclusion which may (or may not) follow from the passage and state of affairs described. You are required to indicate whether or not the conclusion does follow by ticking one of two boxes.
>
> It is vitally important that you read the passage carefully before responding ..."

The final page of the booklet explained the purpose of the experiment by way of debriefing, and is reproduced in Appendix [8c]. No formal timing was undertaken, but subjects all took between 15 and 25 minutes to complete the task.

3.5.2 : Results

The dependent variable is the endorsement or rejection of the four inferences across the eight contexts. The raw data are summarised in Table 3.24 which also shows the column rank for each cell (8=highest, 1=lowest).

Inference:	[MP] (P->Q)		[DA] (~P->~Q)		[AC] (Q->P)		[MT] (~Q->~P)		total endorsed	
	%	(rank)	%	(rank)	%	(rank)	%	(rank)	%	(rank)
Context:										
Temporal	90	(5)	66	(4)	81	(8)	74	(6)	78	(5)
Causal	93	(6.5)	90	(8)	71	(4.5)	88	(8)	85	(8)
Promise	89	(4)	74	(5)	78	(7)	54	(3)	73	(4)
Threat	93	(6.5)	84	(7)	71	(4.5)	84	(7)	83	(7)
Tip	60	(2)	40	(2)	41	(2)	38	(2)	45	(2)
Warning	75	(3)	55	(3)	55	(3)	64	(4)	62	(3)
Universal	56	(1)	38	(1)	34	(1)	33	(1)	40	(1)
Intentional	94	(8)	81	(6)	74	(6)	66	(5)	79	(6)
overall %	81		66		63		62		68	

Table 3.24 : Experiment 8:
Inferences endorsed x context (%)
[N per cell = 80 = 80 S x 1 exemplar]

Looking firstly at the MP inference, a 1-way ANOVA showed a highly significant effect of context ($F = 21.52$; $df = 7$; $p < 0.0001$). Consideration of the ranked pairs of means revealed two distinct clusters, such that MP endorsements on Temporal+Causal+Promise+Threat+Intentional (92% overall) are significantly greater than those on Tip+Warning+Universal (64% overall) ($t = 8.054$; $p < 0.0001$; 1-tailed). Additionally, it should be noted that the predicted natural pairings of the four inducement contexts was

refuted in that MP endorsements in the Tip context (60%) differed significantly from those in the Warning context (75%) ($t = 3.04$; $p = 0.007$; 2-tailed).

For the DA inference, a 1-way ANOVA showed a highly significant effect of context ($F = 21.394$; $df = 7$; $p < 0.0001$). Consideration of the ranked pairs of means revealed two distinct clusters, such that DA endorsements on Temporal+Causal+Promise+Threat+Intentional (79% overall) are significantly greater than those on Tip+Warning+Universal (44% overall) ($t = 7.437$; $p < 0.0001$; 1-tailed). Additionally, it should be noted that the predicted natural pairings of the four inducement contexts was refuted in that DA endorsements in the Tip context (41%) differed significantly from those in the Warning context (55%) ($t = 2.259$; $p = 0.036$; 2-tailed).

For the AC inference, a 1-way ANOVA showed a highly significant effect of context ($F = 16.369$; $df = 7$; $p < 0.0001$). Consideration of the ranked pairs of means revealed two distinct clusters, such that AC endorsements on Temporal+Causal+Promise+Threat+Intentional (75% overall) are significantly greater than those on Tip+Warning+Universal (43% overall) ($t = 6.959$; $p < 0.0001$; 1-tailed). Additionally, it should be noted that the predicted natural pairings of the four inducement contexts was refuted in that AC endorsements in the Tip context (41%) differed significantly from those in the Warning context (55%) ($t = 2.463$; $p = 0.024$; 2-tailed).

Finally, for the MT inference, a 1-way ANOVA showed a highly significant effect of context ($F = 23.526$; $df = 7$; $p < 0.0001$). Consideration of the ranked pairs of means revealed two distinct clusters, such that MT endorsements on Temporal+Causal+Promise+Threat+Intentional+Warning

(72% overall) are significantly greater than those on Tip+Universal (43% overall) ($t = 9.85$; $p < 0.0001$; 1-tailed). Additionally, it should be noted that the predicted natural pairings of the four inducement contexts was refuted not only in that MT endorsements in the Tip context (45%) differed significantly from those in the Warning context (62%) ($t = 4.098$; $p = 0.001$; 2-tailed), but also in that MT endorsements in the Promise context (73%) differed significantly from those in the Threat context (83%) ($t = 4.329$; $p < 0.001$; 2-tailed).

As for Experiment 5, there appears to be a consistent effect of context upon the overall tendency to endorse an inference - that is, there is a similarity in the column ranks across all four inferences. In particular, there is a consistent clustering together of the endorsement rates on the Tip, Warning and Universal contexts, with the Universal context receiving fewest endorsements regardless of the type of inference, closely followed by the Tip context.

Given this extremely strong effect of context across all four inferences, it is possible that any particular effects of context upon individual inferences would have been masked. So as to investigate this possibility, Table 3.25 represents these data in a normalised form - that is, each cell is shown as a percentage of the total number of endorsements of all four inferences for the context in question so as to show the relative effect of each context on each particular inference.

Inference:		[MP] (P->Q) %	[DA] (~P->~Q) %	[AC] (Q->P) %	[MT] (~Q->~P) %	total (100%) [n]
Context:	Temporal	29	21	26	24	[249]
	Causal	27	26	21	26	[273]
	Promise	30	25	26	18	[235]
	Threat	28	25	22	25	[265]
	Tip	34	22	23	21	[143]
	Warning	30	22	22	26	[199]
	Universal	35	23	22	20	[128]
	Intentional	30	26	23	21	[252]
	mean %	30	24	23	23	[100%]

Table 3.25 : Experiment 8:
Inferences endorsed x context (%)
[normalised wrt total endorsements in each context]
[N per cell = 80 = 80 S x 1 exemplar]

Before considering the result of the data normalisation it is necessary to consider how any effects remaining should be interpreted. To the extent that the tendency to endorse inferences in a given context is raised with respect to other contexts due to a general effect of that context (that is, an effect acting upon all four inferences equally), then the effect of normalisation would be to render the endorsement levels of a given inference equal across all eight contexts. To the extent that any contextual differences then remained, then it is clear that these would be the result of factors other than context. More specifically, to the extent that the *normalised* endorsement rate of a given inference in a given context was significantly lower than the rate for that inference in other contexts, then this would indicate that whilst the *actual* rate for that inference was stimulated by a

general effect of context, other factors were simultaneously depressing the rate, but this was masked by the general stimulating effect of context. Similarly, to the extent that the normalised endorsement rate of a given inference in a given context was significantly higher than the rate for that inference in other contexts, then this would indicate that whilst the actual rate for that inference was depressed by a general effect of context, other factors were simultaneously stimulating the rate, but this was masked by the general depressing effect of context.

It is apparent from Table 3.25 that there is, overall, relatively little effect of context on particular inferences - that is, that the context effect shown in Table 3.24 is very largely a general effect of the tendency to endorse the *whole range* of inferences within a given context. 1-way analyses of variance for each inference do, however, show that significant effects of context do remain in the normalised data for the MP ($F = 3.246$; $df = 7$; $p < 0.0033$) and MT ($F = 2.583$; $df = 7$; $p = 0.0158$) inferences, although these are obviously somewhat weaker than for the non-normalised data.

The particular effects of context on these two inferences, whilst significant, are rather hard to characterise, however, since no pairwise comparison of adjacently ranked normalised endorsement rates reaches significance - and there are no apparent clusters amongst the contexts. An important observation might, however, be made about MP inferences in the Tip, Warning and Universal contexts. Whereas the *actual* MP rates in these three contexts were considerably lower than in the other five contexts (Tip 60%, Warning 75%, Universal 56% *vs* 92% overall for the other contexts), the *normalised* MP rates in these three contexts are amongst the highest (Tip 34%, Warning 30%, Universal 35% *vs* 29% overall for the other contexts).

As stated above, this should be interpreted as meaning that the general depressing effect of these three contexts masked a stimulating effect of other factors. If it may be argued that the MP inference for double-affirmative conditionals is so easy to make as almost not to require reasoning at all, then it is not surprising that the general depressing effect of these three contexts exerted rather less influence upon MP endorsements than upon the other types of inference.

Additionally, given the 'natural pairing' hypothesis for the four inducement contexts, it is relevant to note that the normalised rate of MT endorsement in the Promise context (18%) differed significantly from that in the Threat context (25%) ($t = 3.109$; $p = 0.006$; 2-tailed) and that despite there being no overall effect of context upon the normalised AC endorsement rates, the rate in the Promise context (26%) differed significantly from that in the Threat context (22%) ($t = 3.164$; $p = 0.005$; 2-tailed).

3.5.3 : Discussion

The primary purpose of this experiment was to provide corroborative support for the overall patterns of truth tables found in Experiments 5, 6 and 7: specifically, it was hypothesised that a high level of DA and AC endorsement would be found in those contexts in which a high combined level of E-tables and X-tables was seen in those earlier experiments and that the four inducement contexts would form natural pairs.

Hypothesis H1a (that there would be a higher level of DA endorsement in the Temporal, Promise, Threat and Intentional contexts than in any of the

other four contexts) was supported reasonably strongly in as much as these four contexts produced very considerably higher rates than the Tip, Warning and Universal contexts. Hostile to the hypothesis, however, is the fact that the Causal context produced by far the highest DA endorsement rate of all (94%). Similarly, Hypothesis H1b (that there would be a higher level of AC endorsement in the Temporal, Promise, Threat and Intentional contexts than in any of the other four contexts) was strongly supported, with the exception of the Causal context having produced as many (but no more) AC endorsements as the Threat context (71%). Taken together, the general hypothesis that those contexts which produce high levels of E-tables and X-tables (both of which see all four inferences as necessarily valid) would also produce high levels of DA and AC endorsement was borne out by the present findings. With regards to the Causal context, whilst this context did not produce high levels of E-tables and X-tables in Experiments 5 and 6, it did do so in Experiment 7: furthermore, as will be seen below, the high levels of DA and AC endorsement in the Causal context are wholly consistent with the inference task of Experiment 5.

As regards the natural pairing of the four inducement contexts, Hypotheses H2a and H2b (that there would be similar DA and AC endorsement rates on Tip/Warning and Promise/Threat respectively) were strongly refuted in that all four inferences were significantly more often endorsed in the Warning context than in the Tip context, whilst the MT inference was significantly more often endorsed in the Threat context than in the Promise context. Further differences were revealed in the normalised data, in which the normalised MT endorsement rate in the Threat context was significantly higher than in the Promise context, whilst the reverse was the case for the AC rate.

In addition to these general hypotheses with respect to consistency with the earlier truth table data, it was also hypothesised that since the materials were identical to those used in Experiment 5, the findings of the inference endorsement task in that experiment would be broadly replicated. Hypothesis H3a (that the Temporal, Causal, Promise and Threat contexts would produce almost universal endorsement of the MP inference, the other contexts less so) was very strongly supported. Similarly, Hypothesis H3b (that the Temporal, Causal, Promise and Threat contexts would produce a high level of endorsement of the DA inference, the Threat and Warning contexts a rather lower rate and the Universal context an especially low rate) received strong support, with the trifling exception that the rate in the Tip context was very nearly as low as that in the Universal context.

The considerable support for Hypotheses H3a and H3b constitutes a strong replication of the findings of Experiment 5. The fact that these two hypotheses are couched only in terms of *rank*, however, does obscure the fact that the MP and DA endorsement rates overall for this experiment are rather lower than those for Experiment 5 (92% *vs* 81% for MP; 74% *vs* 66% for DA). Whilst the use of a fresh and differently motivated subject pool[56], together with a switch to manual presentation, renders this observation relatively uninteresting, it is surprising to note that the MP endorsement rate (seen earlier to be typically almost universal) was particularly low in three contexts: Tip (60%), Warning (75%) and Universal (56%). Furthermore, it is clear from the normalised rates for these contexts that the general depressing effect of context may actually have been underestimated due to the ease of making this particular inference.

[56] In Experiments 5, 6 and 7, subjects took part for course credits, whilst in the present experiment they received no extrinsic reward.

Overall, this experiment lends considerable support to the notion arising in the earlier truth table experiments that the context in which a conditional is uttered is a strong determinant of how the conditional relation will be interpreted. As has been seen, when the FT case is seen more often as determinately false (that is, in those contexts producing higher levels of E-tables and X-tables), the DA and AC inferences are necessarily valid - and those contexts (notably the Temporal, Promise, Threat and Intentional contexts) did indeed produce the highest rates of DA and AC endorsements. Conversely, when the FT case is rated more often as indeterminate, (that is, in those contexts producing higher levels of D-tables and Y-tables), then the the DA and AC inferences have been argued to be of indeterminate validity - and those contexts (notably the Tip, Warning and Universal contexts) produced the lowest rates of DA and AC endorsements.

It was argued earlier that the production of truth tables other than the E-table in a given context is a result of perceived uncertainty about the logical status of the FT and/or the FF truth cases. That the inference endorsement data in this experiment are so consistent with the truth table data of the preceding experiments strongly supports the notion that there must be certain features intrinsic to particular contexts which lead to more or less uncertainty about each particular truth case. Just what these features might be might will be considered in the general discussion of Experiments 5 to 8, which follows.

3.6 : General Discussion of Experiments 5 to 8

A number of interesting findings have emerged from these four experiments. The most important aspect of the truth table tasks of Experiments 5, 6 and 7 is, without doubt, the observation of substantial proportions of two novel truth tables, dubbed 'X-tables' and 'Y-tables'. The essence and reality of these tables will be considered at some length later in this section. Before doing so, various findings particular to just one or two of the experiments will be briefly summarised, having been discussed at some length earlier as and when they occurred.

3.6.1 : The four 'inducement' contexts

Experiment 5 constituted a systematic, if largely prospective, investigation of the effects of each of a coarse-grained taxonomy of natural language uses of double-affirmative conditionals phrased in 'if...then'. The experiment - and the associated taxonomy - was largely inspired by the work of Fillenbaum (for example, 1975; 1976) on conditionals phrased as 'inducements', and it was upon this work that such predictions as were made were based. Fillenbaum's work strongly implied that all inducements should give rise to a tendency to view the conditional relation as one of material equivalence, but this was seen to be far from the case, in that Tips and Warnings were thus evaluated somewhat seldom. It did appear, however, that Promises/Threats and Tips/Warnings were interpreted in a very similar fashion to each other, so forming respective 'natural pairs'.

The inference task of Experiment 5 was intended to provide corroborative

evidence for the findings of the inference task and was largely successful in that endeavour, in that the distribution of E-tables on the truth table task was generally consonant with the distribution of the DA inference necessarily endorsable under such tables. In addition, strong support was provided for the notion of the natural pairing of Promises/Threats and Tips/Warnings.

The general leaning of Promises and Threats - but not Tips and Warnings - towards an equivalence reading was strongly supported by the truth table tasks of Experiments 6 and 7, and by the inference task of Experiment 8. The natural pairing of Promises/Threats and Tips/Warnings did not, however, fare so well in those later experiments, a number of comparisons between the respective partners revealing significant differences on various specific measures - although it was always the case that Promises/Threats differed very significantly from Tips/Warnings on every kind of comparison.

3.6.2 : The subjective non-contingent likelihood of the consequent

The construction of the taxonomy of conditional usage identified a number of ways in which particular groups of contexts might be distinguished, and one in particular - the subjective non-contingent likelihood of the consequent - appeared to have a considerable effect upon the truth tables produced in Experiment 5. On this basis, Experiment 6 was designed to manipulate this feature systematically, as well as to provide a replication of the truth table findings of Experiment 5 with a fresh set of materials. Whilst Experiment 6 provided a strong replication of the differential occurrence of X-tables and Y-tables, it failed signally to show any effect of this

manipulation, despite its having been relatively gross. It was considered that the most likely reason for this lack of effect was that subjects had ignored the texts in which the manipulation was made, responding solely on the basis of the conditional itself.

An attempt was made in Experiment 7 to remedy this problem by obliging subjects to carry out an additional memory task about which they were warned in advance. Whilst it appeared from the results of the memory task that subjects had indeed paid close attention to the texts, the manipulation of the subjective non-contingent likelihood of the consequent was no more effective than in Experiment 6, thus leading to the conclusion that the probabilistic features introduced into particular scenarios had been overriden by features typical of particular contexts in general - that is, that subjects' responses were determined by the contexts themselves.

3.6.3 : Speaker's and hearer's control

In addition to re-examining the effect of manipulating the subjective non-contingent likelihood of the consequent, Experiment 7 sought (by the introduction of the Intentional context) to investigate factorially the effect of an alternative distinguishing feature in the taxonomy of contexts - the speaker's and hearer's control over the truth of the propositions.

The speaker's control over the truth of the consequent (which is present in the Promise, Threat and Intentional contexts) was seen to lead strongly to the conditional relation being interpreted as one of equivalence. Whilst the hearer's control over the truth of the antecedent (which is present in the Promise, Threat, Tip and Warning contexts) did not have any great effect

on its own, there was some suggestion that it might be interacting with the speaker's control, so that when both parties were in a position to exercise their respective kinds of control (in the Promise and Threat contexts), then the tendency to equivalence was very much higher. Whilst the absence of the Intentional context precluded this factorial analysis of the data from Experiments 5 and 6, *post hoc* pairwise analyses of the three out of four cells available provided considerable support for the varying effects of control by the respective parties over the truth of the propositions.

3.6.4 : X-tables and Y-tables - genesis and reality

The findings summarised above are, in the main, related specifically to particular experiments. What is unquestionably the most important finding to come out of this series of experiments, however, is the highly reliable occurrence of the two novel truth tables - X-tables and Y-tables.

The remainder of this section will be concerned with the extent to which it is reasonable and proper to view these as genuine reflections of subjects' underlying interpretations rather than as quirks of the data. It is not sufficient for this purpose merely to have demonstrated that the appearance of these tables is reliable: if they do reflect genuine interpretations, then an account must be offered of how they might arise.

The appearance of X-tables and Y-tables has been extraordinarily reliable throughout this series of experiments, accounting between them for 35% of the truth tables in Experiment 5, 21% in Experiment 6 and 31% in Experiment 7. Quite apart from the considerable proportion of the data for

which they account, their reality is supported by the observation that the level at which they are seen is highly, and reliably, dependent upon the context in as much as the X-table is by far at it most prolific in the Temporal context, constituting around one third of all response patterns in that context and occurring almost as often as the E-table in Experiments 5 and 6 - and more often than any other table in Experiment 7. Furthermore, these effects were seen to be highly replicable not only with different subjects but also with different materials.

If it is to be argued that the X-table is a genuine reflection of some subjects' interpretation of the situation obtaining in the Temporal context, then it must be shown that such an interpretation is a reasonable one to make in such scenarios. In the X-table, the FF case is indeterminate, whilst the FT case is assigned the determinate value 'false'. Consider the example "If the next train is for Wimbledon then the one after it will be for Ealing". The FT case here is that the next train is for somewhere other than Wimbledon, but the one following that train is indeed for Ealing.

It was argued in the discussion of Experiment 5 that this particular linguistic form would be unlikely to be selected by the speaker unless they wished to communicate their belief that neither train was likely to be paired with any other: on this basis, the FT case might quite reasonably be assigned the value 'false', since we have a situation wherein one of the trains (that for Ealing) is temporally paired with a train other than one for Wimbledon. The FF case is that the next train is for somewhere other than Wimbledon, and the one following that is for somewhere other than Ealing. It was argued in the discussion of Experiment 5 that in order to be a relevant utterance the selection of this particular linguistic form would also imply

that the hearer would be likely to observe the relationship in the relatively near future. This is by no means to say, however, that the relationship is likely to be observed on *precisely* the next two trains to arrive, but rather that two suitable trains are likely to be along shortly. By the use of the words "... the *next* train ..." and "... *the one after it* ...", it is argued that the speaker is not limiting their prediction to the *first two* trains to arrive but rather is, in fact, communicating two related pieces of information: (1) that the temporal pairing reliably exists and (2) (in virtue of the Gricean requirement to be relevant) that it is likely to be become observable in the near future. The arrival of a disparate pair does not, therefore, contradict the statement, since it by no means excludes the arrival of the pair in question in the near future. It is clear that on this basis, the FF case (the arrival of the disparate pair) is merely *irrelevant:* and as such leaves the subject uncertain as to the value which should be assigned to it. It is argued that the X-table is, accordingly, an eminently reasonable response pattern in the Temporal context.

It is also clear from such an analysis, however, that the Temporal context is of a somewhat different kidney from the other contexts in terms of the point at which the hearer might expect to be able to observe the relationship between antecedent and consequent. In every other context, the subsequent assertion that the antecedent did not happen implies that *for the purposes of testing the truth of the conditional statement* the antecedent should be taken as being false. The utterance, for example, in the Causal context of "The lorry was not over the legal weight limit" or, in the Promise context, of "He did not wash the car" would simply not be relevant unless it related specifically to the testing of the conditional statement.

As has been seen, however, the assertion that "The first train was not for Wimbledon", whilst relevant to the *exchange* (given the fact that the

conditional statement referred to "... the first train ...") is not relevant to the *testing of the statement,* since it does not preclude the imminent opportunity to test the statement by virtue of the arrival of the train in question. The X-table has, then, been shown to be a reasonable interpretation when the antecedent and consequent are seen as reliably co-occurring, but the assertion that the antecedent did not occur does not imply that there will not be an imminent opportunity to test the statement - and thus does not imply that the antecedent is thereby 'false' for the purposes of testing the truth of the conditional.

It is also clear that an important distinction might be made between the Temporal context and each of the other seven contexts in terms of what it means to deny the antecedent. In the Temporal context, to deny that an event occurred is to imply that some other similar event did occur: for example, to deny that the first train was for Wimbledon is to imply that a train did arrive but that it happened to be for somewhere other than Wimbledon. In each of the other contexts, however, the denial of the antecedent implies nothing whatever about the occurrence of any other event which would be relevant to the exchange: for example, to deny that Henry washed the car is by no means to imply that Henry performed any other act which would be relevant to the exchange.

In order for this account in terms of the effect of context upon the uncertainty of particular truth cases to hold water, however, similar analyses must be offered not only for the particular uncertainty in D-tables and Y-tables, but also for the lack of such uncertainty in E-tables. It is clear, however, that the uncertainty present in truth tables might be of two distinct types. The first type of uncertainty - which might be termed 'antecedent

uncertainty' - was described above for the Temporal context and is uncertainty about whether a given (antecedent) event should necessarily be taken to mean that the conditional is about to be tested. The second type of uncertainty - which might be termed 'consequent uncertainty' - reflects 'fuzziness' in the consequent. The E-table - which has been argued to be the 'canonical' interpretation of 'if P then Q', implies the strict co-occurrence of P and Q. It is clear that in certain contexts (notably Tips and Warnings) the relationship between antecedent and consequent might be rather less strict, or fuzzy, in that whilst we might take the observation of P to imply Q, we will not take the observation of Q in the absence of P to mean that the conditional is untrue, or that its utterance is in breach of the Gricean maxims.

The four truth tables might be distinguished on the basis of which, if either, of these types of uncertainty is introduced through the effects of context (although to distinguish D-tables and Y-tables will be seen to require in addition a consideration of the degree of consequent uncertainty). Where no uncertainty of either type is introduced by the context, then this leads to the E-table, since P and Q are expected strictly to co-occur. Where antecedent uncertainty alone arises then this leads, as we have seen, to the X-table - and this particular sort of scenario appears intuitively quite familiar where the conditional is about a purely temporal relationship.

Let us now consider the effect of consequent uncertainty, which will be shown to lead to D-tables or Y-tables. The D-table occurred consistently most often in the Tip and Universal contexts, and an example will be considered from each to demonstrate how the D-table might come about. Consider the Universal statement "If the student if doing Economics then he's a

Socialist", which is a conditional expression of the categorical universal "All Economics students are Socialists". It was argued earlier that the selection by the speaker of the conditional form to express a categorical universal strongly implies that there is fuzziness in the consequent, such that the statement might be seen as meaning something like "one way of knowing that a student is a Socialist is to ascertain that they are doing Economics". In these circumstances, the FT case (the observation of a Socialist who is not studying Economics) is no more relevant to the truth of the conditional than is the FF case (the observation of a non-Socialist who is not studying Economics), since non-Economics students simply do not permit the statement to be tested. Where both the false-antecedent cases are perceived as irrelevant due to perceived fuzziness in the consequent, then the D-table results.

For the Tip context, consider a statement made in a crowded bar "If you stand by the pillar then you'll be served immediately", for which the FT case is that the hearer stayed put but was nevertheless served immediately thereafter, whilst the FF case is that they stayed put and were not served immediately thereafter. It was argued earlier that if the speaker wished to convey that standing by the pillar was the *only* way of getting served immediately, then they would be more likely to have phrased the statement along the lines of "the only way to get served immediately is to stand by the pillar". Here again, it is argued that the very selection of the conditional form is sufficient strongly to imply that there is fuzziness in the consequent so that the FT and FF will be perceived as irrelevant to the truth of the statement and the D-table will result.

It is proposed that the Y-table will arise in exactly the same way as the D-table, but only where the degree of consequent uncertainty is especially

slight. Since this table arose consistently more often in the Warning context than in any other, an example will be taken from that context (although it should be noted that even in that context the Y-table occurs relatively seldom, by far the most frequent truth table being the D-table). Consider a statement made by a dog owner to someone who had pulled their dog's tail "If you pull his tail again then he'll bite you", in which the FT case is that the hearer didn't pull the dog's tail yet was bitten, whilst the FF case is that the hearer didn't pull the dog's tail and wasn't bitten. It was argued earlier that such a statement conveys not only the speaker's belief about one relevant way in which the hearer might get bitten but also that the dog is disposed to bite people: it was also argued that the selection by the speaker of the conditional form strongly implies that other ways of getting bitten might exist, since otherwise a more suitable statement would be something like "he *only ever* bites people when ...". In these circumstances, observation of the FT case (that the hearer was bitten without having pulled the dog's tail) will be perceived as irrelevant to the truth of the conditional.

Whether the D-table or Y-table will result, however, depends upon the assessment of the FF case (that the hearer didn't pull the dog's tail and wasn't bitten): in order for the Y-table to result in place of the D-table, the FF case must be perceived as relevant (and as confirming the truth of the conditional). It is proposed that when the fuzziness is perceived as being particularly low (in this case, when the statement is taken to mean that it is highly unlikely that anything will dispose the dog to bite the hearer other than getting his tail pulled) then the relationship between antecedent and consequent will be considered to be one of equivalence *by default,* so that the FF case confirms the truth of the conditional (and is thus relevant). In that a very small degree of fuzziness is perceived, however, FT counterexamples,

whilst not *expected*, will still be *tolerated*, so that the FT case is not seen as falsifying the statement, but rather as merely incapable of positively determining its truth.

3.6.5 : E-tables - the canonical interpretation?

The foregoing account of how the four truth tables might arise in terms of uncertainty is clearly consistent with the notion that the E-table is the canonical interpretation of the conditional whilst D-tables, X-tables and Y-tables are its variously 'defective' progeny. Under this account, E-tables were argued to arise in the absence of perceived uncertainty either in the antecedent (in terms of whether the statement had been put to the test) or in the consequent (in terms of fuzziness). D-tables were argued to arise when there was a non-trivial degree of consequent uncertainty, such that the FT and FF cases were interpreted as no longer determinate and thus as useless for the purpose of testing. Y-tables, on the other hand, were argued to arise in the presence of an especially small degree of consequent uncertainty, such that FT cases would not be expected, but would nevertheless be tolerated. Finally, X-tables were argued to arise in the absence of consequent uncertainty, but where there was uncertainty about whether an (antecedent) observation signalled a test of the conditional.

In these terms, the effect of uncertainty might be characterised as causing the usefulness of one or more of the truth cases of the canonical E-table to 'decay'. Where the level of uncertainty is especially low (as in the Y-table) or is not about the conditional relation itself (as in the X-table), then the decay produces truth tables which are only *partially* defective; where the

level of uncertainty about the conditional relation is rather higher, then the *wholly* defective D-table results.

If this account is correct, then the use of the label "defective implication" to describe the D-table would appear to be somewhat dubious. Furthermore, if the D-table were to be a defective or decayed version of the I-table of material implication, then one would expect subjects to produce not only I-tables, but also a *partially* defective I-table which had the FT case as true and the FF case as irrelevant - and such is simply not the case. It has been seen that in the present series of experiments, the I-table only occurs twice out of 2000 response patterns and it is thus extremely reasonable to suppose that these two occurrences are not reflections of a genuine underlying interpretation at all, but rather just an instance of an *unclassifiable* response pattern. Since it seems that subjects never (or effectively never) respond on the basis of the I-table, it is clearly a nonsense to suppose that the D-table is its defective progeny, since it is inconceivable that the defectiveness would only be relieved in 0.1% of the trials. On these grounds it is proposed that to suggest that D-tables (or, indeed, Y-tables) have anything to do with I-tables is perhaps somewhat bizarre.

It is not, however, defensible merely to knock down an established view without providing a more attractive alternative. The alternative suggestion, that all three variously defective truth tables are allied to the E-table, arose originally out of the observation that X-tables and Y-tables might be simply rather less defective versions of the D-table. In so far as X-tables could only be allied to the E-table (due to the FT case being evaluated as false), however, the notion that all three uncertain truth tables were more or less defective versions of one 'pure' interpretation necessarily meant that this 'pure' interpretation must be the E-table. At this level, however, the argument is

purely descriptive, and might be viewed as having no more validity than the established view of the I-table as the 'pure' interpretation, except in so far as it avoids the embarrassment of the X-table being inconsistent with the I-table and of the I-table never occurring.

The explanation of behaviour in terms of various types of uncertainty inherent in particular contexts has, however, been shown to support the proposal of a *mechanism* whereby the E-table might 'decay' into the variously uncertain D-tables, X-tables and Y-tables. It has thus been shown to be possible within this framework to explain the genesis of each of these three truth tables rather than merely to describe them. On these grounds, it is clearly reasonable to argue that the suggestion that E-tables are the 'naturally canonical' interpretation of conditionals phrased in 'if P then Q' in these real world contexts is on rather firmer ground than the established (largely descriptive) view under which I-tables are cast in this role. Furthermore, the notion that people may tend towards an E-table interpretation by default is consistent with Geis and Zwicky's (1971) suggestion that "conditionals are understood to be perfected[57] unless the hearer has reason to believe that the converse is false" (*ibid.*, p.562).

3.6.6 : Uncertainty and control

Finally, it should noted that the uncertainty account is also consistent with the observed effects of the hearer's control over the truth of the antecedent and the speaker's control over the truth of the consequent. It has been seen

[57] Geis and Zwicky's principle of 'Conditional Perfection' holds that a sentence of the form P -> Q suggests, or invites the inference of, ~P -> ~Q

that the speaker's control over the truth of the consequent (in the Promise, Threat and Intentional contexts) does appear to lead to a greater tendency towards equivalence, but that there is no significant effect of the hearer's control over the truth of the antecedent.

The first type of uncertainty identified was in terms of the perceived relevance of an antecedent event for the purposes of testing the truth of the conditional statement, and was argued to arise mainly in the Temporal context. The second type of uncertainty was in terms of fuzziness in the consequent. Since the speaker is in control of the truth of the consequent in the Intentional, Promise and Threat contexts, however, it is clear that the utterance of a conditional would be grossly misleading where plausible and relevant alternative events might follow the antecedent. This is particularly true in the Promise and Threat contexts, in which the statement is uttered in order to modify the hearer's behaviour, so that the omission from the utterance of alternative relevant consequences of the antecedent would make the speaker less likely to succeed in their purpose. On this basis, it is reasonable for the hearer to assume that no such alternative consequences exist (or, at least that no such alternative consequences are plausible and relevant), and thus to interpret the consequent as free of fuzziness. The effect of the speaker's control over the truth of the consequent is, accordingly, explicable in terms of the reduction of uncertainty rather than in terms of control *per se*.

Chapter Four

Synthesis and Conclusion

4 : Synthesis and Conclusion

On the face of it, the two sets of experiments reported in Chapters 2 and 3 could hardly have been more disparate, in that Experiments 1 to 4[58] employed abstract materials (with the aim of avoiding any possible effects of context), whereas Experiments 5 to 8 used richly thematic materials (with the aim of evoking particular contexts). The purpose of this final chapter is not, therefore, to identify the differences between the two sets of experiments, since these would be signally unsurprising. The aim is, rather, to identify certain common ground in the respective findings and conclusions so as to integrate the two sets of experiments. In addition, consideration will be given to what the results of these experiments have to say about the various issues raised in Chapter 1.

There appear to be two major areas of common ground between the two sets of experiments. The first is the question of fuzziness, which was addressed explicitly in the abstract experiments, and which played a central part in the explanation advanced for the context effects found in the thematic experiments; the second is the common ground in the explanations advanced for non-logical performance in these tasks in terms of task difficulty - or, more specifically, in terms of the extent to which a subject's familiarity with the whole scenario will facilitate performance.

58 These will be referred to hereafter simply as the 'abstract experiments', whilst Experiments 5 to 8 will be referred to as the 'thematic experiments'.

4.1 : Fuzziness

The abstract experiments sought to investigate the extent to which subjects would expect or tolerate fuzziness. Whilst there was little evidence of subjects having spontaneously generated fuzzy representations of the conditional in the abstract construction tasks they did, however, appear to tolerate a certain (small) level of fuzziness in the abstract evaluation tasks - and it was clear that such tolerance was not merely the result of subjects failing to notice the counterexamples. As was argued earlier, the fact that people recognise that a conditional statement is reasonably made about a large array of items even though it is shown to be incorrect very occasionally is by no means to say that those people will spontaneously generate items about which the conditional is false. Indeed, Steve Newstead (personal communication) has suggested that this is true of much of language: whilst in evaluating situations, people will quite happily tolerate deviations from what is supposedly 'correct', they will nonetheless endeavour to produce what they see as the 'correct' forms when required to generate examples.

In the discussion of the contexts experiments, it was argued that the Y-table (which constitutes around one in six response patterns overall) was the result of the scenario somehow suggesting that a small degree of fuzziness in the consequent should be tolerated. It should be noted, however, that the counterexamples of interest in the abstract experiments were the TF case, whilst those responsible for the Y-table are the FT case. In terms of their effect, this meant that in the abstract experiments, the tolerated TF fuzziness caused an array to be evaluated as 'less than absolutely true': in the thematic experiments, however, fuzziness was argued to cause the Y-table by making

the FT case unexpected yet tolerable (and thus indeterminate) whilst still allowing the FF case to prove the truth of the conditional.

Despite these differences, the notion of fuzziness is similar in each case in that it refers to an unexpected, yet tolerable, degree of deviation of the situation from what was expected. That people are happy to tolerate - or even to expect - such fuzziness is intuitively unsurprising. As Kyburg (1983) points out, if an abstract conditional is perceived as a universal generalisation, then peoples' knowledge that the real world supports relatively few genuine universal generalisations will lead them to treat 'if something is an X then it is a Y' as representing 'almost all Xs are Ys'. As was argued earlier, a similar case may be argued for thematic conditionals in 'real life': the statement "if you go through to the Public Bar then you'll be served straight away" seems to be a 'shorthand' for "if you go through to the Public Bar then *it is, in my opinion, statistically probable that* you'll be served straight away", such that the hearer will often be aware that the logical truth of such a statement is less than 100% certain, but will nonetheless act upon it *as if* it were absolutely true.

4.2 : Familiarity and Facilitation

It was argued earlier that the context in which the abstract experiments was presented was necessarily simply that of reasoning with abstract symbols in an experimental laboratory situation. The advantages and disadvantages of investigating conditional reasoning in this way have been presented at length already and so will not be rehearsed again here. It is clear, however, that in the absence of formal training in the conditional, a subject must rely

for the solution of these tasks upon knowledge which has been acquired by experience of usage of the conditional in their everyday life. To the extent that such knowledge might provide a number of alternative interpretations of the situation - as implied, for example, by Cheng & Holyoak's (1985) 'pragmatic reasoning schemas' account - then the subject must somehow be able to decide which of these (if any) might properly be applicable. It was argued earlier that reasoning with abstract symbols in the absence of any familar context might thus be expected to be somewhat haphazard.

In the abstract experiments, the absence of any context is intended to convey that the relationship between antecedent and consequent is strictly one of set inclusion: in particular, great care is taken to ensure that no question of any temporal interdependence is implied. The taxonomy offered in Chapter 3 of the contexts used in the thematic experiments, however, makes it clear that in all contexts except one, the relationship between the antecedent and consequent is strictly temporal in that the antecedent event always precedes the consequent event. The exception to this general rule is the Universal context, in which great pains were taken to make it clear that the relationship between antecedent and consequent was strictly one of set inclusion.

On this basis, it is probably only the Universal context which might sensibly be compared across the two sets of experiments. Whilst the difference in the types of task in each set of experiments means that results are not directly comparable, it is clear from the abstract experiments that subjects' responses were frequently erroneous and that they often appeared to be determined by extra-logical factors - and it is, of course, well-established in the literature that people are simply not particularly good at abstract reasoning tasks.

If a simple lack of thematicity is behind the poor performance on abstract tasks, then the difficulty ought to be eradicated by merely making the antecedent and consequent refer to members of realistic sets in a realistic scenario - as in the Universal context of the thematic experiments. Whilst there is no single 'correct' answer for the truth table tasks of the thematic experiments, it is clear from the results of the inference tasks of Experiments 5 and 8 that considerable difficulty appears to remain. In Experiment 5, for example, only 81% of MP inferences were correctly endorsed, whilst in Experiment 8 this was a mere 56% - and in both cases, the failure to endorse MP was by far at its highest in this context. It is interesting to note, however, that by far the lowest endorsement rate was seen for each of the other inferences in the Universal context also, suggesting that subjects were somehow particularly uncertain about the relationship which obtained.

This finding is, of course, entirely consistent with Cheng and Holyoak's (1985) account in terms of facilitation being effected by the evocation of a suitable 'pragmatic reasoning schema' from which the 'correct' inferences will follow from application of the operations contained therein. In the case of failure to evoke such a schema, subjects are obliged to rely upon extra-logical features of the task (as in 'matching bias'), or to guess - and when they must say whether a given inference necessarily follows, then it is unsurprising that they will be more likely to say that it does not.

As was seen in Chapter 1, Cheng and Holyoak (1985, p.396) specifically stated that arbitrary rules about abstract symbols would fail to evoke any reasoning schemas. More important to the present discussion, however, is their finding that the use of realistic thematic content in a realistic scenario

only facilitated performance where the relationship between antecedent and consequent was explicity declared to be one of 'permission' or 'obligation' and thus was able to evoke a suitable schema.

This is entirely consistent with the argument advanced earlier that the conditional is simply not a natural way in which to express universal relationships: whilst grammatically correct so to do, it might be argued that their unnaturalness means that subjects simply do not know how to deal with them. This is, of course, strongly corroborated by the observation that it was particularly difficult in constructing the materials for this context to produce statements and scenarios which sounded natural and as if they had actually occurred. It may be be argued that the fact that the conditional appeared to be a somehow unnatural form in which to express many of the scenarios used in this series of experiments suggests that its use in preference to some more natural form of expression might sometimes have been perceived to communicate something beyond the mere antecedent-consequent relationship. As was seen in Chapter 1, it is well established that the making of an utterance can often constitute a piece of behaviour which goes beyond the communication of its propositional content. Of particular relevance here is Grice's (1975) notion that conversation proceeds according to a series of culturally shared maxims and that apparent failure to observe these maxims - for example, by appearing to to say something irrelevant, or untruthful, or which gives more or less information than is required - will often lead to the hearer perceiving the communication of additional information through the mechanism of 'conversational implicature'. In these terms, it seems reasonable to suppose that the use of an 'unnatural' linguistic form might be taken as an apparent failure to be relevant, or as stating more than is required, such that a subject might then interpret the

relationship between antecedent and consequent in a very different way from that which was intended, or from that which their experience would lead them to expect in such a scenario.

There is no reason to suppose that subjects in reasoning experiments do not generally strive sincerely to make sense of the materials with which they are presented. Where such materials involve the use of an intuitively unnatural expression, then in coming to an overall understanding of the situation described, we might expect them to bring to bear the full panoply of their logical and linguisitic knowledge - and, where thematic materials are involved, their personal knowledge of the propositions and scenario described, or of analogous propositions and scenarios. The linguistic knowledge brought to bear is not, however, limited to the formal rules of syntax and semantics - it necessarily includes experiential knowledge of how people typically express typical relationships between propositions of a given type in a particular sort of scenario. Where such knowledge suggests that the conditional is an unexpected form of expression, then subjects must come to an interpretation of how its use might suggest that the relationship between antecedent and consequent is somehow special, or unexpected - and, in as much as subjects' experiential knowledge will be highly idiosyncratic, then we should not be surprised by the considerable range of responses between subjects.

Just what the results of these experiments have to say about how people reason in everyday life is, of course, open to question. As was seen earlier, Braine (1978) argues that reasoning experiments place additional demands upon subjects: firstly, to "compartmentalize" information by restricting that used in their reasoning to just that explicitly contained in the premises and, secondly, actively to disregard any implicit information which arises out of

such mechanisms as Grice's (1967) 'conversational implicature' or Geis and Zwicky's (1971) 'invited inference'. Furthermore, Politzer (1986) suggested that there may be "domains of application of the individual's logical-linguistic equipment" such that, for example, one would not wish to use the same rules when testifying in court (when logical conventions might lead to deception) as when taking part in a reasoning experiment (when linguistic conventions might lead to fallacy); similarly, Scribner (1977) suggested that when the format in which material is presented does not "belong to the subjects' repertoire", then they will tend to answer entirely on the basis of pragmatic laws, irrespective of their logical ability. That the manipulation in the thematic experiments of just those things which Braine suggested that subjects should ignore *did* have such a reliable effect, however, makes it clear that subjects did bring a rich panoply of their world knowledge to bear upon the tasks - and there might, thus, be grounds for arguing for some degree of ecological validity for those particular experiments. There is, however, no escaping the fact that any study carried out under laboratory conditions is necessarily highly removed from subjects' everyday life. Perhaps the most serious flaw in any such study (one which is seldom acknowledged in the literature) is the fact that if people are left uncertain about the meaning of an utterance in everyday life then they will typically have (and grasp) the opportunity to ask for clarification. The absence of such an opportunity in the laboratory necessarily means that they are reasoning with less information than they would normally require and there is no reason to suppose that those very different circumstances would lead to use of the same mechanisms of reasoning: indeed, the mechanisms of everyday reasoning might thus be simply unable to operate in such circumstances.

4.3 : Logical Competence and Rationality

As was seen in Chapter 1, the study of human reasoning has given rise to considerable debate about the extent to which people are logically competent - and whether their behaviour on reasoning tasks has anything to say about their rationality. It was argued in that chapter that logical competence might only be tested in a non-circular fashion where the performance system which produces the actual behaviour might be fully specified - failure to solve a problem on some occasion cannot be taken to imply that the subject is not competent to solve it; and neither may be it be assumed that a correct solution has not arisen through the application of entirely extra-logical heuristic strategies, or that is not merely a lucky guess.

The results from these experiments are far from unusual amongst such studies in that they abound with responses which are at odds with formal logic. In the abstract experiments, very few subjects correctly evaluated 'fuzzily true' statements as false and almost all subjects were happy to entertain the notion of degrees of truth, a notion which is anathema to formal logic; similarly, the selection of items in the construction task was shown to be highly subject to extra-logical factors such as availability and matching status, leading at times to the selection of items which were logically incorrect. Similarly, the thematic experiments showed that subjects were highly - and systematically - susceptible to extra-logical factors such as the linguistic context in which the utterance was stated to have been made in both the truth table and inference tasks.

It was recognised earlier that can be no absolute resolution of the 'rationality

debate', since it is ultimately a matter for philosophical rather than scientific, argument. Nonetheless, the results of the thematic experiments have clearly demonstrated that a rich panoply of world knowledge is brought to bear in the solution of problems in scenarios familiar to subjects. It seems difficult to reject the notion that in such circumstances subjects may very well not be reasoning at all but, rather, merely fitting the scenario to what they already know. It matters not whether the mechanism for so doing is couched in terms of domain-specific 'schemas' (Mandler, 1980), domain-independent 'pragmatic reasoning schemas' (Cheng and Holyoak, 1985), or 'reasoning by analogy' (Griggs, 1983): common to all such approaches is the idea that reasoning as such may often simply be unnecessary in everyday life. Indeed, to fly in the face of our experiential knowledge about the likely entailments of a given situation by applying logical reasoning may often be highly maladaptive.

4.4 : Conclusion

There is little virtue in repeating the detailed conclusions contained in Chapter 2 (with respect to the abstract experiments) and Chapter 3 (with respect to the thematic experiments). One or two brief comments might, however, be made to characterise the general flavour of the findings. Firstly, it is clear that people simply are not particularly good at, or particularly consistent in their behaviour on, conditional reasoning tasks in the laboratory. It is clear, however, that both accuracy and consistency can be enhanced by the provision of rich contexts capable of evoking conditional situations which are at least analogous to subjects' own experience.

Secondly, questions have been raised about the advisability and generalisablity of investigating conditional reasoning using contingent universal statements even within rich and realistic contexts, since it seems likely that even these are somewhat unrelated to subjects' own experience, in so far as the conditional form is simply not generally used to refer to such relationships and might even appear somewhat bizarre. Ironically, it has often tended to be the more natural contexts (such as Promises and Threats) which the literature has caused to be tinged with doubt, whereas the reality of the situation seems to be that it is in situations such as these that the conditional is most often actually used in everyday life.

Finally, the fact that subjects appear to be so happy to accept, and to reason about, situations in which fuzziness obtains in the antecedent, or the consequent, or in the relationship between them, tends to suggest what the writer suspected all along - that life is inherently an altogether fuzzy business.

References

ACKERMAN, B.P. (1978a). Children's comprehension of presupposed belief: Logical and pragmatic inferences to speaker belief. *Journal of Experimental Child Psychology,* **26**, 92-114.

ACKERMAN, B.P. (1978b). Children's understanding of speech acts in unconventional directive frames. *Child Development,* **49**, 311-318.

ACKERMAN, B.P. (1979). Children's understanding of definite descriptions. *Journal of Experimental Child Psychology,* **28**, 1-15.

ANDERSON, J.R. (1981). Memory for logical quantifiers. *Journal of Verbal Learning and Verbal Behaviour,* **20**, 306-321.

ATLAS,J.D. & LEVINSON, S.C. (1981). It-clefts, informativeness and logical form: Radical pragmatics (revised standard version). In P. Cole (ed.), *Radical pragmatics.* New York: Academic Press.

AUSTIN, J.L. (1962). *How To Do Things With Words.* Oxford: Oxford University Press.

AYER, A.J. (1936). *Language, truth and logic.* London: Gollancz.

BAR-HILLEL, Y. & CARNAP, R. (1953). Semantic information, *British Journal for the Philosophy of Science,* **4**, 147-157.

BASS, B.M., CASCIO, W.F. & O'CONNOR, E.J. (1974). Magnitude estimations of expressions of frequency and amount. *Journal of Applied Psychology,* **59**, 313-320.

BATES, E. (1976). *Language and context: The acquisition of speech pragmatics.* New York: Academic Press.

BETH, E. & PIAGET, J. (1966). *Mathematical epistemology and psychology.* Dordrecht: Reidel.

BLOOM, A. (1981). *The linguistic shaping of thought: A study in the impact of language on thinking in China and the West.* Hillsdale, NJ: Erlbaum.

BOCK, J.K. & HORNSBY, M.E. (1981). The development of directives: How children ask and tell. *Journal of Child Language,* **8**, 151-163.

BOOLE, G. (1854). *An investigation of the laws of thought.* London: Walton & Maberly.

BOOLOS, G. (1984). On 'syllogistic inference'. *Cognition,* **17**, 181-182.

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

BRAINE, M.D.S., REISER, B.J. & RUMAIN, B. (1984). Some empirical justification for a theory of natural propositional logic. In G.H. Bower (ed.), *The psychology of learning and motivation (Vol. 18).* New York: Academic

Press.

BROWN, J.S. & DE KLEER, J. (1981). Towards a theory of qualitative reasoning about mechanisms and its role in troubleshooting. In J. Rasmussen & W.B. Rouse (eds.), *Human Detection and Diagnosis of System Failures.* New York: Plenum Press.

BROWN, C., KEATS, J.A., KEATS, J.C. & SEGGIE, I. (1980). Reasoning about implication: a comparison of Malaysian and Australian subjects. *Journal of Cross-Cultural Psychology,* **11**, 395-410.

BUCCI, W. (1978). The interpretation of universal affirmative propositions. *Cognition,* **6**, 55-77.

BYRNE, R.M.J. & JOHNSON-LAIRD, P.N. (1989). On the spontaneous use of propositional connectives. Mimeograph, MRC Applied Psychology Unit, Cambridge.

BYRNE, R.M.J. & JOHNSON-LAIRD, P.N. (1990). Models and deductive reasoning. In K.J. Gilhooly, M.T.G. Keane, R.H. Logie & G. Erdos (eds.), *Lines of Thinking: Reflections on the Psychology of Thought. Volume 1: Representation, Reasoning, Analogy and Decision Making.* Chichester: Wiley.

CHAPMAN, L.J. & CHAPMAN, J.P. (1959). Atmosphere effect reexamined. *Journal of Experimental Psychology,* **58**, 220-226.

CHENG, P.W. & HOLYOAK, K.J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology,* **17**, 391-416.

CHENG, P.W., HOLYOAK, K.J., NISBETT, R.E. & OLIVER, L.M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology,* **18(3)**, 293-328.

CHENG, P.W., NISBETT, R.E. & OLIVER, L.M. (1987). The use of causal schemas in the evaluation of evidence for hypotheses. Unpublished manuscript, University of California, Los Angeles.

CHOMSKY, N. (1957). *Syntactic Structures.* The Hague: Mouton.

CLEMENT, C.A. & FALMAGNE, R.J. (1986). Logical reasoning, world knowledge, and mental imagery: Interconnections in cognitive processes. *Memory & Cognition,* **14 (4)**, 299-307.

COHEN, G. (1986). Everyday Memory. In G. Cohen, M.W. Eysenck & M.E. Le Voi, *Memory: A Cognitive Approach.* Milton Keynes: Open University Press.

COHEN, L.J. (1979). On the psychology of prediction: whose is the fallacy? *Cognition,* **7**, 385-407.

COHEN, L.J. (1981). Can human rationality be experimentally demonstrated? *The Behavioral and Brain Sciences,* **4**, 317-370.

COHEN, L.J. (1982). Are people programmed to commit fallacies? Further thought about the interpretation of experimental data on judgement. *Journal for the theory of Social Behavior,* **12,** 251-274.

CONANT, R.C. & ASHBY, W.R. (1970). Every good regulator of a system must be a model of that system. *International Journal of Systems Science,* **1,** 89-97.

COX, J.R. & GRIGGS, R.A. (1982). The effects of experience on performance in Wason's selection task. *Memory and Cognition,* **10,** 496-502.

CRAIK, K. (1943). *The Nature of Explanation.* Cambridge: Cambridge University Press.

DAVIES, D.J.M. & ISARD, S.D. (1972). Utterances as programs. In D. Michie (ed.), *Machine Intelligence, Volume 7.* Edinburgh: Edinburgh University Press.

DE KLEER, J. & BROWN, J.S. (1981). Mental models of physical systems and their acquisition. In J.R. Anderson (ed.), *Cognitive Skills and their Acquisition.* Hillsdale, NJ: Erlbaum.

DE KLEER, J. & BROWN, J.S. (1983). Assumptions and ambiguities in mechanistic mental models. In D. Gentner & A.L. Stevens (eds.), *Mental Models.* Hillsdale, NJ: Erlbaum.

DUCROT, O. (1972). *Dire et ne pas dire.* Paris: Hermann.

EINHORN, H.J. & HOGARTH, R.M. (1981). Behavioral decision theory: processes of judgement and choice. *Annual Review of Psychology,* **32,** 53-88.

ENNIS, R.H. (1976). An alternative to Piaget's conceptualisation of logical competence. *Child Development,* **17,** 903-919.

EVANS, J.St.B.T. (1972a). On the problems of interpreting reasoning data: logical and psychological approaches. *Cognition,* **1,** 373-384.

EVANS, J.St.B.T. (1972b). Reasoning with negatives. *British Journal of Psychology,* **63,** 213-219.

EVANS, J.St.B.T. (1972c). Interpretation and "matching bias" in a reasoning task. *Quarterly Journal of Experimental Psychology,* **24,** 193-199.

EVANS, J.St.B.T. (1975). On interpreting reasoning data: a reply to Van Duyne. *Cognition,* **3,** 387-390.

EVANS, J.St.B.T. (1977). Linguistic factors in reasoning. *Quarterly Journal of Experimental Psychology,* **29,** 297-306.

EVANS, J.St.B.T. (1982). *The Psychology of Deductive Reasoning.* London: Routledge & Kegan Paul.

EVANS, J.St.B.T. (1983a). Selective processes in reasoning. In J.St.B.T. Evans (ed.), *Thinking and Reasoning: Psychological Approaches.* London: Routledge & Kegan Paul.

EVANS, J.St.B.T. (1983b). Linguistic determinants of bias in conditional reasoning. *Quarterly Journal of Experimental Psychology,* **35A,** 645-644.

EVANS, J.St.B.T. (1984). Heuristic and analytic processes in reasoning. *British Journal of Psychology,* **75,** 451-468.

EVANS, J.St.B.T. (1987). Reasoning. In H. Beloff & A.M. Coleman (eds.), *Psychology Survey 6.* Leicester: British Psychological Society.

EVANS, J.St.B.T. (1989). *Bias in human reasoning: Causes and consequences.* Hove: Erlbaum.

EVANS, J.St.B.T., BARSTON, J. & POLLARD, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition,* **11,** 295-306.

EVANS, J.St.B.T. & BECK,M.A. (1981). Directionality and temporal factors in conditional reasoning. *Current Psychological Research,* **1:2,** 111-120.

EVANS, J.St.B.T. & LYNCH, J.S. (1973). Matching bias in the selection task. *British Journal of Psychology,* **64,** 391-397.

EVANS, J.St.B.T. & NEWSTEAD, S.E. (1977). Language and reasoning: a study of temporal factors. *Cognition,* **8,** 265-83.

EVANS, J.St.B.T. & WASON, P.C. (1976). Rationalization in a reasoning task. *British Journal of Psychology,* **67,** 479-486.

FALMAGNE, R.J. (ed.) (1975). *Reasoning: Representation and Process.* New York: Wiley.

FALMAGNE, R.J. (1980). The development of logical competence: a psycholinguistic perspective. In R.H. Kluwe and H. Spada (eds.), *Developmental Models of Thinking.* New York: Academic Press.

FILLENBAUM, S. (1975). If: some uses. *Psychological Research,* **37,** 245-60.

FILLENBAUM, S. (1976). Inducements: on phrasing and logic of conditional promises, threats and warnings. *Psychological Research,* **38,** 231-250.

FILLENBAUM, S. (1977). A condition on plausible inducements. *Language and Speech,* **20,** 136-141.

FILLENBAUM, S. (1986). The use of conditionals in inducements and deterrents. In E.C. Traugott, A.T. Meulen, J.S. Reilly & C.A. Ferguson (eds.) *On Conditionals.* Cambridge: Cambridge University Press, pp.179-195.

FODOR, J.A. (1980). Fixation of belief and concept acquisition. In M. Piatelli-

Palmarini (ed.), *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*. Cambridge, Massachusetts: Harvard University Press.

FONG, G.T., KRANTZ, D.H. & NISBETT, R.E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, **18**, 253-292.

FORBUS, K.D. (1983). Qualitative reasoning about space and motion. In D. Gentner & A.L. Stevens (eds.), *Mental Models*. Hillsdale, NJ: Erlbaum.

FRASE, L.T. (1966). Validity judgements in relation to two sets of items. *Journal of Educational Psychology*, **57**, 539-545.

FRASE, L.T. (1968). Effects of semantic incompatibility upon deductive reasoning tasks. *Psychonomic Science*, **12**, 64.

FREGE, G. (1879). *Begriffsschrift, eine der Arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle: Nebert.

GEIS, M.C. & ZWICKY, A.M. (1971). On invited inferences. *Linguistic Enquiry*, **2**, 561-566.

GENTNER, D. & STEVENS, A.L. (eds.) (1983). *Mental Models*. Hillsdale, NJ: Erlbaum.

GENTZEN, G. (1935/64). Investigations into logical deduction. In M.E. Szabo (ed. and trans.), *The Collected Papers of Gerhard Gentzen*. Amsterdam: North-Holland (1969: originally published 1935).

GILHOOLY, K.J., KEANE, M.T.G., LOGIE, R.H. & ERDOS, G. (eds.), *Lines of Thinking: Reflections on the Psychology of Thought. Volume 1: Representation, Reasoning, Analogy and Decision Making*. Chichester: Wiley.

GIROTTO, V., LIGHT, P. & COLBOURN, C. (1988). Pragmatic schemas and conditional reasoning in children. *Quarterly Journal of Experimental Psychology*, **40A**, 469-482.

GOLDING, E. (1981). The effect of past experience on problem solving. Paper read to the British Psychological Society at Surrey University.

GRICE, H.P. (1967). Logic and Conversation. Unpublished lectures, Harvard University.

GRICE, H.P. (1975). Logic and conversation. [extract from Grice (1967)]. In P. Cole & J.L. Morgan, *Speech Acts*, pp. 45-58. Syntax And Semantics 3. New York: Academic Press.

GRIGGS, R.A. (1983). The role of problem content in the selection task and in the THOG problem. In J.St.B.T. Evans (ed.), *Thinking and Reasoning: Psychological Approaches*. London: Routledge & Kegan Paul.

GRIGGS, R.A. (1984). Memory cueing and instructional effects on Wason's

selection task. *Current Psychological Research and Reviews,* **3,** 3-10.

GRIGGS, R.A. & COX, J.R. (1982). The elusive thematic materials effect in Wason's selection task. *British Journal of Psychology,* **73,** 407-420.

GRIGGS, R.A. & COX, J.R. (1983). The effect of problem content and negation on Wason's selection task. *Quarterly Journal of Experimental Psychology,* **35A,** 519-533.

HARNISH, R.M. (1976). Logical form and implicature. In T.G. Bever, J.J. Katz & D.T. Langendoen (eds.), *An integrated theory of linguistic ability.* New York: Crowell.

HARRIS, R.J. (1978). The effects of jury size and judge's instructions on memory for pragmatic implications from courtroom testimony. *Bulletin of the Psychonomic Society,* **11,** 129-132

HARRIS, R.J. & MONACO, G.E. (1978). Psychology of pragmatic implication: information processing between the lines. *Journal of Experimental Psychology: General,* **107,** 1-22.

HAYES, P.J. (1979). The naive physics manifesto. In D. Michie (ed.), *Expert Systems in the Micro-electronic Age.* Edinburgh: Edinburgh University Press.

HENLE, M. (1962). On the relation between logic and thinking. *Psychological Review,* **69,** 366-378.

HENLE, M. & MICHAEL, M. (1956). The influence of attitudes on syllogistic reasoning. *Journal of Social Psychology,* **44,** 115-127.

HERSH, H.M. & CARAMAZZA, A. (1976). A fuzzy set approach to modifiers and vagueness in natural language. *Journal of Experimental Psychology: General,* **105,** 254-276.

HOLLAND, J.H., HOLYOAK, K.J., NISBETT, R.E. & THAGARD, P.T. (1986). *Induction: Processes of inference, learning and discovery.* Cambridge, MA: Bradford Books/MIT Press.

HOLYOAK, K.J. & GLASS, A.L. (1978). Recognition confusions among quantifiers. *Journal of Verbal Learning and Verbal Behaviour,* **17,** 249-264.

INHELDER, B. & PIAGET, J. (1958). *The Growth of Logical Thinking.* New York: Basic Books.

JAGACINSKI, R.J. & MILLER, R.A. (1978). Describing the human operator's internal model of a dynamic system. *Human Factors,* **20,** 425-433.

JANIS, I.L. & FRICK, F. (1943). The relationship between attitudes towards conclusions and errors in judging logical validity of syllogisms. *Journal of Experimental Psychology,* **33,** 73-77.

JASKOWSKI, S. (1934). On the rules of supposition in formal logic. *Studia*

Logica, 1, 5-32.

JEPSON, C., KRANTZ, D.H. & NISBETT, R.E. (1983). Inductive reasoning: Competence or skill? *Behavioural and Brain Sciences,* **6,** 94-101.

JOHNSON-LAIRD, P.N. (1975). Models of deduction. In R.J. Falmagne (ed.), *Reasoning: Representation and Process.* New York: Wiley.

JOHNSON-LAIRD, P.N. (1980). Mental models in cognitive science. *Cognitive Science,* **4,** 71-115.

JOHNSON-LAIRD, P.N. (1981). Comprehension as the construction of mental models. *Proceedings of the Royal Society of London, B,* **295,** 353-374.

JOHNSON-LAIRD, P.N. (1983a). *Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness.* Cambridge: Cambridge University Press.

JOHNSON-LAIRD, P.N. (1983b). Thinking as a skill. In J.St.B.T. Evans (ed.), *Thinking and Reasoning: Psychological Approaches.* London: Routledge & Kegan Paul.

JOHNSON-LAIRD, P.N. (1983c). Models of conditionals. Unpublished draft, September, 1983. (earlier version of Johnson-Laird, 1986).

JOHNSON-LAIRD, P.N. (1986). Conditionals and mental models. In E.C. Traugott, A.T. Meulen, J.S. Reilly & C.A. Ferguson (eds.) *On Conditionals.* Cambridge: Cambridge University Press, pp.55-76.

JOHNSON-LAIRD, P.N. & BARA, B.G. (1984). Syllogistic inference. *Cognition,* **16,** 1-61.

JOHNSON-LAIRD, BYRNE, R.M.J. & SCHAEKEN, W. (1989). Reasoning by rule or by model: the case of propositional inference. Mimeograph, MRC Applied Psychology Unit, Cambridge.

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. & TAGART, J. (1969). How implication is understood. *American Journal of Psychology,* **82,** 367-373.

JOHNSON-LAIRD, P.N. & WASON, P.C. (1970). Insight into a logical relation. *Quarterly Journal of Experimental Psychology,* **22,** 49-61.

JUST, M. (1974). Comprehending quantified sentences: The relation between sentence-picture and semantic memory verification. *Cognitive Psychology,* 6, 216-236.

KAHNEMAN, D. & TVERSKY, A. (1972). Subjective probability: a judgement of representativeness. *Cognitive Psychology,* **3,** 430-454.

KAHNEMAN, D. & TVERSKY, A. (1973). On the psychology of prediction.

Psychological Review, **80,** 237-251.

KAHNEMAN, D. & TVERSKY, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic & A. Tversky (eds.) *Judgement Under Uncertainty: Heuristics and Biases.* Cambridge: Cambridge University Press.

KALUZA, J. (1986). A new look at conditional sentences in English. *English Language Teaching Journal,* **40 (1),** 59-60.

KAUFMAN, H. & GOLDSTEIN, S. (1967). The effect of emotional value of conclusions upon distortion in syllogistic reasoning. *Psychonomic Science,* **7,** 367-368.

KELLEY, H.H. (1971). Causal schemata and the attribution process. In E.E. Jones, D.E. Kanouse, H.H. Kelley, R.E. Nisbett, S. Valins & B Weiner (eds.), *Attribution: Perceiving the causes of behaviour.* Morristown, NJ: General Learning Press.

KELLEY, H.H. (1973). The process of causal attribution. *American Psychologist,* **28,** 108-128.

KELLY, G.A. (1955). *The Psychology of Personal Constructs. (2 vols).* New York: Norton.

KIERAS, D.E. & BOVAIR, S. The role of a mental model in learning to operate a device. *Cognitive Science,* **8,** 255-273.

KUHN, D. (1977). Conditional reasoning in children. *Developmental psychology,* **13,** 342-353.

KUTSCHERA, F. von. (1974). Indicative conditionals. *Theoretical Linguistics,* **1,** 257-269.

KYBURG, H.E. (1983). Rational belief. *The Behavioural and Brain Sciences,* **6,** 231-273.

LABOV, W. (1973). The boundaries of words and their meanings. In C.-J.N. Bailey & R.W. Shuy (eds.), *New Ways of Analyzing Variation in English.* Washington, D.C.: Georgetown University Press.

LAKOFF, G. (1973). Hedges: A study in the meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic,* **2,** 458-508.

LEHMAN, D.R., LEMPERT, R.O. & NISBETT, R.E. (1988). The effects of graduate training on reasoning. *American Psychologist,* **43 (6),** 431-442.

LEVINSON, S.C. (1983). *Pragmatics.* Cambridge: Cambridge University Press.

LEWIS, D. (1973). *Counterfactuals.* Oxford: Blackwell, 1973.

LONGUET-HIGGINS, H.C. (1972). The algorithmic description of natural

language. *Proceedings of the Royal Society of London,* **B, 182,** 307-322.

McCAWLEY, J.D. (1981). *Everything That Linguists Have Always Wanted To Know About Logic But Were Ashamed To Ask.* Oxford: Basil Blackwell.

MACKIE, J.L. (1973). *Truth, Probability and Paradox.* Oxford: Clarendon Press.

MANDLER, J.M. (1980). Structural invariants in development. Center for Human Information Processing Technical Report no. 96. La Jolla: University of California, San Diego.

MANKTELOW, K.I. (1979). The role of content in reasoning. Unpublished PhD thesis, Plymouth Polytechnic.

MANKTELOW, K.I. & EVANS, J.St.B.T. (1979). Facilitation of reasoning by realism: effect or non-effect? *British Journal of Psychology,* **71,** 227-231.

MARCUS, S.L. & RIPS, L.J. (1979). Conditional Reasoning. *Journal of Verbal Learning and Verbal Behaviour,* **18,** 199-229.

MARKOVITS, H. (1984). Awareness of the 'possible' as a mediator of formal thinking in conditional reasoning problems. *British Journal of Psychology,* **75,** 367-376.

MARKOVITS, H. (1985). Incorrect conditional reasoning among adults: competence or performance? *British Journal of Psychology,* **76,** 241-247.

MARKOVITS, H. (1986). Familiarity effects in conditional reasoning. *Journal of Educational Psychology,* **78 (6),** 492-494.

MARKOVITS, H. (1988). Conditional reasoning, representation and empirical evidence on a concrete task. *Quarterly Journal of Experimental Psychology,* **40A (3),** 483-495.

MILL, J.S. (1874). *A system of logic* (8th ed.). New York: Harper.

MITCHELL-KERNAN, C. & KERNAN, K.T. (1977). Pragmatics of directive choice among children. In S. Ervin-Tripp & C. Mitchell-Kernan (eds.), *Child discourse.* New York: Academic Press.

MORGAN, J.I.B. & MORTON, J.T. (1944). The distortions of syllogistic reasoning produced by personal connections. *Journal of Social Psychology,* **20,** 39-59.

NEWELL, A. (1980). Reasoning, problem solving and decision processes: the problem space as a fundamental category. In R.S. Nickerson (ed.), *Attention and Performance VIII.* Hillside, N.J.: Erlbaum.

NEWSTEAD, S.E. (1988). Quantifiers as fuzzy concepts. In T. Zétényi (ed.), *Fuzzy Sets in Psychology.* Amsterdam: Elsevier (North-Holland).

NEWSTEAD, S.E. & COLLIS, J. (1987). Context and the interpretation of

quantifiers of frequency. *Ergonomics,* **30 (10),** 1447-1462.

NEWSTEAD, S.E. & GRIGGS, R.A. (1984). Fuzzy quantifiers as an explanation of set inclusion performance. *Psychological Review,* **46,** 377-388.

NEWSTEAD, S.E. & GRIGGS, R.A. & CHROSTOWSKI, (1984). Reasoning with realistic disjunctives. *Quarterly Journal of Experimental Psychology,* **36A,** 611-627.

NISBETT, R.E., KRANTZ, D.H., JEPSON, D. & KUNDA, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review,* **90,** 339-363.

NISBETT, R.E., FONG, G.T., LEHMAN, D.R. & CHENG, P.W. (1987). Teaching reasoning. *Science,* **238,** 625-631.

NISBETT, R.E. & WILSON, T.D. (1977) Telling more than we can know: Verbal reports on mental processes. *Psycholgical Review,* **84,** 231-259.

NORMAN, D.E. (1983). Some observations on mental models. In D. Gentner & A.L. Stevens (eds.), *Mental Models.* Hillsdale, NJ: Erlbaum.

OAKHILL, J.V. & JOHNSON-LAIRD (1985). The effects of belief on the spontaneous production of syllogistic conclusions. *Quarterly Journal of Experimental Psychology,* **37A,** 553-569.

O'BRIEN, D.P., COSTA, G. & OVERTON, W.J. (1986). Evaluations of causal and conditional hypotheses. *Quarterly Journal of Experimental Psychology,* **38A,** 493-512.

O'BRIEN, D.P. & OVERTON, W.J. (1980). Conditional reasoning following contradictory evidence: A developmental analysis. *Journal of Experimental Child Psychology,* **30,** 44-60.

O'BRIEN, D.P. & OVERTON, W.J. (1982). Conditional reasoning and the competence-performance issue: A developmental analysis of a training task. *Journal of Experimental Child Psychology,* **34,** 274-290.

OSHERSON, D.N. (1975). Logic and models of logical thinking. In R.J. Falmagne (ed.), *Reasoning: Representation and Process in Children and Adults.* Hillsdale, NJ: Erlbaum.

PIAGET, J. (1972). Intellectual development from adolescence to adulthood. *Human Development,* **15,** 1-12.

POLITZER, G. (1986). Laws of language use and formal logic. *Journal of Psycholinguistic Research,* **15 (1),** 47-92.

POLLARD, P. (1979). Human reasoning: logical and non-logical explanations. Unpublished PhD thesis, Plymouth Polytechnic.

POLLARD, P. (1981). The effect of thematic content in the "Wason selection task". *Current Psychological Research,* **1,** 21-29.

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

POLLARD, P. & EVANS, J.St.B.T. (1980). The influence of logic on conditional reasoning performance. *Quarterly Journal of Experimental Psychology*, **32**, 605-624.

POLLOCK, J. (1976). *Subjunctive Reasoning.* Dordrecht: Reidel.

POPPER, K. (1959). *The Logic of Scientific Discovery.* London: Hutchinson.

QUINE, W. Van O. (1952). *Methods of logic.* London: Routledge & Kegan Paul.

RAMSEY, F.P. (1931). General propositions and causality. In F.P. Ramsey, *The Foundations of Mathematics and Other Logical Essays.* London: Routledge and Kegan Paul.

REICH, S.S. & RUTH, P. (1982). Wason's selection task: Verification, falsification and matching. *British Journal of Psychology*, **73**, 395-405.

REVLIN, R. & LEIRER, V.O. (1978). The effect of personal biases on syllogistic reasoning: rational decisions from personalised representations. In R. Revlin and R.E. Mayer (eds.), *Human Reasoning.* New York: Wiley.

REVLIN, R. & MAYER, R.E. (eds.) (1978). *Human Reasoning.* New York: Wiley.

RIPS, L.J. (1983). Reasoning as a central intellective ability. In R.J. Sternberg (ed.), *Advances in the study of Human Intelligence, Volume 2.* Hillsdale, NJ: Erlbaum.

RIPS, L.J. (1986). Mental muddles. In M. Brand & R.M. Harnish (eds.), The *Representation of Knowledge and Belief.* Tucson, AZ: University of Arizona Press.

RIPS, L.J. & MARCUS, S.L. (1977). Suppositions and the analysis of conditional sentences. In M.A. Just & P.A. Carpenter (eds.), *Cognitive Processes in Comprehension.* Wiley: New York.

ROBERGE, J.J. (1971). Some effects of negation on adults' conditional reasoning abilities. *Psychological Reports*, **29**, 839-844.

ROBERGE, J.J. (1974). Effects of negation on adults' comprehension of fallacious conditional and conjunctive arguments. *Journal of General Psychology*, **91**, 287-293.

ROBERGE, J.J. (1978). Linguistic and psychometric factors in propositional reasoning. *Quarterly Journal of Experimental Psychology*, **30**, 705-716.

ROUSE, W.B. & MORRIS, N.M. (1986). On looking into the black box:

Prospects and limits in the search for mental models. *Psychological Bulletin,* **100 (3),** 349-363.

RUMAIN, B., CONNELL, J. & BRAINE, M.D.S. (1983). Conversational comprehension processes are responsible for reasoning fallacies in children as well as adults: If is not the biconditional. *Developmental Psychology,* **19,** 471-481.

RUMELHART, D.E. (1979). Analogical processes and procedural representation. Center for Human Information Technical Report No. 81, La Jolla, University of California, San Diego.

RUMELHART, D.E. (1980). Schemata: the building blocks of cognition. In R.J. Spiro, B.C. Bruce & W.F. Brewer (eds.), *Theoretical Issues in Reading Comprehension.* Hillsdale, NJ: Erlbaum.

SCRIBNER, S. (1977). Modes of thinking and ways of speaking: Culture and logic reconsidered. In P. Johnson-Laird & P.C. Wason (eds.), *Thinking: Readings in cognitive science.* Cambridge: Cambridge University Press.

SEARLE, J. (1969). *Speech Acts.* London and New York: Cambridge University Press.

SKINNER, B.F. (1938). *The Behaviour of Organisms.* New York: Appleton Century-Crofts.

SKINNER, B.F. (1972). *Beyond Freedom and Dignity.* Cape: London.

SMEDSLUND, J. (1970). On the circular relation between logic and understanding. *Scandinavian Journal of Psychology,* **11,** 217-219.

SPERBER, D. & WILSON, D. (1983). Mutual knowledge and relevance in theories of comprehension. In N.V. Smith (ed.), *Proceedings of the SSRC Colloquium on Mutual Knowledge.* London: Academic Press.

STALNAKER, R.C. (1968). A Theory of conditionals. In N. Rescher (ed.) *Studies in Logical Theory.* Oxford: Blackwell.

STAUDENMAYER, H. (1975). Understanding conditional reasoning with meaningful propositions. In R.J. Falmagne (ed.), *Reasoning: Representation and Process.* New York: Wiley.

STAUDENMAYER, H. & BOURNE, L.E. (1978). The nature of denied propositions in the conditional reasoning task: interpretation and learning. In R. Revlin & R.E. Mayer (eds.), *Human Reasoning,* New York: Wiley.

STERNBERG, R.J. (1982). Reasoning, problem solving and intelligence. In R.J. Sternberg (ed.), *Handbook of Human Intelligence.* Cambridge: Cambridge University 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,** 219-225.

TAPLIN, J.E. & STAUDENMAYER, H. (1973). Interpretation of abstract conditional sentences in deductive reasoning. *Journal of Verbal Learning and Verbal Behaviour,* **12,** 530-542.

TURING, A.M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society,* **2, 42,** 230-265, (corrections, *ibid.* , 43, 544-546).

TVERSKY, A. & KAHNEMAN, D. (1971). The belief in the law of small numbers. *Psychological Bulletin,* **76,** 105-110.

TVERSKY, A. & KAHNEMAN, D. (1973). Availability: a heuristic for judging frequency and probability. *Cognitive Psychology,* **5,** 207-232.

TVERSKY, A. & KAHNEMAN, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science,* **185,** 1124-1131.

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.

VUYK, R. (1981). *Critique of Piaget's Genetic Epistemology 1965-1980, II.* London: Academic Press.

WASON, P.C. (1966). Reasoning. In B.M. Foss (ed.), *New Horizons in Psychology I.* Harmondsworth: Penguin.

WASON, P.C. (1977). Self-contradictions. In P.N. Johnson-Laird & P.C. Wason (eds.), *Thinking: Readings in Cognitive Science.* Cambridge: Cambridge University Press.

WASON, P.C. (1983). Realism and rationality in the selection task. In J.St.B.T. Evans (ed.), *Thinking and Reasoning: Psychological Approaches.* London: Routledge & Kegan Paul.

WASON, P.C. & GREEN, D.W. (1984). Reasoning and representation. *Quarterly Journal of Experimental Psychology,* **36A,** 597-610.

WASON, P.C. & JOHNSON-LAIRD, P.N. (1972). *Psychology of Reasoning: Structure and content.* Cambridge, Mass: Harvard University Press.

WASON, P.C. & SHAPIRO, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology,* **23,** 63-71.

WETHERICK, N.E. (1970). On the representativeness of some experiments in cognition. *Bulletin of the British Psychological Society,* **23,** 213-214.

WICKENS, C.D. (1984). *Engineering Psychology and Human Performance.*

Colombus, OH: Merrill.

WILKINS, M.C. (1928). The effect of changed material on the ability to do formal syllogistic reasoning. *Archives of Psychology,* **102,** New York.

WILKS, Y. (1972). An artificial intelligence approach to machine translation. In R.C. Schank & K.M. Colby (eds.), *Computer Models of Thought and Language.* New York: Academic Press.

WILLIAMS. M.D., HOLLAN, J.D. & STEVENS, A.L. (1983). Human reasoning about a simple physical system. In D. Gentner & A.L. Stevens (eds.), *Mental Models.* Hillsdale, NJ: Erlbaum.

WILSON, D. & SPERBER, D. (1981). On Grice's theory of conversation. In P. Werth (ed.), *Conversation and discourse.* London: Croom Helm.

WOODS, W.A. (1967). Semantics for a question-answering system. Mathematical Linguistics and Automatic Translation Report NSF-19, Harvard Computational Laboratory.

WOODWORTH, R.S. & SELLS, S.B. (1935). An atmosphere effect in formal syllogistic reasoning. *Journal of Experimental Psychology,* **18,** 451-460.

YACHANIN, S.A. (1986). Facilitation in Wason's selection task: Contents and instructions. *Current Psychological Research and Reviews,* **5,** 20-29.

YACHANIN, S.A. & TWENEY, R.D. (1982). The effect of thematic content on cognitive strategies in the four-card selection task. *Bulletin of the Psychonomic Society,* **19,** 87-90.

ZADEH, L.A. (1979). A theory of approximate reasoning. In J.E. Hayes, D. Michie & L.I. Mikulich (eds.), *Machine Intelligence 9.* Chichester: Ellis Horwood.

Appendices

Appendix 1a: Experiment 1: a typical screen

Appendix 1b: Experiment 1: instructions

[general instructions]:

Welcome to Charlie Ellis' experiment!

Do not worry if you find these instructions hard to follow, as you will have an opportunity to see them again and to practise operating the computer as often as you like before the experiment begins.

You will be presented with an array of 240 coloured symbols above which will be shown a statement about the shape and colour of the symbols in the array. At the very top of the screen will be shown a graduated scale ranging from "ABSOLUTELY FALSE" through "CAN'T TELL" to "ABSOLUTELY TRUE", and you will be required to indicate by moving an arrow on the scale how true you feel the statement to be with respect to the array of symbols: note that you should only use the extreme ends of the scale when you feel that the statement is ABSOLUTELY true or false.

When you have finished moving the arrow on the scale, press RETURN to go on to the next task. As it is very easy to press RETURN by mistake, you will then be asked if you are sure that you have finished and should press Y for yes or N for no: if you press N, then you may continue to move the arrow.

You will be presented with 40 such tasks altogether. Before starting the experiment, you will be given an opportunity to practise moving the arrow on the scale. Please go through the instructions and practise as often as you like: it is far more important that you understand what you are doing than that you make a quick start. If after a couple of runs through you are still in some doubt about what you have to do, then please call the experimenter to help you BEFORE going on to the experiment itself.

Appendix 1b: Experiment 1: instructions

[operational instructions]:

The arrow is moved along the scale using the grey RIGHT-ARROW and LEFT-ARROW keys at the top right of the keyboard. The UP-ARROW key will move the arrow to the exact centre of the scale. Try moving the arrow now.

When you have moved the arrow to where you want it to be, press RETURN to go on to the next task. Pressing RETURN now will terminate this practice and let you go through the instructions again if you so wish. Just in case you may have pressed RETURN by mistake, you will be asked if you are sure that you have finished, and you should press Y for yes or N for no: if you press N, then you will be able to continue to move the arrow.

Appendix 1c: Experiment 1: debriefing text

Thank you very much indeed for taking part in this experiment.

The experiment was designed to investigate the hypothesis that people interpret logical statements of the "if x then y" and "x only if y" variety in a "fuzzy" way: that is, as if the statement was really something like "if x then as often as makes no difference y".

What I was interested in was the degree of "fuzziness" which people would tolerate in an array of symbols containing varying amounts of confirming and disconfirming evidence for the truth of a statement, as well as many symbols not explicitly referred to in the statement. There is no single "correct" answer, and so there is no question of your answers reflecting anything whatever about your intelligence or ability.

Thanks again for your help.

Appendix 2a: Experiment 2: a typical screen

Appendix 2b: Experiment 2: instructions

Welcome to Charlie Ellis' experiment!

Do not worry if you find these instructions hard to follow, as you will have an opportunity to see them again and to practise operating the computer as often as you like before the experiment begins.

You will be presented with an empty 6x6 grid of 36 squares. Above the grid will be shown a statement concerning the shape and colour of symbols which are to be put into the grid, and to the right of the grid will be an instruction to show the statement to be TRUE or FALSE. What you have to do is to fill the grid with symbols so that the statement given at the top of the screen is either TRUE or FALSE with respect to the appearance of the grid.

At the beginning of each task, the top left square of the grid will contain a smaller dotted square, called the "cursor" (ie it indicates which square in the grid your instructions will refer to): the cursor may be moved to any square in the grid using the grey arrow keys at the top right of the keyboard.

There are nine different symbols available, made up of three different shapes each of three different colours and the nine symbols will be shown next to the grid with a number (1 to 9) superimposed upon each of them.

To insert a symbol into the grid, first indicate the square to which the symbol is to go by moving the dotted cursor (the smaller square) to the correct position using the arrow keys. When the desired square is indicated, you can place a symbol in it by pressing the relevant number (1 to 9). To delete a symbol, move the dotted cursor to the square concerned and press the DELETE key. To replace a symbol with another symbol, simply put the new symbol on top of the old one.

Appendix 2b: Experiment 2: instructions

The fact that there are nine symbols available does not mean that they all must necessarily be included in the grid. You may use any of the symbols as often or as little as you wish. When you have filled the grid to your satisfaction, press RETURN to go on to the next task: if you have left any squares blank by mistake, then the computer will point this out, and you will have to fill them in before going on to the next task.

You will be presented with 16 such tasks altogether, and in each case you are required to place symbols into the grid so that the statement given at the top of the screen is either TRUE or FALSE with respect to the appearance of the grid (whether it should be shown to be TRUE or FALSE will be displayed to the right of the grid in each task). When you have read these instructions, you will be given an opportunity to become familiar with the display and with the way in which the symbols are placed into the grid.

You will then be given another chance to see these instructions and to practise moving the symbols if you so wish before going on to the experiment itself. Please go through the instructions and practise as often as you like: it is far more important that you understand what you are doing than that you make a quick start. If after a couple of runs through you are still in some doubt about what you have to do, then please call the experimenter to help you BEFORE going on to the experiment itself.

Appendix 2c: Experiment 2: debriefing

Thank you very much indeed for taking part in this experiment.

The experiment was designed to investigate the hypothesis that people interpret logical statements of the "if x then y" and "x only if y" variety in a "fuzzy" way: that is, as if the statement was really something like "if x then as often as makes no difference y".

What I was interested in was the degree of "fuzziness" which people would incorporate into a constructed example of an array of symbols described (truly or falsely) by a conditional rule. As there is an (almost) infinite number of ways in which the nine symbols could be combined in the grid, there is no single "correct" answer, and so there is no question of your answers reflecting anything whatever about your intelligence or ability.

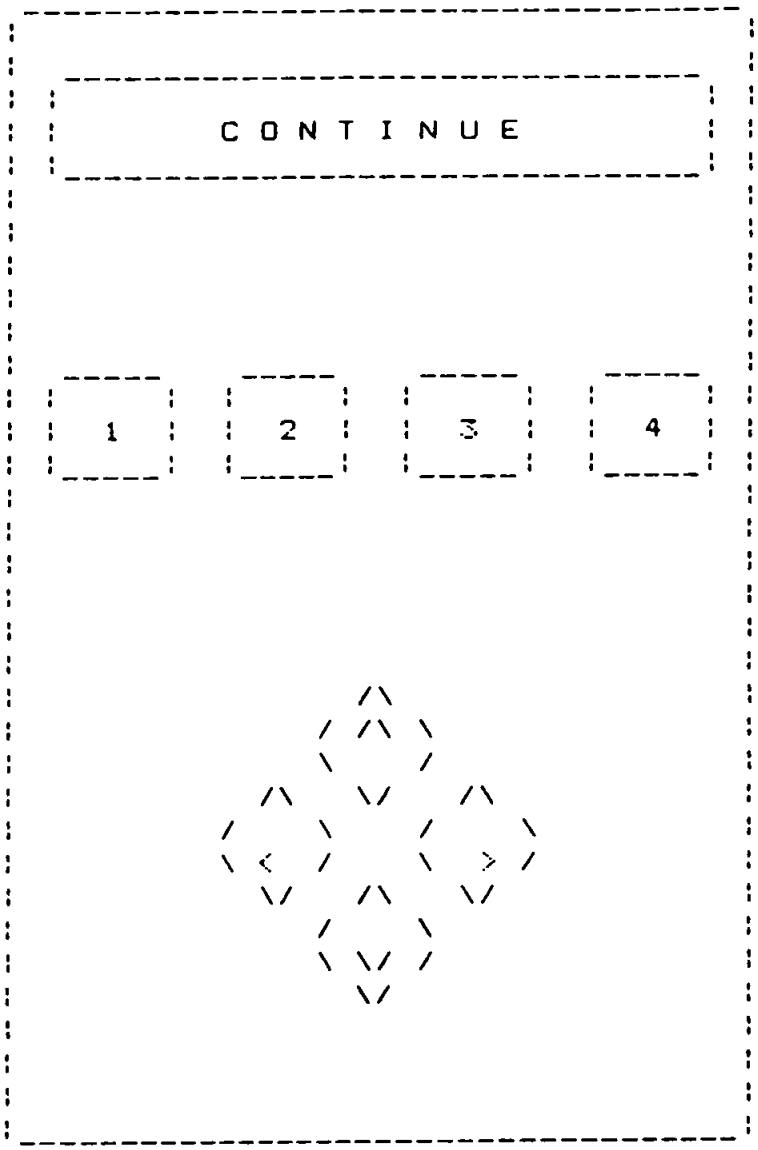
Thanks again for your help.

IF THE SYMBOLS ARE RED THEN THEY ARE SQUARE

FILL THE GRID TO
MAKE THE STATEMENT

TRUE

Appendix 4a: Experiment 4: the keypad



Appendix 4b: Experiment 4: a typical screen

Appendix 4c: Experiment 4: instructions

Thank you for volunteering to take part in this abstract reasoning experiment.

Do not worry if you find these instructions hard to follow, as you will have an opportunity to see them again and to practise operating the computer as often as you like before the experiment begins.

You will be presented with an empty 6×6 grid of 36 squares. Above the grid will be shown a statement concerning the shape and colour of symbols which are to be put into the grid, and to the right of the grid will be an instruction to show the statement to be TRUE or FALSE. What you have to do is to fill the grid with symbols so that the statement given at the top of the screen is either TRUE or FALSE with respect to the appearance of the grid.

At the beginning of each task, the top left square of the grid will contain a smaller dotted square, called the "cursor" (ie it indicates which square in the grid your instructions will refer to): the cursor may be moved to any square in the grid using the arrow keys.

On each task there will be four different symbols available and these will be shown to the right of the grid with a number (1 to 4) beneath each of them.

To insert a symbol into the grid, first indicate the square to which the symbol is to go by moving the dotted cursor (the smaller square) to the correct position using the arrow keys. When the desired square is indicated, you can place a symbol in it by pressing the relevant number (1 to 4). Do not worry if you press the wrong key - to replace a symbol with another symbol, simply put the new symbol on top of the old one.

The fact that there are four symbols available does not mean that they all

Appendix 4c: Experiment 4: instructions

Thank you very much indeed for taking part in this experiment, which was designed to investigate the hypothesis that people interpret logical statements of the "if x then y" and "x only if y" variety in a "fuzzy" way: that is, as if the statement was really something like "if x then as often as makes no difference y".

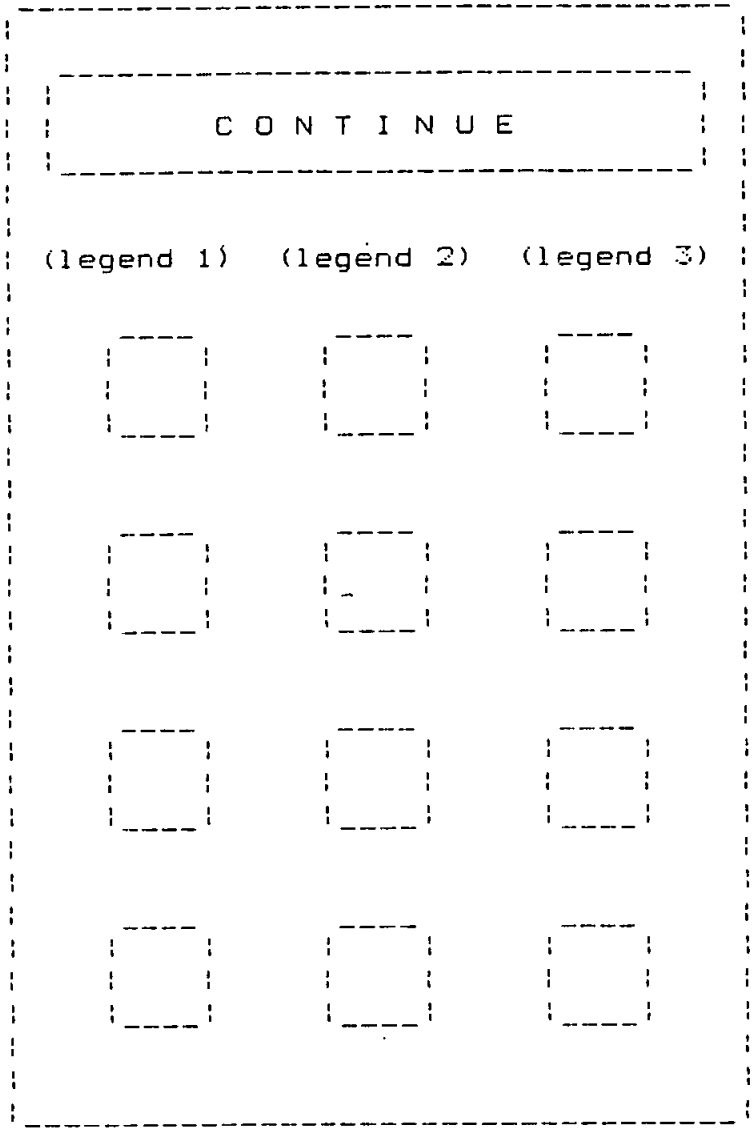
What I was interested in was the degree of "fuzziness" which people would incorporate into a constructed example of an array of symbols described (truly or falsely) by a conditional rule. As there is an (almost) infinite number of ways in which the four symbols could be combined in the grid, there is no single "correct" answer, and so there is no question of your answers reflecting anything whatever about your intelligence or ability.

must necessarily be included in the grid. You may use any of the symbols as often or as little as you wish. When you have filled the grid to your satisfaction, press CONTINUE to go on to the next task: if you have left any squares blank by mistake, then the computer will point this out, and you will have to fill them in before going on to the next task.

You will be presented with 16 such tasks altogether, and in each case you are required to place symbols into the grid so that the statement given at the top of the screen is either TRUE or FALSE with respect to the appearance of the grid (whether it should be shown to be TRUE or FALSE will be displayed to the right of the grid in each task). When you have read these instructions, you will be given an opportunity to become familiar with the display and with the way in which the symbols are placed into the grid.

You will then be given another chance to see these instructions and to practise moving the symbols if you so wish before going on to the experiment itself. Please go through the instructions and practise as often as you like: it is far more important that you understand what you are doing than that you make a quick start. If after a couple of runs through you are still in some doubt about what you have to do, then please call the experimenter to help you BEFORE going on to the experiment itself.

Appendix 5a: Experiment 5: the keypad (truth table task)

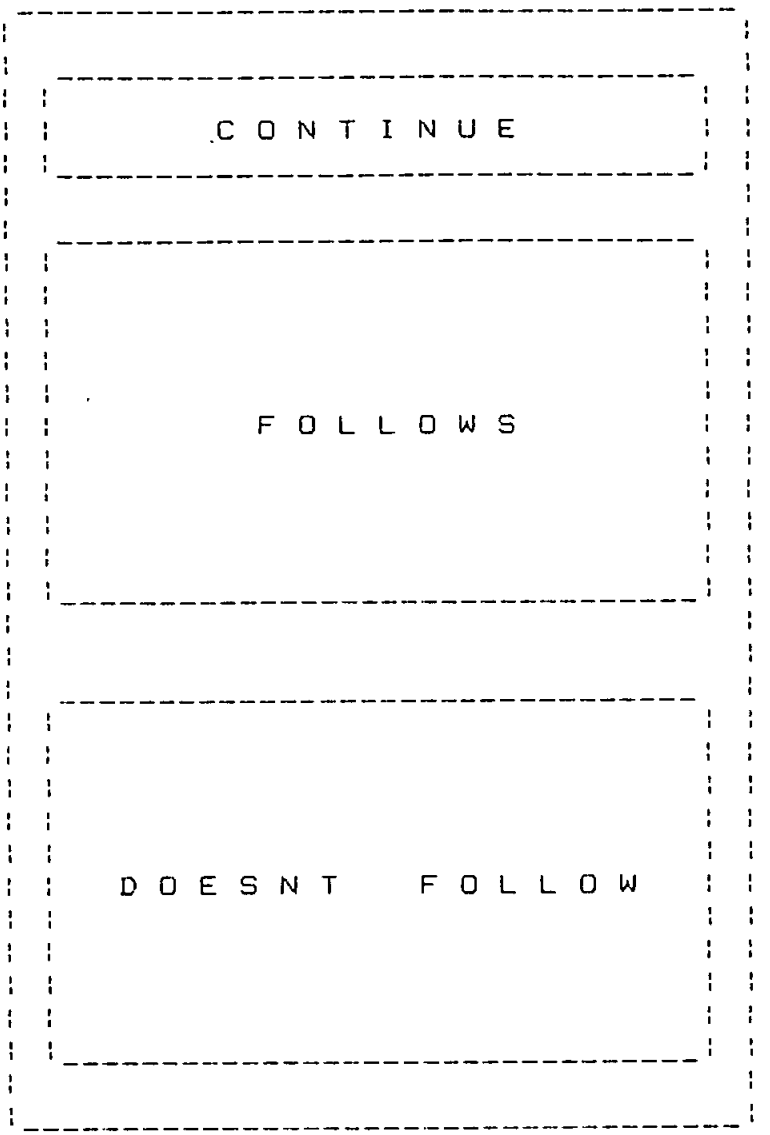


legend 1 : SUPPORTS the Statement

legend 2 : CONTRADICTS the Statement

legend 3 : TELLS US NOTHING ABOUT the Statement

Appendix 5b: Experiment 5: the keypad (inference task)



Appendix 5c: Experiment 5: instructions to expert judges

The materials comprise five each of seven different contexts (total 35 items) presented in a strictly random order. These are to be classified in accordance with "Truth-functional natural language uses of if...then [i]: AA" (01oct85) except that the number of experimental contexts is reduced to seven as follows:

i) [TCFN] and [TCFU] have had to be combined into their super-class [TCF] (Temporal-Causal-Factual). The reason for this is that the necessary/unnecessary distinction is a question of fact which is quite independent of the beliefs and interpretations of the speaker and the hearer: thus, whilst it may be possible to classify an utterance in this way on the basis of its content, such a classification has nothing to do with the way in which it is USED

ii) The two Definitional uses [NDS] and [NDC] are not amenable to manipulation in the same way as the others and thus are not admitted by the particular design of this experiment.

Each item should be coded in one of the ways set out below. The ? code may be selected either where a complete blank is drawn or where two or more codes compete - in this latter case, please identify those codes.

{ N }	Non-causal	[TN]	Temporal Non-Causal	
{ C }	Causal	[TCF]	" Causal Factual	
{ P }	Promise	[TCPC+]	" " Purposive Control	+
{ Th }	Threat	[TCPC-]	" " " "	-
{ Ti }	Tip	[TCPB+]	" " " Belief	+
{ W }	Warning	[TCPC-]	" " " "	-
{ U }	Universal	[NA]	Non-temporal Axiomatic	
{ ? }	dunno!			

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

At the Rugby Club singsong, a new member was curious to note how everyone else always seemed to know which song was coming next. The team captain explained to him that it was a club tradition that certain songs always followed certain others and told him as an example

"If the next song is 'Roll Out The Barrel'
then the one after it will be 'Yellow Submarine'."

[PQ]:
The next song was 'Roll Out The Barrel';
the one after it was 'Yellow Submarine'.

[P~Q]:
The next song was 'Roll Out The Barrel';
the one after it wasn't 'Yellow Submarine'.

[~PQ]:
The next song wasn't 'Roll Out The Barrel';
the one after it was 'Yellow Submarine'.

[~P~Q]:
The next song wasn't 'Roll Out The Barrel';
the one after it wasn't 'Yellow Submarine'.

[text 1 : Temporal 1]

Sally and Nick, the Year 1 reps on the Psychology Course Committee, had each forgotten to bring a copy of the Agenda with them to the meeting. Sally knew that certain items were always dealt with in a particular order, and told Nick

"If the next item is about the Computer Centre
then the one after it will be about the LRC."

[PQ]:
The next item was about the Computer Centre;
the one after it was about the LRC.

[P^Q]:
The next item was about the Computer Centre;
the one after it wasn't about the LRC.

[^PQ]:
The next item wasn't about the Computer Centre;
the one after it was about the LRC.

[^P^Q]:
The next item wasn't about the Computer Centre;
the one after it wasn't about the LRC.

[text 2 : Temporal 2]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Cedric, an ardent transport enthusiast, had memorised the entire London Underground timetable. Standing on Earls Court station, he told his friend Arthur

> "If the next southbound train is for Ealing
> then the one after it will be for Wimbledon."

[PQ]:
The next southbound train was for Ealing;
the one after it was for Wimbledon.

[P~Q]:
The next southbound train was for Ealing;
the one after it wasn't for Wimbledon.

[~PQ]:
The next southbound train wasn't for Ealing;
the one after it was for Wimbledon.

[~P~Q]:
The next southbound train wasn't for Ealing;
the one after it wasn't for Wimbledon.

[text 3 : Temporal 3]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

The editor of 'Metropolitan' magazine was telling a new staff writer that it was the magazine's policy to run features on certain topic areas in strict rotation. By way of example, she told her

> "If the current month's feature is about relationships then the next month's feature will be about fitness."

[PQ]:
That month's feature was about relationships ;
the next month's feature was about fitness.

[P~Q]:
That month's feature was about relationships ;
the next month's feature wasn't about fitness.

[~PQ]:
That month's feature wasn't about relationships ;
the next month's feature was about fitness.

[~P~Q]:
That month's feature wasn't about relationships ;
the next month's feature wasn't about fitness.

[text 4 : Temporal 4]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Sue and Mark were showing their new television to a friend of theirs and complaining about the trouble they were having with it. They told him

> "If it starts to rain
> then the picture goes fuzzy."

[PQ]:
It did start to rain;
the picture did go fuzzy.

[P~Q]:
It did start to rain;
the picture didn't go fuzzy.

[~PQ]:
It didn't start to rain;
the picture did go fuzzy.

[~P~Q]:
It didn't start to rain;
the picture didn't go fuzzy.

[text 5 : Causal 1]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Sarah, a training officer at a nuclear power plant, was explaining the safety precautions to some new recruits. She told them

"If a leak develops in the cooling system
then the reactor will shut down."

[PQ]:
A leak did develop in the cooling system;
the reactor did shut down.

[P~Q]:
A leak did develop in the cooling system;
the reactor didn't shut down.

[~PQ]:
A leak didn't develop in the cooling system;
the reactor did shut down.

[~P~Q]:
A leak didn't develop in the cooling system;
the reactor didn't shut down.

[text 6 : Causal 2]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Harry, a weighbridge operator, was explaining the controls to a visitor. Suggesting that the visitor wait for the next lorry that came in to be weighed, he told him

"If the lorry is heavier than the legal limit
then the alarm bell will ring."

[PQ]:
The lorry was heavier than the legal limit;
the alarm bell did ring.

[P~Q]:
The lorry was heavier than the legal limit;
the alarm bell didn't ring.

[~PQ]:
The lorry wasn't heavier than the legal limit;
the alarm bell did ring.

[~P~Q]:
The lorry wasn't heavier than the legal limit;
the alarm bell didn't ring.

[text 7 : Causal 3]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Bert, a canal lock-keeper, was telling a holidaymaker how the lock worked. Suggesting that they watch the next boat go through the lock, he told the holidaymaker

"If the water in the lock goes above the red line
then it will spill over into the canal."

[PQ]:
The water in the lock did go above the red line;
it did spill over into the canal.

[P~Q]:
The water in the lock did go above the red line;
it didn't spill over into the canal.

[~PQ]:
The water in the lock didn't go above the red line;
it did spill over into the canal.

[~P~Q]:
The water in the lock didn't go above the red line;
it didn't spill over into the canal.

[text 8 : Causal 4]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Susan's mother wanted her to stay in on Saturday night to keep her company. She told her

"If you stay in on Saturday night
then I'll take you sailing on Sunday."

[PQ]:
Susan did stay in that Saturday night;
her mother did take her sailing on the Sunday.

[P~Q]:
Susan did stay in that Saturday night;
her mother didn't take her sailing on the Sunday.

[~PQ]:
Susan didn't stay in that Saturday night;
her mother did take her sailing on the Sunday.

[~P~Q]:
Susan didn't stay in that Saturday night;
her mother didn't take her sailing on the Sunday.

[text 9 : Promise 1]

Bobby's girlfriend wanted him to pick her up from the gym.
She told him

"If you pick me up from the gym
then I'll help you with your essay."

[P&Q]:
Bobby did pick his girlfriend up from the gym;
she did help him with his essay.

[P&~Q]:
Bobby did pick his girlfriend up from the gym;
she didn't help him with his essay.

[~P&Q]:
Bobby didn't pick his girlfriend up from the gym;
she did help him with his essay.

[~P&~Q]:
Bobby didn't pick his girlfriend up from the gym;
she didn't help him with his essay.

[text 10 : Promise 2]

Henry's father wanted him to wash the car. He told his son

"If you wash the car
then I'll let you borrow it tonight."

[P@]:
Henry did wash the car;
his father did let him him borrow it that night.

[P~@]:
Henry did wash the car;
his father didn't let him him borrow it that night.

[~P@]:
Henry didn't wash the car;
his father did let him him borrow it that night.

[~P~@]:
Henry didn't wash the car;
his father didn't let him him borrow it that night.

[text 11 : Promise 3]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

The Director of Norshire Polytechnic was anxious to bring the student sit-in to a close. He told the protesters

"If you leave the building straight away
then the Poly will provide a creche before next term."

[PQ]:
The protesters did leave the building straight away;
the Poly did provide a creche before the next term.

[P~Q]:
The protesters did leave the building straight away;
the Poly didn't provide a creche before the next term.

[~PQ]:
The protesters didn't leave the building straight away;
the Poly did provide a creche before the next term.

[~P~Q]:
The protesters didn't leave the building straight away;
the Poly didn't provide a creche before the next term.

[text 12 : Promise 4]

Little Jenny's parents were having friends around for dinner. Having just set the dining table, her mother told Jenny

"If you mess up the table
then I'll spank you."

[PQ]:
Jenny did mess up the table;
her mother did spank her.

[P^Q]:
Jenny did mess up the table;
her mother didn't spank her.

[^PQ]:
Jenny didn't mess up the table;
her mother did spank her.

[^P^Q]:
Jenny didn't mess up the table;
her mother didn't spank her.

[text 13 : Threat 1]

P.C. Cooper approached a young girl who was waving her fists at passers-by. He told her

"If you threaten anyone else
then I'll arrest you."

[PQ]:
The girl did threaten someone else;
P.C. Cooper did arrest her.

[P~Q]:
The girl did threaten someone else;
P.C. Cooper didn't arrest her.

[~PQ]:
The girl didn't threaten someone else;
P.C. Cooper did arrest her.

[~P~Q]:
The girl didn't threaten someone else;
P.C. Cooper didn't arrest her.

[text 14 : Threat 2]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Jerry caused a stir by turning up for work in the bank wearing jeans and a teeshirt. Sending him home to change, the Manager told him

"If you wear jeans to work again tomorrow
then I shall fire you."

[PQ]:
Jerry did wear jeans to work again the next day;
the Manager did fire him.

[P~Q]:
Jerry did wear jeans to work again the next day;
the Manager didn't fire him.

[~PQ]:
Jerry didn't wear jeans to work again the next day;
the Manager did fire him.

[~P~Q]:
Jerry didn't wear jeans to work again the next day;
the Manager didn't fire him.

[text 15 : Threat 3]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Jane's squash was not going well. Sue, the team captain, told her

"If you lose your next game
then I shall take you out of the team."

[PQ]:
Jane did lose her next game;
Sue did take her out of the team.

[P~Q]:
Jane did lose her next game;
Sue didn't take her out of the team.

[~PQ]:
Jane didn't lose her next game;
Sue did take her out of the team.

[~P~Q]:
Jane didn't lose her next game;
Sue didn't take her out of the team.

[text 16 : Threat 4]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Rupert, a stockbroker, was advising his friend Algernon on his investments. He told him

"If you buy Acme Industries shares this week then you'll become a rich man."

[PQ]:
Algernon did buy Acme Industries shares that week; he did become a rich man.

[P~Q]:
Algernon did buy Acme Industries shares that week; he didn't become a rich man.

[~PQ]:
Algernon didn't buy Acme Industries shares that week; he did become a rich man.

[~P~Q]:
Algernon didn't buy Acme Industries shares that week; he didn't become a rich man.

[text 17 : Tip 1]

Bozo was asking around for the best place to find magic mushrooms. A girl in the pub told him

"If you try the moorland around Budworthy
then you'll find masses of them."

[PQ]:
Bozo did try the moorland around Budworthy;
he did find masses of them.

[P~Q]:
Bozo did try the moorland around Budworthy;
he didn't find masses of them.

[~PQ]:
Bozo didn't try the moorland around Budworthy;
he did find masses of them.

[~P~Q]:
Bozo didn't try the moorland around Budworthy;
he didn't find masses of them.

[text 18 : Tip 2]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Dennis was telling a garage mechanic friend of his that he was on the lookout for an MGB GT. His friend told him

"If you go to London
then you'll find a really cheap one."

[P Q]:
Dennis did go to London;
he did find a really cheap one.

[P ~Q]:
Dennis did go to London;
he didn't find a really cheap one.

[~P Q]:
Dennis didn't go to London;
he did find a really cheap one.

[~P ~Q]:
Dennis didn't go to London;
he didn't find a really cheap one.

[text 19 : Tip 3]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

It was Cathy's turn to buy the drinks at the Rocarola Disco and she was grumbling about how long it took to get served at the bar. Her friend Vicky pointed out that it all depended on where you stood. She told Cathy

"If you stand by the pillar
then you'll be served immediately."

[P Q]:
Cathy did stand by the pillar;
she was served immediately.

[P ~Q]:
Cathy did stand by the pillar;
she wasn't served immediately.

[~P Q]:
Cathy didn't stand by the pillar;
she was served immediately.

[~P ~Q]:
Cathy didn't stand by the pillar;
she wasn't served immediately.

[text 20 : Tip 4]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Sandy, a staunch Everton supporter, was discussing with his father whether it was safe to wear his team's colours when travelling to the forthcoming game with Arsenal. Sandy's father thought not and told him

"If you wear Everton's colours to the match
then you'll be beaten up on the train."

[PQ]:
Sandy did wear Everton's colours to the match;
he was beaten up on the train.

[P~Q]:
Sandy did wear Everton's colours to the match;
he wasn't beaten up on the train.

[~PQ]:
Sandy didn't wear Everton's colours to the match;
he was beaten up on the train.

[~P~Q]:
Sandy didn't wear Everton's colours to the match;
he wasn't beaten up on the train.

[text 21 : Warning 1]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

George, a new employee at a gas depot, was being shown the ropes by a well-established employee. He told George

"If the foreman catches you smoking on the site
then you'll be reprimanded."

[PQ]:
The foreman did catch George smoking on the site;
he was reprimanded.

[P~Q]:
The foreman did catch George smoking on the site;
he wasn't reprimanded.

[~PQ]:
The foreman didn't catch George smoking on the site;
he was reprimanded.

[~P~Q]:
The foreman didn't catch George smoking on the site;
he wasn't reprimanded.

[text 22 : Warning 2]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Diane was advising Norman on his next move during a game of chess. She told him

"If you play your queen
then you'll lose the game."

[P@]:
Norman did play his queen;
he did lose the game.

[P~@]:
Norman did play his queen;
he didn't lose the game.

[~P@]:
Norman didn't play his queen;
he did lose the game.

[~P~@]:
Norman didn't play his queen;
he didn't lose the game.

[text 23 : Warning 3]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Little Jemima was thoroughly enjoying herself pulling the family dog's tail. Her mother could see that the dog was getting annoyed and told Jemima

"If you pull his tail again
then he'll bite you."

[PQ]:
Jemima did pull the dog's tail again;
the dog did bite her.

[P~Q]:
Jemima did pull the dog's tail again;
the dog didn't bite her.

[~PQ]:
Jemima didn't pull the dog's tail again;
the dog did bite her.

[~P~Q]:
Jemima didn't pull the dog's tail again;
the dog didn't bite her.

[text 24 : Warning 4]

Having asked her to help him in his magic show, Paul Daniels asked Anita to pull a coloured counter from a bag without showing it to him. He then told her

> "If the counter is square
> then it is red in colour."

[PQ]:
The counter was square;
it was red in colour.

[P~Q]:
The counter was square;
it wasn't red in colour.

[~PQ]:
The counter wasn't square;
it was red in colour.

[~P~Q]:
The counter wasn't square;
it wasn't red in colour.

[text 25 : Universal 1]

Dr. Fassbender believed that he could tell from the subjects which students took at A-level which options they would take on the Psychology course. Suggesting that his colleague pick out a student at random, he told her

> "If he has an A-level in English
> then he is taking the Linguistics option."

[PQ]:
The student did have an A-level in English;
he was taking the Linguistics option.

[P~Q]:
The student did have an A-level in English;
he wasn't taking the Linguistics option.

[~PQ]:
The student didn't have an A-level in English;
he was taking the Linguistics option.

[~P~Q]:
The student didn't have an A-level in English;
he wasn't taking the Linguistics option.

[text 26 : Universal 2]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Petula was chatting to a stranger on a train about how she reckoned that people in particular jobs tended to have particular hobbies. By way of illustration, she told him

"If you're a civil servant
then you go in for home winemaking."

[PQ]:
The man was a civil servant;
he did go in for home winemaking.

[P~Q]:
The man was a civil servant;
he didn't go in for home winemaking.

[~PQ]:
The man wasn't a civil servant;
he did go in for home winemaking.

[~P~Q]:
The man wasn't a civil servant;
he didn't go in for home winemaking.

[text 27 : Universal 3]

Appendix 5d: Experiment 5: the 28 texts & outcome pairs

Lolita believed that students on certain courses supported particular political parties. Talking to a friend about this, she suggested that he ask the next student who came along which course he was on and which party he supported. She told him

> "If the student is doing Economics
> then he is a Socialist."

[PQ]:
The student was doing Economics;
he was a Socialist.

[P~Q]:
The student was doing Economics;
he wasn't a Socialist.

[~PQ]:
The student wasn't doing Economics;
he was a Socialist.

[~P~Q]:
The student wasn't doing Economics;
he wasn't a Socialist.

[text 28 : Universal 4]

Appendix 5e: Experiment 5: a typical screen (truth table task)

Jane's squash was not going well. Sue, the team captain, told her

"If you lose your next game
then I shall take you out of the team."

Jane did lose her next game; Sue didn't take her out of the team.	TELLS US NOTHING ABOUT the statement
Jane didn't lose her next game; Sue did take her out of the team.	
Jane didn't lose her next game; Sue didn't take her out of the team.	SUPPORTS the statement
Jane did lose her next game; Sue did take her out of the team.	CONTRADICTS the statement

Please indicate whether each state of affairs
SUPPORTS, CONTRADICTS or TELLS US NOTHING ABOUT the Statement ...

(press CONTINUE when you are happy with your decisions)

Do not worry if you find these instructions hard to follow, as you will have an opportunity to see them again as often as you like before the experiment begins.

In each of the 28 trials, you will be presented with a short passage of text on the computer screen, which will end with a statement made by someone who is mentioned in the passage - this statement will be shown in quotation marks. Below the passage of text will be shown four possible states of affairs, and you are required to use the keypad to indicate whether each state of affairs SUPPORTS, CONTRADICTS or TELLS US NOTHING ABOUT the statement in quotation marks.

You will notice that the keypad in front of you has twelve grey keys set out in four rows and three columns. Each of the four rows corresponds to one of the states of affairs shown on the screen: the top row refers to the first state of affairs, the next to the second and so on. The three columns are labelled so as to correspond to the three possible responses. If you think, for example, that the first state of affairs SUPPORTS the statement, you should press the key in the top row of the column labelled SUPPORTS. The response which you have made will appear on the screen next to the state of affairs concerned. Do not worry if you press the wrong key or wish to change your mind - simply press the key which you meant to press and the response on the screen will change accordingly.

When you have indicated what you think about each state of affairs, press CONTINUE to go on to the next trial. You cannot go on to the next trial until you have responded to ALL FOUR states of affairs. If you try to do so then the computer will beep and a question mark will flash against the remaining items to remind you.

Appendix 5f: Experiment 5: instructions (truth table task)

Please go through the instructions as often as you like: it is far more important that you understand what you are doing than that you make a quick start. If after a couple of runs through you are still in some doubt about what you have to do, then please call the experimenter to help you BEFORE going on to the experiment itself.

Appendix 5g: Experiment 5: debriefing text

Thank you very much indeed for taking part in this experiment, which was designed to investigate how people interpret statements of the form "if x then y" in a variety of different contexts.

There is no single "correct" answer in tasks such as this, and so there is no question of your answers reflecting anything whatever about your ability or intelligence.

Appendix 5h: Experiment 5: a typical screen (inference task)

Bozo was asking around for the best place to find magic mushrooms. A girl in the pub told him

"If you try the moorland around Budworthy
then you'll find masses of them.

Bozo did try the moorland around Budworthy.

Therefore,
he did find masses of them.

Please indicate whether or not the conclusion follows ...

(press CONTINUE when you are happy with your decision)

Appendix 5i: Experiment 5: instructions (inference task)

Thank you for volunteering to take part in this verbal reasoning experiment.

Do not worry if you find these instructions hard to follow, as you will have an opportunity to see them again as often as you like before the experiment begins.

In each of the 112 trials, you will be presented with a short passage of text on the computer screen, which will end with a statement made by someone who is mentioned in the passage - this statement will be shown in quotation marks. Beneath the quoted statement will be shown a sentence describing a state of affairs and a conclusion which may or may not follow from the statement and the state of affairs described.

You are required to indicate, using the keypad provided, whether the conclusion FOLLOWS or DOESN'T FOLLOW from the quoted statement and the state of affairs described beneath it. The response which you make will appear on the screen. Do not worry if you press the wrong key or wish to change your mind - simply press the other key and the response on the screen will change accordingly. When you have made your response, press CONTINUE to go on to the next trial. You cannot go on to the next trial until you have made a response. If you try to do so then the computer will beep and a question mark will flash on the screen to remind you.

Please go through the instructions as often as you like: it is far more important that you understand what you are doing than that you make a quick start. If after a couple of runs through you are still in some doubt about what you have to do, then please call the experimenter to help you BEFORE going on to the experiment itself.

Text 1 (items 1 + 2) : Temporal 1

[lo]:

Jim and Sally were waiting on the platform at Oxford Circus tube station, and Jim was showing off his knowledge of the Underground system. He knew, for example, that the trains always ran in a certain order, and that there were roughly the same number of trains every day to Woodford as there were to Leytonstone. To demonstrate his knowledge to Sally, he told her

[hi]:

Jim and Sally were waiting on the platform at Oxford Circus tube station, and Jim was showing off his knowledge of the Underground system. He knew, for example, that the trains always ran in a certain order, and that whilst there were only a few trains a day to Woodford, there were very many trains to Leytonstone. To demonstrate his knowledge to Sally, he told her

"If the next train is for Woodford
then the one after it will be for Leytonstone".

The next train was for Woodford;
the one after it was for Leytonstone.

The next train was for Woodford;
the one after it wasn't for Leytonstone.

The next train wasn't for Woodford;
the one after it was for Leytonstone.

The next train wasn't for Woodford;
the one after it wasn't for Leytonstone.

Text 2 (items 3 + 4) : Temporal 2

[lo]:

The Editor of Teenworld magazine was explaining to a new features editor the magazine's policy about the topic of each month's featured article. She explained that there were several subject areas, such as the pop scene and fitness, and that the magazine liked to cover all of them equally often in a constant order. As an example, she told him

[hi]:

The Editor of Teenworld magazine was explaining to a new features editor the magazine's policy about the topic of each month's featured article. She explained that certain subject areas, such as fitness were featured only occasionally, whilst others, such as the pop scene, were featured very frequently. The order in which the topics appeared, however, was constant, and as an example, she told him

"If next month's feature is on fitness
then the following month's will be on the pop scene."

The next month's feature was on fitness;
the following month's was on the pop scene.

The next month's feature was on fitness;
the following month's wasn't on the pop scene.

The next month's feature wasn't on fitness;
the following month's was on the pop scene.

The next month's feature wasn't on fitness;
the following month's wasn't on the pop scene.

Text 3 (items 5 + 6) : Temporal 3

[lo]:

Dave and Jo had just arrived home. They were particularly anxious to catch a Channel 4 news bulletin, but had no idea what the time was - they did, however, know that they must have missed all but the last Channel 4 News bulletin. As a television addict, Dave knew the order of all the programmes on Channel 4 that night. Turning on the television, he told Jo

[hi]:

Dave and Jo had just arrived home. They were particularly anxious to catch a Channel 4 News bulletin, but had no idea what the time was - they did, however, know that there were several Channel 4 News bulletins still to come. As a television addict, Dave knew that certain programmes were always put on just before a News bulletin. Turning on the television, he told Jo

"If the programme on Channel 4 is The Book Show
then the one after it will be the News."

The programme on Channel 4 was The Book Show;
the one after it was the News.

The programme on Channel 4 was The Book Show;
the one after it wasn't the News.

The programme on Channel 4 wasn't The Book Show;
the one after it was the News.

The programme on Channel 4 wasn't The Book Show;
the one after it wasn't the News.

Text 4 (items 7 + 8) : Temporal 4

[lo]:

Cedric and Amelia were great fans of opera, but had arrived home too late to catch the whole of La Traviata on the radio. Cedric particularly wanted to hear the tenor solo, and knew that it occurred only once in the opera. Amelia knew the order of all the sections of the opera and, switching on the radio, she told Cedric

[hi]:

Cedric and Amelia were great fans of opera, but had arrived home too late to catch the whole of La Traviata on the radio. Cedric particularly wanted to hear the tenor solo, but knew that it was repeated many times throughout the opera. Amelia knew the order of all the sections of the opera and, switching on the radio, she told Cedric

"If the current item is the aria
then the tenor solo will be next."

The current item was the aria;
the one after it was the tenor solo.

The current item was the aria;
the one after it wasn't the tenor solo.

The current item wasn't the aria;
the one after it was the tenor solo.

The current item wasn't the aria;
the one after it wasn't the tenor solo.

Text 5 (items 9 + 10) : Causal 1

[lo]:

A party of M.P.s was visiting the Netherfield Nuclear Power Station. The engineer showing them round was explaining the complex safety system, which had a different type of warning signal for each type of hazardous circumstance which might arise. By way of demonstration, he told them

[hi]:

A party of M.P.s was visiting the Netherfield Nuclear Power Station. The engineer showing them round was explaining the complex safety system, which was designed to sound a warning siren in a whole range of different hazardous circumstances. By way of demonstration, he told them

"If I jam the coolant valve open
then the warning siren will sound."

He did jam the coolant valve open;
the warning siren did sound.

He did jam the coolant valve open;
the warning siren didn't sound.

He didn't jam the coolant valve open;
the warning siren did sound.

He didn't jam the coolant valve open;
the warning siren didn't sound.

Text 6 (items 11 + 12) : Causal 2

[lo]:

A Psychology Department technician was explaining to a group of first year students how to look after their computer disks. He said that they now had a new type of disk which was extremely robust and hard to damage - the disk was, however, still vulnerable in one particular way. To show them what he meant, he picked up a demonstration disk and told them

[hi]:

A Psychology Department technician was explaining to a group of first year students how to look after their computer disks. He said that disks were very fragile and could become unreadable in many different ways. To show them what he meant, he picked up a demonstration disk and told them

"If I put the disk on top of the video monitor
then it will be unreadable when I try to use it."

He did put the disk on top of the video monitor;
the disk was unreadable when he tried to use it.

He did put the disk on top of the video monitor;
the disk wasn't unreadable when he tried to use it.

He didn't put the disk on top of the video monitor;
the disk was unreadable when he tried to use it.

He didn't put the disk on top of the video monitor;
the disk wasn't unreadable when he tried to use it.

Text 7 (items 13 + 14) : Causal 3

[lo]:

A neuro-surgeon was demonstrating to his colleagues a new device for use with stroke patients. The machine was designed to monitor movements of the facial muscles. He explained that the machine could be set up so that a blue light would flash when the patient made one type of facial muscle movement in particular. He wired himself up to the machine and told them

[hi]:

A neuro-surgeon was demonstrating to his colleagues a new device for use with stroke patients. The machine was designed to monitor movements of the facial muscles. He explained that the machine could be set up so that a blue light would flash when the patient made any facial muscle movement at all. He wired himself up to the machine and told them

"If I make a frown
then the blue light will flash."

He did make a frown;
the blue light did flash.

He did make a frown;
the blue light didn't flash.

He didn't make a frown;
the blue light did flash.

He didn't make a frown;
the blue light didn't flash.

Text 8 (items 15 + 16) : Causal 4

[lo]:

Arthur Makepiece was being shown around a security system factory by a salesman. The salesman told Arthur about a device which could recognise individual human faces. A demonstration version of the device, which spoke a greeting to any face it recognised, had been set up by the lift in the foyer of the building but, so far, had only been programmed with the face of the Chairman of the company. The salesman told Arthur

[hi]:

Arthur Makepiece was being shown around a security system factory by a salesman. The salesman told Arthur about a device which could recognise individual human faces. A demonstration version of the device, which spoke a greeting to any face it recognised, had been set up by the lift in the foyer of the building and had been programmed with the face of every person who worked in the building. The salesman told Arthur

"If the next person to pass it is the Chairman
then the device will say 'good morning'."

The next person to pass it was the chairman;
the device did say 'good morning'.

The next person to pass it was the chairman;
the device didn't say 'good morning'.

The next person to pass it wasn't the chairman;
the device did say 'good morning'.

The next person to pass it wasn't the chairman;
the device didn't say 'good morning'.

Text 9 (items 17 + 18) : Promise 1

[lo]:

It was Saturday, and Harry asked his father if he could borrow his car to go out with his mates that evening. His father was not particularly keen to lend Harry his car, but was anxious that the lawn should be cut that day, and saw a chance to get Harry to do it for him. He told Harry

[hi]:

It was Saturday, and Harry asked his father if he could borrow his car to go out with his mates that evening. His father was quite happy to lend Harry his car anyway, but had a lot to do in the garden that weekend and saw a chance to get Harry to give him a hand. He told Harry

"If you cut the lawn
then I will let you borrow the car tonight".

Harry did cut the lawn;
his father did let him borrow the car that night.

Harry did cut the lawn;
his father didn't let him borrow the car that night.

Harry didn't cut the lawn;
his father did let him borrow the car that night.

Harry didn't cut the lawn;
his father didn't let him borrow the car that night.

Text 10 (items 19 + 20) : Promise 2

[lo]:

Nick's girlfriend Jenny had asked him if he would drive her and her girlfriends to the aerobics class. Nick did not really have the time to do so, as he was running out of time on his statistics assignment which was due in quite soon, but told Jenny

[hi]:

Nick's girlfriend Jenny had asked him if he would drive her and her girlfriends to the aerobics class. Nick was always happy to run Jenny about, but was running out of time on several assignments which were due in quite soon, and saw that he could turn the situation to his advantage. He told Jenny

"If you help me with my statistics assignment
then I will drive you to the aerobics class".

Jenny did help him with his statistics assignment;
Nick did drive them to the aerobics class.

Jenny did help him with his statistics assignment;
Nick didn't drive them to the aerobics class.

Jenny didn't help him with his statistics assignment;
Nick did drive them to the aerobics class.

Jenny didn't help him with his statistics assignment;
Nick didn't drive them to the aerobics class.

Text 11 (items 21 + 22) : Promise 3

[lo]:

The workforce at Excelsior Industries had been on strike for several weeks. Despite the fact that negotiations had been going on all day and every day, there was still one point on which they had failed to reach agreement. The strikers were anxious to return to work as soon as possible, but the Union were determined that this last point should be settled in their favour. They told the Directors

[hi]:

The workforce at Excelsior Industries had been on strike for several weeks. Despite the fact that negotiations had been going on all day and every day, there were still a large number of points on which they had failed to reach agreement. The strikers were anxious to return to work as soon as possible, but the Union were determined to settle at least one more point in their favour. They told the Directors

"If you agree to provide shower facilities
then we will call off the strike today".

The Directors did agree to provide shower facilities;
the Union did call off the strike that day.

The Directors did agree to provide shower facilities;
the Union didn't call off the strike that day.

The Directors didn't agree to provide shower facilities;
the Union did call off the strike that day.

The Directors didn't agree to provide shower facilities;
the Union didn't call off the strike that day.

Text 12 (items 23 + 24) : Promise 4

[lo]:

John, the class swot, had asked Julia to go to the Rag Ball with him. Julia did not really want to accept, but saw an opportunity to get John to help her with the design of her mini-project. She replied

[hi]:

John, the class swot, had asked Julia to go to the Rag Ball with him. Julia was quite happy to accept, but also saw an opportunity to get John to help her with the design of her mini-project. She replied

"If you help me design my mini-project
then I will come with you to the Rag Ball".

John did help Julia design her mini-project;
Julia did go with him to the Rag Ball.

John did help Julia design her mini-project;
Julia didn't go with him to the Rag Ball.

John didn't help Julia design her mini-project;
Julia did go with him to the Rag Ball.

John didn't help Julia design her mini-project;
Julia didn't go with him to the Rag Ball.

Text 13 (items 25 + 26) : Threat 1

[lo]:

Dave, who was normally very tolerant and hated physical violence, was getting really wound up by the bloke next to him in the pub belching loudly every few minutes. He told him

[hi]:

Dave, who was known for his intolerance and who loved getting into fights, was getting really wound up by the bloke next to him in the pub belching loudly every few minutes. He told him

"If you belch again
then I'll thump you."

The bloke did belch again;
Dave did thump him.

The bloke did belch again;
Dave didn't thump him.

The bloke didn't belch again;
Dave did thump him.

The bloke didn't belch again;
Dave didn't thump him.

Text 14 (items 27 + 28) : Threat 2

[lo]:

P.C. Jones, the community policeman for the area around the Poly, was famous for his tolerance of the high spirits shown by students leaving the Union Bar on a Saturday night. One thing he would not turn a blind eye to, however, was damage to property. Seeing a student about to climb onto the roof of a parked car, he shouted

[hi]:

P.C. Jones, the community policeman for the area around the Poly, was famous for his intolerance of the high spirits shown by students leaving the Union Bar on a Saturday night. It was as if he would arrest students merely for enjoying themselves. Seeing a student about to climb onto the roof of a parked car, he shouted

"If you get up on that car
then I'll arrest you."

The student did get up onto the car;
P.C. Jones did arrest him.

The student did get up onto the car;
P.C. Jones didn't arrest him.

The student didn't get up onto the car;
P.C. Jones did arrest him.

The student didn't get up onto the car;
P.C. Jones didn't arrest him.

Text 15 (items 29 + 30) : Threat 3

[lo]:

Harry, a warehouse supervisor in a fireworks factory, was generally quite lax about enforcing the company's safety regulations. Since giving up smoking, however, he had become extremely strict about the no-smoking rule, which prohibited staff even from having cigarettes with them in the warehouse, let alone smoking them. Seeing one of the warehouse porters about to light a cigarette, he told him

[hi]:

Harry, a warehouse supervisor in a fireworks factory, was extremely strict about enforcing the company's safety regulations. One such regulation was the no-smoking rule, which prohibited staff even from having cigarettes with them in the warehouse, let alone smoking them. Seeing one of the warehouse porters about to light a cigarette, he told him

"If you light that cigarette
then I'll report you to the management."

The porter did light the cigarette;
Harry did report him to the management.

The porter did light the cigarette;
Harry didn't report him to the management.

The porter didn't light the cigarette;
Harry did report him to the management.

The porter didn't light the cigarette;
Harry didn't report him to the management.

Text 16 (items 31 + 32) : Threat 4

[lo]:

Little Billy's mother strongly disapproved of parents smacking their children for doing something wrong. On one particular day, when his mother was feeling particularly fragile after a heavy evening out, Billy seemed to be in an especially naughty mood. When he turned the radio up full blast, his mother yelled at him

[hi]:

Little Billy's mother, strongly believed that children should be smacked for the merest naughtiness. On one particular day, when his mother was feeling particularly fragile after a heavy evening out, Billy seemed to be in an especially naughty mood. When he turned the radio up full blast, his mother yelled at him

"If you turn that radio up again
then I shall smack you."

Billy did turn the radio up again;
his mother did smack him.

Billy did turn the radio up again;
his mother didn't smack him.

Billy didn't turn the radio up again;
his mother did smack him.

Billy didn't turn the radio up again;
his mother didn't smack him.

Text 17 (items 33 + 34) : Tip 1

[lo]:

Henry Foggett, a financial consultant, was advising a client with a large sum of money to invest. The client had decided to invest the money in the robotics industry. Henry was rather doubtful about the robotics industry in general, but did know of one particular robotics company that looked like a really good investment. He told his client

[hi]:

Henry Foggett, a financial consultant, was advising a client with a large sum of money to invest. The client had decided to invest the money in the robotics industry. Henry was of the opinion that robotics companies in general were good investments. He told his client

"If you use it to buy Rotronics shares
then you'll get a good return on your investment."

The client did use his money to buy Rotronics shares;
he did get a good return on his investment.

The client did use his money to buy Rotronics shares;
he didn't get a good return on his investment.

The client didn't use his money to buy Rotronics shares;
he did get a good return on his investment.

The client didn't use his money to buy Rotronics shares;
he didn't get a good return on his investment.

Text 18 (items 35 + 36) : Tip 2

[lo]:

Stanley, a keen bird watcher, had come to Devon to try to spot the red-breasted grebe. A local enthusiast whom he met in a pub knew that they were extremely hard to find, but told him

[hi]:

Stanley, a keen bird watcher, had come to Devon to try to spot the red-breasted grebe. A local enthusiast whom he met in a pub knew that there were a great many of them about and told him

"If you try Wistman's Wood in the evening
then you'll spot one."

Stanley did try Wistman's Wood in the evening;
he did spot a red-breasted grebe.

Stanley did try Wistman's Wood in the evening;
he didn't spot a red-breasted grebe.

Stanley didn't try Wistman's Wood in the evening;
he did spot a red-breasted grebe.

Stanley didn't try Wistman's Wood in the evening;
he didn't spot a red-breasted grebe.

Text 19 (items 37 + 38) : Tip 3

[lo]:

Little Sarah was delighted by the great dane belonging to her father's friend, Jim, and wanted to ride on his back around the garden. The dog did not normally let people get on his back, but would do almost anything for a chocolate biscuit. Jim told Sarah

[hi]:

Little Sarah was delighted by the great dane belonging to her father's friend, Jim, and wanted to ride on his back around the garden. The dog loved to play with children once he had got to know them, and Jim told Sarah

"If you give him a chocolate biscuit
then he'll let you ride him around the garden."

Sarah did give the dog a chocolate biscuit;
the dog did let her ride him around the garden.

Sarah did give the dog a chocolate biscuit;
the dog didn't let her ride him around the garden.

Sarah didn't give the dog a chocolate biscuit;
the dog did let her ride him around the garden.

Sarah didn't give the dog a chocolate biscuit;
the dog didn't let her ride him around the garden.

Text 20 (items 39 + 40) : Tip 4

[lo]:

Mark was really keen on Maggie, but whenever he invited her out for a drink she seemed to have something else planned. Mark was moaning about the situation to Maggie's friend, Sue, who knew that Maggie did not particularly like Mark. She did, however, know that Maggie was crazy about flat-top haircuts, and told him

[hi]:

Mark was really keen on Maggie, but whenever he invited her out for a drink she seemed to have something else planned. Mark was moaning about the situation to Maggie's friend, Sue, who knew that Maggie quite liked Mark but that she was put off by his untrendy appearance. She told him

"If you get a flat-top haircut
then she will agree to go out with you."

Mark did get a flat-top haircut;
Maggie did agree to go out with him.

Mark did get a flat-top haircut;
Maggie didn't agree to go out with him.

Mark didn't get a flat-top haircut;
Maggie did agree to go out with him.

Mark didn't get a flat-top haircut;
Maggie didn't agree to go out with him.

Text 21 (items 41 + 42) : Warning 1

[lo]:

Towser, the family dog, was remarkably good-tempered, and would allow the children to do virtually anything to him without getting annoyed. Recently, however, his tail had been shut in a door and it was still very sore. Benny, a fearless toddler, started pulling Towser's tail - his mother told him

[hi]:

Towser, the family dog, had become extremely short-tempered as he got older, and it was clear that he didn't like having anything at all to do with the children. Benny, a fearless toddler, started pulling Towser's tail - his mother told him

"If you pull his tail again
then he'll bite you."

Benny did pull Towser's tail again;
Towser did bite him.

Benny did pull Towser's tail again;
Towser didn't bite him.

Benny didn't pull Towser's tail again;
Towser did bite him.

Benny didn't pull Towser's tail again;
Towser didn't bite him.

Text 22 (items 43 + 44) : Warning 2

[lo]:

Cyril Higgins had a large sum of money which he had decided to invest in one of the night-clubs in the city where he lived. He had been offered the chance to invest it in The Pussycat Club and was seeking his bank manager's advice. His bank manager had heard that there had been so many complaints about noise from The Pussycat Club that it was likely to lose its licence. He told Cyril

[hi]:

Cyril Higgins had a large sum of money which he had decided to invest in one of the night-clubs in the city where he lived. He had been offered the chance to invest it in The Pussycat Club and was seeking his bank manager's advice. His bank manager had heard that the city council were going to force all the night-clubs in the city to close at midnight, and could see that this would put many of them out of business. He told Cyril

"If you invest in The Pussycat Club
then you'll lose your money."

Cyril did invest in The Pussycat Club;
he did lose his money.

Cyril did invest in The Pussycat Club;
he didn't lose his money.

Cyril didn't invest in The Pussycat Club;
he did lose his money.

Cyril didn't invest in The Pussycat Club;
he didn't lose his money.

Text 23 (items 45 + 46) : Warning 3

[lo]:

Pete was about to lend his beaten-up old car to Geoff, so that he could go to Exeter for the weekend. He explained to Geoff that although it looked as though it was about to fall to pieces, the car was extremely reliable so long as you didn't go over a certain speed. He told him

[hi]:

Pete was about to lend his beaten-up old car to Geoff, so that he could go to Exeter for the weekend. He explained to Geoff that the car was not particularly reliable, but would probably get him there so long as he took it gently. He told him

"If you go over 50 mph
then it'll break down."

Geoff did go over 50 mph;
the car did break down.

Geoff did go over 50 mph;
the car didn't break down.

Geoff didn't go over 50 mph;
the car did break down.

Geoff didn't go over 50 mph;
the car didn't break down.

Text 24 (items 47 + 48) : Warning 4

[lo]:

Boris Spassky was giving a chess Master Class to a group of young chess players. He approached one particular pair of players, Peter and Paula, for whom the game had almost finished, so that there were only two or three of each player's pieces left on the board. It was Paula's move. After considering the position, Boris whispered in Paula's ear

[hi]:

Boris Spassky was giving a chess Master Class to a group of young chess players. He approached one particular pair of players, Peter and Paula, for whom the game had only just started, so that almost all the pieces were still on the board. It was Paula's move. After considering the position, Boris whispered in Paula's ear

"If you play the queen
then you'll lose the game."

Paula did play the queen;
she did lose the game

Paula did play the queen;
she didn't lose the game

Paula didn't play the queen;
she did lose the game

Paula didn't play the queen;
she didn't lose the game

Text 25 (items 49 + 50) : Universal 1

[lo]:

The Manager of Peach Computers had been investigating the large number of machines that had been returned with a variety of faults. He had come to the conclusion that all the computers which had a fault in the keyboard were from a single batch. It was possible to tell which batch a computer came from by looking at the first three letters of the serial number. Picking up one of the faulty machines at random, he said

[hi]:

The Manager of Peach Computers had been investigating the large number of machines that had been returned with a variety of faults. He had come to the conclusion that all the computers which had a fault in the keyboard were from one of several particular batches. It was possible to tell which batch a computer came from by looking at the first three letters of the serial number. Picking up one of the faulty machines at random, he said

"If the serial number begins with 'XZP'
then there will be a fault in the keyboard."

The serial number did begin with 'XZP';
there was a fault in the keyboard.

The serial number did begin with 'XZP';
there wasn't a fault in the keyboard.

The serial number didn't begin with 'XZP';
there was a fault in the keyboard.

The serial number didn't begin with 'XZP';
there wasn't a fault in the keyboard.

Text 26 (items 51 + 52) : Universal 2

[lo]:

Tom had arrived in Bardsea in search of an old friend. He had lost his address, but knew that he was a member of one of the many sailing clubs in Bardsea and hoped to get his address from them. Tom knew the sort of boat which his friend had, and was asking the Harbour Master which club to try first. The Harbour Master told him that one of the clubs, the Bardwater, tended to specialise in just one type of boat. He told him

[hi]:

Tom had arrived in Bardsea in search of an old friend. He had lost his address, but knew that he was a member of one of the many sailing clubs in Bardsea and hoped to get his address from them. Tom knew the sort of boat which his friend had, and was asking the Harbour Master which club to try first. The Harbour Master told him that one of the clubs, the Bardwater, tended to specialise in several types of boat in particular. He told him

"If your friend owns a Laser
then he'll be a member of the Bardwater."

His friend did own a Laser;
he was a member of the Bardwater.

His friend did own a Laser;
he wasn't a member of the Bardwater.

His friend didn't own a Laser;
he was a member of the Bardwater.

His friend didn't own a Laser;
he wasn't a member of the Bardwater.

Text 27 (items 53 + 54) : Universal 3

[lo]:

Shiela and Mary were on holiday in Athens. They desperately needed to know what day of the week it was, but didn't speak a word of Greek. Mary knew, however, that to combat pollution, there was a law that only cars with registration numbers beginning with certain letters were allowed into the city on any given day. For example, only cars with the letter A were allowed in on Mondays. Looking at a passing car, she told Shiela

[hi]:

Shiela and Mary were on holiday in Athens. They desperately needed to know what day of the week it was, but didn't speak a word of Greek. Mary knew, however, that to combat pollution, there was a law that only cars with registration numbers beginning with certain letters were allowed into the city on any given day. For example, only cars with letters from A to M were allowed in on Mondays. Looking at a passing car, she told Shiela

"If the registration starts with A
then it is Monday."

The registration did start with A;
it was Monday.

The registration did start with A;
it wasn't Monday.

The registration didn't start with A;
it was Monday.

The registration didn't start with A;
it wasn't Monday.

Text 28 (items 55 + 56) : Universal 4

[lo]:

Wisden Poly ran a very large number of different courses. Sue Grub, the Union President, was telling Nick, her deputy, that there was not enough mixing between students from different courses - it seemed to her that the membership of many of the student societies tended to comprise students from just one course. To demonstrate her point, Sue suggested that Nick stop the next student that came along and said

[hi]:

Wisden Poly ran a very large number of different courses. Sue Grub, the Union President, was telling Nick, her deputy, that there was not enough mixing between students from different courses - it seemed to her that the membership of many of the student societies tended to comprise students from just certain particular courses. To demonstrate her point, Sue suggested that Nick stop the next student that came along and said

"If she is on the Biochemistry degree
then she is a member of the Ecology Society."

The student was on the Biochemistry degree;
she was a member of the Ecology Society.

The student was on the Biochemistry degree;
she wasn't a member of the Ecology Society.

The student wasn't on the Biochemistry degree;
she was a member of the Ecology Society.

The student wasn't on the Biochemistry degree;
she wasn't a member of the Ecology Society.

Appendix 7a: Experiment 7: the 4 additional 'Intentional' texts & outcome sentences

---

Intentional text 1 (items 57 + 58)

[lo]:
After three weeks of almost continuous rain, Sara was delighted to hear on the radio that the weather was about to clear up. She told her friend Jonah that it was going to be a fine weekend and asked him how he would spend it. Jonah was a keen angler but hated sitting out in the rain. He told Sara

[hi]:
After three weeks of almost continuous rain, Sara was delighted to hear on the radio that the weather was about to clear up. She told her friend Jonah that it was going to be a fine weekend and asked him how he would spend it. Jonah was a keen angler and was happy sitting out in virtually any weather. He told Sara

"If it is fine over the weekend
then I'll go fishing."

It was fine over the weekend;
Jonah did go fishing.

It was fine over the weekend;
Jonah didn't go fishing.

It wasn't fine over the weekend;
Jonah did go fishing.

It wasn't fine over the weekend;
Jonah didn't go fishing.

[... Jonah the angler and the change in the weather ...]

What effect did the weather tend to have on whether Jonah went fishing ?

[lo]:
He hated sitting out in the rain

[hi]:
He was happy sitting out in virtually any weather

Appendix 7a: Experiment 7: the 4 additional 'Intentional' texts & outcome sentences

---

Intentional text 2 (items 59 + 60)

[lo]:
Bill was talking to his friend Susie, a nurse at the local hospital, about how she would be affected by the Tories being re-elected at the forthcoming General Election. Susie told him that she was convinced that the Tories getting in again would ruin her promotion prospects. She had been reading about the excellent opportunities for nurses in Canada but was unhappy about leaving her friends and family. She told Bill

[hi]:
Bill was talking to his friend Susie, a nurse at the local hospital, about how she would be affected by the Tories being re-elected at the forthcoming General Election. Susie told him that she was convinced that the Tories getting in again would ruin her promotion prospects. She had been reading about the excellent opportunities for nurses in Canada and was quite keen on the idea of emigrating. She told Bill

"If the Tories are re-elected
then I'll emigrate to Canada."

The Tories were re-elected;
Susie did emigrate to Canada.

The Tories were re-elected;
Susie didn't emigrate to Canada.

The Tories weren't re-elected;
Susie did emigrate to Canada.

The Tories weren't re-elected;
Susie didn't emigrate to Canada.

[... the effect of the Tories getting in again on nurses' jobs ...]

How did Susie feel in general about the idea of emigrating to Canada ?

[lo]:
She was unhappy about leaving her friends and family

[hi]:
She was quite keen on the idea

Appendix 7a: Experiment 7: the 4 additional 'Intentional' texts & outcome sentences

---

Intentional text 3 (items 61 + 62)

[lo]:
Colonel "Buffy" Frobisher, Chairman of of the Highsea Officers Club, was asking one of the members, Admiral Cholmondley-Smyth, what he thought about the proposal to allow women to become members of the club. The Admiral disliked women intensely, but going to the club was his only pleasure, and he spent almost all of his time there. He told the Colonel

[hi]:
Colonel "Buffy" Frobisher, Chairman of of the Highsea Officers Club, was asking one of the members, Admiral Cholmondley-Smyth, what he thought about the proposal to allow women to become members of the club. The Admiral hardly ever visited the club, and had been on the point of resigning for some time. He told the Colonel

"If the committee agrees to let women join
then I'll resign."

The committee did agree to let women join;
the Admiral did resign.

The committee did agree to let women join;
the Admiral didn't resign.

The committee didn't agree to let women join;
the Admiral did resign.

The committee didn't agree to let women join;
the Admiral didn't resign.

[... Highsea Officers Club ...]

How keen a member of the club was Admiral Cholmondley-Smyth ?

[lo]:
It was his only pleasure and he spent most of his time there

[hi]:
He hardly ever visited the club and had been on the point of resigning for some time

---

Appendix 7a: Experiment 7: the 4 additional 'Intentional' texts & outcome sentences

---

Intentional text 4 (items 63 + 64)

[lo]:
Sid, who ran a corner shop and off-licence, had heard that Oddbins had applied for planning permission to open an off-licence in the next street. Sid's shop was quite profitable, but mainly because it was the only off-licence in the area, and he realised that his profits would fall dramatically if Oddbins did get permission. Sid's wife had wanted him to retire for years, but so far Sid had always refused. He told her

[hi]:
Sid, who ran a corner shop and off-licence, had heard that Oddbins had applied for planning permission to open an off-licence in the next street. Sid's shop was not particularly profitable, even though it was the only off-licence in the area, and he realised that his profits would fall dramatically if Oddbins did get permission. Sid's wife had wanted him to retire for years, and Sid was on the verge of agreeing anyway. He told her

"If Oddbins do get planning permission
then I'll agree to retire."

Oddbins did get planning permission;
Sid did agree to retire.

Oddbins did get planning permission;
Sid didn't agree to retire.

Oddbins didn't get planning permission;
Sid did agree to retire.

Oddbins didn't get planning permission;
Sid didn't agree to retire.

[... Oddbins hoping to open an off-licence near Sid's corner shop ...]

How had Sid reacted until now to his wife's suggestion that he retire ?

[lo]:
So far he had always refused to retire

[hi]:
He was on the verge of agreeing to retire

[... John wanting Julia to go to the Rag Ball ...]

How did Julia feel about going to the Rag Ball with John ?

[A]:
She was quite happy to accept

[B]:
She did not really want to accept

Press any grey key in the top row for answer A
or any grey key in the bottom row for answer B

< press CONTINUE when you are happy with your decision >

Appendix 7c: Experiment 7: instructions (memory task)

The memory task now follows. You will be asked one question about each of the 32 texts which were presented during the reasoning task. In each case, you will be presented with a "reminder" phrase, to help you bring the relevant passage of text to mind.

Beneath this "reminder" will be shown a question and two possible answers, A and B. Any key in the top row will select answer A, and any key in the bottom row will select answer B. Don't worry if you press the wrong key by mistake - your choice will be noted on the screen and you can change your mind simply by pressing another key. When you are happy with your selection, then you should press CONTINUE to go on to the next trial.

Appendix 8a: Experiment 8: the 32 texts & inferences

---

text 01 : Temporal 1

As the Rugby Club singsong was about to start, the team captain was explaining to a new member that it was a club tradition that certain songs always followed certain others and told him as an example

"If the first song is 'Roll Out The Barrel'
then the second song will be 'Yellow Submarine'."

[MP]:

The first song was 'Roll Out The Barrel'

Therefore,

The second song was 'Yellow Submarine'

[AC]:

The second song was 'Yellow Submarine'

Therefore,

The first song was 'Roll Out The Barrel'

[DA]:

The first song wasn't 'Roll Out The Barrel'

Therefore,

The second song wasn't 'Yellow Submarine'

[MT]:

The second song wasn't 'Yellow Submarine'

Therefore,

The first song wasn't 'Roll Out The Barrel'

Appendix 8a: Experiment 8: the 32 texts & inferences

---

text 02 : Temporal 2

Sally and Nick, the Year 1 reps, were about to walk in late to a Course Committee meeting and were unsure whether they had missed the item about the Computer Centre. Sally knew that certain items were always dealt with in a particular order, and told Nick

"If they're discussing the LRC when we arrive
then the following item will be about the Computer Centre."

[MP]:

They were discussing the LRC when Sally and Nick arrived

Therefore,

The following item was about the Computer Centre

[AC]:

The following item was about the Computer Centre

Therefore,

They were discussing the LRC when Sally and Nick arrived

[DA]:

They weren't discussing the LRC when Sally and Nick arrived

Therefore,

The following item wasn't about the Computer Centre

[MT]:

The following item wasn't about the Computer Centre

Therefore,

They weren't discussing the LRC when Sally and Nick arrived

---

text 03 : Temporal 3

Cedric, an ardent transport enthusiast, had memorised the entire London Underground timetable. Standing on Earl's Court station, he told his friend Arthur

"If the first train is for Ealing
then the second will be for Wimbledon

[MP]:

The first train was for Ealing

Therefore,

The second train was for Wimbledon

[AC]:

The second train was for Wimbledon

Therefore,

The first train was for Ealing

[DA]:

The first train wasn't for Ealing

Therefore,

The second train wasn't for Wimbledon

[MT]:

The second train wasn't for Wimbledon

Therefore,

The first train wasn't for Ealing

---

text 04 : Temporal 4

The editor of 'Metropolitan' magazine was telling a new staff writer that it was the magazine's policy to run features on certain topic areas in strict rotation. By way of example, she told her

"If the current month's feature is about relationships then the following month's will be about fitness."

[MP]:

The current month's feature was about relationships

Therefore,

The following month's feature was about fitness

[AC]:

The following month's feature was about fitness

Therefore,

The current month's feature was about relationships

[DA]:

The current month's feature wasn't about relationships

Therefore,

The following month's feature wasn't about fitness

[MT]:

The following month's feature wasn't about fitness

Therefore,

The current month's feature wasn't about relationships

---

text 05 : Causal 1

Sue and Mark were showing their new television to a friend of theirs and complaining about the trouble they were having with it. They told him

"If it starts to rain
then the picture goes fuzzy."

[MP]:

It did start to rain

Therefore,

The picture did go fuzzy

[AC]:

The picture did go fuzzy

Therefore,

It did start to rain

[DA]:

It didn't start to rain

Therefore,

The picture didn't go fuzzy

[MT]:

The picture didn't go fuzzy

Therefore,

It didn't start to rain

---

text 06 : Causal 2

Sarah, a training officer at a nuclear power plant, was explaining the safety precautions to some new recruits. She told them

"If a leak develops in the cooling system
then the reactor will shut down."

[MP]:

A leak did develop in the cooling system

Therefore,

The reactor did shut down

[AC]:

The reactor did shut down

Therefore,

A leak did develop in the cooling system

[DA]:

A leak didn't develop in the cooling system

Therefore,

The reactor didn't shut down

[MT]:

The reactor didn't shut down

Therefore,

A leak didn't develop in the cooling system

---

text 07 : Causal 3

Harry, a weighbridge operator, was explaining the controls to a visitor. Suggesting that the visitor wait for the next lorry that came in to be weighed, he told him

"If the lorry is heavier than the legal limit
then the alarm bell will ring."

[MP]:

The lorry was heavier than the legal limit

Therefore,

The alarm bell did ring

[AC]:

The alarm bell did ring

Therefore,

The lorry was heavier than the legal limit

[DA]:

The lorry wasn't heavier than the legal limit

Therefore,

The alarm bell didn't ring

[MT]:

The alarm bell didn't ring

Therefore,

The lorry wasn't heavier than the legal limit

---

text 08 : Causal 4

Bert, a canal lock-keeper, was telling a holidaymaker how the lock worked. Suggesting that they watch the next boat go through the lock, he told the holidaymaker

"If the water in the lock goes above the red line
then it will spill over into the canal."

[MP]:

The water in the lock did go above the red line

Therefore,

The water did spill over into the canal

[AC]:

The water did spill over into the canal

Therefore,

The water in the lock did go above the red line

[DA]:

The water in the lock didn't go above the red line

Therefore,

The water didn't spill over into the canal

[MT]:

The water didn't spill over into the canal

Therefore,

The water in the lock didn't go above the red line

---

text 09 : Promise 1

Susan's mother wanted her to stay in on Saturday night. She told Susan

"If you stay in on Saturday night
then I'll take you sailing on Sunday."

[MP]:

Susan did stay in on Saturday night

Therefore,

Her mother did take her sailing on Sunday

[AC]:

Her mother did take her sailing on Sunday

Therefore,

Susan did stay in on Saturday night

[DA]:

Susan didn't stay in on Saturday night

Therefore,

Her mother didn't take her sailing on Sunday

[MT]:

Her mother didn't take her sailing on Sunday

Therefore,

Susan didn't stay in on Saturday night

---

text 10 : Promise 2

Bobby's girlfriend wanted him to pick her up from the gym. She told him

"If you pick me up from the gym
then I'll help you with your essay tomorrow."

[MP]:

Bobby did pick her up from the gym

Therefore,

She did help him with his essay the next day

[AC]:

She did help him with his essay the next day

Therefore,

Bobby did pick her up from the gym

[DA]:

Bobby didn't pick her up from the gym

Therefore,

She didn't help him with his essay the next day

[MT]:

She didn't help him with his essay the next day

Therefore,

Bobby didn't pick her up from the gym

---

text 11 : Promise 3

Henry's father wanted him to wash the car. He told Henry

"If you wash the car this afternoon
then I'll let you borrow it tonight."

[MP]:

Henry did wash the car that afternoon

Therefore,

His father did let him borrow the car that night

[AC]:

His father did let him borrow the car that night

Therefore,

Henry did wash the car that afternoon

[DA]:

Henry didn't wash the car that afternoon

Therefore,

His father didn't let him borrow the car that night

[MT]:

His father didn't let him borrow the car that night

Therefore,

Henry didn't wash the car that afternoon

---

text 12 : Promise 4

The Director of Norshire Polytechnic was anxious to bring the student sit-in to a close. He told the protesters

"If you leave the building immediately
then the Poly will provide a creche before next term."

[MP]:

The students did leave the building immediately

Therefore,

The Poly did provide a creche before the following term

[AC]:

The Poly did provide a creche before the following term

Therefore,

The students did leave the building immediately

[DA]:

The students didn't leave the building immediately

Therefore,

The Poly didn't provide a creche before the following term

[MT]:

The Poly didn't provide a creche before the following term

Therefore,

The students didn't leave the building immediately

---

text 13 : Threat 1

Little Jenny's parents were having friends around for dinner. Having just set the dining table, her mother told Jenny

"If you mess up the table
then I'll spank you."

[MP]:

Jenny did mess up the table

Therefore,

Her mother did spank her

[AC]:

Her mother did spank her

Therefore,

Jenny did mess up the table

[DA]:

Jenny didn't mess up the table

Therefore,

Her mother didn't spank her

[MT]:

Her mother didn't spank her

Therefore,

Jenny didn't mess up the table

---

text 14 : Threat 2

P.C. Cooper approached a young girl who was waving her fists at passers-by. He told her

"If you threaten anyone else
then I'll arrest you."

[MP]:

The girl did threaten someone else

Therefore,

P.C. Cooper did arrest her

[AC]:

P.C. Cooper did arrest her

Therefore,

The girl did threaten someone else

[DA]:

The girl didn't threaten anyone else

Therefore,

P.C. Cooper didn't arrest her

[MT]:

P.C. Cooper didn't arrest her

Therefore,

The girl didn't threaten anyone else

---

text 15 : Threat 3

Jerry caused a stir by turning up for work in the bank wearing jeans and a teeshirt. Sending him home to change, the manager told him

"If you wear jeans to work again tomorrow
then I shall fire you."

[MP]:

Jerry did wear jeans to work again the following day

Therefore,

The manager did fire him

[AC]:

The manager did fire him

Therefore,

Jerry did wear jeans to work again the following day

[DA]:

Jerry didn't wear jeans to work again the following day

Therefore,

The manager didn't fire him

[MT]:

The manager didn't fire him

Therefore,

Jerry didn't wear jeans to work again the following day

---

text 16 : Threat 4

Jane's squash was not going well. Sue, the team captain, told her

"If you lose your next game
then I shall take you out of the team."

[MP]:

Jane didn't lose her next game

Therefore,

Sue didn't take her out of the team

[AC]:

Sue didn't take her out of the team

Therefore,

Jane didn't lose her next game

[DA]:

Jane didn't lose her next game

Therefore,

Sue didn't take her out of the team

[MT]:

Sue didn't take her out of the team

Therefore,

Jane didn't lose her next game

---

text 17 : Tip 1

Rupert, a stockbroker, was advising his friend Algernon on his investments. He told him

"If you buy Acme Industries shares this week
then you'll become a rich man."

[MP]:

He did buy Acme Industries shares that week

Therefore,

He did become a rich man

[AC]:

He did become a rich man

Therefore,

He did buy Acme Industries shares that week

[DA]:

He didn't buy Acme Industries shares that week

Therefore,

He didn't become a rich man

[MT]:

He didn't become a rich man

Therefore,

He didn't buy Acme Industries shares that week

---

text 18 : Tip 2

Bozo was asking around for the best place to find magic mushrooms. A girl in the pub told him

"If you try the moorland around Budworthy
then you'll find masses of them."

[MP]:

He did try the moorland around Budworthy

Therefore,

He did find masses of magic mushrooms

[AC]:

He did find masses of magic mushrooms

Therefore,

He did try the moorland around Budworthy

[DA]:

He didn't try the moorland around Budworthy

Therefore,

He didn't find masses of magic mushrooms

[MT]:

He didn't find masses of magic mushrooms

Therefore,

He didn't try the moorland around Budworthy

---

text 19 : Tip 3

Dennis was telling a garage mechanic friend of his that he was on the lookout for an MGB GT. His friend told him

"If you go to London
then you'll find a really cheap one."

[MP]:

He did go to London

Therefore,

He did find a really cheap one

[AC]:

He did find a really cheap one

Therefore,

He did go to London

[DA]:

He didn't go to London

Therefore,

He didn't find a really cheap one

[MT]:

He didn't find a really cheap one

Therefore,

He didn't go to London

---

text 20 : Tip 4

It was Cathy's turn to buy the drinks at the Rocarola Disco and she was grumbling about how long it took to get served at the bar. Her friend Vicky pointed out that it all depended on where you stood - she told Cathy

"If you stand by the pillar
then you'll be served immediately."

[MP]:

She did stand by the pillar

Therefore,

She was served immediately

[AC]:

She was served immediately

Therefore,

She did stand by the pillar

[DA]:

She didn't stand by the pillar

Therefore,

She wasn't served immediately

[MT]:

She wasn't served immediately

Therefore,

She didn't stand by the pillar

---

text 21 : Warning 1

Sandy, a staunch Everton supporter, was discussing with his father whether it was safe to wear his team's colours when travelling to the forthcoming game with Arsenal. Sandy's father thought not, and told Sandy

"If you wear Everton's colours to the match
then you'll be beaten up on the train."

[MP]:

He did wear Everton's colours to the match

Therefore,

He was beaten up on the train

[AC]:

He was beaten up on the train

Therefore,

He did wear Everton's colours to the match

[DA]:

He didn't wear Everton's colours to the match

Therefore,

He wasn't beaten up on the train

[MT]:

He wasn't beaten up on the train

Therefore,

He didn't wear Everton's colours to the match

---

text 22 : Warning 2

George, a new employee at a gas depot, was being shown the ropes by a well-established employee. He told George

"If the foreman catches you smoking on the site then you'll be reprimanded."

[MP]:

The foreman did catch George smoking on the site

Therefore,

George was reprimanded

[AC]:

George was reprimanded

Therefore,

The foreman did catch George smoking on the site

[DA]:

The foreman didn't catch George smoking on the site

Therefore,

George wasn't reprimanded

[MT]:

George wasn't reprimanded

Therefore,

The foreman didn't catch George smoking on the site

---

text 23 : Warning 3

Diane was advising Norman on his next move during a game of chess. She told him

"If you play your queen
then you will lose the game."

[MP]:

He did play his queen

Therefore,

He did lose the game

[AC]:

He did lose the game

Therefore,

He did play his queen

[DA]:

He didn't play his queen

Therefore,

He didn't lose the game

[MT]:

He didn't lose the game

Therefore,

He didn't play his queen

---

text 24 : Warning 4

Little Jemima was thoroughly enjoying herself pulling the family dog's tail. Her mother could see that the dog was getting annoyed and told Jemima

"If you pull his tail again
then he'll bite you."

[MP]:

She did pull his tail again

Therefore,

The dog did bite her

[AC]:

The dog did bite her

Therefore,

She did pull his tail again

[DA]:

She didn't pull his tail again

Therefore,

The dog didn't bite her

[MT]:

The dog didn't bite her

Therefore,

She didn t pull his tail again

---

text 25 : Universal 1

Having asked her to help him in his magic show, Paul Daniels asked Anita to pull a coloured counter from a bag without showing it to him. He then told her

"If the counter is square in shape
then it is red in colour."

[MP]:

It was square in shape

Therefore,

It was red in colour

[AC]:

It was red in colour

Therefore,

It was square in shape

[DA]:

It wasn't square in shape

Therefore,

It wasn't red in colour

[MT]:

It wasn't red in colour

Therefore,

It wasn't square in shape

---

text 26 : Universal 2

Dr. Fassbender believed that he could tell from the subjects which students took at A-level which options they would take on the Psychology course. Suggesting that his colleague pick out a student at random, he told him

"If the student has an A-level in English
then she is taking the Linguistics option."

[MP]:

She did have an A-level in English

Therefore,

She was taking the Linguistics option

[AC]:

She was taking the Linguistics option

Therefore,

She did have an A-level in English

[DA]:

She didn't have an A-level in English

Therefore,

She wasn't taking the Linguistics option

[MT]:

She wasn't taking the Linguistics option

Therefore,

She didn't have an A-level in English

---

text 27 : Universal 3

Petula was chatting to a stranger on a train about how she reckoned that people in particular jobs tended to have particular hobbies. By way of illustration, she told him

"If you're a civil servant
then you go in for home winemaking."

[MP]:

He was a civil servant

Therefore,

He did go in for home winemaking

[AC]:

He did go in for home winemaking

Therefore,

He was a civil servant

[DA]:

He wasn't a civil servant

Therefore,

He didn't go in for home winemaking

[MT]:

He didn't go in for home winemaking

Therefore,

He wasn't a civil servant

---

text 28 : Universal 4

Lolita believed that students on certain courses supported particular political parties. Talking to a friend about this, she suggested that he ask the next student who came along which course she was on and which party she supported - she told him

"If the student is doing Economics
then she is a Socialist."

[MP]:

She was doing Economics

Therefore,

She was a Socialist

[AC]:

She was a Socialist

Therefore,

She was doing Economics

[DA]:

She wasn't doing Economics

Therefore,

She wasn't a Socialist

[MT]:

She wasn't a Socialist

Therefore,

She wasn't doing Economics

---

text 29 : Intentional 1

After three weeks of almost continuous rain, Sara was delighted to hear on the radio that the weather was about to clear up. She told her friend Jonah that it was going to be a fine weekend and asked him how he would spend it. He told Sara

"If it is fine over the weekend
then I'll go fishing."

[MP]:

It was fine over the weekend

Therefore,

Jonah did go fishing

[AC]:

Jonah did go fishing

Therefore,

It was fine over the weekend

[DA]:

It wasn't fine over the weekend

Therefore,

Jonah didn't go fishing

[MT]:

Jonah didn't go fishing

Therefore,

It wasn't fine over the weekend

---

text 31 : Intentional 3

Colonel "Buffy" Frobisher, Chairman of the Highsea Officers Club, was asking one of the members, Admiral Cholmondley-Smyth, what he thought about the proposal to allow women to become members of the club. He told the Colonel

"If the committee agrees to let women join
then I'll resign."

[MP]:

The committee did agree to let women join

Therefore,

The Admiral did resign

[AC]:

The Admiral did resign

Therefore,

The committee did agree to let women join

[DA]:

The committee didn't agree to let women join

Therefore,

The Admiral didn't resign

[MT]:

The Admiral didn't resign

Therefore,

The committee didn't agree to let women join

---

text 32 : Intentional 4

Sid, who ran a corner shop and off-licence, had heard that Oddbins had applied for planning permission to open an off-licence in the next street. Sid realised that his profits would fall dramatically if Oddbins did get permission. Sid's wife had been on at him to retire for years. Sid told her

"If Oddbins do get planning permission
then I'll agree to retire."

[MP]:

Oddbins did get planning permission

Therefore,

Sid did agree to retire

[AC]:

Sid did agree to retire

Therefore,

Oddbins did get planning permission

[DA]:

Oddbins didn't get planning permission

Therefore,

Sid didn't agree to retire

[MT]:

Sid didn't agree to retire

Therefore,

Oddbins didn't get planning permission

****
PLEASE READ THESE INSTRUCTIONS CAREFULLY BEFORE OPENING THE BOOKLET
****

Thank you for volunteering to take part in this verbal reasoning experiment. This booklet consists of 32 items to each of which you have to respond by ticking a box. Each item consists of a short passage of text which ends with a statement made by someone who is mentioned in the passage.

Beneath the passage will be shown a sentence describing a state of affairs and a conclusion which may (or may not) follow from the passage and state of affairs described. You are required to indicate whether or not the conclusion does follow by ticking one of two boxes.

It is vitally important that you read the passage carefully before responding. If you tick the wrong box by mistake then you may correct it, but please do NOT return to an item once you have turned to the next page. You are not being timed for this experiment, so please take as long as you like in making your responses.

If you have any questions then please raise your hand BEFORE starting the experiment. If you are quite happy about what you are supposed to be doing then you may turn the page and begin the experiment.

****

I M P O R T A N T !!

Will you please check that you have responded to all 32 items before leaving the experiment - if you have missed even one item out, then your data will be of no use. If you have missed the occasional item, then simply respond to it now.

****

Thank you very much indeed for taking part in this experiment, which was designed to investigate how people interpret statements of the form "if x then y" in a variety of different contexts.

There is no single "correct" answer in tasks such as this and so there is no question of your answers reflecting anything whatever about your intelligence or ability.

If you can leave your seat without disturbing the other subjects then you are welcome to go - otherwise, please wait until the row is clear.

---

text 30 : Intentional 2

Bill was talking to his friend Susie, a nurse at the local hospital, about how she would be affected by the Tories being re-elected at the forthcoming General Election. Susie told him that she was convinced that the Tories getting in again would ruin her promotion prospects, and said

"If the Tories are re-elected
then I'll emigrate to Canada."

[MP]:

The Tories were re-elected

Therefore,

Susie did emigrate to Canada

[AC]:

Susie did emigrate to Canada

Therefore,

The Tories were re-elected

[DA]:

The Tories weren't re-elected

Therefore,

Susie didn't emigrate to Canada

[MT]:

Susie didn't emigrate to Canada

Therefore,

The Tories weren't re-elected

---

text 30 : Intentional 2

Bill was talking to his friend Susie, a nurse at the local hospital, about how she would be affected by the Tories being re-elected at the forthcoming General Election. Susie told him that she was convinced that the Tories getting in again would ruin her promotion prospects, and said

"If the Tories are re-elected
then I'll emigrate to Canada."

[MP]:

The Tories were re-elected

Therefore,

Susie did emigrate to Canada

[AC]:

Susie did emigrate to Canada

Therefore,

The Tories were re-elected

[DA]:

The Tories weren't re-elected

Therefore,

Susie didn't emigrate to Canada

[MT]:

Susie didn't emigrate to Canada

Therefore,

The Tories weren't re-elected