Logic and Implementation in Human Reasoning: the Psychology of Syllogisms

Peter Yule

Ph.D.
University of Edinburgh
1996

For Elena

Declaration

I declare that this thesis has been composed by myself and that the research reported here has been conducted by myself unless otherwise indicated.

Peter Yule

Edinburgh, 1 August 1996

Abstract

This thesis presents a novel account of syllogistic reasoning, based on data from a non-standard reasoning task called the Individuals Task. An abstract logical treatment of the system, based on a modalised Euler Circles system (Stenning & Oberlander 1994, 1995) is presented, and it is shown that this can be implemented in a diverse range of notationally distinct ways. The Individual Identification Algorithm, as this method is called, makes use of a logical distinction between the premisses of the syllogism; one has an existential, assertive role, and is called the *source premiss*, whereas the function of the other is to license inference, and so it is called the *conditional premiss*. This distinction is central to the way the IIA employs modal information to make the use of Euler Circles tractable.

The empirical parts of the thesis are concerned with relating the distinction between source and conditional premisses to the Figural Effect (Johnson-Laird & Steedman 1978). It is argued that the Figural Effect is reducible to a tendency for the terms from the source premiss to occur before the terms from the conditional premiss in Individual Conclusions. Since these are comprised of all three terms in the syllogism, it is possible to test new hypotheses concerning the role of the middle term in inference, and the results are shown to be incompatible with all existing theories of the Figural Effect.

Since the Individuals Task is non-standard, it is necessary to compare performance profiles on this task with those on the Standard Task; one result of this comparison is that a primary cause of error in the Standard Task is selection of an appropriate quantifier for the conclusion, a result which concurs with the conclusions of Ford (1994) and Wetherick & Gilhooly (1990), but contradicts those of Mental Models theory (Johnson-Laird 1983).

Certain anomalies in the prediction of term order by the source/conditional distinction lead to the postulation of a second process for conclusion generation, called Minimal Linking. This logically unsound strategy has effects similar to the illicit conversion of A premisses (Chapman & Chapman 1959, Revlis 1975).

Acknowledgements

First of all, I would like to thank Keith Stenning, my supervisor, who knows more about this thesis than anyone else. His contributions, suggestions, criticisms and support have been invaluable.

I am grateful to those who have, in one way or another, contributed directly to the production of this thesis. Cath Ardin, Rosemary Stevenson and Morten Christansen collected data, and Richard Tobin helped with the logic. Lex Holt and Kevin Humphreys gave assistance with computers.

Many thanks to my parents, Alasdair and Sheila Yule, for their support and financial assistance. Thanks also to all my friends and colleagues at Edinburgh, for being there.

This thesis was supported by ESRC Studentship No. R00429124308. It was typeset in LaTeX, and statistics were done using BMDP and Microsoft Excel.

Contents

Declaration **iii**

Abstract **iv**

Acknowledgements **v**

1 Introduction **1**

1.1 The study of reasoning 1
1.2 Definitions 3
1.3 The History of the Syllogism 4
1.4 The psychology of syllogisms 5
1.4.1 Early work 5
1.4.2 Model-theoretic approaches 7
1.4.3 Hybrid theories 12
1.5 Historical overview of thesis 13

2 Description of Experiments **15**

2.1 The Levy (1984) Experiment 15
2.2 Individuals Task Experiment 16
2.3 Durham Standard Task Experiment 17
2.4 The Johnson-Laird & Bara (1984) Experiments 18

3 The Logic of the Syllogism **19**

3.1 Model Theory 19
3.2 Solving syllogisms by identifying critical individuals 21
3.3 Implementations of the Abstract IIA 24
3.3.1 The Graphical Method 24
3.3.2 Individual identification in `prolog` 27
3.3.3 Implementation in Monadic Predicate Calculus 35
3.3.4 Predicate Calculus again, but simpler! 36
3.3.5 Discussion of the Implementations 38

4 Figural Effects and Mental Models Theory 40
4.1 The Mental Models account of the Figural Effect 41
4.2 The Levy (1984) Experiment 43
4.3 Individuals and Figural Effects 48

5 The Individuals Task Experiment 51
5.1 Introduction 51
5.2 Comparison with the Standard Task 52
5.3 Term order in the Individuals Task 57

6 Individuals Task Models 64
6.1 Model 1: The Individuals Task Figural Effect 64
6.2 Model 2: Three-term order effects 68

7 The A-Effect 73
7.1 A-effect in the Standard task 73
7.1.1 Model 3: The Durham Content Experiment 73
7.1.2 Model 4: The Johnson-Laird & Bara (1984) experiments . . 75
7.2 Errors and the A-Effect 77

8 Conclusions 81

References 91

Appendices 93

A Implementation Details 94
A.1 The Graphical Method 94
A.2 The `prolog` program 96

B Raw data tables 100
B.1 Problems with Valid Conclusions 100
B.2 Problems with Valid U-Conclusions 106
B.3 Problems with No Valid Conclusion 108

Chapter 1

Introduction

1.1 The study of reasoning

The study of the psychology of reasoning began in the Twentieth Century, but in earlier times it was a commonplace assumption that logic was the study of the laws of thought (Henle 1962). Philosophers argued about the acceptability of a given pattern of inference on the grounds that they did or did not habitually use that pattern. Modern conceptions of logic have tended to reject this assumption, often simply on the grounds that human reasoning is subject to error, but in recent times the distinction has become blurred again. The rise of Cognitivism has led researchers to consider the possibility of a "mental logic" or a logic which captures the patterns of human inference. The development of non-classical logics such as Fuzzy Logic and Relevant Logic is at least partly concerned with the question whether these variants more accurately capture human intuitions of valid inference (see Read 1988 for an introduction to non-classical logic).

Perhaps it would be reasonable to claim that in the early stages of the development of a logical system, the justification for that system is its relevance to human inference; when that logic becomes adequately formalised, it seems to "look after itself" (Wittgenstein 1922). Of course, logics can also be used in a normative fashion, as a benchmark against which fallible human reasoning can be judged.

The psychological study of reasoning typically lives in the gaps. There is ongoing controversy about the rationality of human reasoning. Some researchers (e.g Woodworth & Sells 1935) have assumed that human performance on logical problems can be characterised as essentially irrational, whereas many more recent approaches (Guyote & Sternberg 1981, Johnson-Laird 1983) characterise human reasoning as essentially rational. In the former case, there is a need to account for the possibility of rational performance, whereas in the latter case, much of the research effort seeks to explain why people make errors.

Understanding why people reason as they do is a worthwhile activity—if we know what people find difficult, and why, it may be possible to take remedial steps. Obviously there could be applications to education, particularly in formal subjects such as mathematics. For example, a recent research program has been studying the use of heterogeneous reasoning systems in the teaching of logic (Stenning, Cox & Oberlander 1995). The psychology of reasoning holds out the possibility of a rigorous theoretical treatment via the tools of logic in a subject (psychology) where theory is often little more than a set of empirical rules of thumb or, conversely, is expressed at such a high level that little or no empirical test is possible.

Psychology has typically been concerned with only a few paradigms of reasoning. There has been a considerable amount of research on conditional reasoning (Wason 1966, Cosmides 1989) and three-term series problems (Hunter 1957), but perhaps the largest literature concerns categorical syllogisms. Syllogisms have certain methodological advantages over other problem types: they produce several robust psychological effects, and since there is only a small finite set of distinct syllogisms, constituting a complete, integrated system, it is easy to compare the results of different studies without sacrificing generality.

The fact that the syllogism was the first branch of logic to be developed suggests that it is of special importance. It deals with a class of inferences concerned with inheritance, so it is fundamental to taxonomy. Since Cantor this role has been supplanted by set theory, but for some 2000 years it was considered one of the main sets of schemata of valid argument; while for most of that time this view was not under serious attack, its durability is a testament to its perceived utility. It is logically interesting because its semantics are expressible in graphical systems of very limited expressive power (Stenning & Oberlander 1994, 1995). This thesis makes the case that a class of algorithms constructed on this basis can help illuminate some aspects of empirical human performance.

	Johnson-Laird's classification					Scholastic classification			
	1	2	3	4		1	2	3	4
Premiss 1	a-b	b-a	a-b	b-a	Major Premiss	M-P	P-M	M-P	P-M
Premiss 2	b-c	c-b	c-b	b-c	Minor Premiss	S-M	S-M	M-S	M-S
					Conclusion	S-P	S-P	S-P	S-P

Figure 1.1: Two classifications of Figure.

1.2 Definitions

Before reviewing the syllogistic reasoning literature, it is helpful to define a number of common terms, since there is no standard usage of certain of these in the literature.

A syllogism consists of two premisses, each of which relate two terms. One of the terms (the *middle term*) occurs in both premisses, while the other two (the *end terms*) each occur in only one premiss. There are 4 *moods* or premiss types, with the quantifiers (and conventional mnemonics) "all" (A), "some" (I), "none" (E) and "some... not" (O).

The other degree of freedom is in the arrangement of terms, known as *Figure*. Unfortunately, in the literature there are two different definitions of Figure: the Scholastic classification, which specifies the positions of terms in both the premisses and the conclusion, and a variant most closely associated with the work of Johnson-Laird (e.g. Johnson-Laird 1983), which specifies only the position of terms in the premisses. Figure 1.1 shows both possibilities. While at first it may appear that the only consequential difference is in the numbering, in fact the Scholastic version's specification of the order of end terms in the conclusion rules out of consideration half of the possible forms of the syllogism. This has the consequence of obscuring the Figural Effect, which is one of the main topics of this thesis (see Chapter 4), so I follow Johnson-Laird's classification.

Since each premiss can be in one of four moods, and each premiss pair can have one of four Figures, there are $4 \times 4 \times 4 = 64$ different premiss pairs. I normally use the word "syllogism" to refer to premiss pairs of this type.

In the text I sometimes use the expressions "stronger premiss" and "weaker premiss"; this refers to the amount that is implied by the premisses, or their relative informativeness. A strong premiss has many consequences, whereas a weak one has few. Practically in the case of the syllogism, Universals are strong, since they imply more, and Particulars are weak, since they imply less.

1.3 The History of the Syllogism

The syllogism was the first branch of logic to be developed, by Aristotle (384-322 BCE), who claimed he invented it all on his own, without precursors (Aristotle 336 BCE/1983). Apparently there is no historical evidence to contradict this. Aristotle also provided a complete proof theory, which proceeded by reducing the various valid moods of the syllogism to the valid moods in the first Figure, which he assumed to be self-evident, although he appears to have been aware that this was only one of many possible ways of formalising the system. The magnitude of Aristotle's intellectual leap is underscored by the fact that not only did he invent logic, but in so doing made the first documented use of variables in the history of science. The work was refined but not significantly advanced upon by his successors in the Peripatetic School, most notably Theophrastus (*circa* 372-288 BCE). Other Ancient Greek schools did produce significant developments in other branches of logic however, particularly the Stoics, who developed systems of sentential reasoning and whose most famous and innovative member was Chrysippus (280-205 BCE) (Mates 1972).

Despite the popular image of syllogising Mediaeval monks, very little of substance was contributed to logic as such, as opposed to the philosophy of language, during the Middle Ages, and students of logic concentrated on preservation and transmission of the doctrines of antiquity, constructing *summulae* (little summaries) of ancient logic. According to Johnson-Laird (1983), the idea of a 'sieve' to catch only valid syllogisms by application of a few rules is attributable to the Mediaeval Scholastics. Five rules are sufficient (reproduced from Johnson-Laird 1983):

1. If both premisses are affirmative, then the conclusion is affirmative.

2. If one premiss is negative, then the conclusion is negative.

3. If both premisses are negative, then there is no valid conclusion.

4. The middle term must be distributed in at least one premiss.

5. No term may be distributed in the conclusion if it is not distributed in the premiss in which it occurs.

A term is said to be distributed in a statement if the entire class of entities to which it refers must be considered in order to determine if the statement is true.

The *summulae* have contributed to some extent to classification schemes; for example, the names by which the Aristotelian moods of the syllogism have come to be known (Barbara, Celarent etc) are apparently due to William of Shyreswood (d. 1249), who used them to compose a mnemonic poem. These names cunningly encode information about mood, Figure and even the method of conversion by means of spelling, so that barbara denotes a syllogism with two A premisses and an A conclusion in the first Figure. The vowels indicate the moods of the premisses and conclusion in order, and the consonants give instructions for reducing the given mood to the first four (Mates 1972).

The method of Euler Circles, which informs much of the work in this thesis, is often attributed to Leonhard Euler, the eighteenth century mathematician, who used it to teach logic to a German princess (Euler 1772). However, the technique was in fact originated by Leibniz (1666), the brilliant polymath who, sadly, has tended not to get due credit for his innovations.

1.4 The psychology of syllogisms

1.4.1 Early work

Before the 1970s, much research concentrated on the explanation of error in syllogistic reasoning. There is a substantial literature on the influence of real world knowledge, beliefs and attitudes on subjects' acceptance of either valid or invalid conclusions (Janis & Frick 1943; Evans *et al* 1983; Oakhill *et al* 1989; Newstead *et al* 1992).

Another approach was to construct unsound inference schemata, the best known of which is the Atmosphere Effect (Woodworth & Sells 1935, Sells 1936).

The Atmosphere Effect Hypothesis postulates that subjects use a few simple rules to assess the validity of conclusions, so that if the premisses are, say, universal and positive, the conclusion will also be universal and positive. Thus the premisses create an "atmosphere", and conclusions are assumed to be valid if they match this atmosphere. Since this scheme corresponds to part of the "Sieve" decision procedure mentioned above, many of the conclusions predicted by the Atmosphere Hypothesis are in fact valid ones. Early versions of the theory were unsatisfactory, since the hypothesis was vaguely formulated, and predictions were tested on a small subset of invalid conclusions, but subsequent researchers (eg Revlis 1975) refined the theory to the stage where it constituted a precisely-specified cognitive procedure, and tested it on a much wider set of data, including valid conclusions. Unfortunately, although the Atmosphere Effect did account for subjects' conclusions quite well, its fit with the data was found to be substantially better for valid conclusions than invalid ones, suggesting that its apparently high degree of confirmation was spurious (Revlis 1975).

Whatever the empirical evidence, the Atmosphere Effect Hypothesis cannot entirely account for human reasoning, if only because it cannot explain valid reasoning in cases where the valid conclusion does not match the atmosphere of the premisses. Moreover, it never predicts a "No Valid Conclusion" response, while subjects often correctly find these. It would need to be supplemented by some sort of valid reasoning process to accommodate subjects' potential rationality, which on the basis of the above findings, would undermine much of the evidence for it.

Other early researchers (e.g. Henle 1962, Ceraso & Provitera 1971) were more alert to the need to give an account of rational reasoning. Although they still omitted to specify the means by which inferences might be drawn, they argued that apparently invalid conclusions could be explained as the valid consequences of non-standard interpretations of premisses. Henle (1962) argued that inferences frequently appear as enthymemes, or incompletely stated syllogisms, and that once the full set of premisses actually assumed by the subject was known, the inferences drawn would be seen to be valid ones. Ceraso & Provitera (1971) showed that by giving more specific interpretations of premisses, subjects' performance improved to near-perfection, but acknowledged that their task was not equivalent to the syllogistic reasoning task, precisely because they had removed the inherent ambiguity of syllogistic premisses, which is a contributory factor to the difficulty of syllogistic inference.

Similarly, Chapman & Chapman (1959) argued that many errors can be attributed to the illicit conversion of A and O premisses, so that given "All As are Bs", subjects may assume "All Bs are As". Such conversion is perfectly legitimate for I and E premisses, since the truth conditions for the unconverted and converted forms are the same, but A and O premisses are not validly convertible, and conversion of these can sometimes lead to errors. Early tests of this hypothesis were equivocal about its usefulness as an explanation of errors (Begg & Denny 1969, Revlis 1975) since many of its predictions were the same as those of the Atmosphere Effect hypothesis, but if the latter is discounted, the Conversion Hypothesis appears more viable in retrospect.

1.4.2 Model-theoretic approaches

The increasing influence of cognitive psychology led to a substantial change of emphasis in the mid-1970s, resulting in greater awareness of the need to account for valid inference. At the same time, the wider availability of computers made it easier both to develop and to test much more intricate models of reasoning processes. Most theories from that period up to the present draw on logical model theory, making the assumption that subjects solve syllogisms by consideration of the space of possible logical models of the premisses. This allows for the potential rationality of subjects' reasoning processes, by basing these on sound logical principles. The two main threads of research in this tradition are based on the so-called Method of Euler Circles and the theory of Mental Models.

The use of the so-called method of Euler Circles in the syllogistic reasoning literature is most commonly associated with the work of Erickson (1974, 1978) and Guyote & Sternberg (1981). The idea is that each premiss can be represented by a set of diagrams which represent alternative models of the premisses, or different "possible worlds" in which the premisses are made true in different ways compatible with their meaning.

For example, there are two different situations compatible with the premiss "All of the Artists are Beekeepers". The Euler Circle representations are given in Figure 1.2. The first shows the situation where every Artist is a Beekeeper, and

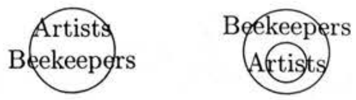


Figure 1.2: Representing the premiss "All of the Artists are Beekeepers" with primitively interpreted Euler Circles

every Beekeeper is an Artist, while the second shows the situation where every Artist is a Beekeeper, but some of the Beekeepers are not Artists. A third possibility, commonly accepted in modern logic, is that there are no Artists, and the generalisation holds because there are no counterexamples in the form of Artists who are not Beekeepers; however in the conventional interpretation of the syllogism, all terms are assumed to be instantiated, so we do not need to consider the possibility.

It is clear then that, in this notation, each diagram represents an interpretation of the premiss or "possible world", and each bounded region within a diagram represents a type of individual which exists in that possible world. The truth conditions or each premiss type can be expressed by a set of between one and four diagrams standing in 1-1 correspondence with distinct logical interpretations. Geometrical properties of the diagrams are unimportant, only topological relationships (patterns of overlap, inclusion and non-overlap) are significant in interpretation.

Ideally, reasoning is assumed to proceed by creating compound diagrams for each possible combination of two diagrams, one for each premiss. Conclusions are checked against all the resulting diagrams, and only those potential conclusions which hold in all of them are accepted as valid. Unfortunately, the resulting set of possibilities can be rather large, to the extent that it is implausible that human reasoners could keep track of them all individually in working memory. Also, the number of diagrams required to solve a problem does not correlate with its observed difficulty—some problems with many possible combinations are solved relatively easily and quickly, whereas other problems, with fewer Eulerian combinations, are much more difficult (Johnson-Laird & Bara 1984).

Ironically, Erickson's solution to the problem of the combinatorial explosion was to suggest that perhaps subjects constructed only one combination diagram, thus rescuing the irrationality of the subject. However, subsequent researchers

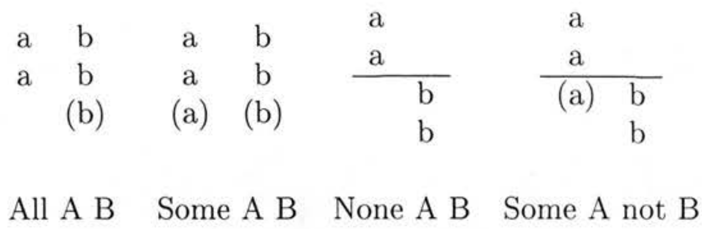


Figure 1.3: Mental Models representations for each premiss type.

have adopted a different strategy, indroducing modal distinctions into representations, to reduce the required number of premiss representations to one per premiss without compromising the soundness or completeness of the system. Johnson-Laird's Mental Models theory (Johnson-Laird & Steedman 1978, Johnson-Laird (1983), Johnson-Laird & Bara 1984, Byrne & Johnson-Laird 1991) marks the distinction between necessary and possible individuals explicitly in the representation. As a consequence, the number of combined representations is reduced to between one and three. The theory has changed significantly over the years, but probably the best-known version is that presented in Johnson-Laird (1983) and Johnson-Laird & Bara (1984).

This theory assumes that premisses are represented by arrays of tokens corresponding to the terms of the syllogism: Figure 1.3 shows the Mental Models representations for the four premiss types. These tokens may represent optional or necessary instances of their referents; optional tokens are bracketed, whereas necessary ones are unbracketed. Rows of tokens represent individuals, and since numbers of any type of individual are insignificant to interpretation, they may be replicated at will. The horizontal lines in the negative cases represent barriers 'fencing off' sets of rows from each other—effectively these restrict the locations where new replicated rows may be added, so that it is impossible to replicate a row across a barrier.

Syllogistic inference is held to consist in integrating the two models and reading off a conclusion. Some pairs of models admit being integrated in only one way, whereas others may have two or three different integrations. This property, the number of distinct integrated models, is used by Johnson-Laird and his coworkers to explain the differential difficulty of syllogisms. Empirically, the cru-

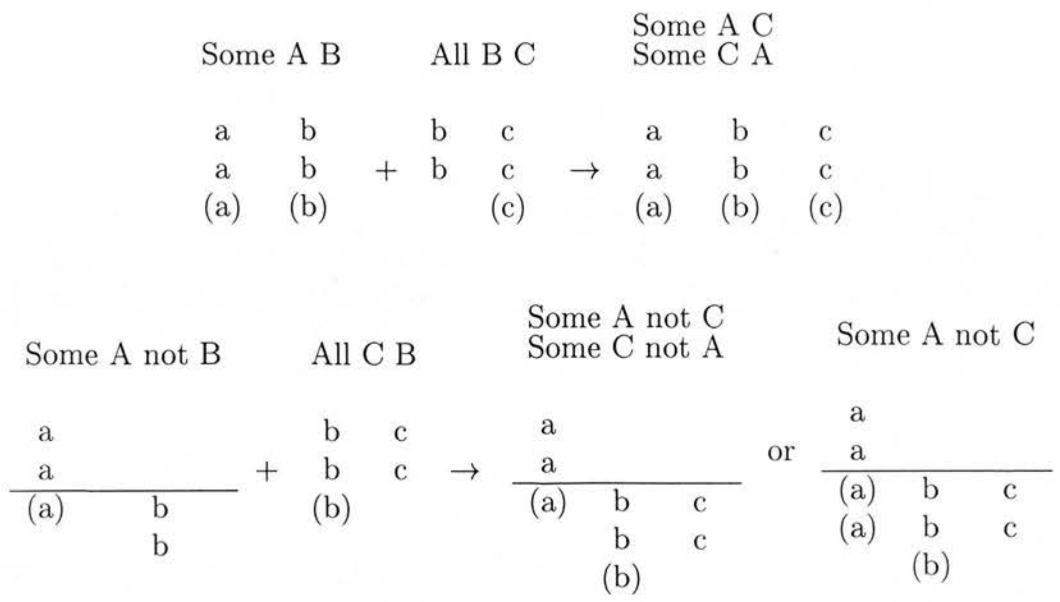


Figure 1.4: Mental Models Theory: examples of a single-model problem (above) and a multiple-model problem (below) (adapted from Johnson-Laird 1983)

cial difference appears to be between 'single-model' problems and 'multiple-model' problems; the early three-way distinction has been dropped in a recent version of the theory (Byrne & Johnson-Laird 1991), since the number of models assumed to be needed for a given premiss pair is implementation-dependent (Johnson-Laird & Bara 1984). Figure 1.4 shows examples of single-model and multiple-model problems. In the first example, it is impossible to modify the initial model in such a way as to refute either of the putative conclusions, so it is a single-model problem; in the second, two models can be constructed, both supporting the conclusion *Some of the As are not Cs*, but one of which supports another conclusion, *Some of the Cs are not As*.

The Byrne & Johnson-Laird (1991) version of the theory uses a different style of representation which employs square brackets to mark terms which have been exhaustively represented ('distributed terms' in the Scholastic terminology), and omits to specify optional elements in initial representations, allowing them to be added later, but the authors do not seem to believe that this difference has many empirical consequences different from the earlier formulations. There is a large literature on the Mental Models theory, which I cannot summarise here; however

Chapter 4 covers a number of points specifically related to the Mental Models account of the Figural Effect.

Johnson-Laird (1983) argues that the discrete nature of Mental Models representations renders them more realistic than Euler Circles, since the latter use a continuous space to represent a set of objects, but as we have seen, such geometrical properties of Euler Circles are irrelevant to their interpretation. Since the syllogism is a subset of Monadic Predicate Calculus without identity, numbers of a type of individual can never be significant (see Section 3.1 in Chapter 3 for more discussion of this point). Stenning & Oberlander (1994, 1995) argue that in fact Mental Models theory is a notational variant of a modalised Euler Circles system. Essentially the same system is presented in this thesis, in Section 3.3.1, and it can be demonstrated that no more than *one* combined diagram need be constructed for any premiss pair, without compromising the completeness of the system, and hence subjects' potential rationality. Recently, Ford (1994) has also developed a modalised Euler Circles system, though the details of premiss combination have not been made explicit.

In these model-theoretic approaches, there are two main ways to accommodate reasoning errors: premisses may be misinterpreted, or the full set of combinations of premiss representations may not be tested. The latter option is favoured by Johnson-Laird, whereas the former is more directly compatible with theories of Conversion and Gricean errors, and the rationalist approach mentioned earlier.

Newstead (1989, 1990) has reported experiments on premiss interpretation using an Euler Circle task and an immediate inference task. In the Euler Circles task, subjects selected primitively (non-modally) interpreted Euler Circle diagrams corresponding to interpretations of a given premiss, and in the immediate inference task, subjects indicated whether or not selected statements were entailed by a given premiss. The results showed a clear tendency to convert A premisses in both tasks, but not O ones, and there was a small tendency to commit Gricean "errors" (Grice 1975). Subjects assessed as A-converters on the immediate inference task made more errors predicted by conversion theory on a syllogistic reasoning task, but the results of the Euler Circles task did not correlate with the results of the immediate inference task, and so the validity of the measures was put in question. Stenning & Cox (1995) argue that Newstead's immediate inference task was logically incoherent, and modified it to permit "Don't know" responses, in order to accommodate the three possible logical relations between statements: entailment, contradiction and logical independence. Their results showed, *contra* Newstead, that subjects who made A-conversion errors in the Euler Circles task did also tend to make A-conversion errors as assessed by the amended immediate inference task. This rectifies the validity problem, making possible a genuine test of the relation between premiss misinterpretation and reasoning errors, but at present the necessary experiments have not yet been completed.

1.4.3 Hybrid theories

While Mental Models theory has enjoyed wide distribution and acceptance in the Cognitive Science community, owing to its descriptive success and relative plausibility, recently several alternative accounts have been proposed. These share a dissatisfaction with the tacit assumption that a single model can usefully be applied to a diverse population of subjects.

As I have mentioned already, Ford (1994) asked subjects to explain their conclusions in a syllogistic reasoning task, and found that almost all subjects could be classified as either verbal or graphical reasoners. Verbal reasoners typically explained their conclusions with reference to the substitution of terms from a universal premiss into the other premiss, whereas graphical reasoners drew diagrams using geometrical shapes, such as circles or squares, depicting set relations. It was found that the two groups showed different error patterns and response biases consistent with reasonable assumptions about how such methods could work.

Wetherick & Gilhooly (1990) argue that a substantial proportion of subjects can be shown to follow a non-logical strategy which they call "Matching". This is related to the Atmosphere Effect, except simpler. Matching subjects simply produce conclusions with the same quantifier as the weaker premiss; this often produces a valid conclusion, but leads to predictable error patterns. Subjects assessed as Matchers on problems without valid conclusions were shown to have good performance on problems whose valid conclusions were derivable by Matching, and poor performance on problems whose valid conclusions were not (i.e. problems whose valid conclusion was not in the same mood as either premiss), whereas the remainder of subjects showed no such effects, and it could be argued that they were following a "logical" strategy.

The account of syllogistic reasoning which is described in this thesis shares some features with both of the above accounts; it is consistent with the idea that "logical" reasoning can be implemented in a variety of phenomenologically distinct ways, but concedes that some error patterns are best explained by a logically unsound strategy. Much of the empirical content is concerned with the explanation of the Figural Effect (Johnson-Laird & Steedman 1978), which is discussed in depth in Chapter 4.

1.5 Historical overview of thesis

My own interest in the psychology of syllogisms was stimulated by reading "Mental Models" (Johnson-Laird 1983) as an undergraduate. But I was dissatisfied by the explanation of the Figural Effect, and felt that the true explanation was that the Figural Effect was a consequence of the logic of the problems rather than of the putative memory systems that were supposed to underlie reasoning.

When I encountered Keith Stennings's work on modalised Euler Circles and the Individuals Task, I was able to work on the subject for my MSc thesis (Yule 1991), which concentrated on the explanation of error in the Individuals Task. At that time, the modalised Euler Circles system was based on a "shading cancellation" algorithm, and as such was neither sound nor complete. The technique of using only the maximal registration diagram (see Chapter 3) was my own innovation, stimulated by the need to construct a computerised scoring system for individual conclusions. The `prolog` implementation in Chapter 3 was developed from the core of this system, and this optimisation of the graphical method was adopted by Stenning & Oberlander (1995). Having gained this crucial insight, the abstract Individual Identification Algorithm and Modal Predicate Calculus implementation were straightforwardly constructed on the same principle.

As a further testament to the usefulness of computer modelling, after finishing my MSc thesis I was struck by the similarity between the ordered term structures generated by the program and the human data from the Individuals Task experiment. This is the origin of the hypothesis that the distinction between source and conditional premisses might account for term order, and consequently the

Figural Effect. Early results, reported in Chapter 5 suggested that the association between the source/conditional distinction and end-term order was worth further investigation.

It was necessary to learn the technique of loglinear modelling to attempt to evaluate the claim that the source/conditional distinction could account for the Figural Effect. Early models, using only valid conclusions with unique source premisses, failed to abolish the main effect of Figure, but with the inclusion of an early version of what came to be known as the A-effect, the Figural Effect was successfully abolished for valid conclusions. Further refinement in the definition of variables, involving defining the source and A-effect variables with three levels (first, second and both/neither), made it possible to construct a model which included all valid and invalid conclusions, and demonstrated that the total (valid and invalid) figural effect could be abolished. This is the end-term model described in Chapter 6.

The recognition of the importance of the A-Effect in the Individuals Task raised the question whether it could also be identified in Standard Task datasets. The models in Chapter 7 are the results of this investigation and show that the phenomenon does appear to extend to the Standard Task. However, the question of the degree of psychological similarity between the Individuals and Standard Tasks was still open to analysis in terms of problem difficulty, prompting the comparative analysis in Chapter 5.

Having gained a reasonable amount of experience in modelling, I was able to experiment with models of three-term order in the Individuals Task. The three-term model in Chapter 6 is the state of the art, but I would like to have been able to experiment for longer, since the space of possibilities in the case of three terms is larger than in the case when only the end terms need be considered.

After establishing the existence and robustness of the A-effect in purely empirical terms, it remained for me to try to find an explanation for it. Work on the comparative problem difficulty analysis led me to reconsider my MSc thesis results, which suggested that the commonest types of error individuals appeared to have their origin in a strategy based on the unification of critical individuals. I call this strategy Minimal Linking, and Chapter 7 argues that this can explain the A-effect.

Chapter 2

Description of Experiments

A variety of experiments are referred to in the course of this thesis. For convenience, these experiments are described here.

2.1 The Levy (1984) Experiment

This experiment was run by Joe Levy in Cambridge in 1984, under the supervision of P.N. Johnson-Laird, for an undergraduate thesis. Since it remains unpublished, it is described in detail here.

Design Subjects were divided into 4 groups, each group receiving problems from a different Figure. They received all 16 problems in their assigned Figure, once in each of two blocks of trials corresponding to single- and double-premiss presentation modes. The order of presentation of blocks was counterbalanced between subjects, to avoid learning effects and permit a repeated measures comparison of presentation types, and problems were presented in a pseudorandom order within blocks. Subjects could respond with any of the 8 quantified conclusion types (4 quantifiers (all, some, none, some... not) × 2 end-term orders (ac or ca)], or 'No Valid Conclusion'.

Each subject also produced two memory-test scores, one for each of the 2+2 and 3+3 tests, both measured as the number of strings transformed in one minute.

Subjects 40 unpaid volunteers, mostly undergraduate students.

Materials Syllogisms were presented on a BBC model B microcomputer, which presented problems and collected responses in both presentation modes. In double-premiss mode, both premisses appeared onscreen together, while in single-premiss mode, premisses appeared singly, the second replacing the first onscreen in response to subjects' keypresses.

Syllogisms were constructed using terms from each of three word cohorts, denoting hobbies, interests and occupations. Word triplets were shuffled so that cohorts were not associated with term types.

Procedure Subjects were given instructions and four practice trials before each block of problems. They were instructed to imagine that the premisses described a group of people in a room, and to draw conclusions relating the end terms, which had to be true of the people in the room given that the premisses were. The practice trials consisted of two three-term series, a syllogism and a syllogism with a numerical quantifier, in order to give the flavour of the problems without unduly affecting subjects' syllogistic performance. The syllogisms followed immediately, then the procedure was repeated for the second block of problems.

After the syllogisms the subjects were given two short memory tests and a questionnaire. The memory tests were similar to those reported by Hamilton *et al* (1977) and were intended to give a measure of the processing capacity of working memory. The tests, known as the 2+2 and 3+3 tests, involved transforming a string of letters into another by counting upwards a number of alphabetic places, so that the 2+2 test entailed transforming two letters by two places (e.g. CP $\rightarrow$ ER), and the 3+3 test entailed transforming three letters by three places (e.g. HQB $\rightarrow$ KTE). In each test, subjects were given one minute to transform as many strings as possible. The questionnaire was intended to elicit introspective reports of subjects' solution strategies.

Overall session lengths ranged from 40 to 65 minutes.

2.2 Individuals Task Experiment

The Individuals Task experiment was designed and run by Cath Ardin in 1991, with assistance from myself and Morten Christiansen. Some results have been reported in Ardin (1991) and Yule (1991); however the analyses reported here are original.

Design Subjects produced descriptions of necessary individuals for each of the 64 syllogisms presented in random order. Subjects could respond with any of 48 different individual conclusions [8 distinct individual types (+++, ++-, +-+, +–, -++, -+-, –+, —) × 6 distinct term orders (ABC, BAC, ACB, CBA, BCA, CAB)], or 'No Valid Conclusion'. Each subject received materials with one of two different random assignments of vocabulary to problems.

Subjects 22 Edinburgh University students took part. None had any prior training in syllogistic logic.

Materials Each subject received a set of 64 slips of paper, on each of which was printed a different pair of premisses. The vocabulary used was selected from sets of nouns denoting nationalities, professions and interests, for example *None of the musicians are chessplayers. All of the musicians are Italians.* Each vocabulary item appeared in two syllogisms, and two different random assignments of vocabulary to syllogisms were used.

Procedure Subjects were instructed to imagine that the premisses on each slip of paper described a group of people at a party. They were instructed to assume that some people corresponding to each of the three terms existed.

Subjects were instructed to decide whether any kind of person who could be described with certainty, in terms of either positive or negative values of all three features, had to be present in the room, and to describe the individual on the slip of paper, or if there was no such individual, to write "No valid conclusion".

Subjects worked individually in quiet surroundings, and were given as much time as they needed to finish all the problems.

2.3 Durham Standard Task Experiment

This experiment was designed and prepared by myself, and run with the assistance of Rosemary Stevenson in Durham during 1992.

Design Separate groups of subjects received Contentful and Contentless materials sets, and each materials set presented all 64 syllogisms in random order, each using a different random assignment of vocabulary to problems. Subjects could respond with any of the 8 quantified conclusion types (4 quantifiers (all, some, none, some... not) × 2 end-term orders (ac or ca)], or 'No Valid Conclusion'.

Subjects 24 Durham University students took part. None had any prior training in syllogistic logic.

Materials Each subject received a set of 64 slips of paper, on each of which was printed a different pair of premisses.

In the Contentful condition, the vocabulary used for the terms was selected from sets of nouns denoting nationalities, professions and interests, for example *None of the musicians are chessplayers. All of the musicians are Italians.* In the Contentless condition, the vocabulary used for the terms was selected from three cohorts of capital letters, for example *None of the Ms are Cs. All of the Ms are Is.* In both conditions, each vocabulary item appeared in two syllogisms, and assignments of vocabulary to syllogisms were randomised.

The vocabulary used in the Contentful condition was the same as that used in the Individuals Experiment.

Procedure Subjects were instructed to imagine that the premisses on each slip of paper described a group of people at a party. They were instructed to assume that some people corresponding to each of the three terms existed.

Subjects were instructed to decide whether there was any statement relating the end terms of the premisses which must hold of the people in the room, and to write the statement on the slip of paper, or if there was no such statement, to write "No valid conclusion".

Subjects worked individually in one large group, and were given as much time as they needed to finish all the problems.

2.4 The Johnson-Laird & Bara (1984) Experiments

The two datasets reanalysed here are from experiments 1 and 3 in Johnson-Laird & Bara (1984). The details are published, and so not repeated here, but in summary, in experiment 1 subjects performed the Standard Task and were required to respond within 10s, whereas in experiment 3 subjects again performed the Standard Task but with no time restriction.

Chapter 3

The Logic of the Syllogism

3.1 Model Theory

Semantics The syllogism can be viewed as a fragment of the monadic (one-placed) predicate calculus. Because the monadic calculus has no identity relation (it lacks relational predicates entirely), its semantics concerns *types* of individuals defined by combinations of properties. Since there is no way to express identity or distinctness, there is no way to discriminate between individuals with the same properties; therefore it is never of semantic significance how many of a type of individual exist.

We define *Maximal* types as combinations of all three predicates of a syllogism. There are eight such types: ABC, AB¬C, A¬BC,A¬B¬C, ¬ABC, ¬AB¬C, ¬A¬BC, and ¬A¬B¬C. *Interpretations* of the syllogism consist of sets of these maximal types of individual. Conventionally, the syllogism is interpreted under the assumption that none of the three sets A, B and C are empty. Subjects naturally adopt this assumption[1]. This 'no-empty-sets' axiom reduces the number of possible models somewhat (Stenning & Oberlander 1994, 1995).

Inferential Structure Of the 64 syllogisms 27 have conventionally valid conclusions which can be formulated by applying one of the four quantifiers to the

[1]As witnessed for example by their readiness in the task situation to conclude from *All A are B* that *Some A are B* - see Stenning, Cox & Oberlander (1995), Newstead (1989, 1990) for discussion of 'immediate inference' tasks

two end terms. The remaining 37 syllogisms do not allow any valid conclusions within the vocabulary of conclusions thus specified, though as we will see, there is another group which allow valid conclusions if a further quantifier is available.

Examination of the valid syllogisms reveals that all syllogisms which have valid conclusions entail the existence of at least one *maximal* type of individual. So, for example, *Some B are* $\neg A$. *All B are C.* entails that there are individuals of the maximal type ¬ABC (and therefore that *Some C are not A*). There is obviously a direct relation between the entailment of the existence of maximal types and the entailment of existential conclusions.

Considering the converse of this generalisation is revealing. It turns out that for all but one small group of syllogisms, if they entail the existence of a maximal type, they also have a valid conventional conclusion. The exceptions are a group of syllogisms with two negative premisses (according to the "sieve" decision procedure given in the Introduction, two negative premisses have no valid conclusion). For example, *No A are B. No B are C* establishes that there are $\neg AB \neg C$ individuals. The quantificational apparatus of the conventional syllogism will not allow expression of the conclusion *Some* $\neg A$ *are* $\neg C$ (these can be called 'U' conclusions following the common notation of the four conventional quantifiers). Stenning & Oberlander use the expression "case-identifiability" to describe the property that any valid conclusion depends crucially on the existence of only one individual, as in the conventional syllogism.

These model-theoretic properties of the syllogism are highly unusual. Even other small finite logical fragments (such as disjunctive syllogisms, for example *A or B. Not A. Therefore, B*) are generally not case-identifiable. Case-identifiability plays a crucial role in determining what representations and algorithms can be employed on this fragment of logic; in particular, it makes the syllogism susceptible to a graphical strategy. There is no need to represent *partially* specified individuals in the diagrammatic representation of the conjunction of two premisses (Stenning & Oberlander 1995). Having an *efficient* graphical strategy which need construct only one diagram per syllogism is further dependent on the even stronger property that the *maximal* model for any syllogism is unique, and consists of all the individuals consistent with the premisses (i.e. there are no contingencies between the presence/absence of types in a model). For more on the graphical strategy, see Section 3.3.1 below.

3.2 Solving syllogisms by identifying critical individuals

The aim of this section is to describe an abstract algorithm for syllogistic reasoning, the IIA. Initially, this algorithm is specified in model-theoretic terms. The algorithm turns on the identification of *critical individuals*, individuals whose existence is necessitated by the premisses, which form a basis for the formulation of quantified conclusions. The Abstract Individual Identification Algorithm can be implemented in a variety of superficially quite different ways, which are described in the remaining sections of this chapter. The aim is to reveal commonalities between what have been taken to be distinct processes (graphical methods, rule based methods and computer program implementations) defining opposing psychological theories of reasoning.

Premiss representation The meaning of premisses is represented in terms of sets of individuals which satisfy them. Model-theoretically, each premiss type is satisfiable by any of several different sets of individual types. However, the present method needs to make use of only two of these sets for each premiss, these being the set of possible individuals whose existence is consistent with the premiss, known as the *maximal model (Max)*, and the set of individuals whose existence is entailed by the premiss, known as the *minimal model (Min)*. These are respectively the largest and smallest sets of types compatible with the premiss. The representation can thus be characterised as the ordered pair $< Max, Min >$. Since $Min \subset Max$, it is feasible to represent each premiss type using a single integrated representation with individuals marked as necessary or possible (cf. Mental Models).

Types of individuals are notated by feature structures, where features are constructed from the terms appearing in the premisses of the syllogism, prefixed by "+" or "−", which indicate whether the individual concerned is or is not a member of the set denoted by the term. For example "+P+Q", denotes a type of individual which is P and is Q, i.e. $(\exists x)(Px\&Qx)$. Similarly, "+Q−R" denotes a type of individual which is Q and is not R, i.e. $(\exists x)(Qx\&\neg Rx)$. Feature structures are unordered owing to the commutativity and associativity of conjunction.

Case Identification Case identification depends on the integration of information from both premisses. The aim is to identify the minimal model of the premiss pair, the set of feature structures which are specified with respect to all three terms in the problem, which denote those individuals whose existence is entailed by the premisses taken together. This can be achieved using unification, an operation which takes two feature structures and returns a third which is their set-union, provided the two do not contain different values for any of their members (see Shieber 1986 for an introduction to unification). So, for example, the unification of +A+B with +B−C is +A+B−C, but the unification of +A+B with −B−C is undefined, since the two do not have the same values for the B feature.

Given two premisses' characteristic representations, the set of all unifications of each member of one maximal model with each member of the other maximal model (the Cartesian product of the two sets) is the maximal model of the premiss pair, containing all types compatible with the premisses. Finding the minimal model for the premiss pair turns on unifying members of the minimal model of one premiss (*critical individuals*) with members of the maximal model of the other premiss. We attempt to unify each critical individual i from one of the premisses (P) with all the members j of the maximal model of the other premiss (Q). If only one such unification is possible for some i, then the feature structure which is formed must denote a member of the minimal model of the premiss pair, *because it is the only possible completion of a type which itself must exist.* For example, supposing i is +A+B and $j \in$ {+B+C,−B+C,−B−C} (the maximal model of "All the Bs are Cs"), the only possible unification is +A+B+C. We know that +A+B must exist, and since +A+B−C is impossible, it follows that +A+B+C must exist. We call the premiss P, which supplies the critical individual i, the *Source premiss*, and since the other premiss Q is always a universal (see below), we will refer to it as the *conditional premiss*—this distinction has consequences in the empirical investigation in later chapters.

If every critical individual from both premisses has been tested but no necessary individuals have been found, then the syllogism has no valid conclusion relating the end terms.

Informally, the abstract IIA can be justified as follows: the critical individual denotes an individual, specified with respect to two properties, whose existence is entailed by the truth of the source premiss. We know *a priori* that this individual either does or does not have some third property mentioned in the conditional premiss, and the maximal model of the conditional premiss constrains the set of such possible completions, by virtue of the shared middle term. If the set of possible completions is constrained to one, then that individual must exist, since the other *a priori* possibility has been ruled out; in other words, the critical individual, which we know to exist, must have the specified value (whatever that is) for the third property, because if it didn't, then its existence would be incompatible with the truth of the conditional premiss.

As was mentioned above, the conditional premiss is always a universal; this is because the conditional premiss must license an inference about the critical individual. Put another way, the conditional must be the stronger premiss, and the source the weaker. If both premisses are particular, of course there is no valid conclusion.

Drawing Quantified Conclusions Feature structures denoting necessary individuals form the basis for quantified conclusions. Since these relate only the end terms, the middle term feature can be deleted from the feature structure. Of the remaining features, if one is positive, a conclusion with that term as subject is warranted, since the grammar of conclusions does not permit negation in subject position[2]. The sign of the remaining feature determines whether the conclusion is positive or negative. This is all that is required for a particular conclusion, so that given e.g. +A+B−C, we delete +B giving +A−C, the only positive term is +A so A is the subject, and since the remaining term −C is negative, the conclusion is "Some of the As are not Cs".

For universal conclusions, a further test is required. This can be characterised as follows: take the subject feature, and unify it with all members of the maximal model for the premiss pair. If only one such unification results, a universal conclusion is warranted, so in the example just given, provided the test was positive, we

[2]U-conclusions, which have only negative end terms, are barred by this condition - see Stenning & Oberlander (1994, 1995)

could conclude "None of the As are Cs". This test is equivalent in graphical terms to checking that the circle labelled with the subject term is not bisected in the registration diagram. Note, however, that this procedure can be simplified: it is sufficient merely to check that the subject term has only one possible unification with the maximal model of the source premiss, or in graphical terms, that the circle labelled by the subject term is not bisected in the source premiss' characteristic diagram. Since we know that the critical individual has only one unification with the maximal model of the conditional premiss, it follows that there can be only one unification with the maximal model of the premiss pair.

3.3 Implementations of the Abstract IIA

This section describes a variety of implementations of the Abstract Individual Identification Algorithm (IIA). Despite superficial differences (a graphical method, a computer program and a rule-based method), each implementation uses clear, straightforward analogues of minimal and maximal models and the unification operation, to draw individual conclusions. Consequently, each implementation can provide a reconstruction of the distinction between source and conditional premisses.

3.3.1 The Graphical Method

The graphical algorithm was presented in Stenning & Oberlander (1995) where it is shown that Johnson-Laird's (e.g. 1983) 'mental models' method is an alternative implementation of the same abstract algorithm. It is restated here for the reader's convenience.

Premiss representation Individual types are represented as regions in a diagram delineated by circles. The maximal model of a premiss is represented as a diagram containing two circles, one for each of the terms. Regions in the diagram correspond to individual types which are compatible with the premiss (the maximal model), while the minimal model is represented by marking with an "×"

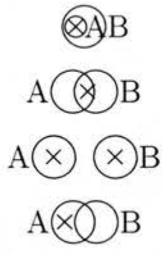
All As are Bs	
Some As are Bs	
No As are Bs	
Some As are not Bs	

Table 3.1: Characteristic diagrams for each premiss type. Regions representing necessary individuals are marked with a ×.

the regions of this diagram which correspond to necessary individuals[3]. Table 3.1 shows the characteristic diagram for each premiss type. The circles are assumed to be of variable size and position. Only their topological relationships (patterns of containment, overlap and exclusion) are significant. This premiss representation is a direct translation of the set-theoretic relations into graphical terms.

Case Identification A new diagram is constructed by 'registering' the two characteristic diagrams by superimposing the B circles. Depending on the problem, the topological constraints of the graphics may be sufficient to determine the final diagram, or there may be more than one way of doing this consistent with the premisses, as in Figure 3.1, making it a "multiple models" problem (Johnson-Laird & Bara 1984). However in the present method only one way is chosen, namely the one which creates the maximum number of sub-regions in the diagram. This is achieved by overlapping the circles representing the end terms if this can be done while remaining consistent with the premisses. The result is the *registration diagram*, which represents the maximal model of the premiss pair.

Now each marked region deriving from the original characteristic diagrams will either have been bisected by the third circle or not. The marked region represents i, a critical individual, so it can be called a *critical region*. If the critical region is bisected, then the mark is removed. The attempt to overlap the end-term circles ensures that all possible unifications of i with each j have been attempted, so if a critical region is not bisected, and consequently remains marked in the registration

[3]This refinement of the "Shading" version of Stenning & Oberlander 1994 is due to Robert Inder.

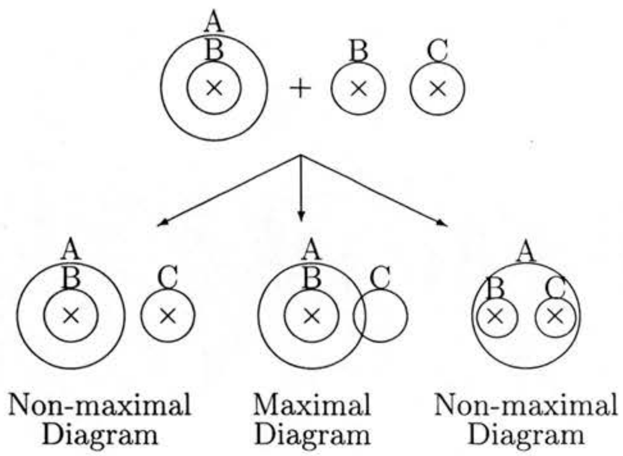


Figure 3.1: Constructing a registration diagram for the example "All of the Bs are As, None of the Bs are Cs". While any of the three possible registration diagrams is consistent with the premisses, the central one is chosen since it has the largest number of regions.

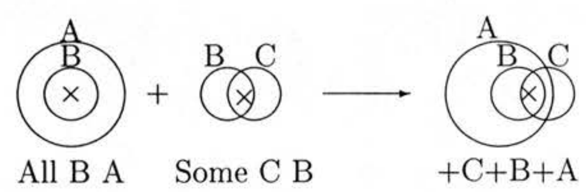


Figure 3.2: Registration diagram for the problem "All of the Bs are As, Some of the Cs are Bs".

diagram, only one unification is possible, so it represents a critical individual.

In Figure 3.1, the only mark which occurs in all the final diagrams is the one in the 'B' circle. The 'C' circle is bisected in the central diagram, so both −A−B+C and +A−B+C are possible, so neither is necessary. The central diagram is the maximal diagram, the most general case, and only the marks which occur in maximal diagrams are guaranteed to correspond to necessary individuals. The marked 'B' region corresponds to the necessary individual +A+B−C, and since this region appears marked in both the characteristic diagrams, both premisses are source premisses in this case.

Another example is shown in Figure 3.2. The registration diagram to the right of the arrow is the maximal diagram this time, since this is the only one which need be drawn. In this case the marked region from the first premiss, "All of the

Bs are As" is removed, and the only one which remains in the final diagram is the one from the second premiss, "Some of the Cs are Bs", so only the second premiss is the source premiss and the necessary individual is +A+B+C.

Registration diagrams can form the basis for a logical classification of problem types. The labelling of the end-term circles is immaterial, since premisses can be reordered without affecting the logic of the problem, so the classification abstracts over such differences. Similarly, the lack of topological distinctions between representations of convertible premisses (e.g. *No As are Bs* $\leftrightarrow$ *No Bs are As*) means that the classification abstracts over these differences too. Figure A.1 in Appendix A.1, shows the resulting set of 21 distinct registration diagrams, for problems with and problems without valid conclusions. It is useful to list positive and negative syllogisms separately.

Note that this graphical algorithm, which remains faithful to Euler's practice, is a deterministic procedure for syllogistic reasoning. It does not suffer from the combinatorial explosion of cases which some misinterpretations of Euler's method exhibit.

3.3.2 Individual identification in `prolog`

The case of two premisses

It is quite easy to implement the Individual Identification Algorithm in the logic programming language `prolog` (Clocksin & Mellish 1984). The code assumes standard versions of the predicates `member/2`, `nth/3` and `select/3`. Appendix A.2 presents a full `prolog` program, based on the Individual Identification code presented in this section, and which can draw quantified as well as individual conclusions.

Premiss Representation In prolog, premiss types can be mapped to characteristic representations using the predicate `characteristic/2`, whose first argument is a premiss and whose second is a characteristic representation, represented as a `Max-Min` pair (see Figure 3.3). Feature structures are represented as lists of features, and models as lists of feature structures. Note that each maximal model includes the type −A−B. In graphical terms this corresponds to the background

```
characteristic(all(A,B),[[-A,-B],[-A,+B],[+A,+B]]-[[+A,+B]]).
characteristic(some(A,B),[[-A,-B],[-A,+B],[+A,-B],[+A,+B]]-[[+A,+B]]).
characteristic(none(A,B),[[-A,-B],[-A,+B],[+A,-B]]-[[+A,-B],[-A,+B]]).
characteristic(somenot(A,B),[[-A,-B],[-A,+B],[+A,-B],[+A,+B]]-[[+A,-B]]).

syll(Premiss0,Premiss1,Conc,Source) :-
        characteristic(Premiss0,Max0-Min0),
        characteristic(Premiss1,Max1-Min1),
        nth(Source,[Min0,Min1],Min),
        nth(Source,[Max1,Max0],Max),
        member(I,Min),
        unification_test(I,Max,Conc).

unification_test(I,Max,K) :-
        bagof(Unified,J^(member(J,Max),unify(I,J,Unified)),[K]).

unify([],X,X).
unify([H|T],X,[H|NT]) :-
        select(H,X,Y),
        !,
        unify(T,Y,NT).
unify([H|T],X,[H|NT]) :-
        opp(H,Opp),
        not member(Opp,X),
        unify(T,X,NT).

opp(+A,-A).
opp(-A,+A).
```

Figure 3.3: Prolog code for individual identification

of the corresponding characteristic diagram.

Case Identification The predicate `syll/4` is the top level of the syllogism solution procedure. It retrieves two characteristic representations, one for each premiss, selects the minimal model (`Min`) from one and the maximal model (`Max`) from the other, selects a critical individual from `Min`, and subjects it to the unification test. If the unification test succeeds, it returns `Conc`, which is instantiated as a necessary individual. Standard prolog backtracking ensures that all critical individuals for both premisses are tested in turn, so the predicate finds all necessary individuals for the premiss pair. Whenever the predicate succeeds, `Source` is incidentally instantiated as the premiss number of the source. If there is no valid conclusion the predicate fails immediately.

`unification_test/3` tests the critical individual `I` by attempting to unify it with all members of `Max`. If the `bagof` goal returns a singleton list, only one unification is possible, so that unification denotes a necessary individual.

`unification_test/3` requires a unification predicate `unify/3`, whose first and second arguments are the feature structures to be unified, which succeeds with the third argument instantiated as their unification, if this is possible, and fails otherwise. Note that in the event of a successful unification, the features from the first argument appear before any from the second in the result, so that in its use above, terms from the source premiss occur before terms from the conditional premiss.

Examples Figure 3.4 shows two examples of the `prolog` Individual Identification program in action. The first example shows the result of giving it the usual problem, "Some of the As are Bs, All of the Bs are Cs". The program correctly concludes ABC, and identifies the source as premiss 1. Note that the terms from the source premiss appear first, in order. When backtracking is forced, the program finds no other solutions to the problem.

In the second example, the program is given a problem with two valid individuals. This time, after finding one of the solutions, when backtracking is forced, the program finds the other solution.

To make the operation of the program clearer, it may be useful to supply a Byrd Box Trace of a third example. The problem is the same as the first example

```
?- syll(some(a,b),all(b,c),Conc,Source).

Conc = [+a,+b,+c]
Source = 1
More (y/n)? y

no
?- syll(all(a,b),none(b,c),Conc,Source).

Conc = [+a,+b,-c]
Source = 1
More (y/n)? y

Conc = [-b,+c,-a]
Source = 2
More (y/n)? y

no
```

Figure 3.4: Examples of use of the prolog code

in Figure 3.4, but with premiss order reversed; consequently the program finds the conclusion only after some backtracking.

```
?- trace,syll(all(b,a),some(c,b),Conc,Source).
   (1) 0 CALL: syll(all(b,a),some(c,b),_426,_460)?
   (2) 1 CALL: characteristic(all(b,a),_698-_702)?
   (2) 1 EXIT: characteristic(all(b,a),[[-b,-a],[-b,+a],[+b,+a]]-[[+b,+a]])
   (3) 1 CALL: characteristic(some(c,b),_906-_910)?
   (3) 1 EXIT: characteristic(some(c,b),[[-c,-b],[+c,-b],[+c,+b],[-c,+b]]-
[[+c,+b]])
```

Since the first two arguments of `syll/4` are instantiated as the two premisses, `all(b,a)` and `some(c,b)`, the first two subgoals of `syll/4` simply retrieve the unique characteristic representations for each of the premisses.

```
   (4) 1 CALL: nth(_460,[[[+b,+a]],[[+c,+b]]],_1162)?skip
 > (4) 1 EXIT: nth(1,[[[+b,+a]],[[+c,+b]]],[[+b,+a]])?
```

Having translated the premisses into characteristic representations, and since the variable `Source` is uninstantiated, one premiss must be selected as the putative source premiss. The standard predicate `nth/3` is used to select one minimal model from a list of two, one for each premiss in order, and to instantiate Source as the source premiss number. Since the minimal models are *in premiss order* in the input list, and the minimal model of the first premiss ([[+b,+a]]) is selected, Source is instantiated as 1.

```
   (5) 1 CALL: nth(1,[[[-c,-b],[+c,-b],[+c,+b],[-c,+b]],[[-b,-a]
,[-b,+a],[+b,+a]]],_1322)?skip
 > (5) 1 EXIT: nth(1,[[[-c,-b],[+c,-b],[+c,+b],[-c,+b]],[[-b,-a]
,[-b,+a],[+b,+a]]],[[-c,-b],[+c,-b],[+c,+b],[-c,+b]])?
```

In order to pair the minimal model of one premiss with the maximal model of the other, the input to the second nth/3 subgoal is a list of the two premiss' maximal models *in reverse premiss order*. Since Source is instantiated, the first maximal model in the list is selected, corresponding to the second premiss. Thus the maximal model is [[-c,-b],[+c,-b],[+c,+b],[-c,+b]].

```
   (6) 1 CALL: member(_1458,[[+b,+a]])?skip
 > (6) 1 EXIT: member([+b,+a],[[+b,+a]])?
```

Since a minimal model may have more than one member (i.e. in the case of 'None'), one of them is chosen as the critical individual. In this example, there is only one, [+b,+a].

```
   (7) 1 CALL: unification_test([+b,+a],[[-c,-b],[+c,-b],[+c,+b],[-c,+b]],_426)?s
 > (7) 1 FAIL: unification_test([+b,+a],[[-c,-b],[+c,-b],[+c,+b],[-c,+b]],_426)?
```

Because there is more than one unification of the critical individual with members of the maximal model, unification_test/2 fails, forcing backtracking.

```
   (6) 1 REDO: member([+b,+a],[[+b,+a]])?skip
 > (6) 1 FAIL: member(_1458,[[+b,+a]])?
   (5) 1 REDO: nth(1,[[[-c,-b],[+c,-b],[+c,+b],[-c,+b]],[[-b,-a]
,[-b,+a],[+b,+a]]],[[-c,-b],[+c,-b],[+c,+b],[-c,+b]])?skip
 > (5) 1 FAIL: nth(1,[[[-c,-b],[+c,-b],[+c,+b],[-c,+b]],[[-b,-a]
,[-b,+a],[+b,+a]]],_1322)?
   (4) 1 REDO: nth(1,[[[+b,+a]],[[+c,+b]]],[[+b,+a]])?skip
 > (4) 1 EXIT: nth(2,[[[+b,+a]],[[+c,+b]]],[[+c,+b]])?
```

Both the member/2 and second nth/3 goals admit no more solutions, so the system backtracks to the first nth/3 goal, the selection of source premiss and minimal model. This resatisfies, and instantiates Source as 2 and Min as the minimal model of the second premiss, [[+c,+b]].

```
   (8) 1 CALL: nth(2,[[[-c,-b],[+c,-b],[+c,+b],[-c,+b]],[[-b,-a]
,[-b,+a],[+b,+a]]],_1334)?skip
 > (8) 1 EXIT: nth(2,[[[-c,-b],[+c,-b],[+c,+b],[-c,+b]],[[-b,-a]
,[-b,+a],[+b,+a]]],[[-b,-a],[-b,+a],[+b,+a]])?
```

Since the putative source is now 2, the second nth/3 goal succeeds and instantiates Max as the second member of the list, the maximal model of the first premiss, [[-b,-a],[-b,+a],[+b,+a]]. This time, the unification test will succeed, since there is only one member of the maximal model ([+b,+a]) with which the only critical individual ([+c,+b]) can unify.

```
   (9) 1 CALL: member(_1482,[[+c,+b]])?skip
 > (9) 1 EXIT: member([+c,+b],[[+c,+b]])?
  (10) 1 CALL: unification_test([+c,+b],[[-b,-a],[-b,+a],[+b,+a]],_426)?skip
 >(10) 1 EXIT: unification_test([+c,+b],[[-b,-a],[-b,+a],[+b,+a]],[+c,+b,+a])?
   (1) 0 EXIT: syll(all(b,a),some(c,b),[+c,+b,+a],2)
Conc = [+c,+b,+a]
Source = 2
More (y/n)? y
```

unification_test/2 is satisfied, and so the top-level goal succeeds; the conclusion is [+c,+b,+a] and the source is premiss 2. Backtracking is forced, to search for further valid conclusions, but no subgoals can resatisfy, and so the top-level goal fails.

```
   (1) 0 REDO: syll(all(b,a),some(c,b),[+c,+b,+a],2)?
  (10) 1 REDO: unification_test([+c,+b],[[-b,-a],[-b,+a],[+b,+a]],[+c,+b,+a])?s
 >(10) 1 FAIL: unification_test([+c,+b],[[-b,-a],[-b,+a],[+b,+a]],_426)?
   (9) 1 REDO: member([+c,+b],[[+c,+b]])?skip
 > (9) 1 FAIL: member(_1482,[[+c,+b]])?
   (8) 1 REDO: nth(2,[[[-c,-b],[+c,-b],[+c,+b],[-c,+b]],[[-b,-a],
[-b,+a],[+b,+a]]],[[-b,-a],[-b,+a],[+b,+a]])?skip
 > (8) 1 FAIL: nth(2,[[[-c,-b],[+c,-b],[+c,+b],[-c,+b]],[[-b,-a],
[-b,+a],[+b,+a]]],_1334)?
```

```
multsyll(Premisses,Conc,Source) :-
        select(Premiss,Premisses,Rest),
        nth(Source,Premisses,Premiss),
        characteristic(Premiss,_-Min),
        member(Ind,Min),
        msyll(Ind,Rest,Conc).

msyll(Conc,[],Conc).
msyll(Ind,Premisses,MConc) :-
        select(Premiss,Premisses,Rest),
        characteristic(Premiss,Max-_),
        unification_test(Ind,Max,Conc),
        msyll(Conc,Rest,MConc).
```

Figure 3.5: Prolog code for multiple premisses.

```
   (4) 1 REDO: nth(2,[[[+b,+a]],[[+c,+b]]],[[+c,+b]])?skip
 > (4) 1 FAIL: nth(_460,[[[+b,+a]],[[+c,+b]]],_1162)?
   (3) 1 REDO: characteristic(some(c,b),[[-c,-b],[+c,-b],[+c,+b],
[-c,+b]]-[[+c,+b]])?
   (3) 1 FAIL: characteristic(some(c,b),_906-_910)
   (2) 1 REDO: characteristic(all(b,a),[[-b,-a],[-b,+a],[+b,+a]]-[[+b,+a]])?
   (2) 1 FAIL: characteristic(all(b,a),_698-_702)
   (1) 0 FAIL: syll(all(b,a),some(c,b),_426,_460)
no
?-
```

The case of multiple premisses

Program code The prolog code can be adapted to the case of multiple syllogistic premisses, as shown in Figure 3.5. As before, the task is to identify maximally specified necessary individuals, so that each term occurring in the premisses must have a determinate value in the conclusion.

The two predicates `multsyll/3` and `msyll/3` replace the old top-level predicate `syll/4`. `multsyll/3` is the new top-level predicate, and its first argument is a list of premisses. It selects a premiss from the list to function as source, retrieves a member of the minimal model for this premiss (`Ind`), and passes this, along with the list of the remaining premisses (`Rest`) to msyll/3. In the event that `msyll/3` fails, backtracking ensures that each premiss is selected in turn as putative source.

`msyll/3` is a recursive predicate to handle the rest of the premisses. Its first

```
?- multsyll([some(a,b),all(b,c),none(c,d)],Conc,Source).

Conc = [+a,+b,+c,-d]
Source = 1
More (y/n)? y

no
?- multsyll([all(a,e),some(a,b),all(b,c),all(c,d)],Conc,Source).

Conc = [+a,+b,+e,+c,+d]
Source = 2
More (y/n)? y

Conc = [+a,+b,+c,+e,+d]
Source = 2
More (y/n)? y

Conc = [+a,+b,+c,+d,+e]
Source = 2
More (y/n)? y

no
```

Figure 3.6: Multiple premiss examples.

argument is the critical individual entailed by the premisses processed so far, and its second argument is a list of premisses yet to be processed. The first clause defines the base case: if the list of premisses is empty, it returns the value of the first argument as the conclusion. In the recursive case, `msyll/3` selects one of the premisses, retrieves its maximal model, and performs the unification test on the input critical individual `Ind`. Provided this is successful, the resulting necessary individual (`Conc`) is passed recursively to `msyll/3` along with the rest of the premisses.

Examples and discussion The operation of `msyll/3` can be summarised as the repeated application of conditionals to an input individual derived from the source, which in turn is identified by `multsyll/3`. Thus the multiple-premiss case also crucially depends on the source/conditional distinction. The program can handle premisses in any order: the logic of the problem determines the output term order.

Figure 3.6 shows two example problems. In the first example, the order of terms in the conclusion is fully determined since the order of application of conditional premisses is determined. As in the two-premiss case, the terms from the source premiss appear first. The second example shows a problem where term order is not fully determined, since there is a choice in the order in which conditionals may be applied. Nevertheless, all conditional applications presuppose the source premiss, so its terms always appear first.

As far as I know, the program is complete for its domain, but since the properties of the domain have not been fully explored, this is at present unproven. It constitutes an extension of the two-premiss program, since it generates exactly the same results as `syll/4` when given a list of two premisses.

3.3.3 Implementation in Monadic Predicate Calculus

This section demonstrates how the Individual Identification Algorithm can be implemented in Monadic Predicate Calculus (MPC).

Representation in MPC In MPC, each premiss can be represented as the conjunction of two sentences, which function analogously to the minimal and maximal models. The first is a conjunction of sentences, each of which establishes the existence of a member of the minimal model. Thus this component carries the existential implications of the premiss under its standard syllogistic interpretation. The second sentence is a universally quantified representation of the truth conditions of the premiss in *canonical disjunctive normal form* (see Lemmon 1965), which has the property that every elementary conjunction in it specifies a unique interpretation of the premiss. It is therefore very like a truth-table. These *characteristic sentences* are given in Figure 3.7.

Case Identification in MPC Figure 3.8 shows an example proof in Monadic Predicate Calculus for the example *Some of the As are Bs, All of the Bs are Cs*. Lines (1) and (2) are the assumptions, which are representations in our chosen form for the two premisses. In line (3) the sentence corresponding to the minimal model for premiss 1 is derived; since this is existentially quantified, (4) assumes the corresponding arbitrarily instantiated sentence. Lines (5) and (6) derive the maximal model for premiss 2, and instantiate it with the chosen arbitrary name.

All A B	$(\exists x)(Ax\&Bx)\&$
	$(\forall y)((Ay\&By) \vee (\neg Ay\&By) \vee (\neg Ay\&\neg By))$
Some A B	$(\exists x)(Ax\&Bx)\&$
	$(\forall y)((Ay\&By) \vee (\neg Ay\&By) \vee (Ay\&\neg By) \vee (\neg Ay\&\neg By))$
No A B	$((\exists x)(\neg Ax\&Bx)\&(\exists y)(Ay\&\neg By))\&$
	$(\forall z)((\neg Az\&Bz) \vee (Az\&\neg Bz) \vee (\neg Az\&\neg Bz))$
Some A not B	$(\exists x)(Ax\&\neg Bx)\&$
	$(\forall y)((Ay\&By) \vee (\neg Ay\&By) \vee (Ay\&\neg By) \vee (\neg Ay\&\neg By))$

Figure 3.7: Characteristic sentences for the MPC implementation

Most of the remainder of the proof is an $\vee$-elimination, so each of the disjuncts in (6) must be assumed in turn, as in lines (7), (11) and (15). The strategy is to find one of these disjuncts which has the same formula containing the B predicate as in (4) (which denotes a member of the minimal model for the first premiss), then to unify it with the formula in (4), to give a formula denoting the necessary individual (9). When the disjunct contains a contradictory formula for the B term, then we derive a contradiction (as in (13) and (17)), and using *Reductio Ad Absurdum* we are then permitted to derive anything ($P\&\neg P \vdash Q$), so we derive the same formula as we did by unification, as in (14) and (18). Since each of the disjuncts in (5) entails the same conclusion, $\vee$-elimination is permitted, discharging the auxiliary assumptions (7), (11) and (15) so that the conclusion now follows from (2) and (4). Since the name in (4) does not occur in (2), $\exists$-elimination is permitted, so the existentially quantified conclusion now follows from (1) and (2).

When more than one of the disjuncts is unifiable with the member of the minimal model, we can only derive a disjunctive conclusion by $\vee$-elimination.

The natural deduction system is a variant on Lemmon's "Beginning Logic", with &-elimination and $\vee$-elimination generalised to handle multiple conjuncts and disjuncts respectively, for convenience and to shorten the proofs.

3.3.4 Predicate Calculus again, but simpler!

The MPC model above is quite complex, since it is a literal translation of the graphical method. It is worthwhile to show that the source/conditional distinction also applies to a more conventional implementation of the IIA in Monadic Predicate Calculus. Figure 3.9 shows the premiss representations.

1	(1)	$(\exists x)(Ax\&Bx)\&(\forall y)((Ay\&By)\vee$ $(\neg Ay\&By)\vee(Ay\&\neg By)\vee(\neg Ay\&\neg By))$	A
2	(2)	$(\exists x)(Bx\&Cx)\&$ $(\forall y)((By\&Cy)\vee(\neg By\&Cy)\vee(\neg By\&\neg Cy))$	A
1	(3)	$(\exists x)(Ax\&Bx)$	1 &E
4	(4)	$(Aa\&Ba)$	A
2	(5)	$(\forall y)((By\&Cy)\vee(\neg By\&Cy)\vee(\neg By\&\neg Cy))$	2 &E
2	(6)	$(Ba\&Ca)\vee(\neg Ba\&Ca)\vee(\neg Ba\&\neg Ca)$	5 $\forall$E
7	(7)	$(Ba\&Ca)$	A
7	(8)	Ca	7 &E
4,7	(9)	$(Aa\&Ba\&Ca)$	4,8 &I
4	(10)	Ba	4 &E
11	(11)	$(\neg Ba\&Ca)$	A
11	(12)	$\neg Ba$	11 &E
4,11	(13)	$(Ba\&\neg Ba)$	10,12 &I
4,11	(14)	$(Aa\&Ba\&Ca)$	13 RAA
15	(15)	$(\neg Ba\&\neg Ca)$	A
15	(16)	$\neg Ba$	15 &E
4,15	(17)	$(Ba\&\neg Ba)$	9,16 &I
4,15	(18)	$(Aa\&Ba\&Ca)$	17 RAA
4,2	(19)	$(Aa\&Ba\&Ca)$	6,7,9,11,14,15,18 $\vee$E
1,2	(20)	$(\exists x)(Ax\&Bx\&Cx)$	3,4,19 $\exists$E

Figure 3.8: An example proof in MPC.

All A B	$(\forall x)(Ax \to Bx)$
Some A B	$(\exists x)(Ax\&Bx)$
No A B	$(\forall x)(Ax \to \neg Bx)$
Some A not B	$(\exists x)(Ax\&\neg Bx)$

Figure 3.9: Direct translation of premisses into Predicate Calculus.

Consider any valid syllogism. A proof of either an individual or quantified conclusion must crucially use at least one of the premiss representations for a conditional inference (either *modus ponens* or *modus tollens*). This is the conditional premiss, and it must be an "All" or "No" premiss. The other provides the input to this main conditional, in the form of an atomic sentence representing the middle term. This premiss may be either particular or universal. If it is particular, the proof strategy is ∃-elimination, followed by &-elimination, followed by the main conditional inference, whereas if it is universal, after ∀-elimination, it is necessary to assume an atomic sentence and apply a conditional inference to derive the middle-term sentence for input into the main conditional.

One of the premisses, the source, serves to assert the existence of something, either directly or indirectly, whereas the other, the conditional, serves as a license to infer something else from that initial assertion. Processing the source premiss is a logical prerequisite for the application of the conditional. It should be obvious that the distinction is likely to apply to quite a wide range of methods—this may help to motivate the use made below of the distinction in analysing term order effects in human reasoning data.

3.3.5 Discussion of the Implementations

While the parallel with the Euler Circles method informed the development of the `prolog` program and the first MPC method, these differ from the Euler Circles implementation in one major respect. The complete maximal model is never constructed at any point during the processing cycle; instead, the source/conditional distinction is assumed from the outset. This results in a kind of focussing on critical individuals, and so only the parts of the maximal model which directly intersect with critical individuals are constructed. By contrast, the graphical nature of the Euler Circles method imposes specificity even where the status of the individuals represented could never be of consequence (Stenning & Oberlander 1994, 1995).

The corollary of this is that in the `prolog` and MPC models, for particular premisses, the part of the characteristic representation representing the maximal model is redundant. Since these premiss types only ever function as source, the maximal model component could in principle be omitted entirely. It is easy to implement this "optimisation", by simply replacing the maximal model components with empty lists, as below.

```
characteristic(some(A,B),[]-[[+A,+B]]).
characteristic(somenot(A,B),[]-[[+A,-B]]).
```

While it would presumably be possible to strip down the Euler Circles representations in some way to achieve the same effect, the graphical method would lose its immediacy, and in any case the method as it stands is arguably more efficient, since it tests all critical regions in a single diagram.

Chapter 4

Figural Effects and Mental Models Theory

Figural effects in human syllogistic reasoning are response biases in the order of end terms in conclusions, which are correlated with the Figure of the premisses (Johnson-Laird & Steedman 1978). The effect can be summarised roughly as the tendency for end terms in conclusions to retain their grammatical status from the premiss in which they occur. Thus when the Figure is ab/bc, ac conclusions are commonest, ba/cb problems give rise to ca conclusions, and there is less overall bias when the Figure is ab/cb or ba/bc. These are highly reliable effects, which have been shown to occur in both valid conclusions (where the conclusion is convertible) and invalid conclusions.

This Chapter first describes and critically analyses Johnson-Laird *et al*'s Mental Models account of the Figural Effect, and presents the results of an unpublished experiment which bears on some of the issues raised by the critical analysis. A novel alternative account of the Figural Effect, based on the Individual Identification Algorithm and the source/conditional distinction, is then introduced and elaborated (first published in an early version by Yule & Stenning 1992). This account forms the basis for the statistical analyses of the Figural Effect presented in subsequent Chapters.

a	b		b	c		a	b	c
a	b	+	b	c	→	a	b	c
(a)	(b)			(c)		(a)	(b)	(c)

Figure 4.1: Constructing an integrated Mental Model for the problem *Some of the As are Bs, All of the Bs are Cs.* Because the b tokens are contiguous, the models can be joined by superimposing the b columns.

4.1 The Mental Models account of the Figural Effect

Johnson-Laird's explanation for the Figural Effect (Johnson-Laird 1983, Johnson-Laird & Bara 1984, Byrne & Johnson-Laird 1991) depends on a presumed First In-First Out (FIFO) characteristic of working memory, in which Mental Models are held to be created and integrated. The order in which mental models are entered into working memory is constrained by the need to link the models via the tokens corresponding to the middle term. Consequently, in ab/bc problems, the mental model for Premiss 1 enters working memory, followed by the mental model for Premiss 2, allowing the b tokens to be brought into contiguity, and the FIFO rule ensures that the a term is read off before the c term, giving an a-c conclusion (see Fig. 4.1). In ba/cb problems, the need to bring the b tokens into contiguity entails that the mental model for Premiss 2 must be entered into working memory first, followed by that for Premiss 1, so the FIFO rule produces a c-a conclusion.

It is necessary to extend the range of processes available in order to handle problems with the ab/cb and ba/bc Figures, which on the account so far cannot be solved, because the b tokens cannot be made contiguous. Johnson-Laird proposes that mental models can be "switched round", i.e. the term order is inverted. If one of the mental models is switched round, the other mental model can be integrated with it. The term order of the conclusion thus depends on which of the mental models is switched round, and since either route is available, both possible term orderings can occur. However, Johnson-Laird makes some predictions of the relative likelihood of each term ordering in each Figure on the basis of the complexity of operations required to form an integrated model. He assumes the

Operations	ab/bc		ba/cb		ab/cb		ba/bc	
Switch round	—	*Pr. 2*	—	*Pr. 1*	—	—	Pr. 1	Pr. 2
Build model	Pr. 1	*Pr. 2*	Pr. 2	*Pr. 1*	Pr. 1	Pr. 2	Pr. 1	Pr. 2
Renew	—	*Pr. 1*	Pr. 1	—	—	Pr. 1	—	Pr. 1
Switch round	—	*Pr. 1*	—	*Pr. 2*	Pr. 2	Pr. 1	—	—
Integrate	Pr. 2	*Pr. 1*	Pr. 1	*Pr. 2*	Pr. 2	Pr. 1	Pr. 2	Pr. 1
Predicted								
Response bias:	a-c	*c-a*	c-a	*a-c*	a-c	c-a	a-c	c-a
Obtained								
Response bias:								
Experiment 1	89%	*11%*	91%	*9%*	50%	50%	64%	36%
Experiment 2	93%	*7%*	86%	*14%*	55%	45%	62%	38%
Experiment 3	94%	*6%*	66%	*34%*	56%	44%	73%	27%

Table 4.1: A modification of Johnson-Laird's (1983) Table 5.3 (p109) in which the operation 'switch round' is permitted in the ab/bc and ba/cb Figures, permitting counterfigural responses in these Figures (italicised columns).

simplest way is the one which is preferentially followed.

Unfortunately the addition of the switching operation to the theory must be restricted to the ab/cb and ba/bc Figures (the *symmetric* Figures) only, since the possibility of switching both mental models for the ab/bc or ba/cb Figures (i.e. the *diagonal* Figures) would permit counterfigural responses in these Figures, and the theory would then permit any term ordering in any Figure. Johnson-Laird (1983) does not discuss the reason for this restriction explicitly, but it is worth noting that it makes the explanation of the Figural Effects different in different Figures, weakening the explanatory power of the theory. A more adequate theory of Figural Effects would explain ordering effects in a uniform fashion across the Figures, instead of relativising explanation to specific groups of Figures.

The most obvious variant of Johnson-Laird's (1983) explanation which meets this criterion would permit switching round, and consequently counterfigural responses, in the diagonal Figures, but argue that subjects will typically use the most efficient available method for forming integrated mental models, so the methods which result in counterfigural responses in these Figures will rarely be used. A variant of Johnson-Laird's (1983) Table 5.3 (p109) based on this proposal is shown here in Table 4.1.

According to Johnson-Laird (1983), the complexity of operations required should affect the ease with which valid conclusions can be drawn, so the four

Figures should show a trend of increasing difficulty, since ab/bc requires a minimum of two operations, ba/cb additionally requires renewal[1] of a premiss, and ab/cb and ba/bc require a 'switching round' operation. Assuming that renewal $\simeq$ switching in difficulty, then error frequency in the four Figures should be distributed thus: $e(ab/bc) < e(ba/cb) \simeq e(ab/cb) \simeq e(ba/bc)$. Johnson-Laird & Bara (1984) provide evidence that $e(ab/bc) < e(ba/cb) < e(ab/cb) < e(ba/bc)$, which could be accommodated by assuming that switching round has more effect on the error rate than does renewal, a plausible assumption given the possibility of conversion errors due to switching and the likely reliability of renewal in circumstances where both premisses are available for inspection throughout. Also, as Table 4.1 shows, counterfigural responses do occur in reasonable numbers in the diagonal Figures, relatively infrequently in ab/bc, where the counterfigural route is the most complex of all, and slightly more frequently in ba/cb, where the counterfigural route is presumably the next most complex route before the c-a routes in the symmetric Figures. So overall, the revised model shows reasonable consonance with the data.

The rarity of counterfigural responses in the diagonal Figures is thus explained by the hypothesis that the simplest available way of forming an integrated model is preferentially followed, so the Figural Effect is due to the optimization of the model-construction process. While the operation of switching is available for diagonal Figure problems, it is unnecessarily complicated to use it, so it rarely happens. However, there is a possibility of observing the effects of changing the relative difficulty of switching and renewal by experimental manipulation, so for example decreasing the attractiveness of the renewal option might increase the likelihood of a switching strategy in ba/cb, resulting in more counterfigural responses. The Levy (1984) experiment explores one means by which such circumstances might be created.

4.2 The Levy (1984) Experiment

In the syllogistic reasoning task normally employed by Mental Models researchers, subjects draw conclusions from pairs of premisses which are presented *together*. In

[1] Accessing the premiss a second time, i.e. rereading it or recovering it from memory.

	ab/bc	ba/cb	ab/cb	ba/bc	Overall
Single-premiss	50	52	48	58	53
Double-premiss	60	60	57	60	59
Mean	55	56	53	59	56

Table 4.2: % correct valid conclusions in each Figure and Presentation condition. Reproduced from Levy (1984)

the present case, however, the standard *double-premiss presentation (D)* is compared with *single-premiss presentation (S)*, in which premisses are presented singly, with removal of the first from the subject's field of view on presentation of the second. Since single-premiss mode makes renewal more difficult (the first premiss must be kept in memory, which is more fallible than paper or a VDU), we expect more errors in single-premiss mode than double-premiss mode. If subjects' selection of model-construction strategies is sensitive to this disadvantage, then there should be an increased tendency for them to choose non-renewing model construction strategies in single-premiss mode, resulting in an increase in the number of a-c responses in ba/cb, ab/cb and ba/bc, compared with double-premiss mode. Double premiss mode should replicate Johnson-Laird's results, since it is the control condition.

Since Working Memory span measures are available for each subject, it is also possible to test for correlations between memory span and performance. Mental Models theory is committed to the importance of memory capacity in determining performance, and thus presumably predicts positive correlations between memory span and performance in all Figures.

Table 4.2 shows the percentage correct valid conclusions in each Figure and presentation condition. Overall, the single-premiss condition produced fewer correct valid conclusions than did the double-premiss condition as predicted, ($Wilcoxon's$ $T(25) = 83$, $p < 0.025$, $one - tailed$), but the differences were not significant within individual Figure groups, possibly due to the small sample sizes. There was no significant effect of Figure on percentage correct valid conclusions (Jonckheere Trend Tests on S and S+D pooled).

Table 4.3 shows the percentages of a-c and c-a responses in each Figure and Presentation condition. Double-premiss mode is a straightforward replication of

	ab/bc		ba/cb		ab/cb		ba/bc	
	S	D	S	D	S	D	S	D
a-c	43.1	46.8	2.5	0.6	11.3	16.9	23.1	28.8
c-a	0.6	1.3	39.4	35.0	18.1	19.4	21.9	18.1
a-c Bias	98.6	97.3	6.0	1.7	38.4	46.6	51.3	61.4

Table 4.3: % of a-c and c-a responses in each Figure and Presentation condition. Reproduced from Levy (1984)

	ab/bc		ba/cb		ab/cb		ba/bc	
WM Measure	2+2	3+3	2+2	3+3	2+2	3+3	2+2	3+3
Single premiss	-0.12	0.18	**0.62**	**0.61**	0.24	0.20	0.21	-0.02
Double premiss	-0.02	-0.08	0.38	0.16	-0.13	-0.13	-0.04	-0.25

Table 4.4: Pearson's Correlations between % correct valid conclusions, for single-premiss and double-premiss presentation modes, and the 2+2 and 3+3 working memory measures, in each Figure. Boldface indicates significance at the 5% level, one-tailed. N=10 in each Figure.

the standard paradigm, so we predict strong a-c and c-a biases for ab/bc and ba/cb respectively, and smaller a-c biases for the symmetric Figures. Although the other three Figures are in line with prediction, ab/cb shows a slight c-a bias. We expect single-premiss mode to increase the likelihood of a-c responses in all Figures. Insignificant increases are observed in the diagonal Figures, but decreases in the numbers of a-c responses are observed in the symmetric Figures, contrary to prediction.

Table 4.4 shows correlations between percentage correct valid conclusions and memory span, for both presentation modes and memory span measures, in each of the four Figures. We expect positive correlations, but find them only for ba/cb subjects, with single-premiss presentation, where % correct valid conclusions correlates significantly with both 2+2 and 3+3 working memory test scores.

Discussion of Levy (1984) While single-premiss presentation made reasoning more error-prone overall, there was no sign of a change in model-construction strategy to avoid renewal-based strategies, evidenced by the lack of increase in the ac bias for ba/cb, ab/cb and ba/bc in single-premiss mode. Possibly switching is unavailable in ba/cb, as Johnson-Laird (1983) assumes, but the failure of subjects to take advantage of switching in the symmetric Figures is surprising, given that Johnson-Laird (1983) assumes that they must in any case use it at least once to succeed in those problems.

The correlation results show no overall effect of Working Memory span, as measured by the 2+2 and 3+3 tests, on performance. Perhaps the tests do not measure the right kind of memory, but the presence of strong correlations, as predicted but in ba/cb only, suggests that single-premiss mode does impose a memory load in this Figure. While the unamended account of the Figural Effect can accommodate this finding on the grounds that ba/cb is the only Figure which *requires* the "renew" operation, whereas in all the other Figures it is possible to build the model on the basis of the first premiss, this would entail an increase in the number of ac conclusions in single premiss mode in the symmetric Figures, which does not occur.

The absence of the predicted trend of increasing difficulty with Figure in this experiment may be due to the between-subjects design, but it is worth mentioning that the analysis used by Johnson-Laird & Bara (1984) to test this hypothesis is rather gross, insofar as no attempt was made to control for problem difficulty among the Figures. It is harder to dismiss Johnson-Laird & Bara's latency results, which also showed an increasing trend with Figure, since in this case only single-model problems were used in the analysis.

Johnson-Laird is left with the dilemma that either the exclusion of switching in the diagonal Figures in his theory is arbitrary but has a degree of confirmation in practice, or the version of the theory without such restriction is not arbitrary, but not reliably confirmed either. Moreover, there is very little direct evidence for the involvement of FIFO memory in syllogistic reasoning. A study by Gilhooly *et al* (1993), which did directly address the issue, failed to show expected suppression effects, though this may have been due to methodological problems of incomplete suppression.

In any event, such appeal to details of memory implementation seems premature when we there is so wide a range of types of implementation which are not distinguishable on the basis of input/output mappings. Natural implementations of the IIA in both graphical and sentential modes generate ordered term structures without making any highly specific assumptions about the substrate of term-order memory. It is clear from the published data (e.g. Johnson-Laird & Steedman 1978, Johnson-Laird & Bara 1984) that the grammatical organisation of conclusions interacts powerfully with the quantifiers as well as with Figure. What is needed is an explanation of the how the grammatical structuring effects result from the processes of reasoning rather than a characterisation of the substrate which preserves order information, especially when it is clear that several substrates may equally well achieve the same effect.

Before moving on to present a novel account of the Figural Effect, a recent proposal by Wetherick (1989, Wetherick & Gilhooly 1990) must also be considered. This relates the Figural Effect to discourse understanding phenomena. In short, subjects seek to find the topic of the syllogistic mini-discourse, and tend to pick a subject end term for this purpose. In the ab/bc and ba/cb, this leads to ac or ca conclusions only, respectively, while in ab/cb there is a choice of subject end terms, so an approximately 50/50 split is expected. Only ba/bc lacks a subject end term, but again there is nothing to choose between the premisses, so again we expect a 50/50 split. Expressed this way, we expect only the gross Figural Effect, so this theory predicts exactly the same as does Mental Models.

If a theory can predict specific term orderings individually for syllogisms, the logical structure of the Figures can result in an emergent gross Figural Effect of the type that has been considered up to now. This point can be clarified by consideration of the fact that the problems in ba/cb are identical to those in ab/bc, except for the premiss reversal of course, while inside each of the symmetric Figures, half the problems are equivalent to the other half, except for premiss reversal. Thus any theory that makes problem-by-problem predictions of term order will automatically produce an overall 50/50 split in the symmetric Figures. In the diagonal Figures, a within-Figure imbalance in the number of problems predicted to have each of the term orders will result in the characteristic overall inversion of preferred orderings between these Figures. Any such theory would have far greater explanatory power than theories, such as those discussed above, which simply predict the gross Figural Effect without making any within-Figure discriminations.

To address this challenge, Section 4.3 elaborates an alternative account of the Figural Effect, in terms of the source/conditional distinction. This generates predictions which are tested in Chapter 5 using data from the novel Individuals Task. The results of this experiment are used to guide the construction of statistical models of term ordering in Chapter 6.

4.3 Individuals and Figural Effects

The IIA casts the syllogistic reasoning task as one of constructing three-term descriptions of critical individuals. In the graphical version of the algorithm, these descriptions correspond to $\times$-marked regions in a finished diagram. In the simple predicate calculus version of the algorithm, they are the three-conjunct statements completed by application of MP or MT followed by conjoining of the concluded clause.

As has been demonstrated, the source premiss is identified as the premiss contributing the $\times$ which persists into the final diagram. In the simple predicate calculus implementation, the source premiss is the one which must be processed first, in preparation for the main conditional inference licensed by the conditional premiss. One logical generalisation about source premisses is particularly noteworthy—if there is an existential premiss, then *it* is the source premiss—but since universals have existential consequences in the conventional interpretation of the syllogism, these can be source premisses too. In general, the source premiss for some valid individual is any premiss which makes an existential assertion, from which the existence of that individual can be inferred by using the other (*conditional*) premiss as a rule of inference. For each valid individual for each syllogism, that individual may specify a unique source (e.g. Figure 4.2), or fail to because more than one derivation of that individual from those premisses is possible.

The source premiss property of problems is empirically interesting because it is distributed unevenly among the four Figures, in a way that appears to parallel the Figural Effect. That is, there is a tendency for ab/bc conclusions to have source Premiss 1, ba/cb conclusions to have source Premiss 2, and for the symmetric Figures to have equal numbers of conclusions with each source premiss. This is a reflection of the *logical* Figural Effect in the Standard Task, where there is an

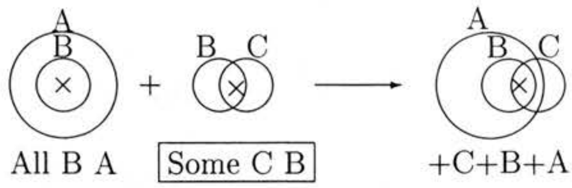


Figure 4.2: Registration diagram for a problem with source Premiss 2

imbalance in the numbers of valid conclusions with each possible term order in each Figure. In ab/bc, for example, there are more unconvertable ac valid conclusions than ca ones. The distribution of source parallels this logical Figural Effect, so that, usually, unconvertable ac conclusions have source premiss 1, and unconvertable ca conclusions have source premiss 2. The exceptions are the problems which have unconvertible, counterfigural[2] conclusions but no unique source (see diagram vn2 in Figure A.1 in Appendix A.1).

However, there are also convertible conclusions with unique source, and in these cases it appears that the source premiss predicts human reasoners' preferences quite well. In the second premiss source problem in Figure 4.2, for example, the overwhelmingly most popular conclusion is "Some of the Cs are As" (Johnson-Laird 1983)

The confounding influence of unconvertible valid conclusions makes the Standard Task an unsatisfactory testbed for hypotheses about source premisses. However in the Individuals Task, in which all conclusions are convertible, we expect source to give rise to term order phenomena comparable to the Standard Task Figural Effect, both for problems whose Standard Task conclusions are convertible, and those which are not. In particular, *we expect that the end term from the source premiss will tend to precede the conditional premiss end term in conclusions*, but given that the individual descriptions contain the middle term b as well as the end terms, we expect the terms from the source premiss to constitute the head of the description, possibly in their premiss order, followed by the remaining end term. The uneven distribution of problem types should then give rise to an end-term order pattern similar to the overall Figural Effect. Thus we might be

[2]As used by Johnson-Laird (1983), this refers to problems whose only valid conclusions have term order opposite to the figurally preferred order, e.g. only valid ca conclusions in ab/bc problems

able to reduce the Figural Effects to the consequence of a strategy of beginning the construction of the individual description from the source premiss.

The next Chapter introduces the Individuals Task, and presents some comparisons of Standard Task and Individuals Task performance results, before proceeding to analyses of term order phenomena in order to evaluate the hypothesis that the source/conditional distinction is the underlying cause of the Figural Effect.

Chapter 5

The Individuals Task Experiment

5.1 Introduction

The case-identifiability property of syllogisms, and the IIA constructed on its semantic basis, suggest both an explanation of ordering effects at an appropriate level of abstraction, and a method of gaining much richer data about mental processes. The IIA is a method for deciding whether there is any maximally specified individual which must exist in any model of the premisses, and of constructing a specification if any does exist. What distinguishes the algorithm is precisely the division of the reasoning process into a part which builds up specifications of fully determined individuals, and a part which draws conclusions by abstracting away from these specifications. The task which most directly operationalises the first half of this process is the task of giving descriptions of maximally specified individuals whose existence is entailed by the premisses.

This task has two main virtues for studying Figural Effects in reasoning. The sequence of terms in the description of such individuals is always logically immaterial—existential conjunctions of terms are indifferent to reordering of conjuncts. In the conventional task of drawing quantified conclusions, not all conclusions are validly convertible, and so sequence of terms is confounded with validity of conclusion. Additionally, the modified task yields data on all three terms of the syllogism, rather than just on end terms.

It should be obvious that this *Individuals Task*, having uniform quantificational structure in conclusions, should be better suited to evaluate claims about the relative difficulty of single- and multiple-model problems than is the *Standard Task* of drawing quantified conclusions relating the end terms. The evidence for a difference in difficulty is confounded, in the Standard Task, by the quantificational structure of conclusions, since multiple-model problems also tend to have conclusions in a different mood from either of the premisses, whereas single-model problems usually do not (Ford 1994).

Finally, the Individuals Task allows us to investigate whether naive reasoners agree with Aristotle that nothing follows from two negative premisses. If subjects can find the maximal individuals entailed by the premisses which licence U-conclusions, that provides evidence that the IIA underlies their reasoning.

In Section 5.2 below, the Durham Standard Task dataset is used for a problem-by-problem comparison with the Individuals Task data, to establish the relation between the performance profiles of the two tasks. The aim is to assess the degree of psychological, as opposed to logical similarity between the tasks. One basis for comparison is the difference in difficulty between single- and multiple-model problems; this can serve as a benchmark for evaluating performance on U-conclusion problems, against performance on conventionally valid multiple-model problems.

The remaining analyses, in Section 5.3, concern term ordering phenomena in the Individuals Task. The aim is to examine the novel data on middle term position in individual descriptions, and to test predictions of the effects of source premisses on term order.

5.2 Comparison with the Standard Task

In order to investigate the relation between performance on the Individuals Task and performance on the Standard Task, two scattergrams were constructed, showing data from problems without Valid Conclusions (Figure 5.1) and with Valid Conclusions (Figure 5.2) respectively[1]. Each point represents a single problem, and its X and Y values represent the percentage of subjects who responded correctly to it on each task. Moreover, each problem is labelled with the name of its diagram type (see Figure A.1 in Appendix A.1). Since the U-conclusion problems have valid conclusions in the Individuals Task but not in the Standard Task,

[1]One vn3 problem has been omitted, owing to an error in the materials set.

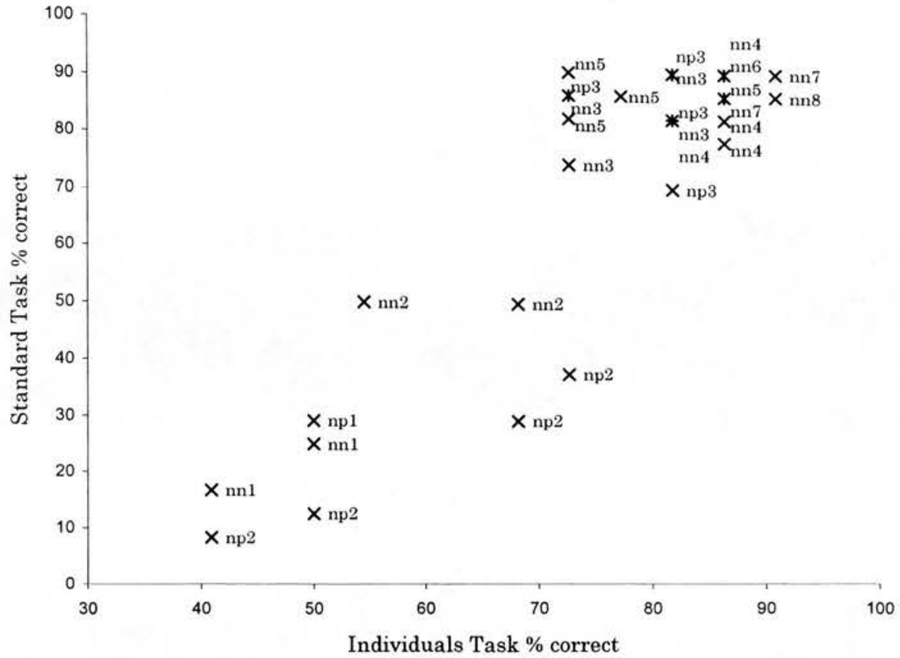


Figure 5.1: Scatter plot showing scores on NVC problems in both tasks.

results for these problems on the two tasks are not directly comparable, so they have been omitted from the main analysis. However, they have been plotted on Figure 5.2 using a different marker type for convenience.

Figure 5.1 shows problem-by-problem accuracy for problems without valid conclusions (see Tables B.11-B.21 in Appendix B.3 for complete tabulation of the raw data for NVC problems). Although scores on the Individuals Task are higher than those on the Standard Task, problem difficulty is highly correlated between the two tasks ($r = 0.88$, 26 $d.f.$ $p < 0.0005$). By inspection, the problems fall into two groups, one high-scoring and one low-scoring; moreover, no diagram type represented in either of the groups occurs in the other group. The high-scoring group is comprised of problems from diagram types np3, nn3, nn4, nn5, nn6, nn7 and nn8, while the low-scoring group comprises diagram types np1, np2, nn1 and nn2. This pattern of results is consistent with the hypothesis that subjects use the Aristotelian generalisations,

- No problem with two particular premisses has a valid conclusion
- No problem withtwo negative premisses has a valid conclusion.

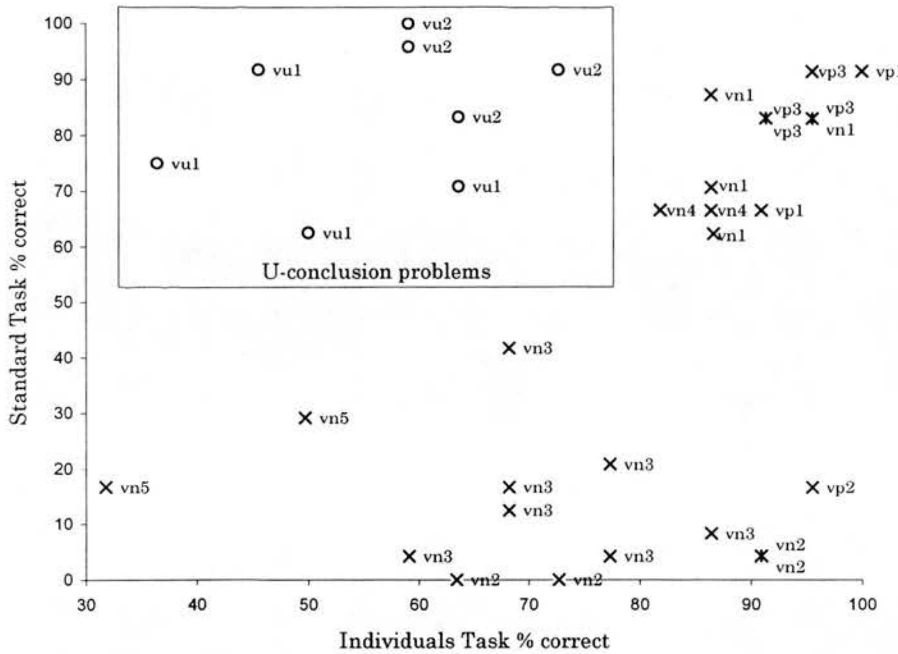


Figure 5.2: Scatter plot showing scores on VC problems in both tasks. Also shown are the U-conclusion problems, marked with circles in the box at top left - since these have valid conclusions in only the Individuals Task, their Standard Task scores are not comparable with the rest of the Standard Task scores on this chart.

The set of problems which can be decided by these rules is exactly the high-scoring group just identified; the rest, which cannot be decided by these rules, correspond to the set identified as the low-scoring problems. From Figure 5.2 it can be verified that the Standard Task scores for U-conclusion problems, where the correct response is No Valid Conclusion, fall in the same high-scoring region as would be expected on this hypothesis.

Note, however, that this 'Aristotelian' partitioning of problems can be equivalently described by reference to the presence or absence of an 'All' premiss: problems with two particular or two negative premisses are *ipso facto* those without an 'All' premiss, so the high-scoring problems are those without an 'All' premiss, whereas the low-scoring problems are those with an 'All' premiss. This is relevant to the discussion of the A-Effect in Section 7.2 below.

Figure 5.2 is a plot of performance on both tasks for problems with valid conclusions (see Tables B.1-B.8 in Appendix B.1 for complete tabulation of the raw data for problems with valid conclusions). For convenience, the U-conclusion problems are represented with circles appearing in the top left region (see Tables B.9 and B.10 in Appendix B.2). The valid conclusion results are less highly correlated than the NVC results ($r = 0.56$, 24 $d.f.$ $p < 0.005$); this is due to a substantial subset of problems in the bottom right region with high Individuals Task scores but very low Standard Task scores. As in Figure 5.1 there is a group of problems with high scores on both tasks: these are instances of diagram types vp1, vp3, vn1 and vn4, and as before no other instances of these diagram types appear anywhere else on the chart. All but one of the single-model diagram types occur in this group. In bottom left region the two multiple-models vn5 problems have low scores on both tasks. Finally, in the uncorrelated region at the bottom right, there are the vn3, vn2 and vp2 problems, with medium to high scores on the Individuals Task, but low scores on the Standard Task.

Since vp2 is a single-model problem, Mental Models theory would predict it to have a high level of Standard Task performance similar to the other single-model problems, rather than being as difficult as the multiple-models problems vn2 and vn3 as observed. Ford (1994) notes that vp2 is reliably difficult in other Standard Task experiments, and observes that it is the only single-model problem whose valid conclusion is in a different mood from either of the premisses, a property it shares with vn2 and vn3 problems. Choosing a suitable quantifier for the conclusion is a source of difficulty which occurs in the Standard task but not in the Individuals task, so it is reasonable to suppose that this different-mood hypothesis accounts for the observed difference in difficulty between the tasks for these problems. However, since most of the Standard Task multiple-models problems are vn2 and vn3 problems, the different-mood hypothesis undermines most of the evidence for a difference in difficulty between single-model and multiple-models problems.

In terms of the Individuals task alone, there is overlap between single-model and multiple-model performance; some vn3 and vn2 problems show similar high levels of performance to single-model problems. However the variation in multiple-model problem scores is much greater than in single-model scores. From the graph it is clear that while generating valid U-conclusions is fairly difficult, there are other multiple-model problems with comparable levels of difficulty. One of the

	Error conclusion		*Error NVC*	
	Inds Task	Std Task	Inds Task	Std Task
single-model	2.48%	5.47%	5.62%	16.51%
multiple-models	10.62%	29.32%	17.91%	58.00%
U-multiple-models	11.98%		29.55%	

Table 5.1: Mean Percentages of error conclusions and error NVC responses in both the Individuals Task (N=22) and Standard Task (N=24), for single-model, multiple-models and U-multiple-models problems (Individuals Task only).

vn5 problems shows lower scores than any U-conclusion problem, and the spread of U-conclusion problems overlaps to a considerable extent the distribution of vn2 and vn3 problems. It is not unreasonable, therefore, to consider the U-conclusion problems as multiple-models problems.

We can use a mixed-model Analysis of Variance to investigate the effect of number of models (repeated measures) on percentage error rates, for both Individuals Task and Standard Task subjects. Two ANOVAs were conducted, using percentage of error conclusion responses and error NVC responses respectively as the dependent variables. Table 5.1 shows for both Tasks, how mean percentages of error conclusions and error NVC responses vary between single-model, multiple-models and U-multiple-models problems.

Percentages of error conclusions are higher in multiple-models problems than in single-model problems, for both tasks ($Overall: \ F(1,44) = 34.84, \ p < 0.0001$). There is also a main effect of Task, such that there are more error conclusions in the Standard Task than in the Individuals Task ($F(1,44) = 11.07, \ p < 0.0018$), and there is a significant interaction between number of models and task ($F(1,44) = 8.40, \ p < 0.0058$). Although the difference between single-model and multiple-models problems is smaller in the Individuals Task, it is nonetheless significant ($t(21) = 2.67, \ p < 0.01$); however there is no significant difference in percentages of Error Individuals between multiple-models and U-multiple-models problems ($t(21) = 0.35, \ n.s.$).

Similarly for Error NVC responses, there are main effects of number of models ($F(1,44) = 95.22, \ p < 0.0001$), task ($F(1,44) = 27.23, \ p < 0.0001$) and a task by number of models interaction ($F(1,44) = 28.08, \ p < 0.0001$). Again, despite the interaction, there are significantly more error NVC responses to multiple-models problems than to single-model ones in the Individuals Task ($t(21) = 4.14,\ p < 0.0005$), but also there are significantly more error NVC responses to U-multiple-models problems than to conventional multiple-models problems ($t(21) = 2.30,\ p < 0.025$).

So the Individuals Task shows an effect of the number of models which is similar to, but less extreme than, the effect familiar from the literature on the Standard Task: multiple-models problems cause more errors of both types than do single-model problems. The U-multiple-models problems, which only occur in the Individuals Task, differ from conventional multiple-models problems only in having a higher rate of Error NVC responses. This may be explicable as the consequence of subjects' (over-hasty) assumption that the Aristotelian generalisation, that nothing follows from two negative premisses, holds for the Individuals Task.

5.3 Term order in the Individuals Task

As was discussed above, the Individuals Task makes an excellent testbed for hypotheses about term ordering phenomena. This section presents preliminary analyses of Individuals Task term order phenomena. In part, these preliminary analyses reproduce work which has already appeared in Ardin (1991). However, the main focus of the analyses presented here is on the novel question whether the source/conditional distinction affects term order, and if so, what relation this effect has to the Figural Effect. This material has been published in Yule & Stenning (1992).

If a subject correctly draws a valid conclusion to a problem, we can determine its source, providing that a unique source is defined for that conclusion to that problem. We predict an association between source premiss and end-term order, such that when the source is Premiss 1, a will precede c in the conclusion, and when the source is Premiss 2, c will precede a. When either premiss could have been the source, we make no prediction.

There is a set of problems which have two different individual conclusions, each with a different source. These have an A premiss whose subject is an end term, and an E premiss (e.g. AabEbc - Figure 5.3 shows the registration diagram).

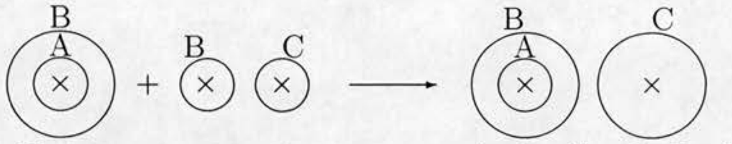


Figure 5.3: Registration diagram for a problem with two valid individuals, each with a different source.

Each of these individuals supports a different quantified conclusion, having only one positive feature corresponding to an end term (e.g. +a+b−c supports Eac (or Oac) and +c−b−a supports Eca (or Oca)). Since we cannot state *a priori* which premiss will be source, we cannot predict an overall Figural Effect in these problems, but we can test for an association between the source and term order, since the source can be determined *post hoc* on the basis of the conclusion drawn. Specifically, conclusions whose source is the A premiss should have the end term from that premiss before the end term from the other, and conclusions whose source is the E premiss should have the opposite end term order.

The dependency between source and end-term order should hold for problems in all the Figures. In the ab/bc and ba/cb Figures the imbalance in numbers of problems with Source 1 and 2 should give rise to an apparent effect of Figure: ab/bc has more Source 1 problems, so it should show an overall ac bias, whereas ba/cb, with more Source 2 problems, should show a ca bias. There are problems in both the ab/cb and ba/bc Figures which have unique source, but owing to the existence of complementary pairs of problems, with the same logic but inverted premiss order, there should be no overall bias due to the distribution of source in these Figures. The size of the diagonal effect can be estimated by assuming all correct conclusions to be equally likely, then the relative numbers of conclusions with each source premiss in the diagonal Figures predicts a small Figural bias in the region of ±7-10% (depending how we count conclusions to the two-individual problems).

The consequence of preservation of term order from the source premiss (or indeed any founding premiss) in individual descriptions is that in each Figure, we predict two dominant three-term orders out of the possible six, one founded on each of the premiss orders which jointly constitute that Figure. Since these always

			Figure				
			ab/bc	ba/cb	ab/cb	ba/bc	Overall
	valid	ac %	65.9	49.2	45.5	57.1	55.8
		N	138	122	88	217	565
Individuals	invalid	ac %	73.0	39.7	64.9	72.2	61.2
Task		N	63	63	57	36	219
	Overall	ac %	68.2	45.9	53.1	59.3	57.3
		N	201	185	145	253	784
	valid	ac %	88.2	8.8	54.2	51.2	50.4
		N	68	68	48	86	270
Standard	invalid	ac %	90.0	15.5	53.7	67.7	56.7
Task		N	90	84	82	65	321
	Overall	ac %	89.2	12.5	53.8	58.3	53.8
		N	158	152	130	151	591

Table 5.2: The Figural Effect in both tasks: percentages of valid and invalid conclusions with ac term order in each Figure.

have the middle term b in first or second position, there should be few conclusions with b final.

By contrast, Mental Models theory would presumably predict that *all* conclusions will be in the forms abc and cba, with no instances of the middle term occurring in either initial or final position. Also, the end-term order predictions by Figure should specify mostly ac conclusions in ab/bc, mostly ca conclusions in ba/cb and approximately equal numbers of ac and ca conclusions in ab/cb and ba/bc. Crucially, Mental Models theory does not discriminate among problems within Figures, since its explanation is solely in terms of Figure, whereas the source-founding predictions are on a problem-by-problem basis

The most straightforward way to compare term order in the Individuals Task with that in the Standard Task is to consider only the order in which the end terms occur in individual descriptions. Tabulating the percentages of conclusions in which a precedes c in the four Figures, for both valid and invalid conclusions, in each Task, we derive Table 5.2, which shows the Individuals Task counterpart of the Figural Effect. Taking account of a slight overall bias towards ac conclusions, it is clear that there is an overall Individuals Task Figural Effect ($\chi^2(3) = 20.88$, $p <$

	source			
	Premiss 1	Premiss 2	Both	*Overall*
ac %	77.9	28.0	65.7	55.8
N	213	218	134	565

Table 5.3: The Source Effect in the Individuals Task: Percentages of valid conclusions with ac order in first- and second-premiss unique source problems, and non-unique source problems.

.0001) similar to the type reliably observed in Standard Task experiments[2]. It is less pronounced for valid than for invalid conclusions, but is significant in both cases [Valid: ($\chi^2(3) = 11.90$, $p < .0077$), Invalid: ($\chi^2(3) = 18.16$, $p < .0004$)]. For valid conclusions in the diagonal Figures, after taking account of an overall ac bias of 6-8%, the size of the effect (±8-10%) is similar to that predicted by the distribution of source in these Figures. The Standard Task Figural Effect is more extreme than this. While the Figural Effect in valid conclusions arises in part from the imbalance in numbers of valid conclusions with each term order in the diagonal Figures, the size of the effect in invalid conclusions is also larger.

Table 5.3 shows how end-term order varies with source premiss. The Table shows ac order for problems with unique source, problems in which both premisses are source, and invalid conclusions. End-term order is strongly associated with source ($\chi^2(2) = 118.1$, $p < 0.0001$), such that a tends to precede c when the first premiss is the source, and c tends to precede a when the second premiss is the source, as predicted.

Table 5.4 shows the relation between end-term order and source premiss for all of the the problems which establish two different individuals (e.g. Figure 5.3). On the basis of the conclusion, we can determine whether the A or E premiss is the source, and we predict that the end-term from the source premiss should precede the end-term from the conditional premiss. As the Table shows, there are more conclusions with the A premiss as source than with the E premiss as source, and the end-term order is strongly predicted by the source premiss ($Yates'$ $\chi^2(1) =$

[2]The use of $\chi^2$with repeated measures data, as here, is liable to error; however, data are treated as frequency tabulations here for comparability with the loglinear models in later chapters.

	source		
End-term order	*A*	*E*	Total
AE	64	0	64
EA	4	10	14
Total	68	10	78

Table 5.4: Association between end-term order and source premiss for valid conclusions to problems which establish two individuals. End-term order is AE if the end term from the A premiss precedes that from the E premiss.

	Figure			
	ab/bc	ba/cb	ab/cb	ba/bc
source 1 ac%	93.8	92.6	90.0	50.7
source 2 ca%	80.0	78.1	91.7	47.8

Table 5.5: Prediction of Individuals Task end-term order by source in each Figure

46.24, $p < .0001$).

Table 5.5 shows how well source predicts end-term order in each Figure. In the first three Figures, source predicts end-term order very well, with more than 90% of problems with unique first-premiss source having ac conclusions; second-premiss source predicts ca conclusions in these Figures slightly less well, but much better than chance. Possibly this difference is attributable to the small overall preference for ac conclusions. More seriously however, in ba/bc source does not predict end-term order better than chance.

	order						
Figure	bac	bca	abc	cba	acb	cab	*N*
ab/bc	7.5	**23.9**	**59.2**	3.5	1.5	4.5	201
ba/cb	**37.3**	6.5	4.3	**44.3**	4.3	3.2	185
ab/cb	13.1	4.1	**36.6**	**35.2**	3.4	7.6	145
ba/bc	**52.6**	**35.6**	5.5	3.2	1.2	2.0	253
Overall %	**30.1**	**19.9**	**24.7**	**18.9**	2.4	4.0	784

Table 5.6: Percentages of Individuals Task conclusions with each possible term order in each Figure. Cells predicted to be the largest in each row are emboldened.

The final predictions concern the novel data on the position of the b term in individual descriptions. The hypothesis that individual descriptions are founded on one of the premisses, followed by the end term of the other premiss, predicts two orders in each Figure. Table 5.6 shows the percentages of responses with each possible term order in each Figure, with the cells corresponding to each predicted order emboldened. It is clear that in each case as predicted, the emboldened cells do contain the highest proportions of responses, by a considerable margin. As expected, there are few responses with b final, and responses with b initial are slightly more frequent than those with b medial.

The preceding analyses show that for most of the relevant data, source premiss is a good predictor of end-term order. The exception cases are in Figure ba/bc, where the association is at the level of chance. Closer examination of individual problems in this Figure reveals that when the conditional premiss has an A quantifier, slightly more individuals appear to be founded on the terms from the A premiss, followed by the source end term, than are founded on the source premiss in the manner predicted. Table B.3 in the Appendix shows an example set of problems, two in the diagonal Figures and two in ba/bc. The first two, on the left, are paradigmatic examples of the source effect (and the Figural Effect), while the two on the right appear to invert the predictions of the source effect, for both the Individuals and Standard Task data.

We call this tendency to found conclusions on the terms of the A premiss the A-Effect, after that reported by Lee (1987), and as with source premiss, it can be treated as a variable with three levels, (only) first and (only) second, and either (covering the cases both and neither).

Table 5.7 shows the effects of having an "all' quantifier in the first and second premiss respectively, on end-term order. The effect is highly significant and in the predicted direction: when only premiss 1 has an "all" quantifier, there is a majority of ac conclusions, when only Premiss 2 does, there is a majority of ca conclusions, and there is a smaller ac majority in the remaining cases ($\chi^2(2) = 60.2$, $p < 0.0001$). Essentially the same observation has also been made by Ardin (1991).

Clearly there are cases where a source premiss is also an A premiss, as in Table 5.4, where both premisses are sources for different individuals, and as we

	A-effect			
	first	second	either	Overall
ac%	71%	34%	62%	57%
N	190	193	401	784

Table 5.7: Experiment 2: Percentages of ac conclusions for each level of **A-effect**.

have seen, despite the strong association between source and end-term order in these problems, most individuals specify the A premiss rather than the E premiss as source here, so in this case the effects operate together. In other cases, as in ba/bc problems, the effects are in opposition to one another. But these preliminary analyses cannot resolve questions about the independence or otherwise of figural, source and A-effects. To address such issues it is necessary to consider the variables together—this can be done using loglinear modelling.

Chapter 6

Individuals Task Models

6.1 Model 1: The Individuals Task Figural Effect

Loglinear modelling (see Dixon *et al* 1990) permits analyses involving multiple variables simultaneously, to establish which variables account best for the observed data, and which can be discarded. This allows a direct test of the hypothesis that the source/conditional distinction accounts for end term order effects, and whether explanations in terms of Figure are also required. However, one major *caveat* is in order here, since strictly the technique is not applicable to repeated-measures data such as the data described in this thesis. As a consequence, significance levels in these analyses may be overestimated. Unfortunately, powerful statistical modelling techniques which can take account of possibly correlated data are not yet generally available, so strict adherence to the path of statistical righteousness in this case would have made it impossible to consider the questions addressed here at all. Readers can decide for themselves whether the results are sufficiently interesting to warrant their inclusion.

The primary objective of Model 1 is to discover whether any main effect of Figure remains, after taking into account the source effect and the A-effect. Modelling can also test whether the observed breakdown of the source effect in ba/bc can be explained by the A-effect. In this case there should be no interaction between source and Figure. We also want to know whether source and A-effect interact or are independent, and finally whether the A-effect interacts with Figure.

	partial association				marginal association			
effect	*d.f.*	χ^2	*p*	*iter*	*d.f.*	χ^2	*p*	*iter*
o.	1	16.64	0.0000					
os.	2	165.80	0.0000	5	2	121.31	0.0000	1
of.	3	5.03	0.1694	12	3	21.05	0.0001	1
oa.	2	122.27	0.0000	5	2	60.46	0.0000	1
osf.	6	20.37	0.0024	12	6	83.47	0.0000	5
osa.	4	0.91	0.9238	15	4	1.25	0.8695	10
ofa.	6	15.64	0.0158	8	6	30.90	0.0000	4
osfa.	7	16.48	0.0211					

Table 6.1: Model 1: Partial and marginal associations for all terms involving **ordac**. (*Key:* o=**ordac**; s=**source**; f=**Figure**; a=**A-effect**)

The multiway frequency tabulation used the dependent variable **ordac**, with the levels ac and ca, and the independent variables **Figure** (levels ab/bc, ba/cb, ab/cb and ba/bc), **source** (levels first, second and either), and **A-effect** (levels first, second and either). Both **source** and **A-effect** collapse the "both" and "neither" categories into the single category "either", for which no overall bias is predicted. This allows the model to include all the individual conclusions, both valid and invalid, and permits a multiway tabulation with few structural zeros.

Table 6.1 shows the partial and marginal associations for all terms involving the dependent variable **ordac**[1]. Marginal associations give the significance of the effect in marginal subtables, after summing over the levels of other variables not included in the term. The marginal associations for the second-order terms **os**, **oa** and **of** thus restate the findings of the preliminary analysis. The partial association for a term of order n gives the difference in fit between the full model of order n and the model not including that term.

For the main effects, then, we are interested in partial associations, and both partial associations are highly significant for the terms **os** and **oa**, the main effects of **source** and **A-effect** respectively. The partial association for **of**, the main effect of **Figure** is not significant. We conclude that there is no main effect of **Figure**. Deleting either of the terms **os** and **oa** from the model does affect its goodness of fit, so even if **of** is included, it is still necessary to include **os** and **oa**.

[1] All interactions between independent variables are included in models automatically

source	first	second	either	A-effect	first	second	either
ac λ	0.915	-1.003	0.088	ac λ	0.577	-0.677	0.100

Table 6.2: Model 1: Loglinear parameters for **source** × **ordac** and **A-effect** × **ordac**.

These effects therefore do appear to account for the apparent Figural Effect.

However, **Figure** does enter into higher-order interactions, with **source** and with **A-effect**. Although the partial association of the term **osfa** is significant, the simplest model with no significant lack of fit with the data is **osf, ofa, sfa** ($LR\ \chi^2(11) = 17.38,\ p < 0.0971$)[2].

Loglinear parameters combine linearly to give the logarithm of the expected value in each cell, so positive values increase the expected value, and larger values increase the expected value more than do smaller values. As before, only parameters for ac are presented (ac λ), since ca parameters can be derived by negating the ac ones.

The loglinear parameters for the model **osf, ofa, sfa** broadly confirm the hypotheses. There is an overall bias toward ac conclusions (**ordac** ac λ: 0.198), and **source** strongly affects **ordac** in the predicted direction (Table 6.2), with **source**=first biasing for ac conclusions, and **source**=second biasing for ca. Notably, when **source**=either, there is negligible ac bias. Similarly **A-effect** is strong and in the predicted directions, and as with **source**, the ac bias when **A-effect**=either is small.

The term **osa** does not occur in the model, so the **source** effect is apparently independent of the A-effect. We had hypothesised that the apparent interaction between **source** and **Figure** could be accounted for in terms of the A-effect. However, even after taking account of **A-effect**, a significant interaction between **source** and **Figure** still remains (**osf**).

The *insignificant* main effect of **Figure** on **ordac** is shown in Table 6.3. The significant **osf** interaction is shown in Table 6.4. On the basis of the size of the parameters, the interaction is most pronounced in ba/bc, where there is a

[2]Since this is a hierarchical model, it includes all component lower-order terms, i.e. **o, os, of, oa, osf, ofa** and all the independent variable terms.

Figure	ab/bc	ba/cb	ab/cb	ba/bc
ac λ	0.218	-0.065	-0.120	-0.034

Table 6.3: Model 1: Loglinear parameters for the insignificant **Figure** $\times$ **ordac** effect.

ac λ	**Figure**			
source	ab/bc	ba/cb	ab/cb	ba/bc
first	0.208	0.163	0.139	-0.510
second	-0.026	0.094	-0.326	0.258
either	-0.182	-0.258	0.187	0.253

Table 6.4: Model 1: Loglinear parameters for the **source** $\times$ **Figure** $\times$ **ordac** effect.

substantial countereffect to the main effect of **source**. It appears that the A-effect has failed adequately to account for this. The next largest discrepancy is in ab/cb, where there appears to be an enhancement of the predicted **source** effect. In comparison with the overall **source** effect, parameter sizes are small.

Finally, Table 6.5 shows the interaction between **Figure** and **A-effect**. The largest deviations occur in Figures ab/cb and ba/bc, when **A-effect** is first or second, and can be summarised as a reduction in the normal A-effect in ab/cb and an increase in it in ba/bc.

Discussion In the model of end-term order, there are highly significant main effects of **source** and **A-effect**, but no main effect of **Figure**. This indicates that the apparent Figural Effect in the Individuals Task data can be accounted for in

ac λ	**Figure**			
A-effect	ab/bc	ba/cb	ab/cb	ba/bc
first	0.117	-0.180	-0.397	0.460
second	0.063	0.037	0.339	-0.439
either	-0.180	0.143	0.058	-0.021

Table 6.5: Model 1: Loglinear parameters for **A-effect** $\times$ **Figure** $\times$ **ordac**.

terms of the **source** effect and **A-Effect**. However **Figure** interacts significantly with both **source** and **A-effect**, indicating discrepancies in the symmetrical Figures.

One possible explanation for the discrepant results in the symmetrical Figures is that they are in some way connected with the position of the middle term in individual conclusions. Note that the forms of the **source** and **A-effect** × **Figure** interactions are approximate inverses of one another, and that the relevant difference between the Figures concerned is that the middle term is never subject when **source** predicts **ordac** best and is always subject when **source** predicts **ordac** worst. It was shown in Chapter 5 that the position of the middle term in conclusions is related to its position in premisses, such that it is usually first in ba/bc problems, usually second in ab/cb problems, and either first or second in the diagonal Figures. The next model examines the relation of premiss grammar to the variables already considered, in the determination of three-term order.

6.2 Model 2: Three-term order effects

The next model adopts a very different approach to modelling the effects of **source**. The **source** × **Figure** × **ordac** interaction discovered in the previous analysis suggests that the main **source** and **A-effect** are dependent on premiss grammar, since this varies appropriately with **Figure**. Consequently two new variables, **srcgram** (source premiss grammar) and **condgram** (conditional premiss grammar) are defined, to replace **Figure**. The aim of this analysis is to give an account of the ordering of all three terms in individual conclusions.

The analysis uses only valid conclusions to problems with unique source premisses (diagram types vp1, vp3, vn1, vn3, vn4, vn5 and vu2 - see Table A.1 in Appendix A.1). The multiway frequency tabulation uses the independent variables **source** (first/second), **srcgram** (source grammar: endsubj/midsubj), **condgram** (conditional premiss grammar: endsubj/midsubj) and **condq** (the quantifier of the conditional premiss: all/none) against the dependent variable **order** (order of all three terms in the individual: abc, bac, cba, bca). Owing to sparsity of data, the 9 valid conclusions with term b final are omitted.

The main overall hypothesis is that individuals will be constructed by taking

		partial association				marginal association		
effect	*d.f.*	χ^2	*p*	*iter*	*d.f.*	χ^2	*p*	*iter*
o.	3	2.13	0.5460					
os.	3	183.57	0.0000	19	3	186.46	0.0000	1
og.	3	321.16	0.0000	8	3	316.50	0.0000	1
oc.	3	5.69	0.1275	20	3	10.60	0.0141	1
oq.	3	4.30	0.2311	21	3	1.51	0.6809	1
osg.	3	48.40	0.0000	35	3	59.75	0.0000	11
osc.	3	15.92	0.0012	68	3	42.88	0.0000	8
osq.	3	40.13	0.0000	33	3	69.00	0.0000	7
ogc.	3	2.89	0.4084	42	3	1.60	0.6601	17
ogq.	3	2.20	0.5317	49	3	6.57	0.0870	17
ocq.	5	8.74	0.1198	10	5	26.60	0.0001	4
osgc.	1	0.39	0.5343	100	8	12.41	0.1340	22
osgq.	1	1.96	0.1610	100	6	7.38	0.2870	19
oscq.	1	0.09	0.7704	100	4	5.25	0.2629	25
ogcq.	0	0.00	1.0000	75	4	4.88	0.3000	7
osgcq.	0	0.02	1.0000					

Table 6.6: Model 2: Partial and marginal associations for all terms involving **order**. (*Key:* o=**order**; s=**source**; g=**srcgram**; c=**condgram**; q=**condq**)

the terms from one (founding) premiss in order, and adding the remaining end term from the other premiss last. Therefore, similar main effects of **srcgram** and **condgram** are predicted, such that midsubj biases for conclusions with the middle term first, while endsubj biases for conclusions with the middle term second. There should be a tendency for the founding premiss to be the **source** premiss, so when the **source** is the first premiss, biases towards conclusions in which a precedes c are expected, and when the **source** is the second premiss, there should be biases towards conclusions in which c precedes a. The presence of the variable **condq** allows us to take account of the A-Effect, by allowing the effect of **source** to interact with the quantifier of the conditional premiss. In view of the considerations raised at the end of the last section, interactions between the grammar variables and **source** should also be evident.

Table 6.6 shows the marginal and partial associations for this analysis. The simplest model with no significant lack of fit is **osg, osc, osq, sgcq** ($LR\ \chi^2(21) = 30.17$, $p < 0.0886$). This includes the interactions **srcgram** × **source** × **order**,

source	**order**			
	abc	*bac*	*cba*	*bca*
first	**0.751**	**0.105**	-0.698	-0.157
second	-0.751	-0.105	**0.698**	**0.157**

Table 6.7: Model 2: Loglinear parameters for the **source** × **order** effect.

condq	**order**			
	abc	*bac*	*cba*	*bca*
all	0.021	-0.042	0.173	-0.153
none	-0.021	0.042	-0.173	0.153

Table 6.8: Model 2: Loglinear parameters for the **condq** × **order** effect.

condgram × **source** × **order** and **condq** × **source** × **order**, as well as their component lower-order effects.

Table 6.7 shows the main effect of **source**. While the parameters predicted to be positive do turn out to be positive, the association is greater when the middle term occurs in medial position than when it occurs in initial position.

No prediction is made for the insignificant **condq** × **order** effect (Table 6.8), and the observed parameters are small. However, Table 6.9 shows how **condq** interacts with the effect of **source**—when the conditional quantifier is "none", the pattern of parameters resembles that predicted for the main effect of **source**, whereas when the conditional quantifier is "all", this pattern is inverted. This suppression of the predicted **source** × **order** effect when the conditional quantifier is "all" is an expression of the A-effect.

condq	**source**	**order**			
		abc	*bac*	*cba*	*bca*
all	first	-0.558	-0.562	0.717	0.403
	second	0.558	0.562	-0.717	-0.403
none	first	**0.558**	**0.562**	-0.717	-0.403
	second	-0.558	-0.562	**0.717**	**0.403**

Table 6.9: Model 2: Loglinear parameters for the **source** × **condq** × **order** effect.

srcgram	**order**			
	abc	*bac*	*cba*	*bca*
endsubj	**0.722**	-1.020	**1.130**	-0.832
midsubj	-0.722	**1.020**	-1.130	**0.832**

Table 6.10: Model 2: Loglinear parameters for the **srcgram** × **order** effect.

condgram	**order**			
	abc	*bac*	*cba*	*bca*
endsubj	**0.263**	-0.399	**0.565**	-0.429
midsubj	-0.263	**0.399**	-0.565	**0.429**

Table 6.11: Model 2: Loglinear parameters for the **condgram** × **order** effect.

Tables 6.10 and 6.11 show the main effects of **srcgram** and **condgram** respectively. In both cases, endsubj biases for abc and cba, and midsubj biases for bac and bca, as predicted. However the parameters for the **srcgram** × **order** effect are considerably larger than those for the **condgram** × **order** effect.

The presence of the **source** × **srcgram** interaction (Table 6.12) shows that **source** does not combine linearly with **srcgram** to determine term order, but rather each combination picks out one particular term order.

Table 6.13 shows the **source** × **condgram** interaction. I interpret this as showing the non-linearity of the A-effect, in combination with the grammar variables.

source	**srcgram**	**order**			
		abc	*bac*	*cba*	*bca*
first	endsubj	**1.333**	-0.576	-0.961	0.204
	midsubj	-1.333	**0.576**	0.961	-0.204
second	endsubj	-1.333	0.576	**0.961**	-0.204
	midsubj	1.333	-0.576	-0.961	**0.204**

Table 6.12: Model 2: Loglinear parameters for the **source** × **srcgram** × **order** effect.

source	**condgram**	**order**			
		abc	*bac*	*cba*	*bca*
first	endsubj	-0.511	0.489	**0.399**	-0.378
	midsubj	0.511	-0.489	-0.399	**0.378**
second	endsubj	**0.511**	-0.489	-0.399	0.378
	midsubj	-0.511	**0.489**	0.399	-0.378

Table 6.13: Model 2: Loglinear parameters for the **source** × **condgram** × **order** effect.

Discussion The three-term model can be interpreted as showing that conclusions are usually founded on one of the premisses, such that the terms from that premiss tend to occur first in order, followed by the remaining end term. There is a tendency for the founding premiss to be the source, however this effect is reduced when the conditional premiss has an 'All' quantifier.

The coverage of this model is rather restricted; the previous model covered all Individual conclusions, whereas here only those valid conclusions with a determinate source premiss were included.

Chapter 7

The A-Effect

7.1 A-effect in the Standard task

7.1.1 Model 3: The Durham Content Experiment

As with the Individuals Task data, it is possible to analyse Standard Task data by constructing loglinear models of the effects of **A-effect** and **Figure** on **ordac**, so a multiway frequency tabulation is constructed. However, owing to the unconvertible nature of many Standard Task conclusions, the effects of **source** cannot be separated from the effects of validity, so **source** is not included in the analysis. Since we are primarily interested in cases with only one A premiss, **A-effect** has the levels first, second and none, the final level covering cases with two A premisses and cases without any A premiss, included for better overall parameter estimates. Since this data set includes the manipulation of contentfulness, the variable **content** is included to increase the power of the analysis.

Table 7.1 shows the marginal and partial associations for all effects in the tabulation. The largest effect is **of**, the Figural Effect, but there are also significant partial associations for **oa** and **oaf**. The **ofc** partial association is barely significant at the 0.05 level, so two models were fit, with and without this effect. The first, **oaf, afc** has no significant lack of fit with the data, ($\chi^2(12) = 19.24,\ p = 0.083$), as does the second, **oaf, ofc, ofc** ($\chi^2(8) = 10.40,\ p = 0.2381$), so the effect **ofc** is not required. Consequently it is necessary only to consider the model **oaf, afc**.

Table 7.2 shows the loglinear parameters for the Figural Effect. This is large

	partial association				marginal association			
effect	*d.f.*	χ^2	*p*	*iter*	*d.f.*	χ^2	*p*	*iter*
o.	1	3.43	0.0640					
oa.	2	11.18	0.0037	6	2	15.69	0.0004	1
of.	3	204.26	0.0000	3	3	208.81	0.0000	1
oc.	1	0.31	0.5802	5	1	0.31	0.5786	1
oaf.	6	53.86	0.0000	7	6	53.50	0.0000	5
oac.	2	0.16	0.9239	5	2	0.01	0.9949	3
ofc.	3	8.17	0.0426	8	3	7.86	0.0489	6
oafc.	6	10.24	0.1148					

Table 7.1: Model 3: Partial and marginal associations for all terms involving **ordac**. (*Key:* o=**ordac**; a=**A-effect**; f=**Figure**; c=**content**)

Figure	ab/bc	ba/cb	ab/cb	ba/bc
ac λ	0.951	-1.065	-0.037	0.151

Table 7.2: Model 3: Loglinear parameters for **ordac** × **Figure**.

and it closely replicates the Figural effect described by Johnson-Laird *et al*: ab/bc problems strongly predispose ac conclusions, whereas ba/cb problems strongly predispose ca conclusions, and there is little overall bias in the symmetrical Figures ab/cb and ba/bc.

Table 7.3 shows the overall A-effect. This is relatively small in comparison with the Figural Effect, and it shows a tendency for the end term from the A premiss to precede the other end term in conclusions, i.e. if the A premiss is first, the conclusion tends to be ac whereas if the A premiss is second, the conclusion tends to be ca. In the 'none' condition there is little overall effect.

Table 7.4 shows the **ordac** × **A-effect** × **Figure** effect. Most of this effect is concerned with the symmetrical Figures, when there is a first or second premiss A-

A-effect	first	second	none
ac λ	0.119	-0.202	0.082

Table 7.3: Model 3: Loglinear parameters for **ordac** × **A-effect**.

ac λ	Figure			
A-effect	ab/bc	ba/cb	ab/cb	ba/bc
first	-0.016	-0.033	-0.603	0.652
second	0.021	0.110	0.489	-0.620
either/both	-0.005	-0.077	0.114	-0.032

Table 7.4: Model 3: Loglinear parameters for **ordac** $\times$ **A-effect** $\times$ **Figure**.

	partial association				marginal association			
effect	*d.f.*	χ^2	*p*	*iter*	*d.f.*	χ^2	*p*	*iter*
o.	1	54.83	0.0000					
oa.	2	12.14	0.0023	9	2	10.40	0.0055	1
of.	3	517.08	0.0000	4	3	512.40	0.0000	1
ox.	1	15.69	0.0001	7	1	9.74	0.0018	1
oaf.	6	98.53	0.0000	7	6	99.88	0.0000	6
oax.	2	2.60	0.2723	6	2	2.68	0.2616	3
ofx.	3	8.10	0.0440	10	3	11.79	0.0081	9
oafx.	6	5.90	0.4349					

Table 7.5: Model 4: Partial and marginal associations for all terms involving **ordac**. (*Key:* o=**ordac**; a=**A-effect**; f=**Figure**; x=**experiment**

effect. In short, the A-effect is suppressed or reversed in ab/cb, but it is enhanced in ba/bc. This effect is quite substantial, and considerably outweighs the overall A-effect in size.

7.1.2 Model 4: The Johnson-Laird & Bara (1984) experiments

We can make similar tabulations with the data tabulated in Johnson-Laird & Bara (1984). These tables omit 'outliers' in the form of unpredicted (by Mental Models theory) incorrect single responses to particular problems, but otherwise are complete. The data are from the 10 second condition of experiment 1, and the third experiment. The variables are **ordac**, **A-effect**, **Figure** and **experiment** (10s/exp3).

Table 7.5 shows the partial and marginal associations for all terms in the anal-

Figure	ab/bc	ba/cb	ab/cb	ba/bc
ac λ	1.103	-1.042	-0.211	0.150

Table 7.6: Model 4: Loglinear parameters for **ordac** × **Figure**.

A-effect	first	second	none
ac λ	-0.159	0.001	0.158

Table 7.7: Model 4: Loglinear parameters for **ordac** × **A-effect**.

ysis. There are significant partial associations for the main **ordac**, the interactions **ordac** × **Figure**, **ordac** × **A-effect** and **ordac** × **experiment**, and the higher-order **ordac** × **A-effect** × **Figure**. Although there is a marginally significant interaction **ordac** × **Figure** × **experiment**, this is not required for a good fit: the model **oaf, ox, afx** has no significant lack of fit ($LR\ \chi^2(11) = 17.19,\ p < 0.1024$).

The parameters for this analysis show an overall bias toward ac responses (ac:0.301,ca:-0.301), and as is well known, there is also a clear Figural Effect (see Table 7.6). The **ordac** × **xpt** effect has only one degree of freedom, and so is untabulated, but it shows that the 10s condition has a higher bias toward ac responses than the untimed experiment 3 (ac λ=0.178).

The effects involving **A-effect** are shown in Tables 7.7 (overall A-effect) and 7.8 (interaction with **Figure**). Unlike the Durham experiment, there is no clear overall A-effect; while there is a difference between conditions, it is between 'first' and 'none', with an A-premiss first tending to increase the number of ca responses (Cf the Durham experiment, where the end term from the A premiss precedes that from the other overall).

ac λ		**Figure**		
A-Effect	ab/bc	ba/cb	ab/cb	ba/bc
first	-0.045	0.073	-0.771	0.743
second	-0.032	-0.114	0.784	-0.637
none	0.078	0.041	-0.012	-0.106

Table 7.8: Model 4: Loglinear parameters for **ordac** × **A-effect** × **Figure**.

However, Table 7.8 shows a very similar pattern to that which was observed in the Durham experiment (Cf Table 7.4). As before, in Figure ba/bc the end term from the A premiss tends to precede the other, whereas in Figure ab/cb the end term from the non-A premiss tends to precede that from the A premiss. The other parameters in the table are comparatively small.

7.2 Errors and the A-Effect

The results of the preceding two analyses suggest that the largest and most important effects of A-premisses on end term order occur in the symmetrical Figures, such that in ba/bc the end term from an A premiss tends to occur before the other end term, whereas in ab/cb this effect is reversed. This effect is considerably larger than any overall A-effect, and replicates reliably between experiments. The overall A-effect does not even appear to have the same direction in the two analyses. While it might be expected that ab/bc and ba/cb would show complementary cancelling effects, thus reducing the overall A-effect, the interactions with Figure do not bear this out. The interaction is similar to that found in the end-term order model of the Individuals Task data.

The three-term model shows that as well as the predicted founding of conclusions on the source premiss, there is also a competing tendency to found conclusions on A premisses. We note that one way of discriminating the difficult NVC problems is in terms of whether they have A premisses (see the disussion of problem difficulty in the two tasks, in Section 5.2). It is possible that a single mechanism could explain both these phenomena, for example a solution strategy which is unsound in general but yields valid conclusions in many cases.

While the unsound heuristics most frequently discussed in the literature, such as the Atmosphere Effect (Woodworth & Sells 1935; Revlis 1975) and Matching (Wetherick 1989; Wetherick & Gilhooly 1990) cannot be easily applied to the case of the Individuals Task, since a major component of these concerns quantifier selection, the Conversion Hypothesis (Chapman & Chapman 1959; Revlis 1975; Newstead 1989, 1990) is straightforwardly applicable. The Individual Identification Algorithm can be adapted to make conversion errors simply by changing the characteristic representations of premisses. In graphical terms, this is achieved by using a single circle to represent both terms of the A premiss—this strategy would generate correct answers to problems with A premisses and valid conclusions, but would yield extra invalid conclusions in np1, np2, nn1 and nn2. However, the IIA Conversion Hypothesis cannot, in itself, account for the A-effect on term order, because there is no reason to expect that processing would not proceed in the usual, source-first fashion; even in a natural deduction system, since conversion changes the order of the terms of the *conditional* premiss, it is not obvious that this would lead to any term order effects different from those predicted by the source-founding hypothesis.

Another possible explanation derives from previous work on the Individuals Task dataset. Yule (1991) demonstrated that the majority of error individuals could be construed as simple unifications of the critical individuals from the two premisses. This tendency was most marked in problems with at least one A premiss. Thus with premisses "All As are Bs, Some Bs are Cs", the A premiss provides the critical individual +A+B, which is simply unified with the critical individual from the I premiss, if possible (i.e. provided the middle term features match in polarity), to give the erroneous conclusion +A+B+C. We can call this operation *Minimal Linking*, and obviously it will often return the valid conclusion. For example, provided the middle terms match, and the conditional is an A premiss, Minimal Linking will usually return the valid conclusion if one exists. In such cases, if no valid conclusion exists, the Minimal Linking conclusion will be the same as that predicted by conversion.

The Minimal Linking Hypothesis is distinguishable from the Conversion Hypothesis because it does not predict errors in nn2 problems (e.g. All B A, Some C not B), since the middle terms do not match in polarity, but in the other NVC problem types its predictions are the same. Furthermore, it can naturally accommodate the term-order effects, since the strategy does not depend on a source/conditional distinction, so all else being equal, conditional-founded individuals would be expected to be as likely as source-founded ones.

Unfortunately, nn2 differs in another respect from the rest of the problem types in which conversion theory predicts errors; we can think of the two categories as *Modus Ponens* problems (np1, np2 and nn1) and *Modus Tollens* problems (nn2). In valid *Modus Ponens* problems, the terms in the conclusion have the same polarity

Problem Type	Error Inds	Conversion Attributable	A founded	Source founded
np1 (1)	11	9	(7)	
np2 (4)	36	36	17	4
nn1 (2)	22	21	19	0
nn2 (2)	15	4	0	3

Table 7.9: Errors attributable to conversion.

as their occurrences in the A premiss, whereas in *Modus Tollens* problems, the terms in the conclusion are the negations of their occurrences in the A premiss. Minimal Linking concludes that there is no valid conclusion for *Modus Tollens* problems on the grounds that the polarity of the middle term is different in the two premisses, which is why it doesn't find the conversion conclusion in nn2.

Valid *Modus Tollens* problems are generally much harder for human subjects than *Modus Ponens* problems; the only valid cases are the vn5 problems and the E-source conclusions to the vn1 problems, which are among the hardest conclusions for human subjects to find (see Figure 5.2 in Section 5.2). Consequently, the invalid conversion conclusions in nn2 might be expected, on purely empirical grounds, to be "harder" too. Subjects who think conversion of A premisses is legitimate are hardly likely to appreciate the validity of *Modus Tollens.*

With this in mind, we can look at the relevant data. Table 7.9 shows, for each of the problems concerned, the total number of error individuals, numbers of error individuals which can be attributed to conversion (and therefore to Minimal Linking where appropriate), and finally the numbers of conversion-attributable individuals which are either A-founded or source-founded (in the usual sense that the terms from the appropriate premiss occur in order before the remaining end term).

For the first three problem types, almost all error individuals (96%) are attributable to conversion, whereas for the *Modus Tollens* problem type nn2, only 27% of error individuals are attributable to conversion. In the case of problem type np1, A-founding is indistinguishable from source-founding, but for problem types np2 and nn1, 63% of conversion-attributable problems are A-founded, whereas only 7% are source-founded. By contrast, no nn2 conversion-attributable conclusions are A-founded, and 75% are source-founded.

Despite the large difference between nn2 and the rest of the problem types in terms of the number of error individuals which are attributable to conversion, as we have seen the *Modus Ponens/Modus Tollens* distinction is a possible confound. However the fact that a majority of conversion-attributable conclusions are A-founded is inexplicable in terms of the Conversion Hypothesis, while it is compatible with the Minimal Linking Hypothesis, so on balance it would appear that the Minimal Linking hypothesis is preferable.

Supposing that there is some tendency among the subject population to use Minimal Linking can then explain why the *Modus Tollens* problems are so difficult, since they are insoluble by Minimal Linking.

One final observation is that for the core problem types in the diagonal Figures, the invalid conclusions predicted by the minimal linking hypothesis are all Figural ones, so this hypothesis predicts an Invalid Conclusions' Figural Effect. As a consequence, estimates of the overall A-effect in the Standard Task analyses may be inaccurate owing to the confounding of the Figural and A-Effects in these analyses.

Chapter 8

Conclusions

There is a need for an idealised theory of syllogistic competence in the literature; many theories make reference to an unspecified logical process to explain much of the variance, and to provide a backdrop against which to articulate an account of errors. The Atmosphere Effect (Woodworth & Sells 1935, Sells 1936, Revlis 1975), Matching (Wetherick & Gilhooly 1990) and Illicit Conversion (Chapman & Chapman 1959, Revlis 1975) all work this way, and all neglect to specify the putative logical process. If this type of appeal is to be plausible, it must be assumed that some subjects are able to produce most or all valid conclusions easily, quickly and perhaps reliably. The inherent difficulty of primitively interpreted Euler Circles or even Mental Models makes both of them unsuitable candidates, whereas the Individual Identification Algorithm is simple enough to fulfil this role.

The IIA can be viewed as an abstraction of Erickson's (1974, 1978) and Guyote & Sternberg's (1981) Euler Circles, and of Johnson-Laird *et al*'s Mental Models theories. For this reason it tends to subsume these theories—they all expect even error conclusions to be logically consistent with the premisses, for example, and the IIA can degenerate into a Mental Models scenario with multiple models causing performance decrements. Its intimate relation with a realistic construal of the use of Euler Circles is also a major strength; real subjects do not draw lots of combinations of circles, and they don't select premiss representations and combinations randomly (Ford 1994, Newstead 1989, 1990, Yule 1991).

A major aim of this thesis has been to explain how a model like the IIA can account for another well-known empirical effect, the Figural Effect (Johnson-Laird

& Bara 1984). Chapter 3 showed that the Individual Identification Algorithm can be implemented in a variety of superficially different ways, but that in each case a functional distinction can be drawn between source and conditional premisses. The distinction relates to the role played by each premiss in inference: the source asserts the existence of a critical individual, which can be considered the subject of the syllogism, whereas the conditional provides the means to infer something about that critical individual—to predicate something of it.

The source/conditional distinction is closely related to term order effects in the Individuals Task data, in such a way that terms from source premisses tend to precede those from conditional premisses in conclusions. Moreover, this can explain the apparent effect of Figure on end-term order, via the imbalance in numbers of problems with each source/conditional assignment in the diagonal Figures. Model 1 demonstrated that, in conjunction with the A-effect, the source/conditional distinction can account for the figural variance in end-term order. Model 2 showed that three-term order of conclusions can be explained as the result of taking the terms from one (founding) premiss in order, and adding the remaining term last. This founding premiss tends to be the source premiss, but this tendency is reduced when the conditional quantifier is "all".

The source-founding hypothesis not only accounts for gross figural end-term order effects, but in three out of four Figures, predicts term order accurately on a problem-by-problem basis. In ab/cb, although there is indeed an approximately 50/50 split between ac and ca responses in the Individuals Task data, the source-founding hypothesis predicts which problems will produce ac responses and which will produce ca responses. Neither Mental Models theory nor Wetherick's grammatical formulation (Wetherick 1989, Wetherick & Gilhooly 1990), which as we have seen make equivalent predictions, can explain these strong biases.

The A-effect reflects a tendency in some circumstances to found conclusions on conditional premisses. This seems to be most common in ba/bc problems, and evidence for the effect can be found in Standard Task datasets as well as the Individuals Task dataset.

Moreover, comparison of error rates in Standard Task and Individuals Task data suggests that the main cause of difficulty in the Standard Task is selection of an appropriate quantifier for the conclusion, and that this may account for the apparent effect of the number of Mental Models required. This result concurs with the conclusions of Ford (1994) and Wetherick & Gilhooly (1990). While there is a difference in difficulty between single- and multiple-models problems in the Individuals Task data, it is possible that this is attributable to an unsound solution strategy called Minimal Linking, which is hypothesised to occur, along with a logically sound strategy, in the data at large. Minimal Linking can in turn explain the A-effect.

As it stands, the Minimal Linking Hypothesis requires further investigation. The strategy is similar to Matching (Wetherick 1989), and the latter has been found to be attributable to a subgroup of subjects. It might be expected that Minimal Linking would follow a similar pattern, and this could be assessed by attempting to identify Linking subjects, for example by selecting those subjects whose Linking scores on the identifying set of NVC problems is highest. Linking subjects should score badly on *Modus Tollens* problems such as vn5 problems, but quite well on *Modus Ponens* problems, whereas Logical subjects should score better on vn5 problems. Such a partitioning of the subject group would allow the fitting of different models to the Logical and Linking subjects; the hypothesis would be that Logical subjects would not show evidence of an A-effect, whereas Linking subjects should show less evidence of a source effect.

While the success in predicting end term order by the source/conditional distinction might appear to lend confirmation to the Individual Identification Algorithm, it should be stressed that the distinction appears to be more general than this. Its relation to the much more universal logical distinction between existential, assertive statements and conditional, inference-licensing statements tends to undercut the suggestion that subjects really use the IIA. It is quite possible that even the Logical subjects use a variety of different reasoning systems; the depth and universality of the source/conditional distinction could result in an overall source effect and hence a Figural Effect nevertheless.

On the other hand, provided that sources of error can be accounted for by things like the Linking strategy, the lack of evidence for a single/multiple models distinction may license a stronger interpretation of the involvement of the IIA. It is, after all, a one-model theory in Mental Models terms, and so it might appear to make a prediction of no differences in difficulty, all else being equal. However, as Section 3.3.1 showed, the situation is not really as simple as this; even a strongly-held IIA hypothesis need not predict uniform performance—a version of the IIA which permited non-maximal registration would behave very like the Mental Models algorithm, since it would then be necessary to consider multiple distinct combined representations of premisses, with a consequently heavier memory load. Indeed, the number of individuals in a *single* registration diagram might cause differential memory load. It would be necessary to take account of Minimal Linking before such questions could be addressed for putative "Logical" subjects alone.

While this study has concentrated primarily on explanation of the results of the Individuals Task, it is possible to extrapolate to the case of the Standard Task. Thus the preference for Figural convertible conclusions is straightforwardly explicable by the source/conditional distinction, and this can also explain the facilitation of Figural unconvertible unique-source problems. Unfortunately it is very hard directly to test for source/conditional effects in the Standard Task, owing to the confound with validity effects, and most valid Standard Task conclusions are unconvertible. When they *are* convertible, it can be problematic to assign source and conditional premisses — problems with two valid individuals (vn1 problems) have two "none" conclusions, and since these are validly convertible, in principle each could be derived from either valid individual.

In the Individuals Task, syntax and semantics of conclusions are independent, and assignment of source and conditional premisses to valid individual conclusions depends purely on the semantics of the individual description, not the order of mention of terms, so it is possible to test for the effect of source/conditional assignment on term order without assuming what was to be proven. The methodological assumption that *is* made is that if a conclusion is valid, it is meaningful to assign source and conditional premisses to it. This would not be true if the conclusion was actually reached by means of a reasoning process that did not distinguish source and conditional premisses, such as Mental Models (Johnson-Laird 1983). The procedure of source/conditional assignment is justified only to the extent that it works empirically, to predict term order.

For valid individual conclusions, notwithstanding these reservations, it is fairly straightforward to assign source and conditional premisses. However, if the distinction was to be applied to invalid conclusions, it would be necessary to amend the definitions somewhat; strictly speaking, source is defined only for valid conclusions. Unfortunately, there is no satisfactory principled way of assigning source and conditional premisses for invalid conclusions, since allowing for both non-maximal registration and premiss misrepresentation renders almost all such assignments indeterminate. It would be necessary to restrict the range of possible sources of error, for example to permit premiss misrepresentation but assume perfect combination of premisses, before a suitable extended definition would be useful. Because this would introduce empirical commitments into the definitions, thereby robbing them of their *a priori* character, I am reluctant to do this.

As we have seen, however, the statistical model of end-term order effects accounted for the *total* (valid and invalid) Figural Effect. This was achieved by pooling all invalid conclusions, as well as valid conclusions with indeterminate source, into a single category ("either") for which no prediction was made on the basis of source/conditional assignment. This allowed the variable **A-effect** to range over the largest possible dataset, since this variable is well-defined for both valid and invalid conclusions. On the basis of that analysis, it appears that the A-effect must have been sufficient to account for the Figural Effect in invalid conclusions.

Indeed, term ordering in invalid Linking conclusions is more strongly biased towards A-founding than would be expected on the basis of the Minimal Linking hypothesis given in Chapter 5.7, which simply predicts an approximately equal number of source-founded and conditional-founded conclusions. Empirically, it seems that Linking subjects prefer to start with the A premiss. Ardin (1991) has reported evidence from a syllogism recall study that an A premiss will tend to be recalled as the first premiss in the syllogism, irrespective of its original position, and argued that subjects appeared to prefer to start processing with a positive or universal premiss. The Individuals Task term ordering data was explained as the consequence of this—positive terms tended to appear first, followed by negative ones. Unfortunately Ardin's account falls short of explaining Figural Effect phenomena in syllogisms with two positive premisses, which are usually given as paradigms of the effect.

While in the present study the A-effect was an unexpected discovery of the model building process, characterised primarily in terms of performance deviations from the predictions of the source/conditional hypothesis, it has appeared before in the literature on syllogistic reasoning. Johnson-Laird & Bara (1984) discuss a similarly unpredicted phenomenon concerning the ab/cb and ba/bc Figures, writing

> "where the conclusion was in the same mood as just one of the premises, the end term of the premise tended to play the same grammatical role in the conclusion as it did in the premise itself" (p22)

We can call this the *same-mood/same-position effect.*

If we restrict ourselves, for now, to considering only valid conclusions, it should be clear that the premiss with the same mood as the conclusion must be the source premiss, and the conditional premiss must have an "all" quantifier, or else the mood of the conclusion would differ from that of the source. So the effect translates into the claim that in ab/cb, the subject (i.e. the end term) from the source premiss becomes the subject of the conclusion, as would be expected on the basis of the source-founding hypothesis, whereas in ba/bc, the predicate of the source (again the end term) becomes the predicate of the conclusion, instead of becoming the subject of the conclusion as would be predicted by source-founding. Thus the end term from the conditional becomes the subject of the conclusion in ba/bc. This is exactly what was found in the present study—in the Individuals Task and both Standard Task datasets, there was an interaction between A-effect, Figure and end term order such that an end term from an A premiss was more likely to precede the other end term in the conclusion in ba/bc, but less likely to do so in ab/cb.

Both Lee (1987) and Ford (1994) argue that extending the same-mood/same-position effect to the diagonal Figures would generate the Figural Effect, so perhaps the symmetrical- and diagonal-Figure effects have a common cause. Again, this leads to predictions which parallel those of source-founding to a large extent, for the same reasons. The same-mood/same-position account has the strength that it correctly predicts end term order in conclusions in all four Figures, unlike the source-founding account which needs to be supplemented by the A-effect. However, source-founding also covers cases where the mood of the source is not the same as the mood of the conclusion (*different-mood conclusions*), but in the

Some B A	+B+A	Some A not B	+A-B
All B C	+B+C	All C B	-C-B
	(+B+C)+A		+A(-C-B)

Figure 8.1: Individual Identification by "substitution".

Standard Task, as we have seen, subjects typically find it difficult to draw different-mood conclusions, so such cases would be expected to have less effect on the overall statistics. In any event, in the Standard Task, valid different-mood conclusions usually have the "some... not" quantifier, so they are not validly convertible anyway, and their end-term order can be attributed to the "validity effect".

Ford's (1994) examination of subjects' protocols revealed that both verbal and spatial reasoners showed the same-mood/same-position effect in the diagonal Figures, although only verbal reasoners showed it in the symmetric Figures. Ford explains the 'verbal' result as the consequence of substitution of conditional premiss terms for the middle term in the source premiss, to generate the same-mood/same-position effect. In ab/cb, substituting into the source keeps the source end term first, while in ba/bc, substituting into the source leaves the source end term in last place. Reliance on substitution in this way also explains the observed difficulty of drawing different-mood conclusions.

While substitution accounts quite well for end-term order results in the Standard Task, it is not straightforwardly applicable to the Individuals Task, since the conclusion is never derivable from either of the premisses by substitution, because there is no cancellation of the middle term. But it is possible to construct a formal analogue which parallels it reasonably closely—the idea is to perform the substitution with individuals, rather than premisses. Simply replace the middle term in the source premiss' critical individual with both terms from the appropriate individual from the conditional premiss' maximal model, in their original order. Note that the logic of the process is just the same as the usual IIA, using the unification test, but the syntax differs. In the examples in Figure 8.1, the conditional premiss individual displayed is the only one which is unifiable with the source premiss critical individual.

Although this approach might appear to account for the observed term order results in ba/bc, since it generates the A-effect in that Figure, it also generates only abc and cba orders in the diagonal Figures, and cab and acb orders in ab/cb. The orders predicted for the Diagonal Figures are commoner than would be expected on the source-founding hypothesis alone, but there are also substantial numbers of conclusions with the middle term first in these Figures, and these are predicted only by the source-founding hypothesis. Conclusions with the middle term last, as predicted by Substitution for ab/cb, are very rare. Moreover, as we have seen, the source-founding hypothesis has problems only when the conditional quantifier is "all"; when the conditional quantifier is "none", the source-founding hypothesis predicts term order quite accurately. So as things stand, Substitution does not fare well as an explanation of Individuals Task term order.

Ford's explanation for the same-mood/same-position effect, as the consequence of substitution, is applicable only to the 'verbal' subjects. Her 'spatial' subjects also showed evidence of a same-mood/same-position effect in diagonal-Figure problems, but not in symmetric-Figure problems. This may be explicable as the consequence of source-founding; unfortunately the published data are insufficient to determine whether this is the case. It would be interesting to conduct a replication of the Individuals Task experiment using Ford's "thinking aloud" methodology.

The middle term position results raise problems not only for substitution, but also, of course, for the Mental Models account of the Figural Effect. It is hard to see how Mental Models theory could predict anything other than abc and cba orders in the Individuals Task. The reason for this is that reordering of terms is crucial to the integration of information from the premisses in the Mental Models account—the middle terms must be brought together to construct the mental model. The present account, in determining term order largely on the basis of surface grammar, makes it clear that term order is not crucial to reasoning; rather the order of terms in conclusions appears to be no more than an accidental trace left by a reasoning process which is fundamentally indifferent to term and premiss order, being driven by logical properties of the premisses instead.

It might be argued that a FIFO store underlies the preservation of order from founding premisses, and that the source premiss terms are first out because they are first in. However this seems spurious; the relevant orders are already specified in the materials, and postulating a mental process to simply maintain an already existent external order multiplies entities unnecessarily. Subjects may report terms in a specific order because they read them off from the premisses when they *report* a conclusion. This is not to say that FIFO stores don't exist, simply that this type of task tells us nothing about them.

As far as it goes, the memory evidence from the Levy (1984) experiment is at least as compatible with the source-founding account as it is with Mental Models. Single premiss presentation caused sufficient memory load that performance was correlated with memory span in ba/cb alone, suggesting that in this Figure it was necessary to maintain the first premiss in memory, so that processing of the problem could be begun with the second premiss. Certainly this is compatible with the Mental Models account, but it is also what would be expected if it was necessary to start processing with the source premiss. The theories predict different results on the single-premiss task in the symmetric Figures: Mental Models would expect a preponderance of ac conclusions, but no memory load problems, whereas source founding would expect no preponderance of ac conclusions but a small memory load problem. The evidence is that there was no increase in the number of ac conclusions, but also no evidence for a memory load problem. The sensitivity of this experiment was lamentably low, owing to the small sample sizes and between-subjects design, so the evidence cannot conclusively distinguish between the hypotheses.

It would be very easy to adapt Mental Models theory to account for the results of the present study; to do so would simply make it completely isomorphic with the IIA, and consequently make it lose most of its distinctive empirical consequences. Indeed, the fatal weakness of Mental Models theory would appear to be that its empirical consequences almost all derive from assumptions about the way the theory is implemented, whereas in the present study the main thrust is to show that robust effects such as the Figural Effect are likely to occur in many implementationally distinct systems, for logical reasons. So while we might conclude that the source/conditional distinction constrains the class of algorithms which might underlie human performance on these tasks, as we have seen from Chapter 3 the form of any intermediate representations (mental imagery, Mental Models or sentential representations) remains completely unconstrained. As a consequence, tasks which merely relate premisses to conclusions, such as the ones described here and in most of the relevant literature, will not provide much information about the implementation of reasoning.

The importance of this point lies in the fact that where the appropriate information has been gathered, as it has been by Ford (1994), there is ample evidence for the diversity of individual phenomenology, and hence arguably a diversity of implementations, in the syllogistic reasoning task. That such a range of reasoning styles can realistically be expected in any experimental sample is evidence that any robust psychological effects such as the Figural Effect must be explained by relatively high-order invariants of human reasoning. The ins and outs of memory stores are likely to vary widely with representational formats, especially in view of Baddeley's (1986) contention that separate systems underlie spatial/imagistic and linguistic processing modes, so it seems unlikely *a priori* that they would accidentally coincide to produce any highly robust overall effect. The advantage of the source-founding account of term order effects is that it explains why phenomenologically different reasoning methods can give rise to the same results. It also indicates that further investigation requires much richer data than premiss-to-conclusion mappings—data from phenomenology and learning, and analysis in terms of individual differences, are required to resolve implementational issues.

References

Ardin, C.L. (1991). *Mental Representations Underlying Syllogistic Reasoning.* PhD Thesis, University of Edinburgh.

Aristotle (336 BCE/1983). Prior Analytics. Tredennik H. (Trans.) Harvard University Press, Cambridge Massachusetts.

Baddeley, A. (1986). *Working Memory.* Oxford: Clarendon Press.

Begg, I. & Denny, J.P. (1969). Empirical reconciliation of atmospheric and conversion interpretations of syllogistic reasoning errors. *Journal of Experimental Psychology, 81,* 351-354.

Ceraso, J. & Provitera, A. (1971). Sources of Error in Syllogistic Reasoning. *Cognitive Psychology, 2,* 400-410.

Chapman, L.J & Chapman, J.P. (1959). Atmosphere Effect Re-examined. *Journal of Experimental Psychology, 38,* 220-226.

Clocksin, W.F. & Mellish, C.S. (1984). *Programming in Prolog.* (2nd edition) Berlin: Springer-Verlag.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31* 187-276.

Dixon, W.J., Brown, M.B., Engelman, L. & Jennrich, R.I. (1990). *BMDP Statistical Software Manual, Vol 1.* Berkeley: University of California Press.

Erickson, J.R. (1974). A set analysis theory of behaviour in formal syllogistic reasoning tasks. In R.Solso (Ed.) *Theories in cognitive psychology: the Loyola Symposium.* Potomac: LEA.

Erickson, J.R. (1978). Research on syllogistic reasoning. In R. Revlin & R.E. Mayer (Eds.), *Human Reasoning.* Washington, D.C.: Winston.

Euler, L. (1772). Lettres a une princesse d'Allemagne. *2* sur divers sujets de physique et de philosophie. Letters 102–108.

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

Ford, M. (1994). Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition, 54,* 1-71.

Gilhooly, K.J., Logie, R.H., Wetherick, N.E. & Wynn, V. (1993). Working memory and strategies in syllogistic reasoning tasks. *Memory and Cognition, 21:1,* 115-124.

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

Guyote, M.J. & Sternberg, R.J. (1981). A transitive-chain theory of syllogistic reasoning. *Cognitive Psychology, 13,* 461-525.

Henle, M. (1962). On the relation between logic and thinking. *Psychological Review, 4,* 366-378.

Hunter, I.M.L. (1957). The solving of three term series problems. *British Journal of Psychology, 48,* 286-98.

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

Johnson-Laird, P.N. (1983). *Mental Models.* Cambridge: CUP.

Johnson-Laird, P.N. & Bara, B. (1984). Syllogistic Inference. *Cognition, 16,* 1-61.

Johnson-Laird, P.N. & Byrne, R. (1991). *Deduction.* Hove: LEA.

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

Lee, J. (1987). *Metalogic and the Psychology of Reasoning.* PhD Thesis, University of Edinburgh.

Leibniz, G.W. (1666/1966). Dissertation on the Combinatorial Art. In G.H.R Parkinson (ed. and trans.), *Logical papers: A selection.* Oxford: Clarendon Press.

Lemmon, E.J. (1965) *Beginning Logic.* Walton-on-Thames: Nelson.

Levy, J. (1984). *The Effect of Presenting Syllogisms One Premise at a Time.* Undergraduate Thesis, Cambridge University.

Mates, B. (1972). *Elementary Logic.* New York: Oxford University Press.

Newstead, S.E. (1989). Interpretational errors in syllogistic reasoning. *Journal of Memory and Language, 28,* 78-91.

Newstead, S.E. (1990). Conversion in Syllogistic Reasoning. In K.J. Gilhooly, M.T.G. Keane, R.H. Logie & G. Erdos (Eds.) *Lines of Thinking: Reflections on the Psychology of Thought, Vol. 1* Wiley: Chichester.

Newstead, S.E., Pollard, P., Evans, J.St.B.T. & Allen, J.L. (1992) The source of belief bias effects in syllogistic reasoning. *Cognition, 45,* 257-284.

Oakhill, J., Johnson-Laird, P.N., & Garnham, A. (1989). Believability and syllogistic reasoning. *Cognition, 31,* 117-140.

Read, S. (1988) *Relevant Logic: a philosophical examination of inference.* Oxford: Blackwell.

Revlis, R. (1975). Two Models of Syllogistic Reasoning: Feature Selection and Conversion. *Journal of Verbal Learning and Verbal Behaviour, 14,* 180-195.

Shieber, S.M. (1986). *An Introduction to Unification-Based Approaches to Grammar.* Stanford: CSLI.

Stenning, K. & Cox, R. (1995). Attitudes to logical independence: traits in quantifier interpretation. Proceedings of the 17th Annual Conference of the Cognitive Science Society of America. Pittsburgh, Erlbaum.

Stenning, K., Cox, R. & Oberlander, J. (1995). Contrasting the cognitive effects of graphical and sentential logic teaching: reasoning, representation and individual differences. *Language and Cognitive Processes, 10,* 333-354.

Stenning, K. & Oberlander, J. (1994). Spatial inclusion as an analogy for set membership: a case study of analogy at work. In K. Holyoak & J. Barnden (Eds.) *Analogical Connections.* Erlbaum.

Stenning, K. & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: logic and implementation. *Cognitive Science, 19,* 97-140.

Stenning, K. & Yule, P. (1996). Logic and Implementation in Human Reasoning: A Syllogistic Illustration. HCRC Research Paper RP-75, Human Communication Research Centre, University of Edinburgh.

Wason, P.C. (1966). Reasoning. In B.M. Foss (Ed.), *New Horizons in Psychology.* Harmondsworth: Penguin.

Wetherick, N.E. (1989). Psychology and syllogistic reasoning. *Philosophical Psychology, 2,* 111-124.

Wetherick, N.E. & Gilhooly, K.J. (1990). Syllogistic Reasoning: Effects of Premise Order. In K.J. Gilhooly, M.T.G. Keane, R.H. Logie & G. Erdos (Eds.) *Lines of Thinking: Reflections on the Psychology of Thought, Vol. 1* Wiley: Chichester.

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus.* Routledge & Kegan Paul.

Woodworth, R.S. & Sells, S.B. (1935). An atmosphere effect in formal syllogistic reasoning. *Journal of Experimental Psychology, 18,* 451-460.

Yule, P. (1991). *An experimental investigation of a new theory of syllogistic reasoning based on Euler Circles.* MSc Thesis, University of Edinburgh.

Yule, P. & Stenning, K. (1992). The Figural Effect and a Graphical Algorithm for Syllogistic Reasoning. Proceedings of the 14th Conference of the Cognitive Science Society. Bloomington, Indiana.

Appendix A

Implementation Details

A.1 The Graphical Method

Table A.1 shows the full set of topologically distinct registration diagrams for the Graphical Individual Identification Algorithm (see Section 3.3.1). These diagrams are used as the basis for the definition of a set of equivalence classes of syllogisms (*problem types*), which abstract over premiss order and the valid conversion of I and E premisses, and are referenced in the text by the labels given to the diagrams (vp1, vp2 etc).

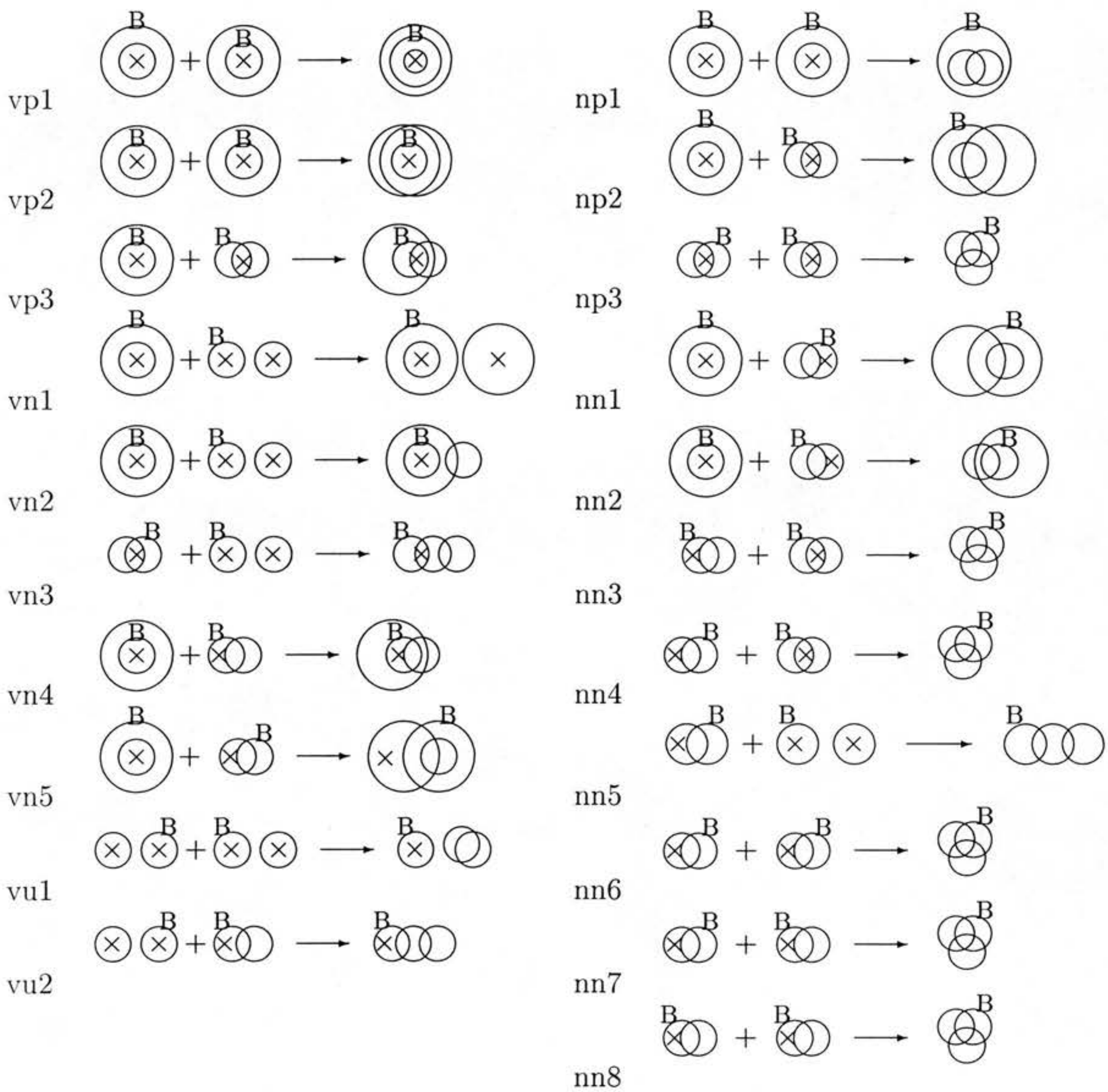


Figure A.1: The full set of topologically distinct registration diagrams. Those on the left have valid individuals; those on the right do not. The unlabelled circles can represent either A or C, depending on the premisses.

A.2 The prolog program

This section presents a complete `prolog` syllogism solution program, based on the Individual Identification code described in Chapter 3. The top-level predicates are `standard/6` and `individual/6`, which produce quantified and individual conclusions respectively. Thus a goal like

```
?- standard(all(a,b),all(b,c),Terms,Conc,Syllno,Source).
```

will instantiate `Conc` as each valid quantified conclusion for the premisses in turn. Replacing `standard` with `individual` in this goal will produce valid individual conclusions as before.

The program does full syntax checking for premisses, using the predicate `gensyll/4`, and as a side-effect this ensures that premisses become instantiated—this means that the program can systematically produce all Standard conclusions to all problems in one run. To do this, use a top-level goal like

```
?- standard(Premiss1,Premiss2,[a,b,c],Conc,Syllno,Source).
```

Note that if premisses are to be left uninstantiated like this, the terms of the problem must be specified. The main function of `gensyll/4` is to produce codings for data scoring (represented in compound form by `Syllno`); however, the quantified conclusions code depends on `gensyll/4` to identify the middle term if terms are not specified.

`standard/6` acts just like `individual/6` up to the `unification_test/3` goal, where instead of returning the necessary individual as the conclusion, it proceeds to convert it into a valid quantified conclusion. First, the middle term feature is removed from `NecInd` to give `EndTerms`. Then `getconc/3` selects a positive end term feature for the subject of the conclusion, and performs another unification test with the maximal model of the source premiss, to see if a universal conclusion is warranted. Failure here forces backtracking to produce a particular conclusion.

The program produces all valid conclusions for a given pair of premisses, and if possible their end-term order preserves the term order in the individual due to source-founding. However, backtracking will also produce the "counterfigural" conclusions, since the program is a complete decision procedure for the syllogism.

```
/***********************************************/
/* Syllogism Solver                            */
/* P Yule 1995                                 */
/* Quantified and Individual conclusions       */
/***********************************************/

/***********************************************/
/* Premiss representations                     */

characteristic(all(A,B),[[-A,-B],[-A,+B],[+A,+B]]-[[+A,+B]]).
characteristic(some(A,B),[[-A,-B],[-A,+B],[+A,-B],[+A,+B]]-[[+A,+B]]).
characteristic(none(A,B),[[-A,-B],[-A,+B],[+A,-B]]-[[+A,-B],[-A,+B]]).
characteristic(somenot(A,B),[[-A,-B],[-A,+B],[+A,-B],[+A,+B]]-[[+A,-B]]).

/***********************************************/
/* The theorem prover                          */

standard(Premiss1,Premiss2,[A,B,C],Conc,Syllno,Source) :-
        gensyll(Premiss1,Premiss2,[A,B,C],Syllno),
        characteristic(Premiss1,Max1-Min1),
        characteristic(Premiss2,Max2-Min2),
        nth(Source,[Min1,Min2],Min),
        nth(Source,[Max2,Max1],Max),
        member(I,Min),
        unification_test(I,Max,NecInd),
        remove_mid(B,NecInd,EndTerms),
        nth(Source,[Max1,Max2],OwnMax),
        getconc(EndTerms,OwnMax,Conc).

individual(Premiss1,Premiss2,Conc,Syllno,Source) :-
        gensyll(Premiss1,Premiss2,Terms,Syllno),
        characteristic(Premiss1,Max1-Min1),
        characteristic(Premiss2,Max2-Min2),
        nth(Source,[Min1,Min2],Min),
        nth(Source,[Max2,Max1],Max),
        member(I,Min),
        unification_test(I,Max,Conc).

/***********************************************/
/* Code for quantified conclusions             */

remove_mid(B,NecInd,EndTerms) :-
        select(+B,NecInd,EndTerms).
remove_mid(B,NecInd,EndTerms) :-
        select(-B,NecInd,EndTerms).
```

```
getconc(EndTerms,OwnMax,Conc) :-                    /* Universal */
        select(+Subj,EndTerms,[Fpred]),
        unification_test([+Subj],OwnMax,Ind),
        quantifier(Conc,[Subj,Pred],u,Fpred,_).
getconc(EndTerms,OwnMax,Conc) :-                    /* Particular */
        select(+Subj,EndTerms,[Fpred]),
        quantifier(Conc,[Subj,Pred],p,Fpred,_).

/************************************************/
/* Syllogism generation, syntax checking        */

gensyll(Premiss1,Premiss2,Terms,Syllno-[N1,N2,F]) :-
        Figure(F,Terms1,Terms2,Terms),
        quantifier(Premiss2,Terms2,_,_,N2),
        quantifier(Premiss1,Terms1,_,_,N1),
        Syllno is 16*(F-1)+4*N2+N1+1.

quantifier(all(A,B),[A,B],u,+B,0).
quantifier(some(A,B),[A,B],p,+B,1).
quantifier(none(A,B),[A,B],u,-B,2).
quantifier(somenot(A,B),[A,B],p,-B,3).

Figure(1,[A,B],[B,C],[A,B,C]).
Figure(2,[B,A],[C,B],[A,B,C]).
Figure(3,[A,B],[C,B],[A,B,C]).
Figure(4,[B,A],[B,C],[A,B,C]).

/************************************************/
/* Unification test code                        */

unification_test(I,Max,K) :-
        bagof(Unified,J^(member(J,Max),unify(I,J,Unified)),[K]).

unify([],X,X).
unify([H|T],X,[H|NT]) :-
        select(H,X,Y),
        !,
        unify(T,Y,NT).
unify([H|T],X,[H|NT]) :-
        opp(H,Opp),
        not member(Opp,X),
        unify(T,X,NT).

opp(+A,-A).
opp(-A,+A).
```

```
/************************************************/
/* standard utilities                            */

select(X,[X|Y],Y).
select(Z,[X|Y],[X|W]) :-
        select(Z,Y,W).

member(X,[X|_]).
member(X,[_|Y]) :-
        member(X,Y).

append([],A,A).
append([A|B],C,[A|D]) :-
        append(B,C,D).

nth(1,[A|B],A).
nth(X,[A|B],C) :-
        nth(D,B,C),
        X is D+1.
```

Appendix B

Raw data tables

In the following tables, unique source premisses and conclusions founded on them are indicated by **bold**. Other valid conclusions are in plain text, while invalid conclusions are indicated by *italics*. Data are from the Individuals Task (N=22) and Durham Standard Task (N=24) experiments. Tables are organised around the equivalence classes defined by the set of distinct registration diagrams (Figure A.1).

B.1 Problems with Valid Conclusions

All **A B**	
All B C	
ABC (21)	
ACB (1)	
NVC (0)	
all A C (21)	some C A (1)
	all C A (1)
NVC (1)	

All B A	
All C B	
ABC (1)	**CBA (9)**
ACB (1)	BCA (1)
BAC (8)	
NVC (1)	
all A C (2)	**all C A (16)**
NVC (2)	

Table B.1: Conclusions in both Tasks for vp1 problems.

All B A	
All B C	
BAC (15)	BCA (5)
ACB (1)	
NVC (1)	
some A C (4)	*all C A (1)*
all A C (9)	
NVC (3)	

Table B.2: Conclusions in both Tasks for vp2 problems.

All B A	
Some C B	
BAC (5)	**CBA (13)**
ACB (1)	BCA (1)
NVC (2)	
some A C (3)	**some C A (17)**
NVC (1)	

All B A	
Some B C	
BAC (9)	**BCA (8)**
ABC (3)	
AB-C (1)	
NVC (1)	
some A C (15)	**some C A (5)**
NVC (4)	

Some A B	
All B C	
ABC (17)	BCA (4)
NVC (1)	
some A C (18)	some C A (2)
NVC (1)	

Some B A	
All B C	
BAC (9)	BCA (10)
	CAB (1)
	CBA (1)
NVC (1)	
some A C (6)	some C A (16)
all A C (1)	
NVC (0)	

Table B.3: Conclusions in both Tasks for vp3 problems.

All A B **None C B**	
AB-C (12)	-CBA (1)
	C-B-A (4)
	C-A-B (2)
AC-B (1)	
NVC (2)	
none A C (5)	none C A (10)
NVC (9)	

All A B **None B C**	
AB-C (19)	
BA-C (1)	B-CA (1)
NVC (0)	
none A C (16)	none C A (4)
NVC (3)	

None A B **All C B**	
A-B-C (3)	
-ABC (1)	**CB-A (12)**
B-AC (1)	BC-A (2)
NVC (2)	
none A C (15)	none C A (2)
NVC (7)	

None B A **All C B**	
-BA-C (1)	
	CB-A (17)
	BC-A (1)
	CBA (3)
NVC (0)	
none A C (2)	none C A (19)
NVC (2)	

Table B.4: Conclusions in both Tasks for vn1 problems. Each premiss is source for a different individual.

None A B All B C	
-ABC (2)	BC-A (11)
B-AC (2)	CB-A (1)
CA-B (1)	*A-B-C (1)*
NVC (3)	
none A C (6)	*none C A (3)*
some A C (1)	
NVC (13)	

None B A All B C	
B-AC (1)	BC-A (17)
	CB-A (2)
-BAC (1)	
NVC (1)	
none A C (6)	somenot C A (1)
	none C A (7)
NVC (10)	

All B A None C B	
BA-C (12)	-CBA (1)
AB-C (1)	*C-A-B (2)*
	C-B-A (1)
NVC (5)	
none A C (2)	*none C A (9)*
NVC (11)	

All B A None B C	
BA-C (18)	
AB-C (2)	
AC-B (1)	*BC-A (1)*
NVC (0)	
somenot A C (1)	*none C A (1)*
none A C (13)	
NVC (6)	

Table B.5: Conclusions in both Tasks for vn2 problems.

None A B
Some B C

-ABC (3)	**BC-A (12)**
	CB-A (2)
	CA-B (1)
	BCA (1)
NVC (3)	

none A C (5)	somenot C A (1)
NVC (18)	

Some A B
None B C

AB-C (17)	
AC-B (1)	*C-B-A*
NVC (3)	

somenot A C (5)	*none C A (1)*
none A C (2)	
NVC (16)	

None B A
Some C B

	CB-A (15)
	-CB-A (1)
B-A-C (1)	*CBA (1)*
NVC (2)	

none A C (1)	somenot C A (10)
	none C A (2)
NVC (9)	

Some B A
None C B

BA-C (12)	
AB-C (1)	
AC-B (1)	
NVC (8)	

somenot A C (1)	*none C A (3)*
none A C (1)	
NVC (19)	

None A B
Some C B

-ABC (1)	**CB-A (14)**
-AB-C (1)	*-CB-A (1)*
NVC (3)	

none A C (5)	somenot C A (3)
	none C A (2)
NVC (14)	

Some A B*
None C B

AB-C (7)	
NVC (4)	

somenot A C (2)	*none C A (3)*
none A C (2)	
NVC (17)	

None B A
Some B C

B-AC (2)	**BC-A (13)**
BAC (1)	
-BA-C (1)	*-CB-A (1)*
A-B-C (1)	
NVC (2)	

none A C (4)	somenot C A (4)
	none C A (2)
NVC (13)	

Some B A
None B C

BA-C (13)	B-CA (3)
AB-C (3)	
NVC (3)	

somenot A C (2)	*some C A (1)*
none A C (1)	*none C A (1)*
NVC (16)	

Table B.6: Conclusions in both Tasks for vn3 problems. 11 Individuals Task subjects did not receive the problem marked * owing to an error in the materials set.

All B A **Somenot B C**	
BA-C (14)	**B-CA (2)**
AB-C (2)	
	CBA (1)
NVC (3)	
---	---
somenot A C (16)	
some A C (1)	
NVC (6)	

Somenot B A All B C	
B-AC (3)	BC-A (15)
	CB-A (1)
BAC (1)	
NVC (2)	
---	---
	somenot C A (16)
	some C A (1)
NVC (6)	

Table B.7: Conclusions in both Tasks for vn4 problems.

All A B **Somenot C B**	
-A-BC (1)	**C-B-A (9)**
	C-A-B (1)
ACB (1)	
AB-C (1)	
NVC (8)	
---	---
somenot A C (1)	somenot C A (7)
some A C (1)	*some C A (2)*
	none C A (1)
NVC (12)	

Somenot A B All C B	
A-B-C (5)	-C-BA (2)
ABC (1)	*CB-A (2)*
B-AC (1)	*BC-A (1)*
-A-C-B (1)	*-C-B-A (1)*
NVC (7)	
---	---
somenot A C (4)	*somenot C A (2)*
some A C (1)	*some C A (2)*
NVC (13)	

Table B.8: Conclusions in both Tasks for vn5 problems.

B.2 Problems with Valid U-Conclusions

None A B
None B C

-AB-C (1)	B-C-A (6)
B-A-C (4)	*CA-B (1)*
	C-B-A (1)
	-C-B-A (1)
NVC (6)	
none A C (9)	
NVC (15)	

None B A
None C B

B-A-C (6)	B-C-A (1)
	-CB-A (1)
A-B-C (1)	*CA-B (1)*
NVC (10)	
none A C (1)	*none C A (3)*
	some C A (1)
NVC (18)	

None A B
None C B

-AB-C (1)	B-C-A (1)
B-A-C (8)	
AC-B (1)	
	C-B-A (1)
-A-B-C (1)	
NVC (9)	
none A C (1)	*none C A (1)*
NVC (22)	

None B A
None B C

-AB-C (1)	
-A-CB (1)	
B-A-C (12)	
-BAC (1)	*CA-B (1)*
-B-A-C (1)	
NVC (4)	
none A C (3)	*none C A (1)*
NVC (17)	

Table B.9: Conclusions in both Tasks for vu1 problems. Standard Task has NVC.

None A B
Somenot B C

-AB-C (1)	**B-C-A (10)**
-A-CB (1)	
B-A-C (1)	
A-B-C (1)	
NVC (7)	
NVC (23)	

Somenot B A
None C B

B-A-C (11)	B-C-A (2)
NVC (8)	
NVC (24)	

None B A
Somenot B C

B-A-C (5)	**B-C-A (9)**
B-AC (2)	
AB-C (1)	
B-A-C (1)	
NVC (4)	
	none C A (1)
	somenot C A (2)
NVC (20)	

Somenot B A
None B C

B-A-C (10)	B-C-A (5)
	-C-AB (1)
BA-C (2)	*CA-B (1)*
NVC (3)	
somenot A C (1)	
NVC (22)	

Table B.10: Conclusions in both Tasks for vu2 problems. Standard Task has NVC.

B.3 Problems with No Valid Conclusion

All A B
All C B

ABC (4)	*CAB (3)*
BAC (2)	
-B-A-C (2)	
NVC (11)	
all A C (10)	*all C A (2)*
some A C (3)	*some C A (1)*
NVC (7)	

Table B.11: Conclusions in both Tasks for np1 problems.

All A B
Some B C

ABC (10)	*CAB (2)*
BAC (1)	
NVC (9)	
some A C (19)	*some C A (1)*
NVC (2)	

Some B A
All C B

ACB (4)	*CBA (4)*
BAC (2)	*BCA (1)*
NVC (11)	
some A C (3)	*some C A (15)*
NVC (3)	

All A B
Some C B

ABC (2)	*CBA (1)*
ACB (1)	*CAB (1)*
BAC (1)	
NVC (16)	
some A C (4)	*some C A (8)*
all A C (1)	
NVC (9)	

Some A B
All C B

ABC (2)	*CBA (1)*
BAC (1)	*CAB (2)*
NVC (15)	
some A C (8)	*some C A (8)*
NVC (7)	

Table B.12: Conclusions in both Tasks for np2 problems.

Some A B	
Some B C	
ABC (3)	
BAC (1)	
NVC (18)	
some A C (5)	*some C A (1)*
NVC (17)	

Some B A	
Some C B	
ABC (1)	*CBA (1)*
BAC (1)	*BCA (1)*
NVC (18)	
some A C (1)	*some C A (2)*
NVC (20)	

Some A B	
Some C B	
ABC (1)	*CAB (1)*
BAC (1)	
NVC (18)	
some A C (2)	
NVC (22)	

Some B A	
Some B C	
BAC (4)	*CBA (1)*
	CAB (1)
NVC (16)	
some A C (3)	
NVC (21)	

Table B.13: Conclusions in both Tasks for np3 problems.

All A B	
Somenot B C	
AB-C (8)	*A-B-C (1)*
BA-C (2)	
NVC (11)	
some A C (1)	
somenot A C (14)	
NVC (6)	

Somenot B A	
All C B	
ACB (1)	*CB-A (11)*
NVC (9)	
none A C (1)	*some C A (1)*
	somenot C A (15)
NVC (4)	

Table B.14: Conclusions in both Tasks for nn1 problems.

Somenot A B	
All B C	
-ABC (1)	*CB-A (1)*
ABC (1)	*BC-A (1)*
A-BC (1)	*CA-B (1)*
A-B-C (2)	
NVC (12)	
somenot A C (6)	*somenot C A (1)*
some A C (3)	
NVC (12)	

All B A	
Somenot C B	
BA-C (3)	*B-CA (1)*
AB-C (1)	*C-B-A (1)*
	C-A-B (1)
NVC (15)	
some A C (1)	*some C A (2)*
	somenot C A (6)
NVC (12)	

Table B.15: Conclusions in both Tasks for nn2 problems.

Some A B	
Somenot B C	
ABC (2)	
AB-C (4)	
NVC (16)	
somenot A C (2)	*somenot C A (1)*
NVC (21)	

Somenot B A	
Some C B	
B-AC (2)	*CB-A (3)*
BCA (1)	
NVC (16)	
	somenot C A (5)
NVC (18)	

Somenot B A	
Some B C	
B-AC (1)	*BC-A (2)*
CB-A (1)	
NVC (18)	
	somenot C A (3)
NVC (20)	

Some B A	
Somenot B C	
BAC (1)	
B-AC (1)	
BA-C (2)	
NVC (18)	
somenot A C (1)	
NVC (22)	

Table B.16: Conclusions in both Tasks for nn3 problems.

Somenot A B	
Some B C	
BAC (1)	*BC-A (1)*
A-BC (1)	
NVC (18)	
some A C (2)	
somenot A C (2)	
NVC (20)	

Some B A	
Somenot C B	
ABC (1)	
BAC (1)	
AB-C (1)	
NVC (19)	
	somenot C A (2)
NVC (22)	

Somenot A B	
Some C B	
ABC (1)	*CB-A (1)*
NVC (19)	
somenot A C (2)	*some C A (1)*
NVC (20)	

Some A B	
Somenot C B	
AB-C (1)	*CBA (1)*
BA-C (1)	
NVC (19)	
somenot A C (1)	*somenot C A (1)*
NVC (19)	

Table B.17: Conclusions in both Tasks for nn4 problems.

Somenot A B	
None B C	
A-BC (1)	*CA-B (1)*
AB-C (1)	*B-CA (1)*
B-A-C (1)	*C-A-B (1)*
NVC (16)	
somenot A C (2)	
NVC (22)	

None B A	
Somenot C B	
-BAC (1)	*CA-B (1)*
B-A-C (2)	*B-C-A (1)*
NVC (17)	
	somenot C A (2)
NVC (21)	

Somenot A B	
None C B	
AB-C (2)	
A-B-C (3)	
NVC (16)	
some A C (1)	
somenot A C (1)	
NVC (20)	

None A B	
Somenot C B	
-AB-C (1)	*BC-A (1)*
B-A-C (1)	
NVC (19)	
	somenot C A (3)
NVC (21)	

Table B.18: Conclusions in both Tasks for nn5 problems.

Somenot A B	
Somenot C B	
A-BC (1)	*CA-B (1)*
B-C-A (1)	
NVC (19)	
	somenot C A (1)
NVC (22)	

Somenot B A
Somenot B C
BAC (1)
B-A-C (1)
NVC (20)
somenot A C (1)
NVC (21)

Table B.19: Conclusions in both Tasks for nn6 and nn8 problems.

Somenot A B	
Somenot B C	
B-A-C (1)	*CA-B (1)*
NVC (20)	
some A C (1)	
somenot A C (1)	
NVC (22)	

Somenot B A	
Somenot C B	
B-A-C (1)	*BC-A (1)*
	CA-B (1)
NVC (19)	
	somenot C A (2)
NVC (21)	

Table B.20: Conclusions in both Tasks for nn7 problems.

Somenot B A
Somenot B C
BAC (1)
B-A-C (1)
NVC (20)
somenot A C (1)
NVC (21)

Table B.21: Conclusions in both Tasks for nn8 problems.