2003

Individual differences and strategies for human reasoning

Bacon, Alison Margaret

http://hdl.handle.net/10026.1/349

http://dx.doi.org/10.24382/4861

University of Plymouth

Individual Differences and Strategies for Human Reasoning

By

Alison Margaret Bacon

A thesis submitted to the University of Plymouth
in partial fulfilment for the degree of

Doctor of Philosophy

Department of Psychology
Faculty of Science

September 2003

Alison M. Bacon

Individual Differences and Strategies for Human Reasoning

Abstract

Theories of human reasoning have tended to assume cognitive universality, i.e. that all individuals reason in basically the same way. However, some research (e.g. that of Ford, 1995) has found evidence of individual differences in the strategies people use for syllogistic reasoning. This thesis presents a series of experiments which aimed to identify individual differences in strategies for human reasoning and investigate their nature and aetiology. Experiment 1 successfully replicated and extended Ford (1995) and provided further evidence that most individuals prefer to reason with either verbal-propositional or visuo-spatial representations. Data from verbal and written protocols showed that *verbal reasoners* tended to use a method of substitution whereby they obtain a value for the common term from one premise and then simply substitute it in the other premise to obtain a conclusion. *Spatial reasoners*, on the other hand, presented protocols which resembled Euler circles and described the syllogistic premises in terms of sets and subsets. Experiment 2 provided some further qualitative evidence about the nature of such strategies, especially the verbal reasoners, showing that within strategy variations occurred. Experiment 3 extended this line of research, identifying a strong association between verbal and spatial strategies for syllogistic reasoning and abstract and concrete strategies for transitive inference (the latter having originally been identified by Egan and Grimes-Farrow, 1982). Experiments 1-3 also showed that inter-strategic differences in accuracy are generally not observed, hence, reasoners present an outward appearance of ubiquity despite underlying differences in reasoning processes. Experiments 5 and 6 investigated individual differences in cognitive factors which may underpin strategy preference. Whilst no apparent effects of verbal and spatial ability or cognitive style were found, reasoners did appear to draw differentially on the verbal and spatial components of working memory.

Confirmatory factor analysis showed that whilst verbal reasoners draw primarily on the verbal memory resource, spatial reasoners draw both on this and on spatial resource. Overall, these findings have important implications for theories of human reasoning, which need to take into account possible individual differences in strategies if they are to present a truly comprehensive account of how people reason.

LIST OF CONTENTS

Chapter 2: Strategies in Syllogistic Reasoning

Chapter 3: Strategies in Other Reasoning Domains

Chapter 4: Reasoning Strategies, Abilities and Cognitive Style

Chapter 5: Reasoning Strategies and Working Memory

Chapter 6: General Discussion

LIST OF TABLES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

LIST OF FIGURES

Chapter 2

Chapter 3

Chapter 4

Chapter 5

LIST OF APPENDICES

Chapter 2

Chapter 3

Chapter 4

Chapter 5

ACKNOWLEDGEMENTS

Firstly, I would like to acknowledge the support of my supervisory team. My Director of Studies, Dr. Simon Handley, has been unfailingly supportive (and patient) throughout this period of study as well as offering invaluable advice and guidance. My second supervisor, Professor Steve Newstead, has also been extremely supportive and this thesis has benefited from his considerable knowledge and experience of reasoning research. Both supervisors have always found time to talk and meet with me whenever necessary and this thesis would never have been completed without their help.

Secondly, this work was supported financially by a grant from the Economic and Social Research Council without which this thesis would not have been possible.

Last, but certainly not least, my gratitude goes to all the friends and colleagues who have supported me throughout, either academically or emotionally (or both). My thanks go to everyone who believed in me at times when I couldn't believe in myself.

Author's Declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other university award except the permitted Postgraduate Diploma in Psychological Research Methods.

This programme of research was supported by funding from the Economic and Social Research Council (Award No. R42200034033) from 1st October 2000 to 31 September 2003.

This thesis is the result of the author's own investigation, carried out under the guidance of her supervisory team. Other sources are explicitly referenced and a bibliography is appended.

During this period of study, the author has also completed the Postgraduate Diploma in Psychological Research Methods, and achieved the following conference contributions and publications.

Refereed Conference Contributions:

Bacon, A.M., Handley, S.J. and Newstead, S.E. (2002). *Individual differences in strategies for human reasoning*. Paper presented at the XV11 BPS Annual Cognitive Section Conference, Kent, 9-11 September 2002.

Refereed Journal Publications

Bacon, A.M., Handley, S.J. and Newstead, S.E. (2003). Individual differences in strategies for syllogistic reasoning. *Thinking & Reasoning, 9(2)*, 133-168.

Other Publications

Bacon, A.M., Handley, S.J. and Newstead, S.E. (in press). Verbal and spatial strategies in reasoning. In Roberts, M. and Newton, E. (Eds.). *Methods of Thought: Individual Differences in Reasoning Strategies*. Hove: Psychology Press.

Seminar Presentations

An extended version of the above conference paper was also presented to University of Plymouth colleagues in a departmental research seminar, Nov. 2001.

Signed..

Date......17|11|03......

CHAPTER 1

Human Reasoning: A Brief Review

1.1 General Introduction

1.1.1 What is reasoning and why study it?

This thesis is about human reasoning, or, how individuals use knowledge and information to make inferences and deduce conclusions about the world. This manipulation and transformation of knowledge allows for decision making, revision of beliefs and achievement of goals. As such, reasoning is a fundamental cognitive activity and integral to everyday thinking (Evans, Newstead and Byrne, 1993; Evans Over and Manktelow, 1993; Galotti, 1989; Johnson-Laird and Bara, 1984; Johnson-Laird and Byrne, 1991, Kuhn, 1991). The ability to draw necessary conclusions based on what is known about the world lies at the core of human intelligence (Stanovich, 1999). As Johnson-Laird and Byrne (1991, page 3) state:

> "A world without deduction would be a world without science, technology, laws, social conventions and culture".

As such human reasoning has deservedly afforded a long history of psychological research dating back to the early 1900s (see for instance Evans et al 1993 for a review). It is a process initiated by premises which can take the form of facts, perceptions or beliefs and which, ideally, can lead to a conclusion which is not explicit in those premises. However, despite decades of research, the intervening processes remain mysterious (Johnson-Laird, 2001). The programme of research presented in this thesis offers a contribution to the understanding of these processes and, more specifically, to our knowledge of the nature of the individual strategies people use during reasoning.

Johnson-Laird, Savary and Bucciarelli (2000) define a strategy as "the sequence of steps that an individual follows in solving, or attempting to solve, a problem" (page 210), referring to each of these steps as a *tactic*. This provides a useful basic working definition which informs the general use of the word strategy throughout this thesis. However, there are also other more complex issues (such as degree of conscious control and the relationship to other cognitive mechanisms) which need to be accounted for in any comprehensive definition of strategies. These will be discussed at a later stage. This first chapter will initially concentrate on reviewing the background to reasoning research, how reasoning is studied and the most significant theories which have attempted to explain reasoning processes.

1.1.2 The study of human reasoning

Most research on the psychology of human reasoning has attempted to use the logical analysis of problems as a basis for the classification of errors and correct responses. It is not the role of this thesis to investigate or comment on the role of formal logic as a normative theory of human reasoning, but a basic understanding of logical terms is desirable for the understanding of many tasks used in experimental studies, most especially syllogistic reasoning with which this thesis will be extensively concerned. Moreover, many terms used in logic occur frequently in the psychological literature on reasoning (and hence also in this thesis) and it is useful to define some of these at the outset.

1.1.3 Syllogistic Reasoning

Categorical syllogisms are deductive arguments comprising three propositions, two premises and a conclusion, and three terms, each of which occur in two of the propositions. The propositions are categorical in that they make claims about class membership of the terms, affirming or denying that one class S (the subject) is included in another class P (the predicate) in whole or part. The third proposition typically is a conclusion which follows logically from the two earlier premises. Syllogisms may be presented with abstract content, for instance:

	All B are A
	Some C are B
Therefore,	Some C are A

Or thematic content,

	All teachers are psychologists
	Some writers are teachers
Therefore,	Some writers are psychologists

Each syllogism contains three terms (A, B and C or psychologists, teachers and writers in the above examples) one of which, the middle term B, is common to both premises. The two syllogisms above are logically identical, and the nature of the deduction remains valid whatever terms are substituted for A, B and C, whilst the problem retains this form. The classic syllogistic inference is to determine, from the information given in the premises, the one relationship which is not explicitly stated, i.e. that between the two end terms, A and C. This forms the conclusion. Each proposition contains one of four possible quantifiers (either *All, Some, None or Some...not*) which describe the relationship between the terms. The *mood* of the syllogism is described in terms of the combination of quantifiers present and is given the traditional designation shown in table 1.1.

Table 1.1: The four syllogistic moods

Quantifier	Mood Description	Designation
All	Universal affirmative	A
Some	Particular affirmative	I
None	Universal negative	E
Some...not	Particular negative	O

Syllogism structure is also described in terms of one of four *figures* based on the arrangement of the terms (see Table 1.2). It should be noted that figural convention varies between sources, but for the purposes of this thesis, any reference to figure relates to the forms illustrated in Table 1.2. Johnson-Laird and Bara (1984) showed that problems tend to increase in difficulty in line with these figures from 1 to 4.

Table 1.2: The four syllogistic figures

Figure 1	Figure 2	Figure 3	Figure 4
A – B	B – A	A – B	B – A
B – C	C – B	C – B	B – C

As there are four possible quantifiers there are 16 distinct premise combinations (moods) associated with each figure, and hence 64 possible syllogistic forms. Given that each conclusion will contain one of the four quantifiers and the end terms may follow the order A-C or C-A, this gives a total of 512 possible syllogisms. However, of these, relatively few have logically valid conclusions – ones which can be logically deduced, given that the premises are true. Johnson-Laird and Byrne (1991) present a set of 27 valid syllogistic problems and these are those considered as valid for the purposes of the present programme of research (Garnham and Oakhill, 1994, Chapter 6, discuss methods of determining syllogistic validity in some detail).

Typically, reasoning tasks centre on the validity of syllogistic conclusions (see Evans et al, 1993, for a review). The task is presented in three possible basic forms; participants may be asked to demonstrate reasoning either by selecting the one valid conclusion from a set of alternatives, evaluating whether the one conclusion presented is valid or invalid, or by generating their own conclusion from the premises given. Error rates are generally high (an often cited study by Dickstein, 1978, found that just 52% of conclusions were correctly evaluated) and systematic biases are consistently reported associated with the interaction between prior knowledge/belief and problem content and these have consistently been shown to adversely influence reasoning performance (see for instance Evans, 1989, for an extensive discussion of biases in reasoning).

Although syllogistic tasks may appear hypothetical and circumscribed, researchers such as Johnson-Laird and Bara (1984) have argued that they in fact have a high degree of ecological relevance. Everyday reasoning commonly involves deciding what conclusion, if any, can be drawn from certain assumptions about category membership and syllogisms provide small scale replications of such problems (Galotti, 1989; Gilhooly, 1996). The identification of assumptions, use of stored knowledge, evaluation of arguments and deduction of conclusions are all processes present in syllogistic reasoning and such problems provide a useful test of these abilities in a format that exists in limited, and hence controllable, structures (Johnson-Laird and Bara, 1984).

1.2 Theories of Human Reasoning

A considerable body of research has attempted to determine the nature of underlying reasoning processes. Two major and opposing schools of thought have emerged and claim that reasoning depends either on the use of mental models or on logical rules. The following sections consider these in more detail. Both viewpoints form general and wide ranging theories which are proposed to apply across all reasoning domains, though both have also been specifically, and extensively, applied to reasoning with syllogisms.

1.2.1. Mental Model Theories

The fundamental tenet of the mental models theory is that when individuals are presented with a problem which requires reasoning, they capture the meaning of the problem by imagining the relationships between terms described in the premises (Johnson-Laird and Byrne, 1991). In doing so, they construct analogous internal models of the state of the world described by those premises. Deductions and inferences depend on three stages of thought (ibid., pages 35-36):

1. Comprehension: reasoners use their knowledge of language and of the world, prior experience, etc to understand the premises and construct an analogical internal model.

2. Description: They attempt to form a parsimonious conclusion by fleshing out the model they have constructed. This description should assert a state of affairs which is not explicitly stated in the premises. Where no such assertion can be found, reasoners conclude that nothing follows from the premises

3. Validation: Reasoners search for counterexamples – alternative models of the premises in which their putative assertion (conclusion) is false, perhaps further fleshing out the initial models to do so. If no counterexample if found, they assume the putative conclusion to be valid. If counter models are discovered, then prudent reasoners return to stage 2 and attempt to discover whether there is any conclusion which is true for all the possible models they have constructed. If so, they then search for further counterexamples, and so on until the search is exhausted. For deductions which rely on quantifiers or connectives, the number of possible models is finite and hence the search can, in principle, be exhaustive.

Recently, Bara and Bucciarelli (2000) have broken these stages down into five phases; construction, integration, conclusion, falsification and response. Their account places greater emphasis on the processes involved in the search for counterexamples which may falsify the putative conclusion. According to mental models theory, errors occur due to incorrect or inadequate fleshing out of the basic model, or failure to find such counterexamples. Hence not all possible models of the premises are considered. Although reasoners are generally *able* to search for counter-examples (e.g. Newstead, Pollard, Evans and Allen, 1992) they frequently fail to do so, preferring to settle for a conclusion consistent with the first model they construct (Handley, Dennis, Evans and Capon, 2000; Newstead, Handley and Buck, 1999). This failure has been attributed to variation in individual abilities (for instance Handley et al 2000 identified a specific cognitive factor which seemed to predict ability to search for counterexamples) and also to limitations in working memory capacity (see for instance Gilhooly, 1998; Johnson-Laird and Byrne, 1991; Johnson-Laird, 2001).

Mental model theory has been extensively applied to syllogistic reasoning, see for example Johnson-Laird and Bara (1984), Bara, Bucciarelli and Johnson Laird (1995), Johnson-Laird and Bucciarelli (1999) to name but a few. Basically, Johnson-Laird and Bara (1984) argued that the difficulty of a syllogism depends on two main factors:

1. The number of models to be constructed in attempting to establish a valid conclusion, given the limited capacity of working memory. Of the 27 valid syllogisms highlighted by Johnson-Laird and Byrne (1991, pp. 107-110), 10 are single model problems, that is the conclusion can be drawn from one initial model and no fleshing out is required. The remaining 17 are all multi-model problems, that is there are two or three possible models of the premises which need to be considered. The data presented by Johnson-

Laird and Byrne clearly shows that the single model problems afford reasoners less difficulty than multi-model.

2. The figure of the syllogism (cf. Table 1.2), given that the order in which information is retrieved from working memory is optimally the order in which it entered working memory. With all but Figure 1 problems (A-B; B-C), in forming an initial model, reasoners are required to carry out initial mental operations on syllogistic premises in order to bring the two middle (B) terms into contiguity in order to integrate the premises. This additional cognitive burden becomes increasingly onerous in line with figure from 2 to 4, with more difficult problems eliciting fewer valid conclusions and a greater number of "no valid conclusion" responses (Johnson-Laird and Bara, 1984; Johnson-Laird and Byrne, 1991; Bara et al, 1995).

In constructing models, the theory proposes that finite sets of entities are represented by finite sets of mental tokens that accommodate a representation of the relationship between entities. Moreover, the initial model will operate according to a principle of economy in that it represents only that information which is necessary to determine premise meaning. Other information which is implied, but not explicit, will be registered in a mental footnote which is retained in memory until it is needed to flesh out the initial model at a later stage. One major criticism of mental models has been that although an arbitrary number of tokens may suffice for syllogistic quantifiers, if sets are indeed represented this way, the theory cannot generalise easily to natural language quantifiers such as "most" or "few".

Models of syllogistic premises are portrayed according to specific notation, as shown in Table 1.3.

Table 1.3: Mental model representations of syllogistic premises (after Johnson-Laird & Byrne, 1991).

Premise		Initial models	Explicit models
A	All A are B	[a] b [a] b ...	[a] [b] [a] [b] [¬a] [b] [¬a] [¬b]
I	Some A are B	a b a b ...	a b a b ¬a b a ¬b ¬a ¬b
E	No A are B	[a] [a] [b] [b[...	[a] ¬b [a] ¬b ¬a [b] ¬a [b] ¬a ¬b
O	Some A are not B	a a a [b] [b] ...	a ¬b a ¬b a [b] ¬a [b] ¬a ¬b

¬ tag for negation
[ ] indicates exhaustive representation
... alternative model without explicit content

As an illustrative example, consider the following syllogism;

All A are B
No B are C

The first premise could be constructed in an initial model of the form:

[a] b
[a] b
...

Here the square brackets indicate that this term is exhaustively represented, that is, it embodies the knowledge that an a cannot appear in any other case (because *All* a are b). The three dots denote implicit information (the mental footnote). However, it is also possible that although there can only be As which are also Bs, there could also be Bs that are not As, and even things which are neither As nor Bs. The fleshed out version of the model introduces the symbol for negation (¬) which allows for these possibilities to be included:

```
[a]   [b]
[a]   [b]
[¬a]  [b]
[¬a]  [¬b]
```

The second premise of the syllogism will be represented as in the E row of Table 1.3, i.e.

```
[b]    ¬c
[b]    ¬c
¬b     [c]
¬b     [c]
¬b     ¬c
```

These two models can be combined, again following the constraint that reasoners will not explicitly represent more information that is necessary, for instance they will not represent b twice. The combined model will look something like:

```
[[a]  b]
[[a]  b]
          [c]
          [c]
```

For this one model problem, this is the only way in which information about the three terms, A, B and C could be combined, so there is no need to find counterexamples. This final model is consistent with the valid conclusion *No A are C.*

Johnson-Laird and Byrne (1991) base their empirical predictions on the number of mental models required to solve a given problem. They present findings from a series of experiments which support this claim, together with details of the percentage of correct conclusions generated to all 27 valid problems which they identify (Table 6.1, pages 107-110). These percentages vary from around 70– 90% correct conclusions for the single model syllogisms, to between 15 and 30% for three-model syllogisms. Moreover, many studies of belief bias in syllogistic reasoning also support model theory. Mental models theory suggests that prior knowledge or belief may prejudice reasoning either by biasing the initial premise interpretation, by acting as a post hoc censor or by influencing the inclinations to search for counter-examples. Reasoners will search more assiduously if their initial putative conclusion is unbelievable, than if it is believable. Studies by Oakhill, Johnson-Laird and Garnham (1989), Newstead, Pollard, Evans and Allen (1992) and Cherubini, Garnham, Oakhill and Morley (1998) have all shown that belief bias is suppressed when previous knowledge is incompatible with the premises and their results are compatible with idea that representation of prior knowledge precedes modelling of the combined premises.

1.2.2 Mental Logic Theories

Formal logic proposes that an argument is deductively correct if, and only if, the conclusion is true in all states of affairs in which the premises are true. The conclusion is said to be *entailed* by the premises and the argument is valid. Two theories of reasoning based on logical rules have been proposed (Rips, 1994a; Braine and O'Brien, e.g. 1998). These both assume that reasoning proceeds like a logical proof, using rules of inference stored in a mental logic. Individuals rely on propositional, or language like, representations and on cognitive processes akin to the natural deduction rules developed by logicians, where separate rules exist for dealing with every connective or quantifier.

In a general sense, a formal proof is a finite sequence of propositions in which each sentence is either a premise, an axiom of the logical system (such as an additional inference rule), or a sentence which follows from a preceding sentence by one of the system rules. An argument is deducible if the final sentence in the proof is the conclusion. One of the most straightforward proofs is that which logicians have termed *modus ponens* (or affirming the antecedent). According to this principle, the premise *If p then q*, and the sentence p, jointly entail the concluding proposition, q. In more thematic terms, and to anglicise one of Rips (1994a) examples;

> If Simon deposits 50 pence, Simon will get a coke
> <u>Simon deposits 50 pence</u>
> Simon gets a coke.

Furthermore, natural deduction methods allow for temporary supposition in order to simplify a proof. For instance, again given the premise *If p then q*, but then given *not q*, then the reasoners can suppose, p and as a result of that, infer q. This temporary conclusion contradicts the categorical premise, *not q*. Reasoners can therefore reverse their original supposition and conclude that *not p* holds. In other words, if we find that Simon does not get a coke, we can we can make the supposition that he has not deposited his 50 pence. This is an example of another mental proof, that of *Modus tollens* (or denying the consequent). This indirect supposition is more prone to errors, compared to a direct inference such as modus ponens. Table 1.4 presents inference rules based on classical logic which Rips (1994a) has incorporated into his mental logic account of human reasoning.

For problems such as syllogisms which include quantifiers such as *All* and *Some*, similar rules are applied, again to borrow an example from Rips (1994a):

> All square blocks are green blocks
> <u>Some big blocks are square blocks</u>
> Some big blocks are green blocks

Here Rips suggests that reasoning may follow the lines of "some big blocks are square blocks; so take an arbitrary big square block and call it b. Block b must be green since b is square and all square blocks are green. Hence some big blocks (b for instance) are green, as stated in the conclusion" (page 50). In other words, the proof proceeds by instantiating premises to get a new sentence with a temporary name (in the above example b), applying the rules for sentence connectives to this instantiation, and then generalising to the quantifiers in the conclusion. Hence his system also supplements the rules shown in Table 1.4 with others for introducing and eliminating quantifiers, as shown in 1.5.

Table 1.4: Inference rules (from Rips, 1994a, page 45)

IF elimination (Modus ponens)
a) If sentence of the form IF P THEN Q and P hold in a given domain,
b) Then the sentence Q can be added to that domain

If introduction (Conditionalisation)
a) If a sentence Q holds in a subdomain whose supposition is P,
b) then IF P THEN Q can be added to the immediate superdomain

NOT Elimination (Modus tollens 1)
a) If the sentences Q and NOT Q hold in a subdomain whose supposition is NOT P,
b) then the sentence P can be added to the immediate superdomain

NOT Introduction (modus tollens 2)
a) If the sentences Q and NOT Q hold in a subdomain whose supposition is P,
b) then NOT P can be added to the immediate superdomain

Double negation Elimination
a) If the sentence NOT NOT P holds in a given domain,
b) then the sentence P can be added to that domain

AND Elimination
a) If the sentence P AND Q holds in a given domain,
b) then the sentences P and Q can be added to that domain

AND Introduction
a) If the sentence P and the sentence Q hold in a given domain,
b) then the sentence P AND Q can be added to that domain

OR Elimination
a) If the sentence P OR Q holds in a given domain D,
b) and the sentence R holds in an immediate subdomain of D whose supposition is P,
c) and the sentence R holds in an immediate subdomain of D whose supposition is Q,
d) then R can be added to D.

OR Introduction
a) If the sentence P holds in a given domain,
b) then the sentences P OR Q and Q OR P can be added to that domain, where Q is an arbitrary sentence.

In the notation P(v) represents a expression containing a variable, v. These quantifier rules are more complex as they need to incorporate variables (e.g. big block, green block) in terms of whether they are bound by a given quantifier (as in *All* v) or are free variables (such as v1 AND v2). Furthermore, some variables may be allocated temporary supposed names (which may subsequently become permanent) and these are indicated by t. Subscripts indicate temporary names.

Table 1.5: Inference rules for elimination and introduction of the quantifiers ALL and SOME (from Rips, 1994a, page 52).

ALL Elimination

a) If (FOR ALL v) P(v) holds in a given domain,
b) and P(t) is the results of replacing all free occurrences of v in P(v) with t,
c) then P(t) can be added to the domain.

ALL Introduction

a) If P(a) holds in a given domain,
b) and a does not occur in subscript in P(a),
c) and a was not produced by FOR SOME elimination,
d) and a does not occur in any suppositions that hold in the domain,
e) and a does not occur within the scope of (FOR ALL v) or (FOR SOME v) in P(a),
f) and P(v) is the result of replacing all occurrences of a in P(a) by v,
g) then (FOR ALL v) P(v) can be added to the domain.

SOME Elimination

a) If (FOR SOME v) P(v) holds in some domain,
b) and b had not yet appeared in the proof,
c) and a_1, a_2..., a_k is a list of the temporary names (possibly empty) that appear in P(v) and that first appeared in a supposition or in an application of FOR ALL elimination,
d) and P ($b_{a1,\ a2,\ldots,\ k}$) is the result of replacing all free occurrences of v in P(v) by $b_{a1,\ a2,\ldots,\ k}$.
e) then P($b_{a1,\ a2,\ldots,\ k}$) can be added to the domain.

SOME Introduction

a) If P(t) holds in a given domain,
b) and t does not occur within the scope of either (FOR ALL v) or (FOR SOME v) in P(t),
c) and P(v) is the result of replacing all occurrences of t in P(t) with v,
d) then (FOR SOME v) P(v) can be added to the domain.

For a more thorough explanation of all these rules see Rips (1994a, Chapter 2). Rips proposes a theory which assumes that when confronted with a reasoning problem, individuals attempt to solve it by generating, within working memory, a set of propositions linking the premises to the conclusion. Each link in this process comprises an inference rule similar to those in Tables 1.4 and. 1.5 which the individual "recognises as intuitively sound" (page 103). Errors occur because individuals may not possess, or may be unable to apply, the appropriate rules for a given proof required for a given problem. They may possess non-standard rules which lead them to non-logical conclusions, but Rips claims that they at least attempt to construct an internal mental proof of the conclusion entailed by the premises. Limitations in working memory capacity also mean that problems which require the application of a greater number of rules, or more complex ones, will prove most intractable.

Working memory plays an important role for it is here that premise information is first stored, then scanned for possible inferences. The newly deduced sentences are added to memory store, rescanned, and followed up with further deductions. The process is repeated until a proof is found or no further rules apply. Rips terms this system PSYCOP (short for Psychology of Proof) and has developed a computer program based on the system. PSYCOP uses forward rules to draw implications from the premises and backward rules to create subgoals based on the conclusion. In this way, some of the computational complexity of formal logical inference rules is reduced. In evaluating a given conclusion, backward rules operate initially, pursuing a chain of reasoning until required the subgoal, and finally the premises, are reached. If no subgoals are reached, then PSYCOP gives up and concludes that no valid conclusion is possible. In situations where PSYCOP is expected to produce conclusions rather than evaluate them, it can use only the forward rules to complete the task. Where no such rules seem to apply, it assumes there is no conclusion.

When working with syllogistic problems, PSYCOP makes use of the fact that syllogistic premises can all be expressed using the basic logical operators together with variables (terms). PSYCOP incorporates three additional rules for syllogistic inference, *transitivity* (that if three terms lie in a linear relationship and A implies B which in turn implies C, then A implies C), *exclusivity* (where although A implies B, B does not imply C hence, neither does A) and *conversion* (which allows for the reversal of terms in a premise, e.g. the assumption that *No A are B* is equivalent to *No B are A*). PSYCOP also incorporates the Gricean implicatures (e.g. that some A are B implies some A are not B) and the existential presupposition that there are some As and Bs present to begin with. The translations of the four quantified expressions and the implicatures proposed by PSYCOP are shown in table

1.6. Once the premises have been expressed in this way, the general mental proof rules can be applied.

Table 1.6: The four syllogistic sentences expressed in words (first lines), their translations into PSYCOP inference rule notation (second lines), and their implicatures (third lines). From Manktelow (1999), adapted from Rips (1994a, Chapter 7).

Mood		
A	All A are B	
	IF A(x) THEN B(x)	[if x is A then x is B]
	A(a) AND B(a)	[there are things a, which are a and b]
I	Some A are B	
	A(b) AND B(b)	[there are things b, which are A and B]
	A(a) AND NOT B(a)	[there are things a, which are A and not B]
E	No A are B	
	NOT (A(x) AND B(x))	[it is not the case that x is A and x is B]
	A(a) AND NOT B(b)	[there are things a, which are A and not B]
O	Some A are not B	
	A(b) AND NOT B(a)	[there are things b, which are A and not B]
	A(a) AND B(a)	[There are things a, which are A and B]

x = a variable (i.e. a label for a class); a and b are temporary names (i.e. labels for possible instances).

An alternative mental logic theory has been presented by Braine and O'Brien and colleagues (e.g. Braine, 1990, Braine, Reiser and Rumain, 1984, Braine and O'Brien, 1998; O'Brien, 1995). Like Rips' theory, it is based on natural deduction and includes many of the same rules. These form part of what Braine and O'Brien term inference schemas which, rather than just comprising mental versions of logic rules for the introduction and elimination of quantifiers, also comprise a reasoning program which controls how the rules are used and includes processes for both inference and conclusion evaluation. Schemas are directly applied by the program wherever appropriate propositions are encountered. For instance, when working memory holds "p or q" and the premise "not p" is encountered, the conclusion "q" is automatically supplied. O'Brien (1998, page 29) claims that such inferences are made "routinely and without apparent effort" and are considered to be fundamental to human deductive competence. Hence, tasks which require only the direct application of these simple rules will tend to be those which are carried out accurately. However, an indirect reasoning routine also allows for individuals to acquire more complex schemas through learning, and facilitates their application in appropriate problem domains. Problems which require more sophisticated reasoning will call upon this indirect reasoning routine and are less likely to be solved accurately by most individuals. Unlike Rips` theory which proposes that reasoning inferences are made in order to proceed towards a goal or subgoal, the schemas allow for reasoners to look ahead to the types of inferences which will be required later in the reasoning process and to consider possible conclusions. Where no clear goal/conclusion is presented, core schemas are directly applied which allows a putative conclusion to be generated.

The theory also includes a pragmatic component based on natural language understanding of quantifiers and connectives. An example is the *invited inference*, where, for instance, in addition to applying a rule such as modus ponens, reasoners may be invited to infer that "if p then q" also implies "if not p then not q". This echoes instances in everyday deontic reasoning (e.g. if John tidies his room then he will receive £5), the pragmatic assumption being that if John does not tidy his room, then the £5 will not be forthcoming. Similarly, Cheng and Holyoak (1985) suggested pragmatic reasoning schemas which facilitate inferences involving conditionals which in natural language express permission and/or obligations. However, as their theory has mainly been applied to conditional and deontic reasoning, rather than to syllogisms, it is less relevant to the current purpose. O'Brien (1995) has suggested that pragmatic schemas also allow for "non-logical" inferences to be made, based for instance on schemas of stereotyped social situations. Overall, some of these pragmatic inferences may not be logically sound, based as they are on prior knowledge and conversational implicatures, but they offer an explanation for some of the inaccuracies, fallacious inferences and biases frequently reported in the reasoning literature.

Rule theories predict errors on the basis of the number and complexity of rules which need to be applied in order to solve a given problem. Rips (1994a, Table 7.1, page 235) presented the results of extensive testing of PSYCOP which seem to support this hypothesis. On simple syllogisms which require the use of just one rule, participants typically perform very well. For instance, the syllogism *All A are B, All B are C*, requires only the transitivity rule to reach the conclusion, *All A are C*. 90% of participants in Rips' studies evaluated this conclusion correctly. Similarly, *All A are B, No B are C* requires only the exclusivity rule and 85% of participants made a correct evaluation on this problem. However, syllogisms which require multiple rules, such as *All A are B; Some C are not B*, produce just 35% correct evaluations. According to Rips, solving this syllogism requires four rules (forward and-elimination, followed by three backward rules, and-introduction, not-introduction and if-elimination). Rule theories suggest that effects of prior knowledge or belief prejudice reasoning either by biasing the initial premise interpretation or by acting as a post hoc censor. As the rules are, by nature, blind to semantic content effects, there cannot be any influence of belief on the deduction process itself. However, Rips does not place emphasis on explaining this feature any further .

1.3 Theories Specific to Syllogistic Reasoning

In addition to rules and mental models, a number of other theories have been specifically developed in an attempt to explain syllogistic reasoning. A brief review of the main theories is presented below. The first two of these offer explanations primarily from the perspective of propositional and spatial processes respectively. Section 1.4 will then present theories which assume reasoning results from non-logical heuristic processes.

1.3.1 Verbal Reasoning Hypothesis (VRH)

Polk and Newell (1995) offer a hypothesis which allows for the presence of both representational and rule-based processes. Polk and Newell claim that the theories discussed thus far are what they term transduction paradigms, that is, individuals initially encode premise information into some form of internal representation (be it a proposition, schema or model), upon which specialised reasoning mechanisms can operate before the results are decoded to produce a conclusion. They proposed that rather than being peripheral processes, coding and encoding actually play the central role in reasoning. Whilst it remains unclear whether untrained individuals possess sophisticated reasoning-specific mechanisms, their linguistic abilities are not in doubt and it is these abilities which are deployed adaptively to deductive reasoning tasks. As the same higher-level cognitive processes which serve to transform verbal to semantic representations and back again on an everyday communicative basis are also those which support deduction, the overall deductive process is termed *verbal reasoning*. However, the process must occur in a way which reflects the needs of deduction, rather than those of everyday communication. When this does not happen, errors associated with problem content, prior belief and Gricean implicature arise. Polk and Newell present a detailed computational model of verbal reasoning. The initial stage of the process is to construct a mental model of a situation in which the premise in question is true. The objects represented in the model are annotated by two additional pieces of information; a *not* flag indicating that the object does not have a specified property, and an *identifying* flag indicating that the object is identified by a specific property. For syllogisms, identifying properties correspond to the grammatical subject of the premise and are distinguished from other, secondary, properties by being more easily assessable, e.g. given the premise *All A are B*, the model distinguishes an A (identifying) who is a B, from a B (identifying) who is an A. However, there are often several ways in which a premise may be represented, and hence an annotated model may contain information which is not inherent in the original premise (an unwarranted model) or may fail to encode information which is inherent (an incomplete model). For instance, given the premise *Some A are B*, the model may contain unwarranted information if it also encodes the idea that S*ome A are not B.* Similarly, a model for the premise *No A are B*, may be incomplete if it fails to encode the fact that the secondary object (*B*) cannot be an *A*. Table 1.7 shows encodings for the VR model for all four premise moods. The table shows a default encoding, plus how they may be changed or augmented when information from a second premise is introduced.

Table 1.7: Encoding for verbal reasoning model (adapted from Polk and Newell, 1995, page 538).

Premise	Model	Augmented model
All A are B	(A' B)	all (A...) → (A' ...B)
Some A are B	(A' B) (A')	MR (A ...) → (A' ... B) (A' ...)
No A are B	(A' –B)	all (A ...) → (A' ... — B)
Some A are not B	(A' –B) (A')	MR (A ...) → (A' ... –B) (A'...)

Key: ' identifying property flag
- not flag
MR most recently assessed
... other properties

Once premises have been encoded, the verbal reasoning process attempts to produce a conclusion based on the annotated model. Conclusions about identifying properties are generated before those about secondary properties. Simple putative conclusions (which are true according to the annotated model) are generated and then tested to ascertain whether they form a legitimate syllogistic conclusion (i.e. whether they relate the two end terms using one of the four quantifiers). If a suitable conclusion is not found, then the premises are re-encoded and the reasoning process repeated. As a default, Polk and Newell assumed that conclusions of the form *All X are Y* or *No X are Y* will be proposed when there is an object with *X* as an identifying property, and all objects with the property *X* also possess the property *Y* or *–Y*. Conclusions of the form *Some X are Y* or *Some X are not Y* will be proposed where there is an object with *X* as the identifying property and at least one other object with the properties *X* and *Y/-Y*.

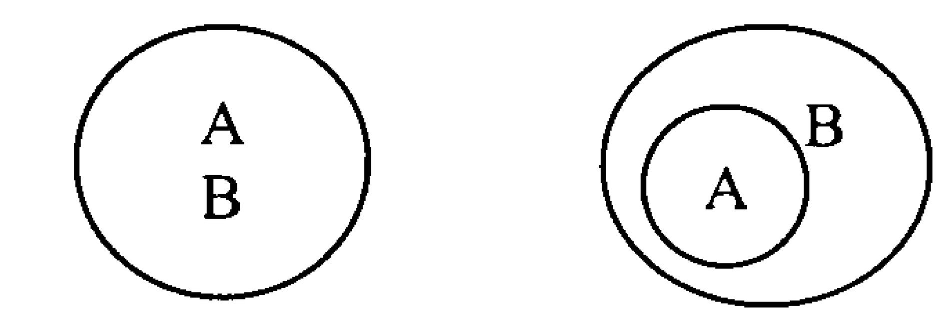
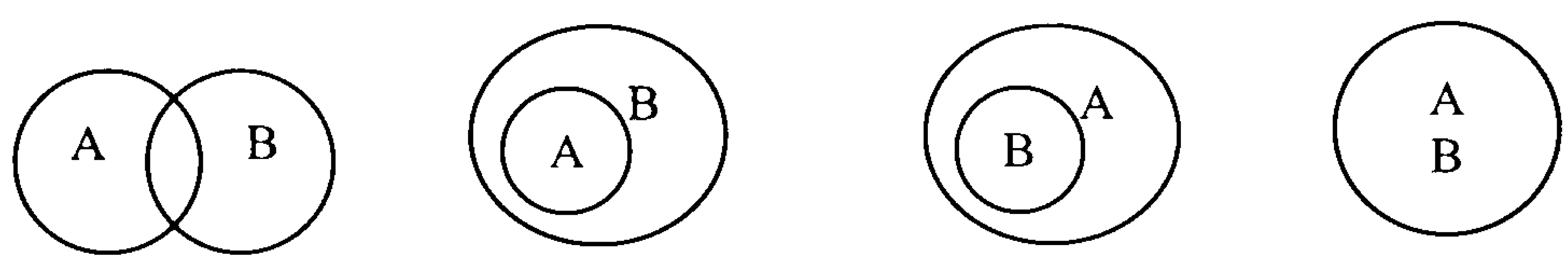
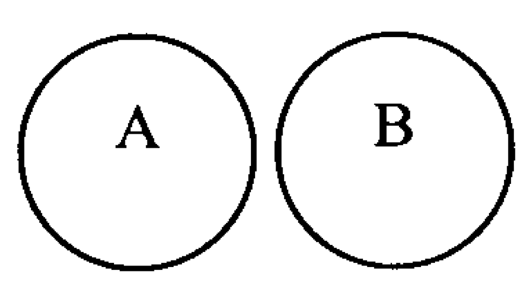
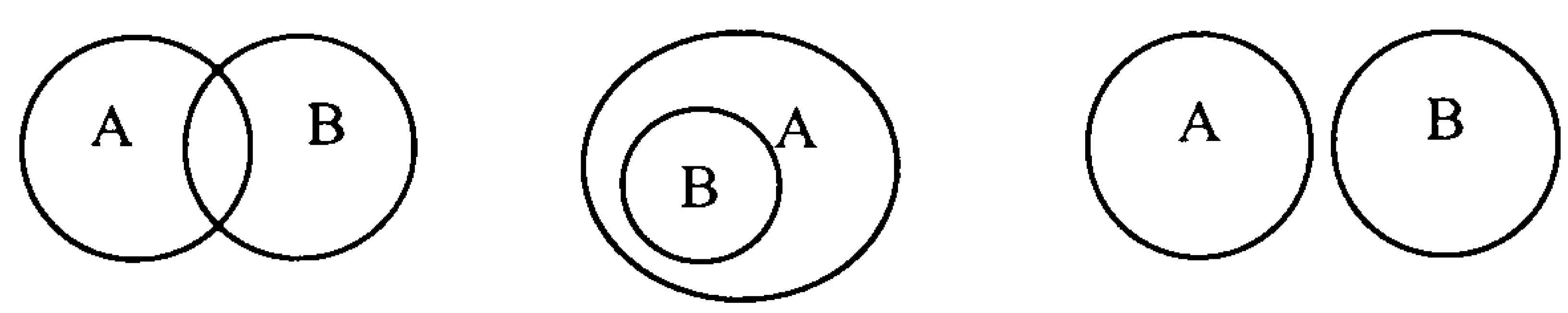
Polk and Newell present data collected over a number of experiments, with over 100 participants, to show that their computational model can account for much of the variation in syllogistic reasoning data produced by human subjects. This experimental data presented an overall accuracy rate of between 40-69%, compared to the 38-75% for the VR model. The constantly observed effects of atmosphere, conversion and figure are also predicted by the VR system, with an effect size within 10% of that shown by human subjects. Other syllogistic reasoning effects such as the number of NVC responses across various problem types, and relative accuracy on valid and invalid problems are also accounted for.

Polk and Newell also present detailed individual differences data to show that VR can account for strategic differences such as substitution and describe the system as building on ideas from both mental model and rules theories. It uses annotated mental models as a basic data structure, but encoding and re-encoding processes resemble the application of formal inference rules. As such, verbal reasoning comprises not only language processing but also supports the use of visuo-spatial representations and does not preclude the use of reasoning-specific mechanisms, especially in trained individuals. However, they deny that VR is a hybrid approach, its most important assumption (the one that distinguishes it from both rules and models theories) is that the central processes are borrowed form language comprehension and generation. Such processes are able to map between verbal and semantic representations, hence although the theory bears some resemblance to mental models (indeed Polk and Newell admit that their annotated models were based directly on mental models) no mental logic is required.

1.3.2 Euler Circles

Like models, Euler circles (Erickson, 1974; 1978) comprise mental analogues of syllogistic premises. Terms are represented by circles and syllogistic moods represented by particular arrangements of these circles as shown in Table 1.8. As such they explain how premises may be both interpreted and combined.

Table 1.8: Euler circle representations for syllogistic premises

Premise	Euler circle representations
A All A are B	 Identity relationship; Subset-set relationship
I Some A are B	 Overlap relationship; Subset-set relationship; Set-subset relationship; Identity relationship
E No A are B	 Disjointed relationship
O Some A are not B	

Erickson proposed a detailed process whereby premises are encoded as Euler circles, but individuals only have the capacity to handle one diagram per premise, hence given the premise *All A are B*, they will encode it either as a subset *or* an identity relationship.

Subjects then combine the two premise representations into one composite. As again it is assumed they can handle only one at time, some possible combinations may be overlooked and such omissions lead to errors. Indeed one of the most common criticisms of Euler circles is that they lead to a combinatorial explosion because many syllogistic premises call for several possible representations, and as such are beyond cognitive processing limitations. However, some studies have found that individuals can and do adopt more than one Euler model of a premise (e.g. Neimark and Chapman, 1975).

Another theoretical approach based on Euler circles was proposed by Guyote and Sternberg (1981). They also suggested set/subset representations were employed and that difficulty should depend on the number of diagrams needed to represent a combination of the premises. Their theory offered an explanation of how the premises are combined. Transitive chains are essentially links between end terms using the middle terms as a connecting device. A positive link is made when the first term has a positive connection with the middle term which then in turn has a passive connection with the third term. A negative (or ambiguous) chain is established when no positive link can be made via the quantifiers. Guyote and Sternberg present a model of transitive chain theory comprising symbolic representations based on Euler circles, as shown in their basic form in Table 1.9. It is notable that these representations bear resemblance to those in the verbal reasoning hypothesis of Polk and Newell (1995, described previously) despite transitive chain theory being based on a spatial approach to reasoning, rather than grounded in natural language. For equivalent Euler circle representations and syllogistic premise moods, see Table 1.8 above. As the Table 1.9 implies, the symbolic representation of each premise is comprised of components relating to the given terms (here, as in Table 1.8, these are A and B).

Table 1.9: Symbolic representations of set relations according to transitive chain theory (Guyote and Sternberg, 1981, page 468)

Set relation	Symbolic representation
Identity	$a_1 \rightarrow B$ | $b_1 \rightarrow A$ $a_2 \rightarrow B$ | $b_2 \rightarrow A$
Subset-set	$a_1 \rightarrow B$ | $b_1 \rightarrow A$ $a_2 \rightarrow B$ | $b_2 \rightarrow - A$
Set-subset	$a_1 \rightarrow B$ | $b_1 \rightarrow A$ $a_2 \rightarrow - B$ | $b_2 \rightarrow A$
Overlap	$a_1 \rightarrow B$ | $b_1 \rightarrow A$ $a_2 \rightarrow - B$ | $b_2 \rightarrow - A$
Disjoint	$a_1 \rightarrow - B$ | $b_1 \rightarrow - A$ $a_2 \rightarrow - B$ | $b_2 \rightarrow - A$

Once both syllogistic premises have been represented, inferences about the relationship between A and C terms can be made by integrating the two transitive chains that are formed from the components of the two representations. A transitive chain is formed from two components in which the first term in a component matches the second term in the other (e.g. an AB component and a BC components form an AB-BC transitive chain). A major criticism of Euler circles (e.g. Johnson-Laird and Byrne, 1991) has been that they cannot represent assertions containing more than one quantifier (e.g. All of the A are the same as some of the B), a situation which is common in everyday reasoning. Nor do they generalise well to natural language quantifiers such as *most*, or *few*. Sternberg and Guyote's version claims to overcome such difficulties, as well as that of combinatorial explosion. They claim that the simplicity of the representations and the rules applied in combining them avoid such difficulties. However, the theory also rests upon a strong assumption that premises are always encoded completely and correctly and cite evidence form an earlier study by Sternberg and Turner (1981) in support. However, given the ubiquity of systematic errors and biases in the literature, it seems unlikely that this is a practical tenet on which to base a competence model.

Stenning and Oberlander (1995) have proposed a theory which suggest that just one Euler circle type representation per premise may be required. They claim that representing premise information graphically reduces abstraction, i.e. it allows reasoners specificity in representing the information in an easily interpretable fashion. Such representations are claimed to be easier to process than abstract rules and hence reasoners which incorporate such methods are thought to be most effective. They compare this with mental models and claim that the two are in fact isomorphic. Stenning and Yule (1997) later claim that equivalent sentential implementations also exist. Model theorists would disagree, claiming that although Euler circles allow reasoners to model premises, they are remote from the actual perceptual structure of the situation. Mental models on the other hand, are claimed to be semantically meaningful discursive structures which may be embedded in a model of the real world (Johnson-Laird, 1983). Overall, experimental evidence does not wholly support Euler circle based approaches generally. Some of the syllogisms which reasoners find easiest actually require a large number of Euler diagrams. For instance, *Some A are B; All B are C*, requires 16 different Eulerian combinations, yet experimental data indicates it to be one of the easiest syllogisms (Johnson-Laird and Byrne, 1991, report a 88% accuracy rate).

1.3.3 Monotonicity Theory

Recently, an alternative and novel theory of syllogistic reasoning based on natural language semantics has been proposed (Geurts, 2003). The key tenet of this theory is that all quantified premises have a fixed pattern of monotonicity. By this, Geurts means that the four possible quantifiers can be seen as lying in a sequence, within which successive quantifiers are either monotone increasing or monotone decreasing. This idea is borrowed from linguistic theory which describes each member of a monotone increasing sequence as greater than or equal to the preceding member, and each member of a monotone decreasing sequence as less than or equal to the preceding member. In terms of syllogistic reasoning, in premises with an "*all*" quantifier, the first term is monotone decreasing and the second increasing (denoted by the symbol -/+), premises with quantifier "*some*" are +/+, with "*none*" are -/- and with "*some...not*" are +/-. From this implicit knowledge, it is therefore possible to substitute terms within syllogisms in order to arrive at the correct conclusion. For example, given the syllogism *All A are B ; No B are C*, in the second premise B is monotone decreasing. B can therefore be replaced by any term which implies it. From premise one, reasoners can assume that A implies B, and therefore can return to premise 2, replacing B with A to produce the correct conclusion, *No A are C*. Geurts also employs a number of other linguistic principles to illustrate how the relative difficulty of syllogisms are influenced by natural language assumptions, such as that of conversion (as described previously), that "*all*" implies "*some*", "*no*" implies "all are not", and also that dealing with "*some...not*" premises presents particular difficulty.

Newstead (in press) has supported the theory in terms of its grounding in linguistic principles, and in that it avoids representing terms by means of an arbitrary number of tokens. However, he also points out that data from the psychological literature confounds some of Geurts' findings. For instance, monotonicity theory predicts that quantifiers such as "most" will present more difficulty than "some" or "at least" because the former is linguistically marked. However, work by Bradon, Capon, Dennis, Evans, Handley and Newstead (2002) has shown no differences between these quantifiers. Although Geurts' linguistic model appears to be supported by the published data he cites, this may be, at least in part, attributed to the fact that those problems where monotonic substitution applies are also those which are generally known to be easiest to solve. The fit is less impressive when more difficult problems are considered. Although the data cited shows a variance of over 60% in accuracy, Geurts' theory suggests that these problems should be of equal difficulty. Overall the monotonicity theory remains underspecified in a number of ways; it does not extend to invalid syllogisms, nor explain figural effects. Nor does Geurts present a detailed and testable processing model which may be tested against research data. Moreover, (though it must be said that in this aspect Geurts is in line with most other theorists) the possibility of individual differences is not considered.

1.4 Heuristic Approaches to Syllogistic Reasoning

Other research suggests that reasoning performance may be explained by various heuristic processes triggered by interpretation of the quantifiers present. For the present purposes three approaches are considered, all of which make similar predictions about performance, but offer contrasting explanations for it, suggesting reasoning is based on either atmosphere, matching of quantifiers or probability.

1.4.1 Atmosphere

One of the earliest heuristic explanations was the atmosphere effect proposed by Woodworth and Sells (1935) and refined by Begg and Denny (1969). They argued that when two premises are of the same logical form (i.e. both are A, I, E or O), then the "atmosphere" of the syllogism makes it likely that a conclusion of that form will be assumed to follow. When the two premises contain different quantifiers, two supplementary hypotheses are required:

1. A negative premise (i.e. quantifier *No* or *Some...not*) creates a negative atmosphere, even when the other premise is affirmative (favouring a negative conclusion),

2. A particular premise (i.e. quantifier *Some* or *Some...not*) creates a particular atmosphere, even when the other premise is universal (favouring a particular conclusion).

So, for the syllogism:

All B are A
No C are B

the theory would predict a prevailing negative atmosphere and hence the conclusion *No C are A*. However, although the atmosphere hypothesis does seem to account for much observed reasoning performance (many valid syllogisms have conclusions which are consistent with atmosphere) it only describes patterns of performance and cannot show differences in difficulty between syllogism (Evans et al, 1993). Moreover, because it can, in theory, predict a conclusion for every problem, the atmosphere hypothesis fails to explain why participant produce "*no valid conclusion*" responses.

1.4.2 Matching

Wetherick (1989) and Wetherick and Gilhooly (1990; 1995) claim that when the logic of a problem is not immediately apparent, reasoners generate a response which has the appearance of logic by choosing a conclusion where the quantifier matches that of one of the premises. Where there is a choice of premise mood, they select that which is the most conservative, i.e. that which makes an assertion about the smaller proportion of whatever entities are represented by the subject term. As such, conclusion terms are preferred in the order E > O = I >>A. An A conclusion (quantifier *All*) may be chosen for the conclusion if both premises contain *All*, but if only one premise contains the quantifier a*ll*, and the other *some, no* or *some...not*, the chosen conclusion will contain one of the latter forms. To consider the syllogism again:

All B are A
No C are B

the matching theory would predict a conclusion with the logical form *No C are A.*

Hence the matching theory predicts the same conclusion for this syllogism as does the atmosphere effect. This is so in all but three instances (moods EI, EO and IO, where atmosphere predicts an O conclusion in each case whereas matching allows for conclusions matching the other moods, E or I, E and I respectively. Wetherick and Gilhooly (1995) compared the two. They re-examined data from a study by Sells (1936) which claimed that the atmosphere was a "potent factor" in deciding on a conclusion. They showed that the atmosphere did not predict the most frequent responses to syllogisms of the mood AA and AE and that in many cases, non-predicted responses were given as frequently as predicted ones. The data support matching just as strongly. Moreover, whereas, atmosphere claimed to account for reasoning of all individuals, matching theory claims only to account for the work of those who are unable, or unwilling to use more logical non-heuristic methods. Matching assumes that some people will attempt these methods, with varying standards of success. Wetherick and Gilhooly present a premise construction task whereby conclusions were presented and participants asked to construct premise pairs from which that conclusion could logically follow. Their findings clearly showed that although a significant number of subjects showed evidence of matching (by generating correct premise pairs, or incorrect but matching pairs), some subjects were also able to construct non-matching premise pairs, which were also logically valid. This suggested that these individuals were employing a more sophisticated logical strategy.

Many of the syllogisms which reasoners find easiest do have at least one premise where the mood matches the that of the correct conclusion. Such a conclusion may be obtained by matching, but, equally, some other strategy may be at work. Wetherick and Gilhooly (1990) report data which indicates that for syllogisms where a non-matched conclusion is correct, 28% of reasoners provided conclusions in accordance with matching theory, and this figure rose to 53% in the data provided by Johnson-Laird and Bara (1984). They offer this as evidence of matching by these participants and suggest that, as it is reasonable to suppose some consistency in the use of this approach, at least a proportion of correct conclusions to matchable syllogisms must also have been obtained by this method. Matching theory predicts most difficulty with syllogisms with non-matching conclusions. However, mental-model theorists have pointed out that the latter problems are also multiple-model. Bucciarelli and Johnson-Laird (1999) argue that if reasoners are responding simply to premise mood, then they should be unaffected by model count. Johnson-Laird and Byrne (1991) have shown that individuals are more likely to draw a conclusion matching one of premises for one model problems, than they are for multiple model problems. But then, as participants could equally be obtaining these conclusions by matching in the majority of cases, this argument seems somewhat circular. However, they also make the more salient point that if reasoners were governed by mood they would never respond "no valid conclusion". Wetherick and Gilhooly (1990) make a similar point, adding that although matchers in their study did produce this conclusion, they had been instructed that over half of the problems had no conclusion. Moreover, they frequently produced this response to valid problems, whilst participants who used a more logical approach did so only when it was the correct response.

1.4.3 Probability Heuristics Model (PHM)

The Probability Heuristics Model (PHM) proposed by Chater and Oaksford (1999) also predicts that reasoners will draw syllogistic conclusions of the same logical form as the least informative (i.e. most conservative) premise. However, rather than simply matching quantifiers, they argue that individuals use probability based heuristics based on their knowledge of the informational strength of the premises. This allows for the generation of the most informative valid conclusion. The theory presents some fairly complex computational detail, but in basic terms it proposes three main heuristics for conclusion generation:

G1: The *min-heuristic* suggests that the quantifier for the conclusion will be the same as that for the least informative premise (the *min*-premise). Almost all valid syllogisms follow this rule as shown by the matching theory above, producing a *min-conclusion.*

G2: P-entailments: some conclusions pr*obabilistically entail* (p-entail) further logical conclusions of a form different to either premise. For example, if *All X are Y*, then it is probable that *Some X are Y* (this follows as long as there are at least some Xs present). P-entailments are the next most preferred conclusions to those predicted by the min-heuristic. P-entailments include a family of heuristics corresponding to the probabilistic relationships between quantifiers.

G3: Attachment heuristic. Whereas heuristic 1 and 2 specify the end quantifier, the attachment heuristic specifies the order of end term in the conclusion. If the min-premise has an end term (i.e. either A are C) as its subject, this becomes the subject of the conclusion.

Chater and Oaksford (1999) present the following illustrations:

All X are Y	(*max*-premise)
Some Z are Y	(*min*-premise)
I type conclusion	(by *min*)
Some Z are X	(by *attachment*)

By the min-heuristic, the conclusion is I (some). The *min*-premise has an end-term (Z) as its subject, therefore by *attachment*, this will also be the subject of the conclusion, resulting in the conclusion, *Some Z are X.* In contrast, were the order of terms in both above premises reversed, the *min*-heuristic would still suggest an *I* conclusion but the *min*-premise would not have an end term as its subject. In this case, the end-term of the *max*-premise would become the end-subject instead, giving the conclusion *Some X are Z.*

A more complex example is shown below. In this case the valid conclusion is of a different form to that of either premise, hence the *min*-heuristic does not suggest the *probabilistically-valid* conclusion (type *O* or, quantifier *Some...not*).

Some X are Y	(*max*-premise)
No Z are Y	(*min*-premise)
O-type conclusion	(by *p-entailment, p-validity* or *logic*)
Some X are not Z	(by *attachment*)
E-type conclusion	(by *min*)
No Z are X	(by *attachment*)

In this case, use of the *min*-heuristic produces the incorrect conclusion, *No Z are X.* In order to generate the correct conclusion, reasoners have to then use *p-entailment* (if No Z are X, then it is probable that Some X are not Z). Alternatively, more skilled reasoners may reach this conclusion by use of logic. Two *test heuristics* are proposed to facilitate the assessment of the probabilistic validity of a putative conclusion:

T1: The *max-heuristic*: Confidence in the conclusion generated by G1 to G3 in proportion to the informativeness of the most informative premise (the max-premise).

T2: The *O-heuristic*: Avoid producing or accepting O conclusions (those with Some...not) as they are so uninformative relative to other forms of conclusion.

Chater and Oaksford conducted a detailed meta-analysis of five experiments that used all 64 syllogistic forms (two studies by Dickstein, 1978; two by Johnson-Laird and Steedman, 1978; and that of Johnson-Laird and Bara, 1984). In testing the predictions of PHM against this experimental data, conclusion types (including NVC) were examined across the individual syllogistic forms. PHM was found to account for over 80% of the variance in the data. The *min*-heuristic accounted for over 90% of conclusions drawn across both valid and invalid problems and a highly significant linear trend for conclusion preference in the order A > I > O = E (as predicted by the model) was also observed. The *min*-heuristic also predicts that the majority of such responses will not rely on *p*-entailment. On appropriate syllogisms across the five studies, over 50 *min*-responses were observed compared to between 5 and 13 *p*-entailed responses. In terms of the *attachment*-heuristic, conclusion order across the three Johnson-Laird studies (which required subjects to generate conclusions) found that in 54 out of 56 cases, the conclusion order conformed to that predicted by the *attachment*-heuristic.

Overall, the PHM argues that reasoning about the likely relationship between terms is based on the relative informativeness of the premises and claims that such simple, but rational, syllogistic reasoning strategies are justified by probability theory rather than formal logic. Chater and Oaksford (1999) present a detailed comparison between the PHM and the rules, mental models and verbal reasoning theories, showing that, in every case, PHM not only fits the data as well as any of these theories, but also provides a more accurate account of details including figural effects, NVC responses and effects of model count.

1.5 Concluding Remarks on Theories of Reasoning

Overall, reasoning is a wide ranging phenomenon, fundamental in everyday cognition, and which has attracted a range of differing theories all attempting to explain both the process people use to reason, and the source of the systematic errors and biases which are reported in the literature. Although reasoning has been investigated by means of a range of established experimental paradigms, this thesis will be fundamentally concerned with syllogistic reasoning, at least at the outset, and hence the theories reviewed above have focused particularly on attempts to explain reasoning on those problems. Where the programme of research diverges into other tasks, full explanations of those will be presented as appropriate. Rule and mental model theories have been designed to provide a unified account of reasoning across domains, but have been extensively applied to syllogistic reasoning. Other heuristic based theories have been developed with the specific goal of explaining syllogistic reasoning. The above theoretical review is not exhaustive, but presents a flavour of some of the main issues and constructs which will be addressed during the course of this thesis.

However, most syllogistic reasoning theories, and any models of reasoning data based around them, have assumed cognitive universality – that all individuals reason in basically the same way, for instance with logical rules, or with mental models, but not both (see for instance, Roberts, 1993; 2000, for further discussion of this point). Conversely, individual differences in *other* areas of cognitive function have been the subject of much research.

This thesis will be concerned with the investigation of such differences with respect to reasoning. Wetherick and Gilhooly (1995) have presented data which suggest such differences exist, certainly with respect to their matching hypothesis, and Section 1.7 later will deal further with strategic differences of this kind. Firstly however, Section 1.6 to follow, will introduce the concept of individual differences research generally, and outline some of the differences which have been observed in some key aspects of cognitive function thought to be central to human reasoning; working memory, verbal and spatial abilities and cognitive style.

1.6 Individual Differences

As the above review reflects, much research has been aimed at supporting the primacy of one reasoning theory over others. However, this current programme of work is concerned not with upholding one particular theory, but in investigating possible individual differences in the ways that people reason. This next section will introduce the concept of individual differences research and will outline some of the major research which has been carried out in this area. It will then develop this line of thinking further by considering evidence for qualitative differences in how individuals choose to reason, i.e. in the strategies they choose to adopt in attempting to solve reasoning problems. Finally, this chapter will conclude with an overview of the remainder of the thesis and the overall aims of the programme of research presented within it.

1.6.1 Introduction to Individual Differences Research

The individual differences approach tries to identify components which can be measured independently and in which there is independent individual variation. It offers an alternative way to examine which components are involved in cognitive tasks such as reasoning, and how they may differentially influence individual performance. Individual differences are readily observed across a range of abilities and capacities, including the ability to perceive and remember events and other stimuli, to use and comprehend verbal and propositional material (including language), to form concepts, to learn complex tasks, to represent and manipulate spatial information and not least, to reason, make inferences and arrive at conclusions. According to Carroll (1983), the study of individual differences is concerned with six basic and interrelated questions:

1. How individual differences can be observed, measured and described,
2. The dimensions of such differences,
3. Whether these dimensions are consistent across different populations,
4. The sources and causes of the differences,
5. The extent that differences can be modified by intervention, training etc,
6. How differences affect an individual in their everyday activities and social status.

Mental models and rules are certainly the two most influential explanations of how people actually attempt to reason, rather than rely on heuristics. Although the two theories present very different approaches and have relied heavily on different methodologies in providing supporting data (rules theorists preferring to present conclusions for evaluation whilst mental modellers tend towards production tasks), they do have some factors in common. One major similarity is the assumption that errors frequently result from limitations in cognitive capacity, particular in working memory (WM). Both schools of thought recognise that some individuals seem to have greater capacity than others, and that some are more competent reasoners than others. This may suggest that individual differences in cognitive capacity may be associated with similar differences in reasoning ability. The following section will present a brief review of research which has shown this to be the case, firstly with regard to working memory, then in terms of verbal and spatial abilities, and finally dispositional and cognitive style factors.

However, although the following section is ostensibly divided into three parts, they cannot be considered mutually exclusive. Many of the cognitive factors under investigation are inextricably linked, not only by association with various facets of WM, but also with that nebulous construct we refer to as *intelligence*. Hence it is worth attempting to define some basic terms concerning intelligence at this point. Intelligence itself is notoriously difficult to capture in a single, simple definition (see for instance, Kline, 1991). Baron (1998, page 110) describes intelligence as "those general mental abilities that help people achieve their goals, whatever their goals may be, in any real environment". As such it can be seen to be constituted from a whole range of cognitive processes, some of which are inherent and fixed (capacities/abilities), whilst others, such as cognitive styles or dispositions (Baron uses the terms interchangeably), may be learned and under conscious control. The psychometric measurement of intelligence was initiated by Spearman (1904) when he introduced the concept of general intelligence or, *g*, a factor thought to be common to all problem solving tasks. This approach remains to the present day, and as factor analytical techniques have developed, a wide range of abilities have been sampled and biological and neurological correlates of the *g* factor identified (e.g. Matarazzo, 1992). Cattell (1971) has broken down *g* into two forms, crystallised and fluid intelligence. Fluid intelligence (*gF*) is the basic reasoning ability of an individual, dependent largely on genetic and neurological factors (and hence is a fairly stable characteristic) whilst crystallised intelligence (*gC*) represents this ability as it is evinced in the particular culture or situation that the individual finds themselves in and may be evidenced by ability to accumulate knowledge through the lifespan and reflect on experience (and is therefore somewhat more malleable).

Fluid intelligence provides the innate ability for crystallised intelligence to develop, whilst crystallised intelligence combines with fluid in activities such as using prior experience to solve novel problems.

1.6.2 Working Memory

Working memory capacity is a factor long associated with reasoning performance (see for instance Gilhooly, 1998; Johnson-Laird, 2001). Baddeley (1986) describes a tripartite model of working memory (WM) comprising a controlling attentional system, termed the central executive which oversees and co-ordinates the operation of two slave systems. The phonological loop (PL) is concerned with processing linguistic material and a visuo-spatial sketchpad (VSSP) which is responsible for the construction and manipulation of visuo-spatial images. A considerable amount of research, primarily using dual-task methodology, has indicated WM involvement in an array of cognitive activities, including reasoning. Generally a decrement in performance is observed on a primary task when performed concurrently with a secondary task. The WM slave systems are thought to play a fairly passive role in higher level cognitive tasks, those which involve processing as well as storage of information. For instance, simple measures of PL capacity, such as digit recall span, do not predict accuracy in reading comprehension. However, more complex measures, which are assumed to also reflect the role of the CE, do predict performance on such tasks (e.g. Daneman and Carpenter, 1980; Gilhooly, 1998; Shah and Miyake, 1996). For a discussion of recent developments in the WM model see Baddeley and Logie (1999). This work will be discussed further in Chapter 5.

Limitations in capacity are believed to arise from the need to simultaneously satisfy both storage and processing demands. Reasoning is typical of tasks which place high demand on both of these aspects. The mental models theory of reasoning suggests a crucial role for working memory in the storage of initial models and of mental footnotes. These must be retained whilst further models are constructed. The information must then be accurately recalled and used during the fleshing out process. If stored material is subject to decay, and/or is inaccurately recalled, then errors will result. Similarly, in mental logic theories, storage of progress on temporary subgoals must be retained whilst the ongoing processing of logical rules takes place. Hence both theories place importance on WM capacity in determining competence. Tasks which tax ability to store verbal and/or spatial information produce data which can inform about the individual's capacity in respect to these systems. Similarly, tasks which require storage and processing concurrently may also tell us about CE capacity. Systematic individual differences in WM capacity have been found to be central factors in explaining individual differences in a array of cognitive tasks, including reading comprehension (Just and Carpenter, 1992), acquisition of logic skills (Kyllonen and Stephens, 1990) and of skills in the workplace (Gitomer, 1988) and more importantly for the current discussion, syllogistic reasoning (e.g. Gilhooly, 1998).

Engle, Kane and Tuholski (1999) describe working memory capacity as being the capabilities of the limited attention mechanism which Baddeley and colleagues have termed the central executive (CE). Kane and Engle (1999) also argued that WM capacity is isomorphic to the capacity for controlled processing, which in turn is shown to be strongly related to fluid intelligence (*gF*). This relates to the ability to solve novel problems and is putatively non-verbal and culture free (e.g. Horn and Cattell, 1967). Engle and colleagues have shown, through confirmatory factor analysis (CFA), that although performance on complex WM measures was related to *g*, performance on simple span measures was not. They argue that the CE is crucial to complex tasks, where processing, rather than storage, is the more important factor.

Kyllonen and Christal (1990) investigated the relationship between WM capacity and reasoning ability. CFA indicated the correlation between WM capacity and reasoning ability to be between .8 and .9 across four independent studies, using a wide range of measures for both constructs. Like Engle et al, their findings support the idea that general abilities reflect availability of attentional resources. Developmental work has also suggested that the early stages of skill acquisition (which requires a high degree of resource limited processing and hence WM capacity) depend on general abilities which are primarily reasoning ability or *g* (e.g. Gustafsson, 1984). Kyllonen and Christal also conclude that these general reasoning abilities reflect WM. However, they concede that their results are equally supportive of the hypothesis that WM capacity is primarily determined by reasoning ability. Capacity may be partly characterised by the size of storage buffers such as the PL. This capacity may in turn be managed through some kind of reasoning process. Hence, success in WM span tasks requires an ability to reason successfully about how to manage short-term storage resources. Such reasoning is likely to be a CE function and, as Baddeley (1986) has argued, many processes within the CE remain underspecified.

Kyllonen (1996) provides an extensive review of empirical findings which suggest that WM capacity is more highly related to reasoning performance and learning than any other cognitive factor. He concludes that WM capacity is, essentially, Spearman's *g*. Reasoning ability is traditionally "considered to be at the core of what is ordinarily meant by intelligence" (Carroll, 1989, page 56). It is also one of the primary determinants of the degree to which an individual benefits from instruction (e.g. Snow and Yallow, 1982) and is successful at accumulating lifelong crystallised knowledge. Overall, it would seem that individual differences in reasoning ability do reflect differences in WM capacity, which in turn seems to support claims made by both rules and mental model theories of reasoning.

Recent work by Capon, Handley and Dennis (2003) and Handley, Capon, Copp and Harper (2002) has indicated that individual differences in verbal and spatial working memory capacity are indeed associated with reasoning performance, and these findings will be discussed in more detail in Chapter 5.

1.6.3 Verbal and Spatial Abilities

Cognitive abilities may be defined as "any non-ephemeral characteristics of an individual that determine the level of performance on a cognitive task when maximal performance is attempted" (Carroll, 1983, page 4). By non-ephemeral, Carroll means that the characteristic is relatively stable, if it does change or develop, the process occurs only very gradually over considerable time. Cognitive tasks can comprise any activity which requires the mental processing of information, from simple everyday decisions to complex experimental tasks presented in the laboratory. Tests of cognitive ability typically consist of a series of tasks, either homogenous repetitions of trials in order to obtain fine data such as response times, repetitions of a task under differing conditions, or batteries of psychometric tests such as those typically used to measure intelligence. Statistical techniques such as factor analysis are frequently used to reduce the resulting data to a handful of cognitive factors which are thought to draw on the abilities being measured. However, such techniques can be limited if it is assumed that the required abilities for a given task are the same for all individuals.

Section 1.5.1 above has already discussed one aspect of cognitive ability, working memory capacity. Another aspect which has been extensively researched is verbal and spatial ability. To refer back again to the two main theories of human reasoning, rule theories emphasise the requirement to reason propositionally (which would seem to be a verbal skill) whilst mental models can be seen to be examples of more spatial representational approach. Other such as Polk and Newell's (1995) verbal reasoning hypothesis, suggest the possibility for both propositional and spatial modes of representation. Abilities in the verbal domain can encompass a wide range of skills from reading (in terms of both speed and comprehension) to vocabulary, grammatical and syntactic knowledge, to the ability to solve verbal analogy problems and manipulate information in a propositional form. Hence, no single task might be said to wholly and uniquely capture or define the verbal capacity possessed by an individual (Carroll, 1993; Pellegrino and Goldman, 1983). However, intelligence and aptitude tests have tended to concentrate primarily on vocabulary, comprehension and verbal analogy. Psychologists have frequently used scores on the Scholastic Aptitude Test (SAT), a test of supposed general intelligence which is used as an assessment of suitability for University entrance in the USA. The SAT includes a measure of verbal reasoning ability, and performance on this measure shows high levels of inter-correlation with other cognitive tasks which emphasise verbal skills.

Individual differences in verbal ability have also been closely associated WM. In an investigation of reading comprehension (which is of course a key skill in many reasoning tasks), Daneman and Carpenter (1980) presented participants with a typical comprehension task, where they were asked to read a passage of text and then answer questions on it's content and structure. They also presented a complex reading span test as a measure of functional WM span for language. This involved reading sentences whilst also keeping track of the final word of each sentence as to-be-remembered items. After a series if sentences, participants were asked to recall these items. Performance on this task has been repeatedly shown to correlate highly with performance on Verbal SAT (see for instance Daneman and Merikle, 1996, for meta-analysis). Daneman and Carpenter also examined results alongside verbal SAT scores. Those subjects who presented higher comprehension scores, also presented higher reading span scores and this WM measure also correlated significantly (r =.6) with SAT scores. The findings have since been replicated (for instance, Just and Carpenter, 1992). This seems to suggest that verbal WM ability (possibly facilitated by the PL) is a strong predictor of reading comprehension. However, this may also be facilitated by factors such as educational level, vocabulary knowledge etc which also involve different aspects of memory.

However, the same reading span test used by Daneman and Carpenter was also employed on a later study by Oakhill, Yuill and Parkin (1986). They tested two groups of children one of which performed poorly on tests of comprehension, but normally on reading and vocabulary, the other group performed at a normal level for their age on all three measures. The children in the latter group performed significantly better on the WM task than those in the low comprehension group. This difference increased in line with the number of to-be-remembered items. This, together with Daneman and Carpenter's (1980) findings, suggest that the two groups of children differed in the additional capacity offered by the CE. Just and Carpenter (1992) have also found similar associations between verbal WM and cognitive ability in other linguistic tasks. The above research informs the current thesis in that it indicates significant individual differences in verbal abilities, these in turn underpin the verbal skills necessarily required for reasoning, especially with linguistic material such as syllogistic premises. Chapter 4 will investigate verbal abilities with regard to individual differences in reasoning strategy.

Similarly, individual differences in the ability to work with material which involves spatial representation have been observed. Carroll (1993, page 309) defines this as the "process of apprehending, encoding and mentally manipulating spatial forms". Just and Carpenter (1985) presented subjects with a test of spatial ability followed by an experimental task which required them to compare two serially presented novel representations of the same object. Eye tracking measures showed that those with low spatial ability frequently had to rotate the object repeatedly during the comparison process, suggesting they had forgotten the previous representation. High ability individuals however rarely had to rotate the object more than once. This suggested that low spatial ability individuals may have difficulty maintaining spatial memory traces whilst concurrently processing transformations Therefore, it would seem that individual differences in spatial visualisation ability may be, at least in part, accounted for by differences in spatial WM capacity (presumably VSSP function). Later work such as that by Shah and Miyake (1996) and Capon, Handley and Dennis (2003) has also used the verbal WM span tasks designed by Daneman and Carpenter, together with spatial WM span measures, in investigating the relative contribution of the three working memory components during reasoning. Their findings strongly suggest individual differences in the roles of the verbal and spatial WM components.

Both verbal and spatial ability measures were employed by MacLeod, Hunt and Mathews (1978) in an investigation of individual differences in reasoning on a sentence-picture verification task (where subjects are first presented with a simple verbal description, followed by a picture, and have to decide whether the picture accurately represents the sentence). This work will be discussed in more detail in Chapter 3, but, in essence, two distinct groups of participants were identified. It was suggested that whilst one adopted a pictorial strategy, the other used a linguistic approach. Although the two differed in spatial ability (the pictorial group presenting higher levels) measures of verbal ability (which incorporated comprehension, vocabulary and verbal analogy) could not discriminate between individuals. Overall, and most importantly for the current thesis, when the research reviewed above is taken together, findings strongly suggest that individual differences in memory, verbal-spatial abilities and reasoning are inextricably linked.

1.6.4 Cognitive Style and Thinking Disposition

In addition to cognitive abilities, thinking dispositions and cognitive style, behavioural concepts which lie further towards the personality end of the spectrum, have also attracted attention as possible sources of individual differences. Sternberg (e.g. 1997) describes how individuals may possess preferred ways of thinking which are not in themselves abilities, but represent preferred ways of using abilities. Sternberg and Grigorenko (1997) state that these preferences are but one aspect of what has been termed "cognitive styles", characteristic, stable and typically preferred modes of processing information. Such styles have also been variously termed "thinking dispositions" (e.g. Stanovich and West, 1998; Perkins, Farady and Bushey, 1991), "intellectual styles" (e.g. Sternberg, 1988) and "information processing styles" (e.g. Klaczynski, Gordon and Fauth, 1997), these terms often being used interchangeably. According to Sternberg (1997), these styles are not in themselves cognitive abilities, but reflect how abilities are used. Hence, individuals may have similar abilities, but their styles of thinking may differ considerably.

Stanovich, West and Sá (1999) prefer the term thinking dispositions and suggest that these tend to generate characteristic behavioural tendencies and tactics during reasoning. They further propose that cognitive capacities and thinking dispositions are constructs at different levels of analysis in cognitive theory, hence their *joint* relationship to performance on thinking and reasoning tasks must be considered in attempting to explain the origin of errors. Baron (1988) and Perkins (e.g. 1993) make a similar distinction between cognitive abilities and style. For example, Baron refers to abilities as the processes underlying performance on psychometric and experimental tasks, such as perceptual speed and WM capacity - what Baltes (1987) has called the "mechanics of intelligence". These are stable characteristics inherent to the individual. In contrast, thinking dispositions are viewed as cognitive styles which are malleable, i.e. they can be improved/changed by practice or instruction. According to Baron (1988), thinking dispositions are those things which relate to the adequacy of belief formation and modification, perseverance, motivation etc and he comments that traditional psychometric measures do not tap into such dispositions. Stanovich (1999) takes a similar view and adds that as dispositions and abilities may differ in the degree of malleability and stability they afford, variance in reasoning performance may in fact be independently motivated by the two. Overall it might seem that whilst cognitive style refers to the preferred way that information is used in a cognitive processing sense, disposition relates to personality and situational factors such as motivation which may themselves influence style.

According to Sternberg and Grigorenko (1997) cognitive style is a broad concept which, although encompassing thinking dispositions, also includes other aspects which characterise preferred ways of processing information. Cognitive styles have been extensively investigated as a dimensional construct, whereby individuals are assessed and placed on a continuum of behaviour, examples include reflection vs. impulsively (e.g. Kagan, 1966), assimilator-explorer (e.g. Martinsen and Kaufman, 2000), rational vs. experimental (e.g. Epstein, Pacini, Denes-Raj and Heier, 1996), adapter vs. innovator (Kirton, 1976), visualiser vs. verbaliser (Paivio, 1971) to name but a few. Cognitive style has been associated with reasoning performance, including as a source of errors and belief bias (e.g. Pacini and Epstein, 1999; Handley, Newstead and Wright, 2000) and also with individual differences in a range of real-world behaviours including educational achievement (e.g. Perkins, Jay and Tishman, 1993), interpersonal conduct (e.g. Salmon, 1991), creativity (e.g. Baron 1988), reasoning about everyday problems (e.g. Perkins, Farady and Bushy, 1991; Schribner, 1986) and occupational choice and performance (e.g. Sternberg, 1997; Armstrong, 2000).

A substantial amount of research in this area has been carried out by Riding and colleagues (see Riding and Rayner, 1998, for a comprehensive review). Riding and Cheema (1991) introduced a computer based measurement tool which they term Cognitive Style Analysis (CSA). The CSA measures cognitive style across two dimensions, wholist-analytic (the tendency to process information wholistically or in parts) and verbal-visual (the tendency to represent information in words or pictures). Hence scores on the CSA place individuals in one of four cognitive style categories, according to their position on the two dimensions. Riding and Rayner (1998) review an immense range of studies in which CSA style category membership has been significantly associated with a variety of factors, primarily concerned with education and learning style across the lifespan. Their concept of cognitive style has also been associated with social behaviours (e.g. Riding, Burton, Rees and Sharratt, 1995), susceptibility to stress (Borg and Riding, 1993), challenging behaviour in school (Riding and Craig, 1995) and many other social and cognitive questions. Sternberg and Grigorenko (1997) present a useful and concise review of cognitive style research and its implications. More about the association between cognitive style and reasoning is presented in Chapter 4 of this thesis.

Thinking disposition is a closely related concept and the term is often used interchangeably with cognitive style (e.g. Baron, 1988). A major programme of research into the relationship between thinking dispositions and reasoning has been carried out by Keith Stanovich and colleagues, (e.g. Stanovich and West, 1997, 1998; Sá, West and Stanovich, 1999; the many studies reviewed in Stanovich, 1999). Stanovich and West (1998) report an extensive investigation of associations between a cognitive ability score (composite of several measures, including SAT, Raven's matrices and reading comprehension) and thinking disposition. The latter was measured by a questionnaire designed to tap into behaviours with potential epistemic significance, such as the tendency to weigh new evidence (or the opinions of others) against a favoured belief, or the disposition to spend a lot/very little time on attempting to solve a problem. Overall the questionnaire subscales measured epistemological absolutism, willingness to switch perspective, willingness to decontextualise and tendency to consider alternative opinions and evidence. Reasoning tasks including syllogisms and an informal argument evaluation task were also included. Stanovich and West report significant correlations between syllogistic reasoning performance and both SAT scores ($r = 0.4$) and thinking disposition composite scores ($r = 0.3$). However, it must be noted that these correlations are of moderate strength only. They are statistically significant, though this may be an artefact of the large number of participants ($N = 546$) in Stanovich and West's study. In some earlier research, Stanovich and West (1997) used an informal argument evaluation task (supposedly more indicative of everyday reasoning) where subjects were asked to assess the strength of rebuttals to arguments about everyday topics such as education and other social issues. They found that individual differences in the ability to evaluate arguments objectively (i.e. independently of prior belief) were reliably associated with cognitive ability (as measured by SAT score) and performance on informal argument evaluation. Moreover, actively open-minded thinking (a disposition connected with thoroughness in searching for possibilities, ability at self-reflection, criticism and revision of beliefs) was also found to predict argument evaluation score, even when individual differences in cognitive ability had been partialled out.

Another dispositional factor studied extensively by Stanovich and colleagues is decontextualisation, what Donaldson (1978) has called "disembedding", the ability to move thought processes away from the immediacy of the present context and apply them within other contexts which are not immediately apparent. In everyday life, such skills are required for thinking which involves abstract rule systems, such as in mathematics and logic (which of course some theorists claim are also required to solve syllogistic problems). The belief bias effect in reasoning occurs when decontextualisation does not take place, in other words, conclusions are placed on prior belief, rather than on logical validity. Details of this bias are beyond the remit of this review, but are well documented in the reasoning literature (see for instance Evans, 1989; Markovits and Nantel, 1989, Newstead, Pollard, Evans and Allen, 1992; Oakhill, Johnson-Laird and Garnham, 1989). Stanovich (1999) reviews a range of developmental studies which emphasise the importance of such skills to the development of higher level thought processes in children. Individuals who showed higher cognitive abilities also showed greater skills appropriate to decontextualisation on a syllogistic reasoning task, were less prone to making errors associated with belief bias (Sá et al, 1999). This supported earlier findings reported by Klaczynski (1997) and Klaczynski, Gordon and Fauth (1997) who also found that thinking disposition could predict belief bias in reasoning problems. These effects are explained in terms of the depth of processing employed by individuals who were able to reason beyond the cognitive boundaries imposed by pre-existing theories and beliefs. A recent study by Newstead, Handley, Harley, Wright and Farrelly (in press) has also examined individual differences in a range of deductive reasoning tasks, including syllogisms, together with measures of thinking disposition, ability to generate alternative representations (which involves decontextualisation) and intellectual ability (the latter included both verbal and spatial ability items among others). They found that intellectual ability was a good predictor of logical performance on syllogistic reasoning, especially where there was a conflict between logic and believability, and the ability to generate alternative representations proved an excellent predictor. However, other aspects of thinking disposition (as measured by a self-report questionnaire) were not associated with reasoning performance.

In summary, individual differences have been observed in a range of cognitive functions, including working memory capacity, verbal and spatial abilities and cognitive/thinking style. Research such as that of Capon et al (2003), Newstead et al (in press) and Stanovich and colleagues, has presented evidence for some clear associations between such differences and reasoning ability. Moreover, there is evidence that individual differences in cognitive ability may also predict differences in reasoning strategy (e.g. MacLeod et al, 1978).

1.7 Individual Differences in Reasoning Strategy

1.7.1 What *exactly* do we mean by strategy?

Section 1 of this Chapter touched briefly on the subject of strategies and presented a simple working definition of strategy as described by Johnson-Laird et al (2000). This idea of a strategy being composed of a series of tactics, echoes that of earlier definitions offered by Sternberg (for instance, 1982) who described the execution of a reasoning strategy as comprising a combination of processes which he termed *components*. Evans (2000) adds that strategy use is dynamic and flexible. Individuals can therefore choose whether or not to adopt a particular strategy, or can change strategy as task demands differ. Hence, he claims, strategies do not result from the operation of hardwired mechanisms, nor are they constrained by the demands of earlier learning. The notion of explicit conscious control is key to Evans' idea of strategy. Processes which are goal-directed and systematic, but which occur rapidly and without conscious thought or planning, are not strategic according to Evans, rather they arise from intuitive and unconscious reactions to stimuli. Evans' account of strategy emphasises awareness, the ability to be conscious of and to articulate a process, and intentionality, having choice and control over the process which is carried out, for instance if a certain strategy proves unsuccessful, they can decide to try another.

The issue of whether knowledge of available strategies, or when to use them, is conscious or not is avoided by other definitions such as that of Siegler and Jenkins (1989). They define strategies simply as being any procedure which is goal-directed and non-obligatory. Roberts (1993) offers an essentially similar explanation, a set of cognitive processes which have been observed in deductive reasoning tasks but for which there is insufficient evidence to suggest that these processes are themselves the fundamental reasoning mechanism. Roberts (2000) also makes an important distinction between *intrastrategic differences*, where all reasoners use a given strategy for a particular task, although they may vary as to the exact processes used (e.g. they all use some form of mental models but vary in exactly how they do so), and *interstrategic differenc*es, where individual differ in the strategies used on a task (e.g. some use mental models whilst others use rules). Research has indeed shown that individuals do adopt different strategies for reasoning which seem to fit Roberts' description quite well, as the following short review will demonstrate.

1.7.2 Rules or Mental Models...?

Chapter 1 has presented a brief review of some the main theories which have attempted to explain human reasoning. Two have been particularly influential; rule based theories (e.g. Rips, 1994a) suggest that individuals possess inherent knowledge of logical rules which can be applied to premises to generate a conclusion. Errors can be accounted for by the types and the number of rules which need to be applied to a given problem; mental models theory (e.g. Johnson-Laird and Byrne, 1991) suggests that reasoning involves the construction of mental representational analogous to the possible states of the world described by the premises. Errors occur as a function of the number of models required for given problems.

The two approaches share the assumption of a universal fundamental reasoning mechanism, used by all reasoners over all reasoning domains. Much research has been oriented to supporting claims of primacy and exclusivity, proposing that any fundamental reasoning process must be based upon either mental models or rules, but not both (Roberts, 1993; 2000). For instance, Johnson-Laird and Byrne (1991, p. x) in the introduction to their classic text on mental models, state "we have formulated the first comprehensive theory in psychology to explain all the main varieties of deductive reasoning". In contrast, Rips (1986) launched a scathing attack on what he termed "mental muddles" describing them as mere thought experiments which are vague and underspecified. Johnson-Laird (1989) in turn claimed that rules theories cannot explain how propositions relate to the world, and therefore the role of semantics in reasoning. Whereas a mental model can be judged to be true if it can be symbolically embedded in a model of the world, according to Johnson-Laird, rules or abstract propositions alone do not allow for an assertion of truth or falsity to be made. Rips however contests this argument by asserting that although mental models claim to account for content effects such as belief bias, because they are comprised of tokens (which have consistent properties) they are just as abstract as rules. In a later article (Rips, 1994b) reiterates his earlier view, and adds that Johnson-Laird and Byrne's (1991) claim that models become semantically meaningful because of their causal link with the world, could equally apply to other forms of representation, including propositional ones. Indeed, Johnson-Laird and Byrne (1991) state that mental models may not always take the form of explicit images and may "transcend the perceptible" (page 39). However, in cases where a premise contains negation, nothing in the perceptual world environment corresponds to the part of the model symbolised by "¬". Rips (1994b) argues that the abstract nature of this notion is evident from the fact that representation of such negated premises are often isomorphic to expressions in formal logic. Johnson-Laird, Byrne and Schaeken (1992) acknowledge as such, but go on to deny that this makes models theory equivalent to rules theory, claiming that any abstraction model requires a semantic interpretation which connect the model to the world.

Moreover, Rips claims that developmental research supports the idea of an inherent mental logic. Rules are likely to be innate as they are just the sort of processes which are thought to be central to cognitive development and learning. He claims that if mental models formed the basis of reasoning, they, or at least the quantifiers and connectives on which they are based, would have to be consciously learned before a great deal of the cognitive development which requires reasoning ability could occur. Indeed some developmental research, e.g. that of Crowley, Shrager and Siegler (1997) and Siegler (1996) has suggested that cognitive development in children can be understood in terms of reasoning ability, and interestingly, in term of the development of strategies, although neither of these studies place particular emphasis on either rules or mental model theories. Whereas rules theorist claim that they explain how individuals can reason across any domain about any subject, regardless of content, Model theory can be said to be more parsimonious, Rips has to introduce additional specialised components to extend his theory to syllogisms and to account for conclusion generation tasks, whereas mental model theory readily incorporate both these aspects. Moreover, model theory can predict not only the kinds of syllogisms where errors will be made, but also the nature of those errors. for instance, where fleshing out of an initial model is required, errors will be consistent with the initial model, and where multiple models are required, errors will be consistent with the final model produced. Johnson-Laird (1995) present evidence for both such effects, but they are beyond the capability of the rule models. Other rule theory proponents (e.g. Braine, 1993)have suggested that individuals may reason by use of both inference-rules and mental models (but exactly how this may occur also remains unspecified).

Several observers have questioned the value of the rules vs. model debate. Stenning and Yule (1997) have claimed that in essence, the two approaches reduce to the same thing and that both rules and models are in fact members of what they term "a family of individual identification algorithms" which are variously implemented, either in diagrams or propositionally. Hence, all reasoners are using the same reasoning algorithm, but individual differences arise in how the algorithm is implemented. Moreover, Oaksford and Chater (1995) have argued that mental models theory simply replaces a logic system based on syntactic proof methods with one based on semantic proofs. Similar claims have been made by Evans, Over and Handley (in press). It has also been suggested that neither theory is sufficiently well developed to be fully testable and therefore attempts to decide empirically between the two are futile (Evans and Over, 1996). Roberts (1993) has highlighted that the research strategies for both camps are similar, both resting on data from a specific task (involving either evaluation of generation of conclusions) and research by a small number of individuals committed to one or other cause. Overall, the models vs. rules debate remains unresolved and continues to present another of the dichotomies of opinion of which psychology seems so fond. See for instance, Johnson-Laird (1999) and Rips (1994a, 1994b) for discussions of the debate from the models and rules perspectives respectively.

1.7.3 ... Or Individual Differences in Strategy?

However, although such processes may indeed underpin deductive reasoning, differences have been observed whereby some individuals appear to use higher-level processes based on spatial models, whilst others employ strategies that are propositional in nature. Two major forms of representational strategy have been identified - spatial strategies whereby information is represented in a spatial configuration which corresponds to the environment (e.g. Gattis and Dupeyrat, 2000; Roberts, Gilmore and Wood, 1997) and verbal strategies whereby information is represented in the form of linguistic or abstract propositions (e.g. Rips, 1994a, b). Individual differences in the spatial-verbal distinction was apparent in reasoning on sentence-picture verification tasks where participants are presented with a series of trials in which simple sentences are followed by a diagram. The task is to determine whether the picture depicts the state of affairs described in the sentence. MacLeod, Hunt and Mathews (1978) found that strategic preferences on such tasks were associated with performance on independent tests of verbal and spatial ability. Moreover, a recent magnetic resonance imaging study (Reichle, Carpenter and Just, 2000) found differential neural correlates for verbal and spatial reasoning during sentence-picture verification.

Individual differences in strategy have also been identified on reasoning with linear syllogisms, which require deductions about the physical relationships between entities. Sternberg and Weil (1980) identified four strategies for this task, spatial, verbal, mixed and algorithmic (a task-specific heuristic approach). Again these were correlated with scores at verbal and spatial ability tests. Egan and Grimes-Farrow (1982) identified two distinct strategies for linear syllogisms which, although not described in terms of a verbal-spatial distinction, presented elements which suggested either pictorial or abstract approaches were being used. Strategies observed during solving of linear syllogisms and in sentence-picture verification will be discussed further in Chapter 3, which presents a specific experimental study involving these tasks.

Similar verbal-spatial strategy differences have been observed in individuals solving a compass-point direction task. On this task, subjects have to decide where a person following a set of compass point directions (e.g. one step north, two steps south ...) will finish up in relation to a given starting point. The natural strategy here appears to be spatial, representing the entire path and hence the spatial relationship between the final location and the start. However, some people adopt a heuristic cancellation strategy whereby they cancel out opposite steps and those which remain constitute the correct answer. A running total of cancellation would appear to be less demanding on WM than remembering the whole path, and this strategy reduces both response time and error rate significantly. Roberts et al (1997) reported not only verbal-spatial differences, but also evidence of ability to develop and switch strategy according to task demands.

However, the above research, like much of that concerned with differences in cognitive ability (see section 1.5.3 previously), have relied on standardised tests of verbal and spatial ability. These may themselves be subject to use of different strategies (e.g. Just and Carpenter, 1985). Lohman and Kyllonen (1983) have observed that individuals with higher spatial abilities as measured by such tests, seem to have a greater range of strategies at their disposal generally, and do not always use a spatial representational approach when solving reasoning problems. Similarly, Roberts et al (1997) found that low spatial ability individuals failed to realise that there were any alternatives to the spatial approach. They found this strategy onerous but continued with it, even though cancellation would have saved them much effort. High spatial ability individuals however, tended *not* to choose the spatial strategy (even though they had the aptitude for it), instead identifying the savings of effort that cancellation afforded. Moreover, on more difficult problems where cancellation was less effective, they could change back to a spatial strategy in order to successfully complete the problem. Overall, high ability individuals seemed to have more choices available to them and possessed a cognitive flexibility not observed in the low spatial ability reasoners.

1.7.4 Strategies for Syllogistic Reasoning

Few studies have attempted to investigate possible inter-strategic differences in syllogistic reasoning and it remains unclear whether a natural or prevalent strategy for solving such problems exists (Roberts, 2000). One significant study involving categorical syllogisms has been conducted by Ford (1995). Participants were asked to generate conclusions to all 27 valid syllogisms as identified by Johnson-Laird and Byrne (1991). Although the main aim was to show that cognitive representations assumed in mental models theory were inadequate to account for reasoning performance, one of the most interesting findings concerns the identification of individual differences in reasoning behaviour. Ford presented extensive verbal and written protocol data to show that, of the 20 subjects, all but four appeared to display one of two types of behaviour. Those which Ford termed *spatial reasoners* used shapes such as circles or squares placed in different spatial relationships to represent the relationship between the terms in the premises. They provided written protocols illustrating such models diagrammatically and their verbal reports frequently described these relationships in terms of group membership, sets and subsets.

Ford understandably likens the spatial diagrams to traditional Euler circles, diagrammatic models representing set membership and the relationship between sets (see Table 1.4). Interestingly, these are an example of reasoning-specific strategies which Polk and Newell (1995) claim are unlikely to be employed by untrained individuals. Erickson (1978) proposed that reasoners use a mental analogue of Euler circles in syllogistic tasks, representing each premise in turn and then combining the two diagrams to represent the relationship between all three terms in the syllogism. However, Ford described how spatial reasoners represent the relationship between terms in the first premise and then add information about the third term to that representation in a way which reflected its relationship with the common term. The other main strategy group Ford identified are referred to as *verbal reasoners*. Ford claims these showed no evidence of using mental models at all. Rather, they displayed various types of substitution behaviour, that is, they replaced the middle term from one premise with the end term from the other premise to reach a conclusion, as if solving an algebraic problem. Their protocols seemed to suggest that they represented the premises propositionally and applied simple heuristic rules to reach a conclusion.

But are the protocols presented in Ford's work actually evidence of individual differences in strategies? To return to Evans (2000) and the issue of conscious control, at some level these participants have selected their strategy, but it seems unlikely (indeed Ford presents no evidence to suggest) that they *consciously* weighed up the benefits of one strategy against another and *consciously* chose to use the strategy they did. Rather they approached the problems they were given in what appeared to them to be the best way they knew. However, equally, there is nothing to suggest that they *could not* have adopted other strategies, had they been aware of that possibility. In this sense the strategies were non-obligatory as suggested by Siegler and Jenkins (1989). This may by the crux of the conscious control issue – reasoners cannot make conscious choices about which strategy to use unless they have had some experience (or training) which makes them aware that there are an array of possibilities to choose from. In logically naive individuals, this may rarely be the case (however, also see findings of Roberts et al, 1997, discussed previously).

Hence, for such individuals, the strategy they present may be instinctive. It may not always be the most logically appropriate for the task, but something within the individual impels them towards that particular way of working. It is these issues that the current thesis aims to address. As Roberts (2000) points out, most of the studies which have considered strategies have been concerned with identifying strategies rather than explaining why people differ. The existence of individual differences in strategies, i.e. qualitative differences in *how* people reason, has important implications for what has at times been a heated debate between proponents of mental model and rules theories, which may help to explain why such differences remain largely under-researched.

Ford's remains the most detailed study into possible inter-strategic differences in syllogistic reasoning and she presents some useful insights into individual differences in strategies. She rightly concludes that the data on syllogistic reasoning are complex and there are many subtleties in the strategies used which have not been explored. Her findings seem to suggest that individual differences are present and that these are not taken into account within the prevalent mental model or rule based theories. Roberts (1993; 2000) agrees with Ford that strategies remain underspecified and remain a matter for investigation. As Roberts (2000) points out, the number of individual differences studies is very small, and replications are desirable. Overall, the way in which individuals differentially represent and manipulate information in reasoning is a neglected area that merits further research. Thus provides the rationale for the programme of work reported in this thesis.

1.8 Overview of Thesis

Roberts (1993, 2000) has argued cogently that the issue of individual differences in the strategies people use for reasoning is a neglected area which merits further research. The experimental programme outlined in the coming chapters of this thesis offers an investigation into this issue. The broad, overall aims are:

1. To establish further evidence for the existence of individual differences in reasoning strategies,
2. To investigate the nature of such differences,
3. To determine whether they are a product of individual reasoning tasks or whether they are inherent to individuals and hence universal across reasoning domains,
4. To investigate some of the factors which may give rise to such differences.

The outcomes of such aims have clear implications for our understanding of how people reason, for the interpretation of experimental data (both past and present) and hence for theories and models of human reasoning based upon such data.

Chapters 2 to 5 present a series of experimental studies which were carried out in pursuit of the four aims outlined above. Chapter 2 presents two major individual differences studies. Experiment 1 replicates and extends the work of Ford (1995) with a view to ameliorating some of the methodological concerns surrounding her original study. Experiment 2, extends this work to consider invalid syllogisms and both the production and evaluation forms of the syllogistic reasoning task. This reveals more about the nature of strategies adopted. In Chapter 3, Experiments 3 and 4 continue this line of research into two further reasoning paradigms, linear syllogisms (transitive inference) and sentence-picture verification. These tasks were briefly outlined earlier, further details and background are presented in Chapter 3. Chapter 4 presents an experiment which begins to consider some of the possible origins of strategic differences, specifically factors of cognitive ability and style. The concluding study in the programme is described in Chapter 5. Here individual strategies are considered in terms of how they may draw differentially on working memory resources. The final chapter, Chapter 6, will present an overall discussion of experimental findings across the six studies, their implications for human reasoning within both applied and theoretical settings and offers suggestions as to future directions for research.

CHAPTER 2

STRATEGIES IN SYLLOGISTIC REASONING

2.1 Introduction to Experiment 1

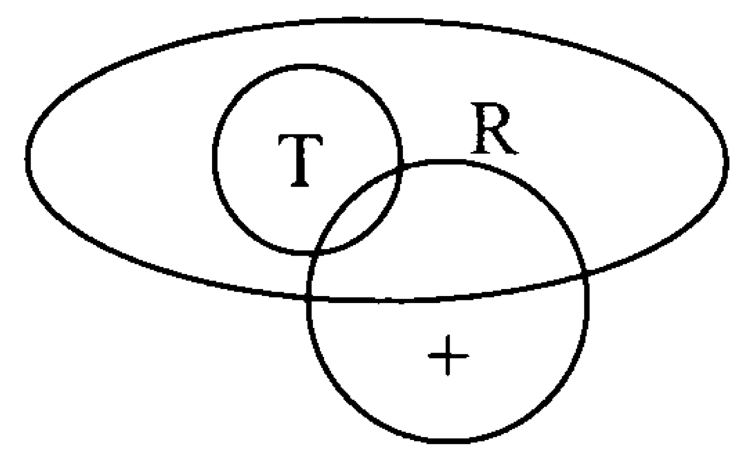
Chapter 1 has reviewed some of the main theories of human reasoning and discussed how individual differences in cognitive abilities may be associated with individual differences in reasoning strategies. In doing so, Section 1.6.4 referred to the work of Ford (1995) which presented detailed written and verbal protocols to show that while some individuals appeared to use a visuo-spatial strategy for generating syllogistic conclusions, others seemed to employ a verbal-propositional approach. As this thesis is primarily concerned with the investigation of individual differences in reasoning strategies, and as Ford's paper is one of the few which explicitly advocates this possibility, her findings deserve further consideration both as an introduction to the current programme of research, and as a foundation for Experiment 1.

Ford proposed that the written and verbal protocols produced by her participants present evidence about the nature of their individual reasoning strategy. Those individuals which Ford termed *spatial reasoners* used shapes such as circles or squares placed in different spatial relationships to represent the relationship between the terms in the premises. They provided written protocols illustrating such models diagrammatically and their verbal reports frequently described these relationships in terms of group membership, sets and subsets. Figure 2.1 below presents Ford's example 73:

Figure 2.1: A typical protocol produced by spatial reasoners (Ford, 1995, page 67)

Syllogism: All of the teetotallers are reporters
Some of the artists are not reporters.

Written protocol:



Verbal protocol:

"so if all of the teetotallers are reporters the they're inside then that that's a smaller subset of reporters then some of the artists are not reporters, that means if some of them are not reporters then there are definitely some artists that are not teetotallers".

Correct conclusion: Some of the artists are not teetotallers.

Ford identified six syllogisms where she claimed there was just one way in which this relationship could be represented. These largely correspond to the single-model syllogisms identified by mental model theory (Johnson-Laird and Byrne, 1991). Ford described these as having *totally constrained* representations and claimed that spatial reasoners perform best on these problems. The remaining syllogisms, Ford described as having *less-constrained* representations and she identified multiple ways in which the premises might be represented. The majority of Ford's less-constrained syllogisms would be described as multi-model according to Johnson-Laird and Byrne with just four (those containing *All-some* premises) classed as single-model. Ford presents data to show that her spatial reasoners performed less well on less-constrained problems though interestingly, the difference was not significant when the comparison involved only single mental model problems.

Ford understandably likens the spatial diagrams to traditional Euler circles, diagrammatic models representing set membership and the relationship between sets (see Table 1.8). Interestingly, these are an example of reasoning-specific strategies which Polk and Newell

(1995) claim are unlikely to be employed by untrained individuals. Erickson (1978) proposed that reasoners use a mental analogue of Euler circles in syllogistic tasks, representing each premise in turn and then combining the two diagrams to represent the relationship between all three terms in the syllogism. However, Ford described how spatial reasoners represent the relationship between terms in the first premise and then add information about the third term to that representation in a way which reflected its relationship with the common term.

The other main strategy group Ford identified are referred to as *verbal reasoners*. Ford claims these showed no evidence of using mental models at all. Rather, they displayed various types of substitution behaviour, that is, they replaced the middle term from one premise with the end term from the other premise to reach a conclusion, as if solving an algebraic problem. A typical example (Ford's example 24) is presented in Figure 2.2.

Figure 2.2: A typical protocol produced by verbal reasoners (Ford, 1995, page 16)

Syllogism: All of the secretaries are football fans
None of the soccer players are secretaries

Written protocol: All of the secretaries are football fans
None of the soccer players are secretaries

Verbal protocol: " ...when you have all or none then you can replace something like if all the secretaries are football fans then you can replace this secretaries with this secretaries and since all the secretaries are football fans means none of the soccer players are football fans..." [excerpt only presented as this protocol was very lengthy].

Correct conclusion: Some of the football fans are not soccer players

These protocols suggested to Ford that verbal reasoners represented the premises propositionally and applied simple heuristic rules to reach a conclusion. Although Ford does highlight cases of what she termed "naive" substitution" (page 25), where terms were replaced literally and without regard for quantifiers or logical relationships (Figure 2.2 above is a good example), she claims that the majority (7 out of 8 verbal reasoners in her sample) showed awareness of logical principles (for instance that *All A are B* is not logically equivalent to *All B are A*) on at least some of the problems. According to Ford, such awareness is evident from statements to that effect in the verbal protocols or by simply generating a correct conclusion to at least one syllogism where such principles apply. Ford concluded that verbal reasoners were employing a "simple substitution" strategy which involved the application of the following logical inference rules (page 21):

A. If a rule exists affirming of every member of the class C the property P then:

- *i. whenever a specific object O, that is a member of C is encountered it can be inferred that O has the property P, and*
- *ii. whenever a specific object O, that lacks property P is encountered it can be inferred that O is not a member of C*

B. If a rule exists denying of every member of the class C the property P then:

- *i. whenever a specific object O, that is a member of C is encountered it can be inferred that O does not have the property P, and*
- *ii. whenever a specific object O, that possesses the property P is encountered it can be inferred that O is not a member of C*

In general terms, Ford compared verbal reasoners' performance on syllogisms where she claimed such simple substitution rules could be directly applied, and those where the logical form of the problem made this more difficult. Here she proposed that verbal reasoners use "sophisticated substitution" - a process which involved the reformulation of one of the premises in order that the simple rules can be applied (for instance, realising that *All A are B* implies that *Some B are A*), or modification of both premises (for instance, combining them into a single statement) in order to establish the relationship between the two end terms and the common term. Ford generally found significant detriment in performance for verbal reasoners on the syllogisms requiring sophisticated substitution (see Ford, pages 25-34 for a more detailed exposition of these analyses). Moreover, Ford made distinctions between problems where the *i* and *ii* versions of her rules apply. She likens these versions to modus ponens and modus tollens respectively – inference rules recognised by logicians for centuries. Much syllogistic research has shown that the latter are notoriously more difficult to apply and Ford reports significant detriment in verbal reasoners' performance on syllogisms which she claims require *Aii* or *Bii*. Throughout this complex analysis Ford commented on the relative homogeneity of spatial reasoner performance, and suggested that the variations in performance observed in the verbal group stemmed from the differing application of rules which are irrelevant for spatial reasoners.

However, some methodological issues are apparent with the Ford (1995) study. Firstly, the small sample size (total N=20) must raise some doubt about generalisability of findings. Some of the variations in performance which Ford presents are very small - for instance, the difference between spatial reasoners' performance on constrained and less-constrained syllogisms does not reach significance. Given the sample size of just 8 participants in each strategy group, such differences might not be strong evidence for strategy-driven disparities in performance. The other major issue concerns the use of verbal protocols. Chater and Oaksford (1999) claim that individual differences such as those described by Ford occur as a function of employing verbal protocol methodology. They suggest that reasoners are forced to change their strategy when a protocol has to be generated and this leads to subtle individual differences in the heuristics used. Ericsson and Simon (1980, 1993) have written at length about the implications of using verbal protocols as data and present evidence that asking subjects to think aloud may produce additional cognitive workload which alters both the course and nature of the processes under observation. Memory limitations may influence the accuracy and completeness of retrospective reports, and data may reflect a subject's inferences about the research question and their own meta-memory, rather than comprise an accurate account of their mental events. Also, as Johnson-Laird et al (2000) point out, some cognitive processes involved in reasoning may occur unconsciously and be unavailable for articulation. Bucciarelli and Johnson-Laird (1999) present written protocol evidence which appears similar to that in Ford's work, however, the corresponding verbal accounts were wholly uninformative.

However, Larkin and Simon (1987) have suggested that is valid to treat internal and external spatial representations as equivalent. They define mental imagery as "...the use of diagrams and other pictorial representations that are not stored on paper but are held in human memory" (page 97) and also claim that "...the creation of a mental image (for instance from a verbal description) employs inference processes like those that make information explicit in the course of drawing a diagram" (page 98). This might suggest that the protocols drawn by participants in Ford's study are, at least in part, representative of the mental processes at work during reasoning about syllogisms.

Overall, Ford presented important evidence for the existence of both verbal and spatial strategies for syllogistic reasoning. Given the paucity of research into individual differences in reasoning strategies generally, her findings provided a useful foundation for the current programme of research. Experiment 1 presents a replication and extension of Ford (1995) aimed at overcoming some of the methodological criticisms levelled at her research. It also offers an incipient investigation into whether insight into strategy choice can be reliably obtained by alternative, less laborious, methods which avoid the possible pitfalls of verbal protocol.

2.2 Experiment 1

2.2.1 Aims

Experiment 1 had two main aims:

1. Given the paucity of individual differences research, the main aim of experiment 1 was to replicate and extend the work of Ford (1995) using a larger sample. It was predicted that evidence of two main reasoning strategies would emerge and that these would be identifiable as being verbal and spatial in nature, according to Ford`s criteria. Strategy was not expected to influence overall task performance.

2. To conduct an initial investigation into whether insight into strategy choice could be reliably obtained by means of alternative, less laborious, methods which avoid the possible pitfalls of verbal protocol. To this end, a second experimental condition was introduced, over and above the replication, where no verbal protocols were collected. Also, a questionnaire was developed to be completed by all participants. This comprised 13 items, gathered both qualitative and quantitative data about participant's strategy choice and was designed to identify the types of reasoning behaviours which Ford associated with verbal and spatial strategies.

2.2.2 Methods

Participants

A total of 51 undergraduate students from the University of Plymouth volunteered to take part in the study in return for course credit. This sample comprised 7 males and 44 females with a mean age of 22.82 years. None had received formal training in logic.

Materials

Syllogistic Reasoning Test Items

In line with the procedure adopted by Ford (1995), participants were presented with a set of 27 thematic syllogisms generated from the 27 valid forms identified by Johnson-Laird and Bara (1984) and Johnson-Laird and Byrne (1991). As in Ford's (1995) study, the common terms were names of people with given occupations (e.g. lawyers or librarians) and the end terms were names given to people with certain hobbies or persuasions (for instance vegetarians or beekeepers). In each case, the two premises were presented and participants asked to generate the appropriate conclusion. Each participant was given a booklet containing the appropriate instructions (see procedures for each condition) followed by two practice items. The test items were presented in a different random order within each booklet, each on a single page, having space below for written notes. The 27 syllogisms are presented, along with their logical conclusions, in Appendix 2A.

Reasoning Behaviours Questionnaire

Participants were also presented with a Reasoning Behaviours Questionnaire. The 13 items comprised a mixture of quantitative and qualitative items designed to identify the types of reasoning processes employed without support from verbal protocols. The items were developed in line with the types of reasoning behaviours which Ford described and associated with verbal and spatial strategies (as discussed previously). The questionnaire is presented in Appendix 2B, together with indication of which items were intended to identify which strategy. Experiment 1 acted as a pilot for the questionnaire.

Procedures

Participants were randomly allocated to one of two experimental conditions.

Condition 1: Replication of Ford (1995) (N=19)

Participants were run individually in a single session lasting around 1 hour. As this condition was designed to be a replication of Ford's (1995) study, her procedure was followed as closely as possible. The written instructions in Table 2.1 duplicate those presented by Ford.

Table 2.1: Written instructions presented to participants in the replication condition

This is an experiment on how people combine information in order to draw conclusions from it. You will be given a series of pairs of statements about different groups of people. You are to read the statements and figure out what, if anything, follows necessarily from these premises about the people. Your conclusion should be based solely on the information in the premises, and not on plausible suppositions or general knowledge. Because we are trying to find out how people solve these problems, it is VITAL that you "think aloud" while working out your answer, so there should not be any silent periods on the tapes. Also if you would like to use pen and paper to help you come to your answer, please do so. When you have reached your conclusion, simply state your conclusion.

There are 27 pages in the booklet you will be given. Each page has a pair of statements about different groups of people on it. You are to read the statements aloud and then proceed to work out what conclusion, if any, follows from the premises. Remember it is important to "think aloud": don't sit there being silent. Feel free to use a pen and paper. You can use the space below the statements. Once you have finished with the first pair of statements go to the next page and the next problem. Take as much time as you need to figure out the problems.

Participants were then asked to complete 2 practice items which comprised generating conclusions to two syllogisms of the type presented in Appendix 2A, whilst also producing concurrent written protocols. Whilst generating conclusions to the 27 test items, they also provided a verbal protocol of their reasoning which was recorded on a standard cassette tape recorder. Following this, they were asked to go through the 27 items a second time, explaining to the researcher how they approached the task in each case, producing written notes in the test booklet illustrating their explanation (regardless of whether or not they had chosen to write anything during their first attempt). These verbal explanations were also tape recorded. Finally, they completed the questionnaire as described previously and shown in Appendix 2B. 20 participants were originally allocated to this condition in line with Ford's sample size. However, one participant was unable to produce any written accounts of her working out, or explain her thinking, and hence her data was dropped from the study.

Condition 2: Written protocol only (N=32)

Again participants were run individually with each single session lasting around 45 minutes. In this condition, participants worked through the same 27 items but no verbal reports were collected. Instead, they were specifically asked to illustrate their reasoning by making written notes in whatever format they felt appropriate. The written instructions for these participants are presented in Table 2.2 below.

Table 2.2: Written instructions presented to participants in the written protocol only condition

This is an experiment on how people combine information in order to draw conclusions from it. You will be given a series of pairs of statements about different groups of people. You are to read the statements and figure out what, if anything follows necessarily from these premises about the people. Your conclusion should be based solely on the information in the premises, and not on plausible suppositions or general knowledge. Because we are trying to find out how people solve these problems, it is VITAL that we have a written record of your work. We would therefore like you to USE PEN AND PAPER to help you come to your conclusion. Feel free to write down anything that helps you. When you have reached your conclusion, simply state that conclusion clearly in *writing.*

There are 27 pages in the booklet you will be given. Each page has a pair of statements about different groups of people on it. You are to read the statements and then proceed to work out what conclusion, if any, follows from the premises. Remember it is important that you write down your workings out. You can use the space below the statements to write in. Once you have finished with the first pair of statements go to the next page and the next problem. Take as much time as you need to figure out the problems.

These instructions were designed to be as consistent as possible with those used in condition 1, differing only as was necessary to accommodate the revised procedure. Participants in this condition also attempted the two practice items and completed the questionnaire.

2.2.3 Results

2.2.3.1 Identification of strategy types

The descriptions of verbal and spatial strategies presented by Ford (1995) suggested that their written protocols show very specific characteristics. Using these characteristics as classification criteria, a visual examination of the written protocols produced in the present study clearly indicated the presence of both verbal and spatial reasoners. In addition, participants in the replication condition (V/WP condition) produced verbal protocols which provided supporting evidence of individual differences in strategy choice. Verbal reasoners frequently referred to actions such as replacing, substituting and cancelling syllogistic terms, while spatial reasoners often described the terms, and their inter-relationships, as groups or subsets. However, unlike the written protocols, not all verbal reports provided substantial evidence in themselves. Quality varied considerably both between individuals and between syllogisms and, in some cases, there was insufficient data to clearly identify the strategy used without additional support from written protocols. Participants in the written protocol only (WP-only) condition produced written protocols of a very similar nature to those in the V/WP condition and hence could be categorised into strategy groups without the necessity for supporting verbal data.

In the majority of cases the data were highly consistent, participants producing typical verbal or spatial style protocols for all syllogisms. However, 5 participants produced protocols which suggested somewhat mixed behaviour, for instance using what appeared to be verbal written representations (arrows, mathematical symbols etc) but presenting the information spatially, or producing verbal style protocols for some problems and spatial style ones for others. The distribution of verbal and spatial reasoners across both experimental conditions in shown in Table 2.3. As so few mixed reasoners were identified, they are not included in this analysis. Pearson's Chi square suggested no association between condition and strategy type employed; $\chi^2 = 0.71$, df = 1, $p > 0.05$.

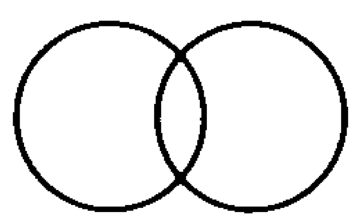
Table 2.3 Distribution of strategy types across two experimental conditions, presented by count and within-strategy percentage.

Strategy	V/WP Condition (N=19)	W-only Condition (N=32)
Verbal	9 (32.1%)	19 (67.9%)
Spatial	8 (44.4%)	10 (55.6%)

2.2.3.2 Evidence from Questionnaire Responses

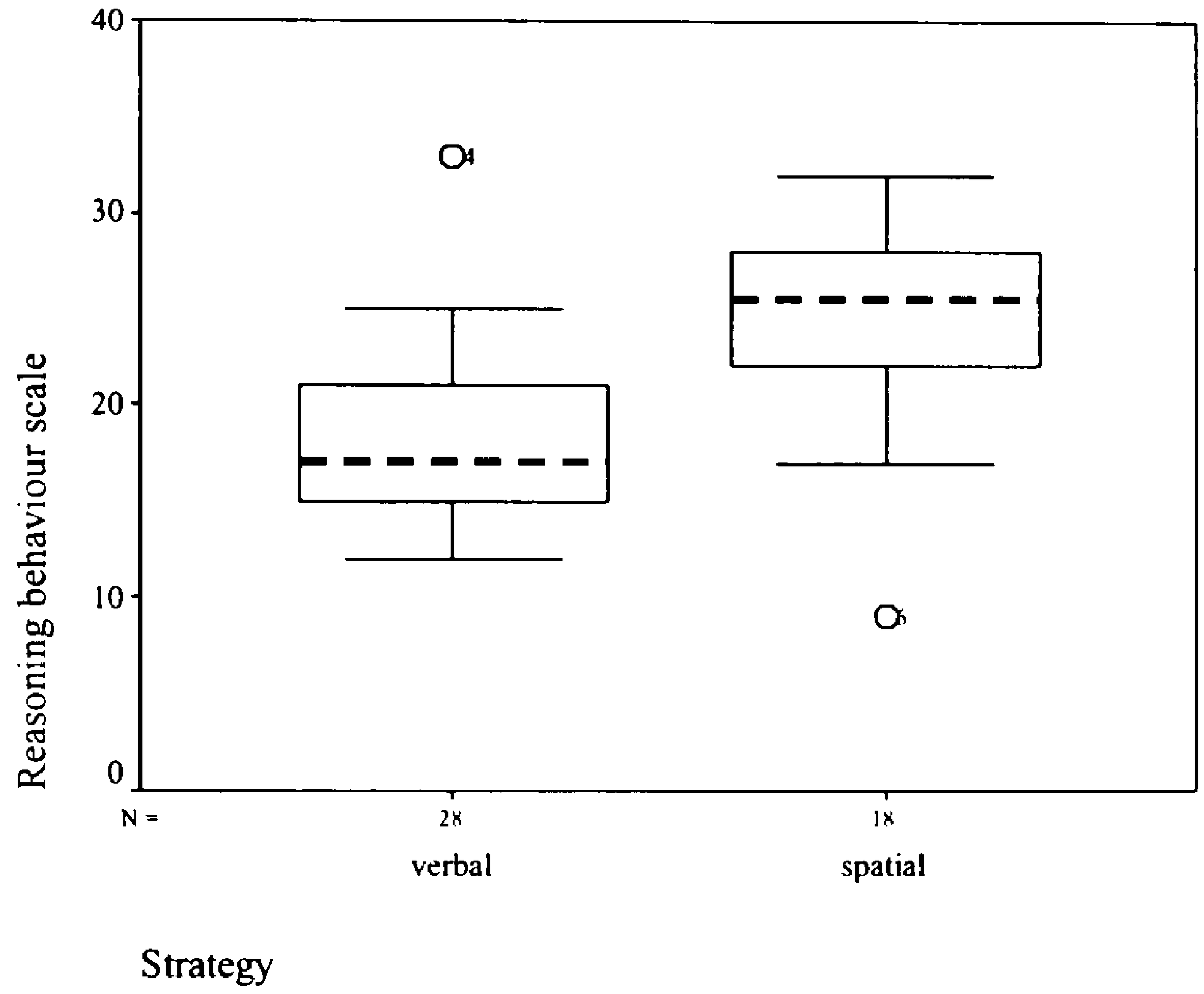
Following data collection, reliability analysis was conducted on the 13 quantitative questionnaire items. Reliability analysis suggested the elimination of 5 items with low item to total correlations resulting in an optimum Cronbach's Alpha of 0.68. SPSS output showing this analysis is presented in Appendix 2C and the final 8 reliable items are shown in Table 2.4 below.

Table 2.4: The final 8 valid questionnaire items. Note that items 1 to 6 were presented with a 5 point response scale as shown in original questionnaire, Appendix 2B. V or S in parentheses indicates whether item intended to identify verbal or spatial reasoning behaviours respectively.

1. To what extent did you attempt to reverse the position of the occupations within the statements? (V)
2. To what extent did you use mental images of shapes (e.g. circles squares) in spatial relationship in deciding on a conclusion? (S)
3. To what extent did you attempt to substitute terms/occupations from one statement to another (i.e. switch the occupations around between statements)? (V)
4. To what extent did you think about the words used in the statements? (V)
5. To what extent did you attempt to combine the two statements into one to form a single, longer verbal description? (V)
6. To what extent did you represent the information in linguistic form(e.g. sentences/words) in deciding on a conclusion? (V)
7. Did you develop a rule, or set of rules, to help you which you then reapplied to each subsequent problems? YES/NO (V)
8. Did you think about sets/Venn Diagrams (similar to that shown below) either using circles or any other shape? YES/NO  (S)

Questionnaire item responses appeared to provide converging evidence for the presence of verbal and spatial strategies, with reasoners whose written protocols identified them as verbal responding positively to verbal questionnaire items and spatial reasoners responding most positively to the spatial items (see Appendix 2B for which items were intended to identify which strategy). Interestingly, participants allocated to the mixed strategy group generally provided mixed or conflicting responses to the eight questionnaire items, for instance, answering yes to both question 3 (a spatial item) and also to one or more of questions 4 to 7 (verbal items). Reverse–scaling the items relating to the verbal strategy resulted in an overall reasoning behaviour scale with a maximum score of 40, on which a high score represented a high degree of spatial reasoning traits and a low score indicated a high degree of verbal reasoning traits. Analysis of variance (ANOVA) comparing mean scale scores for the 3 strategy groups indicated a significant difference between verbal ($\underline{M}$ = 18.18, sd = 4.6) and spatial reasoners ($\underline{M}$ = 24.78, sd = 5.5), with the mixed group falling between (M = 22.20, sd = 2.7), $F(2,48) = 10.35$, $p < 0.001$. Scheffé post-hoc comparisons indicated that the difference between scores for the verbal and spatial groups was highly significant ($p<0.001$). All ANOVA tables for Experiment 1 analyses are shown in Appendix 2E. However, examination of the distribution of scores on this scale revealed that although the majority of verbal and spatial reasoners presented distinct mean scores, some degree of overlap between the two groups was apparent. This is illustrated in the plots shown in Figure 2.3 below.

Figure 2.3: Distribution of reasoning behaviour scale scores across verbal and spatial strategy groups.



When the mixed strategy group were removed from analysis, K-means cluster analysis (see Hartigan and Wong, 1979, for explanation of this procedure) applied to responses to the 8 reliable questionnaire items (as given in Table 2.4) revealed two distinct clusters of participants. These clusters are crosstabulated with the verbal and spatial strategy groups in Table 2.5. Pearson's Chi square suggests a highly significant association between strategy type and questionnaire response, $\chi^2(1) = 12.3$, $p < 0.001$.

Table 2.5: K-means cluster analysis of questionnaire responses: cluster membership by strategy. Figures represent within strategy group numbers and percentages.

Strategy	Cluster 1 (N=25)	Cluster 2 (N=21)	Totals
Verbal	21 (75%)	7 (25%)	28 (100%)
Spatial	4 (22.2%)	14 (77.8%)	18 (100%)

As Appendix 2B showed, the questionnaire also contained a number of qualitative items aimed at obtaining more detail about specific aspects of the strategies. Unfortunately, these yielded very little information, with many reasoners failing to complete them at all, whilst others made very general comments such as "I found it hard". Informal feedback suggested that people found difficulty in articulating their strategy in such depth. Often, individuals were seen to struggle for some time with these items, at the end of what was, for many, an already demanding experimental session. No analysis was therefore conducted on responses to these items.

2.2.3.3 Characteristics of Verbal and Spatial Reasoners

As so few mixed reasoners were identified, the remainder of this chapter will concentrate on the verbal and spatial strategies.

Verbal Reasoners (N=28)

Participants in the replication (V/WP) consistently produced written protocols which presented features typical of those Ford claims describe the behaviours exhibited by verbal reasoners. These are supported by the content of verbal reports as illustrated in the following two protocols.

Protocol 2.1

Syllogism 15:
Participant 3:

Some of the painters are linguists → ∴ Birdwatchers
All of the linguists are birdwatchers
Replace

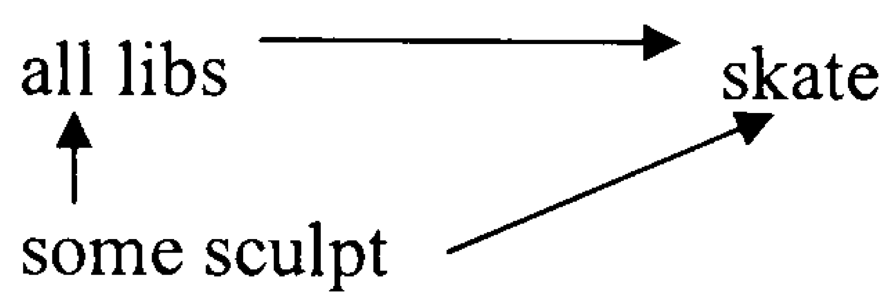
"Some of the painters must be birdwatchers, because again you know that all linguists are birdwatchers, you can just replace there. You can just replace linguists with birdwatchers. So you can just put arrows there and some of the painters are birdwatchers".

Correct conclusion given: Some of the painters are birdwatchers

Protocol 2.2

Syllogism 10: All of the librarians are skaters
Some of the sculptors are librarians

Participant 17:



"All the librarians skate, some of the librarians sculpt. So all of them do that…some of them do that [draws arrows]. Put it all into one sentence…some of the sculptors skate".

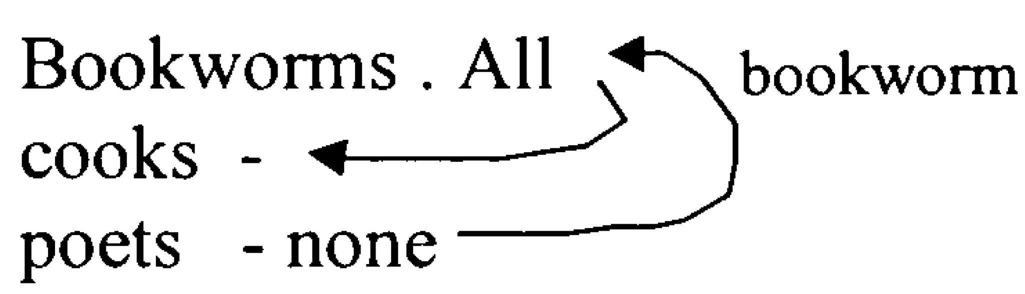
Correct conclusion given: *Some of the sculptors are skaters*

Participants in the written protocol only condition (WP-only) presented written protocols with notable similarities, for instance:

Protocol 2.3

Syllogism 9: All of the bookworms are cooks
None of the poets are bookworms

Participant 22
Written protocol only:



Correct conclusion given: *None of the poets are cooks*

For many of the verbal reasoners, initial examination of the protocols suggests that they apply what Ford termed naïve substitution, that is, they simply obtain a value for the B term from the first universal affirmative premise they encounter, then substitute that value for B in the other premise to obtain a conclusion with the same quantifier as that premise. This substitution seems to be made literally by term and position, irrespective of the quantifier in the premise where the substitution is made.

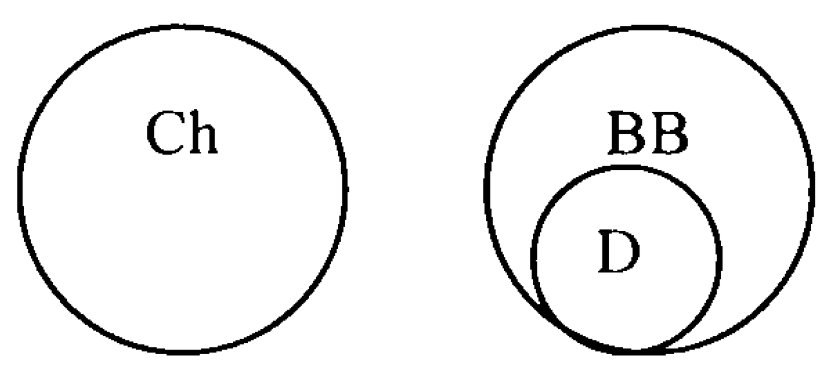
Spatial Reasoners (N=18)

Participants across both conditions produced written protocols which suggest, according to Ford's criteria, that they were using a spatial reasoning strategy. These protocols typically represented terms within shapes (usually circles or ovals) placed in differing spatial relationships to represent the relationship of the terms within and between the premises. Verbal protocols in the V/WP condition suggested they were thinking about the premises in terms of sets or groups, for instance:

Protocol 2.4

Syllogism 27: None of the chessplayers are bookbinders
All of the dancers are bookbinders

Participant 23



"So now we have the chessplayers and the bookbinders as two completely separate groups...and all of the dancers are bookbinders so we have the dancers as a subset of bookbinders...and dancers... none of the dancers are chessplayers".

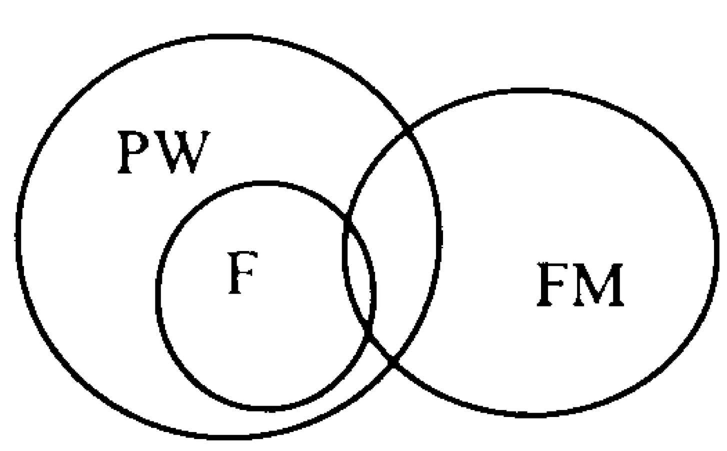
Correct conclusion given: *None of the dancers are chessplayers*

W-only participants' protocols were again comparable, for instance:

Protocol 2.5

Syllogism 21: Some of the farmers are freemasons
All of the farmers are prizewinners

Participant 12:
Written protocol only:



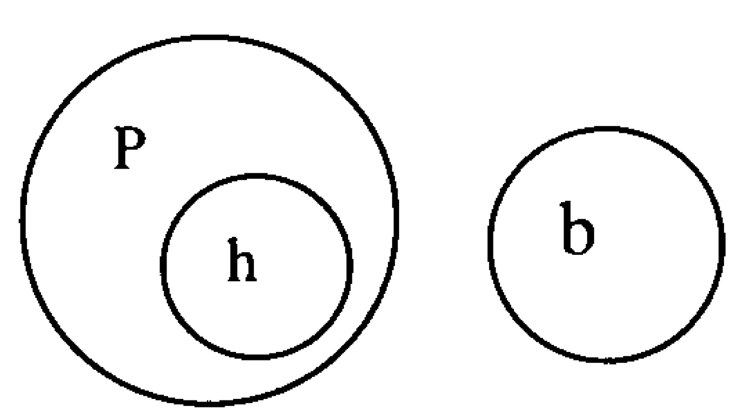
Correct conclusion given: Some of the freemasons are prizewinners

Whereas the verbal reasoners consistently appear to begin their reasoning with the universal premise, spatial reasoners appear to begin by constructing a representation of the first premise they are presented with, irrespective of its mood, and even when the later premise is universal affirmative. Information from the second premise regarding the C term (prizewinners) is then added to this representation in a manner which describes its relationship with the common term (farmers). These representations resemble traditional Euler circles but there is little evidence that spatial reasoners attempt to represent all the logical possibilities of the premises. For instance, Protocol 2.6 below represents the disjointed relationship between beekeepers and historians (premise two), and by implication, the same relationship between beekeepers and philatelists. The verbal protocol suggests that this subject may in fact have had some awareness that an alternative interpretation of the premises was possible, allowing for a subset relationship between beekeepers and philatelists. Had she pursued this line of thought further, it might have led her to the correct conclusion (*Some of the philatelists are not beekeepers*). Instead, she seems almost to dismiss it, preferring to work with her original representation and so generating an incorrect conclusion.

Protocol 2.6

Syllogism 12: All of the historians are philatelists
None of the beekeepers are historians

Participant 33:



"All historians are philatelists so they are in there [draws subset representation of first premise] and none of the beekeepers are historians so I'll put them separate but you could also have it on because it doesn't actually say that none of them are philatelists".

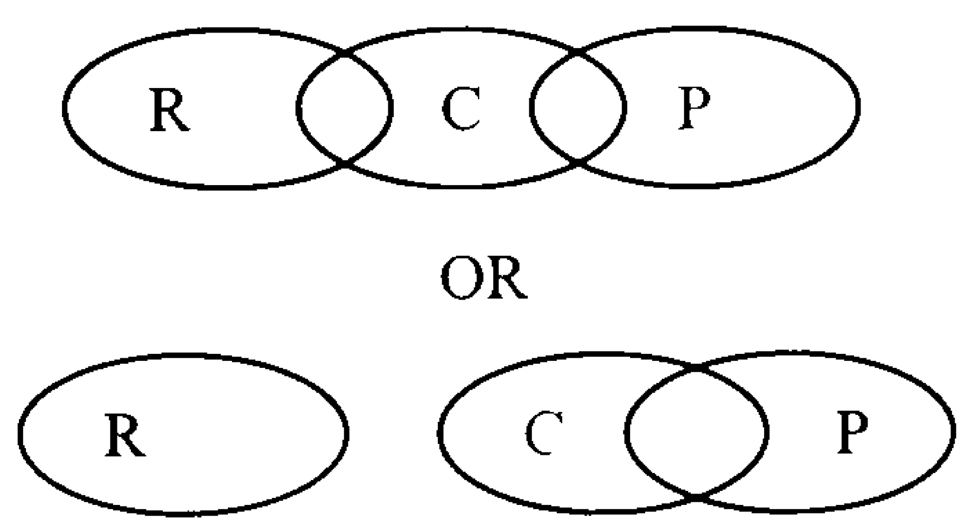
Incorrect conclusion given: *None of the beekeepers are philatelists*

Whereas Ford (1995) cites several examples where she claims reasoners represent more than one possibility in their diagrams, in the present study only two spatial participants attempted this. Neither were from the V/WP condition but their written reports merit consideration and are presented below:

Protocol 2.7

Syllogism 13: Some of the clubbers are pilots
None of the rockclimbers are pilots

Participant 47
Written protocol only

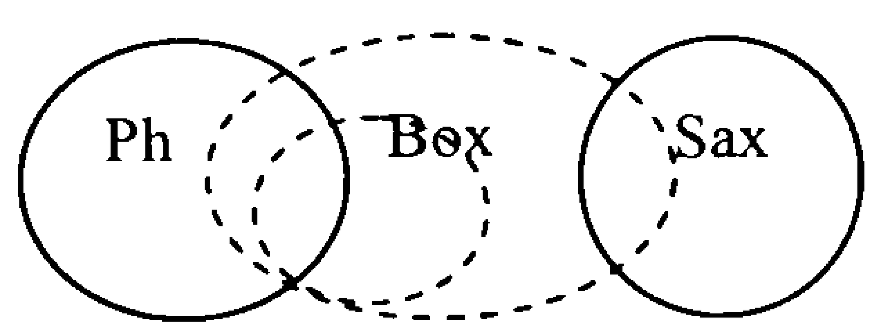


Correct conclusion given: *Some of the clubbers are not rockclimbers*

Protocol 2.8

Syllogism 6: None of the boxers are saxophonists
Some of the boxers are philosophers

Participant 12
Written protocol only



Correct conclusion given: *Some boxers are not saxophonists*

2.2.3.4 Performance of Verbal and Spatial Reasoners

Overall Accuracy

Overall performance on the syllogistic reasoning task was assessed in terms of number of correct conclusions generated. Descriptive statistics are presented in Table 2.6. No significant differences in performance were apparent between the two strategies overall: $t(44) = 1.39$, $p > 0.05$, or between experimental conditions (V/WP $M = 12.1$, W-only $M = 13.6$); $t(41.26) = 0.96$, $p > 0.05$.

Table 2.6: Descriptive statistics illustrating overall syllogistic reasoning performance across the two main strategy groups. Performance assessed according to number of correct conclusions generated out of 27 syllogisms.

Strategy	N	Mean correct	Std. Dev.	Min	Max
Verbal	28	12.00	3.31	7	20
Spatial	18	13.61	4.48	8	23
Total	51	12.69	3.81	7	23

Same and Different Form Syllogisms

For verbal reasoners, the naive substitution described previously can be successfully applied to all same-form (SF) syllogisms (those where the form of the conclusion matches that of one of the premises). However, a correct conclusion may not be as easily obtained for different-form (DF) syllogisms (those where the form of the conclusion does not match that of either premise), even when the syllogism contains an *All* premise. Verbal reasoners' inclination towards naïve substitution would suggest that their performance on DF syllogisms is likely to be considerably worse than on SF problems. Spatial reasoners on the other hand may find syllogisms intractable if there is more than one way in which the premises might be represented. In fact, the DF syllogisms are also those which Johnson-Laird and Byrne (1991) suggest are three-model problems and hence afford greatest difficulty. A 2x2 mixed analysis of variance with repeated measures on the second factor (2 x strategy group; 2 x syllogism form) showed no main effect of strategy ($p > 0.05$). However, a highly significant main effect of syllogism form; $F(1,44) = 144.05$, $MSE = 58909.12$, $p < 0.001$, was observed and the interaction between form and strategy approached significance; $F(1,44) = 2.914$, $MSE = 1191.53$, $p = 0.09$. These effects are illustrated by the means shown in Table 2.7 below.

Table 2.7: Percentage of correct responses to same-form and different-form syllogisms.

Form	Strategy	Mean	Std. Dev.
Same	Verbal	73.0	15.8
	Spatial	71.8	12.8
Different	Verbal	13.7	17.6
	Spatial	27.4	33.1

Simple main effects analysis of the interaction indicated that although the two groups performed similarly on SF syllogisms, spatial reasoners significantly outperformed verbal on DF; $\underline{F}$ (1,44) = 5.0, $\underline{MSE}$ = 408.9, $\underline{p} < 0.05$. When the nature of errors is examined it is apparent that 72.3% of conclusions generated by verbal reasoners to DF syllogisms were of the *same* form as one of the premises. Spatial reasoners presented a similar, if less extreme, trend with 60.7% same-form conclusions.

Replication of Ford's Analyses

Ford's analyses were outlined in the introduction to this chapter and these were applied to data from the present study. Given the significant effects of syllogism form for both strategies, and that the same-different form distinction maps closely onto the levels of difficulty predicted by mental model count, the following sections summarise findings from the present study according to those factors. Comparisons with Ford's findings regarding the reasoning processes observed are also highlighted with protocol evidence as appropriate. All syllogism numbers refer to the present study and means indicate mean percentage correct. Appendix 2D presents a corresponding summary of responses to all 27 syllogisms, also by form and model count and cross referenced with Ford's analysis.

Same-form, Single-model Syllogisms (N = 10)

For verbal reasoners, these comprise all seven of the syllogisms which Ford describes as requiring direct application of her simple substitution rule (syllogisms 3, 8 9, 10, 15, 25 and 27). As in Ford's study, all participants do well on these problems, however, unlike Ford, in no instance did we find significant differences in verbal reasoners' performance between syllogisms requiring the *Ai/Bi* (syllogisms 10, 15, 25/ 8) and *Aii/Bii* (syllogism 9/27) forms of her inference rule or as function of whether the rule required application to the first or second universal premise (syllogism 3) ($p > 0.05$ in every case). The introduction to this chapter presents the rules in detail. Also in contrast to Ford, verbal reasoners appeared to use naive substitution behaviour throughout, obtaining their value for B from the universal premise, irrespective of presentation order. Even on the problems Ford highlights as presenting greater difficulty, they were able to apply naive substitution in the usual way by making the assumption that *All A are B* is equivalent to *All B are A*. For instance on syllogism 27, which Ford suggests requires rule *Bii*;

Protocol 2.9

Syllogism 27:
Participant 3;

None of the chessplayers are ~~bookbinders~~ dancers
All of the dancers are bookbinders
replace

"Again, you can just replace dancers with bookbinders, and put that into the first statement so that changes to none of the chessplayers are dancers"

For two SF, single-model problems (syllogisms 14 and 21), Ford suggests that although naive substitution will obtain a correct conclusion, sophisticated substitution involving rule *Ai* is logically required. She duly reports a detriment in performance as a function of this. Our verbal reasoners actually performed significantly better on these two syllogisms

($\underline{M}$=92.6, sd = 25.2) than on those discussed previously where *Ai* can be directly applied in simple substitution ($\underline{M}$=72.6, sd = 35.5); $\underline{t}(27) = 3.6$, $\underline{p} = 0.001$ or, on syllogism 26 which requires sophisticated substitution ($\underline{M}$= 3.57, sd = 19.9); $\underline{t}(27) = 18.9$, $\underline{p} < 0.001$. Again they appear to use naive substitution rather than attempting more sophisticated reasoning. For example, the two written protocols to follow were produced by the same participant, the first to syllogism 10, a "simple substitution" problem and the second to syllogism 14. The mood is consistent between problems, and in both cases she provides the correct conclusion.

Protocol 2.10

Syllogism 10:	All of the librarians are skaters Some of the sculptors are librarians
Participant 50:	L = S Some Sc = L ∴ Some Sc = S

Correct conclusion given: *Some of the sculptors are skaters*

Syllogism 14:	All of the plumbers are gamblers Some of the plumbers are snowboarders
Participant 50:	P = G Some P = S ∴ Some G = S

Correct conclusion given: *Some of the gamblers are snowboarders*

For spatial reasoners, the SF, single model problems comprise all 6 of Ford's totally-constrained syllogisms (syllogisms 3, 8, 9, 11, 25, 27) and the 4 less-constrained problems with *All-some* premises (10, 14, 15, 21). In the present study, in contrast to Ford, spatial reasoners performed *better* on the less-constrained syllogisms, $\underline{t}$ (17) = 3.2, $\underline{p} < 0.05$ (constrained $\underline{M}$ = 78.7, sd = 38.6; less-constrained $\underline{M}$ = 94.4, sd = 19.8). Protocols 4 and 5, previously, present typical examples of spatial protocols on SF totally-constrained and less-constrained syllogisms respectively. These seem to present a single model of the premises and, unlike Ford, show little evidence of considering any alternative representations in the latter case

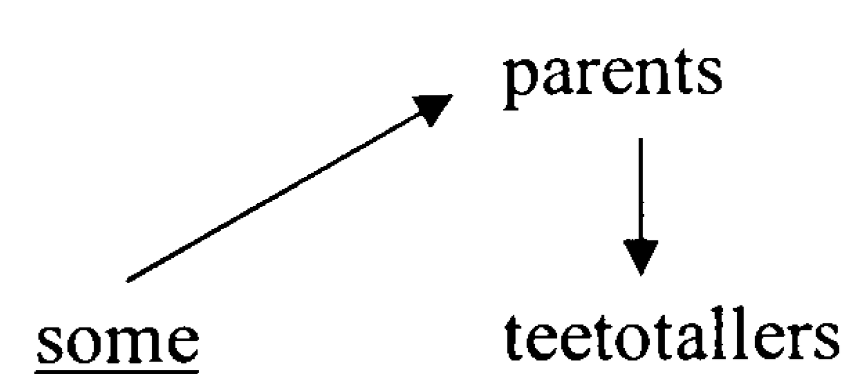
Same-form; multi-model Syllogisms (N = 4)

The four remaining SF syllogisms are two mental model problems and have a premise and conclusion with the form *Some of the X are not Y.* Inspection of the data reveals anomalies in performance on these problems which concur with Ford's findings. As Ford maintains, and the table in Appendix 2D shows, on two syllogisms (numbers 5 and 17) naive substitution works well ($\underline{M}$= 69.7, sd = 46.7). On the other two syllogisms however (numbers 2 and 20), verbal reasoners perform significantly less well ($\underline{M}$ = 19.7, sd = 40.4); $\underline{t}$ (27) = 5.9, $\underline{p} < 0.001$. For example;

Protocol 2.11

Syllogism 2: All of the parents are teachers
Some of the teetotallers are not teachers

Participant 39:



"Well this one was difficult. Had the repeated term at the end but you have the "are not" in. When you have the same words leading up to the repeated occupation it's easier, but when they are different it's harder. This one... I looked at some and then had a look at parents. Then that had to come back down to teetotallers. I think I cancel out teachers in the first one, I have established that all of the parents are teachers so I don't need that any more. Some of the teetotallers are not teachers... so I kept that as it was and think if some of the parents are teachers yet some of the teetotallers are not teachers, that means that some of the parents are teetotallers. I think the key word here is some, it's not all".

Incorrect conclusion given: *Some of the parents are teetotallers*

This reasoner is easily able to establish the relationship between parents and teachers from premise one in the usual way but a difficulty arises in establishing the relationship between parents and *not* teachers – because there is no common B term naive substitution is not applicable. However, by making the Gricean error of assuming that if some teetotallers are not teachers, then some are, the reasoner <u>can</u> substitute into premise 2 as she did on other SF syllogisms, hence drawing the incorrect conclusion: *Some of the parents are teetotallers*. Ironically, participants who just ignored the word *not* and substituted naively, tended to obtain the right conclusion, as in protocol 2.12:

Protocol 2.12

Syllogism 20: Some of the hikers are not politicians
Participant 3: All of the cyclists are politicians ↗Replace

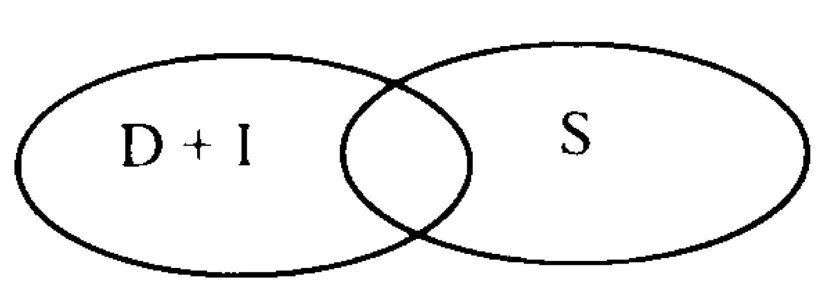
"Because all the cyclists are politicians you can replace cyclists into the first statement so some of the hikers are not cyclists".

Spatial reasoners present a different error profile across these four problems. Syllogisms 17 and 20 present the *some-not* premise first and in both cases the most popular spatial conclusion generated was *Some A are not C*. As in protocol 2.13, below, when the *some-not* premise is presented first, spatial reasoners typically process the *second* premise first. This seems to be the only syllogistic form where spatial reasoners do this and may be reflected in the characteristic conclusion direction errors they present on these two syllogisms. However, this does not lead to a significant detriment in performance for these problems compared to syllogisms 2 and 5 ($p > 0.05$).

Protocol 2.13

Syllogism 17: Some of the doctors are not singers
All of the doctors are intellectuals

Participant 19:



"So everybody's an intellectual but some of them aren't singers, so some of the singers are not intellectuals".

Incorrect conclusion given: *Some of the singers are not intellectuals*

In summary, naive substitution seems to cause problems for verbal reasoners where the end premise contains *not B* and this is reflected in their percentage correct conclusions for syllogisms 5/17 (M = 69.7) and syllogisms 2/20 (M = 19.7). Spatial reasoners however perform similarly on these two pairs of syllogisms, although some variance in their performance is observed (M = 38.9 in both cases, sd = 51.5 for item 5 and 46.1 for item 17). A 2x2 mixed analysis of variance (2 x syllogism pair; 2x strategy) revealed no significant main effects for strategy ($p > 0.05$). However, a highly significant effect of syllogism pair; F (1,44) = 10.96, MSE = 5.48, $p < 0.005$, and a significant two-way interaction between syllogism type and strategy, F (1,44) = 10.96, MSE = 5.48, $p < 0.005$ were indicated.

Different-form; multi-model Syllogisms (N= 13)

Ford claims simple substitution rules apply to just four of these syllogisms (4, 6, 13, 19). In line with her findings, verbal reasoners in the present study perform significantly better on these problems (M =25.9, sd = 44.2) compared to other DF multi-model problems which Ford claims require sophisticated substitution; t (27) = 3.4, $p < 0.005$ for "sophisticated" syllogisms with *All-None* premises (M= 4.3, 15.3); and t (27) = 2.5, $p < 0.05$, for those with *Some-None* premises (M=13.4, sd = 34.6). However, again we found no significant differences in performance between syllogisms requiring *Ai/Bi* and *Aii/Bii* ($p > 0.05$ in each case) and this, together with consistency in protocol content and a strong preference for *None* conclusions throughout, suggests that once again verbal reasoners applied naive substitution for all these syllogisms. Those Ford claims require sophisticated substitution present similar difficulties to those described for the SF two-model problems previously. For instance, where there is no *all* premise, more successful participants use the alternative universal negative premise instead to gain a value for "*not B*". Again participant 3 was accomplished at this.

Protocol 2.14

Syllogism 18:
Participant 3:

Replace
None of the criminals are florists
Some of the florists are trainspotters
^
not

"Because if none of the criminals are florists, you can replace criminals and florists...but put it negative. So some of the criminals are not...make that negative and turn it round the other way, so some of the trainspotters are not criminals".

Correct conclusion given: *Some of the trainspotters are not criminals*

Conversely, participant 17 was typical of many who performed less well on these problems. In every case, she attempted to work with the most conservative premise (some) and substitute into the least conservative (none), giving a *None* conclusion. For instance, in protocol 2.15, she almost gets the correct answer and then decides to go with the more conservative option.

Protocol 2.15

Syllogism 4:
Participant 17:

None of the pianists are mechanics
Some of the experts are mechanics

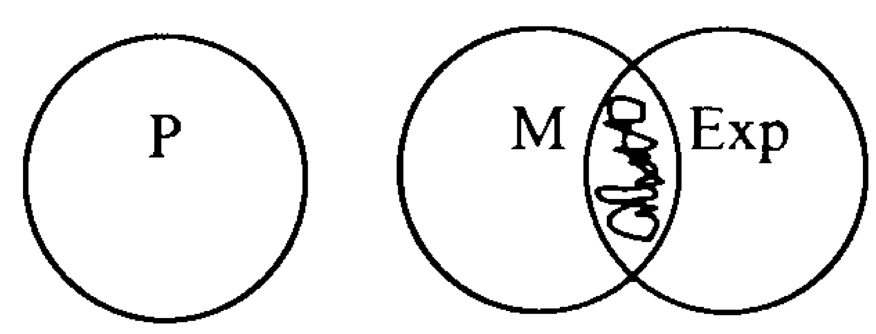
"So some... some of the experts are mechanics but no pianists are. How do I say that in another way... Some experts, are no pianists, no pianists are experts".

For spatial reasoners, Ford described how the most commonly drawn representations present an overlapping relationship between the end terms, but that these are generally interpreted as *Some X are Z*, rather than *Some X are not Z*. As this appears incompatible with the premises, reasoners are led to decide on either a *None* conclusion or that no valid conclusion is possible. The participant in protocol 2.16 also considers such possibilities but presents the disjointed relationship between end terms which was characteristic of reasoners in the present study.

Protocol 2.16

Syllogism 4: None of the pianists are mechanics
Some of the experts are mechanics

Participant 37:



"So we have two separate circles...a population of pianists...and mechanics separate. Some of the experts are mechanics, so it's circle will overlap...E.X.P. So that means...um...no pianists are mechanics, some of the experts are mechanics, so...that means...[long pause] no valid conclusion.... hold on... does that mean that none of the experts are pianists".

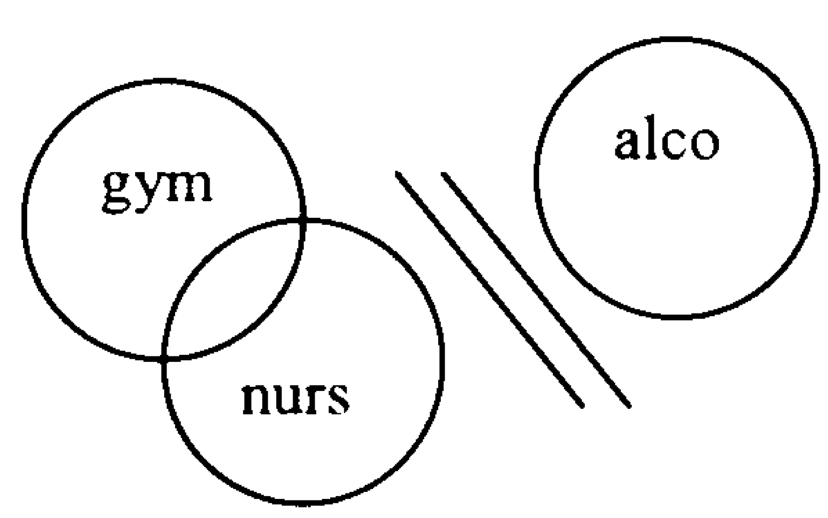
Incorrect conclusion given: *None of the experts are pianists*

In contrast to Ford, this type of written representation is typical of those used by 88.8% of spatial participants in the present study, irrespective of performance. The following is an example from a participant who clearly shows awareness that her overlapping circles may indicate a *some-not* relationship.

Protocol 2.17

Syllogism 19: Some of the gymnasts are nurses
None of the nurses are alcoholics

Participant 13:



"Right we've got nurses, alcoholics and gymnasts. None of the nurses are alcoholics, so you can't be a nurse and an alcoholic [draws two lines between these groups] but some of the gymnasts are nurses so there's a subgroup of gymnasts who are nurses, so that

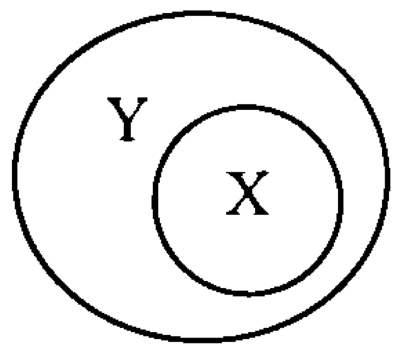
subgroup who are nurses cannot be alcoholics as well...so some of the gymnasts are not alcoholics".

Correct conclusion given: *Some of the gymnasts are not alcoholics*

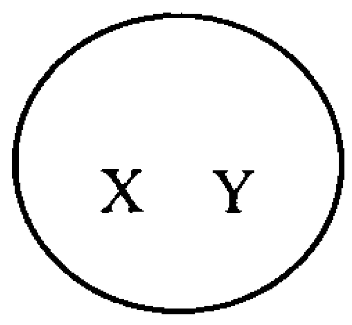
In Ford's terms, all the DF syllogisms have less-constrained representations. She categorises all DF syllogisms with *Some-None* premises together (syllogisms 1, 4, 6, 13, 18, 19, 22, 23) and points out the relative homogeneity of spatial performance as evidence that this categorisation of syllogisms is appropriate for these reasoners. In contrast, the present data indicates ranges of 27.7 (16.7-44.4) for spatial and 11.4 (10.7-22.1) for verbal reasoners. Spatial reasoners perform the better of the two strategies on these problems (spatial $\underline{M}$=31.9, sd = 46.7, ; verbal $\underline{M}$=19.7, sd = 39.4) but the difference is not significant ($\underline{p} > 0.05$). Like the verbal reasoners, spatial people perform less well on these DF *Some-None* syllogisms compared to the SF *All-Some* syllogisms; $\underline{t}$ (17) = 4.8, $\underline{p} < 0.001$, and SF *All-None* syllogisms; $\underline{t}$ (17) = 6.8, $\underline{p} < 0.001$, see also page 84 previously. Similarly, for the four DF *All-None* syllogisms ($\underline{M}$ = 19.4, sd = 45.1), again our findings concur with Ford in that spatial reasoners perform less well than on the SF *All-none* syllogisms; $\underline{t}$ (17) = 6, $\underline{p} < 0.001$ and the SF *All-some* syllogisms; $\underline{t}$ (17) = 8.5, $\underline{p} < 0.001$. However, all these SF syllogisms are single-mental model problems – those which research has repeatedly shown to be least difficult and where the verbal reasoners also perform better than they do on DF problems. Moreover, while stating that "spatial reasoners tend to be at a loss" (page 59) Ford does not appear to directly compare the two strategies. In the present study, verbal reasoners performed noticeably worse than spatial on all these DF problems, and for those with *All-None* premises the difference was significant; $\underline{t}$ (44) = 2.0, $\underline{p}$ = 0.05, suggesting they may be equally at a loss.

Spatial Representations of the *All* Premise

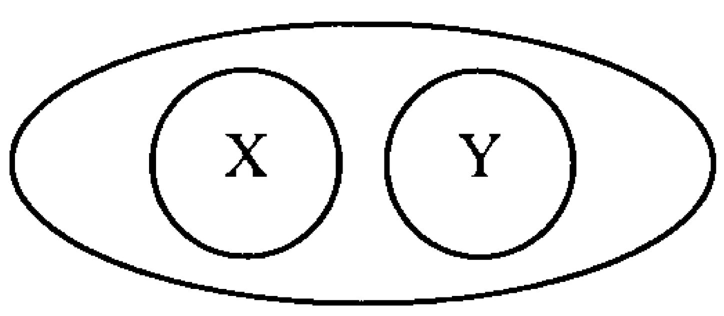
Ford states that, for her spatial reasoners, the preferred representation of the universal affirmative premise, *All of the X are Y,* was as a subset relationship, as below:



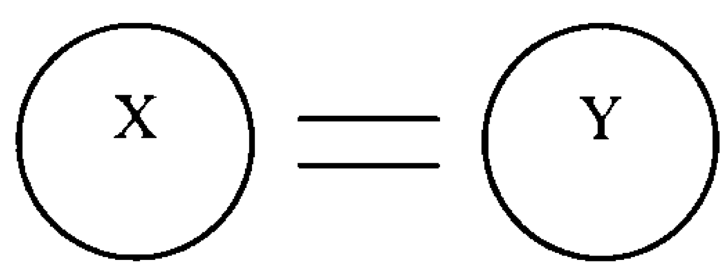
In the present study, just 7 spatial reasoners (38.8% of total) consistently represented such premises in this way. Examples of these can be seen in several of the protocols cited previously. The majority(10 people, 55.6%) represented *All of the X are Y* in the form of an identity relationship, mostly in the classical form:



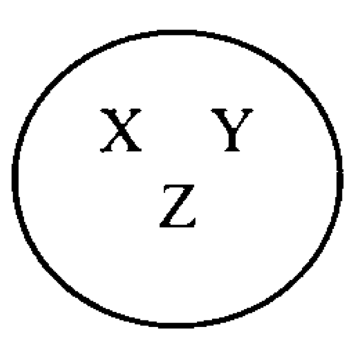
But also as variations of:



or,



The one remaining spatial reasoner used a mixture of the two approaches. Of the 17 people who used one or other representation consistently, there was no significant difference in overall performance for syllogisms with *All* premises as a function of how that premise was represented ($p > 0.05$). On the three problems with two *All* premises (syllogisms 3, 25, 26), 50% of participants represented the two premises as a three-way identity:



Again this was not detrimental to their performance compared with other reasoners ($p > 0.05$). When considered in terms of Ford's syllogism categorisations, the only category where a noticeable difference in performance was observed was the DF *All-None* syllogisms (numbers 7, 12, 16 and 24) where only two identity users got any of the problems correct (identity M = 10.0, sd = 25.4; subset M = 32.5, sd = 44). However, due to the exceptional performance of one identity user, this difference in performance as a function of representation type was still not significant ($p > 0.05$).

2.2.4 Discussion

The general aim of Experiment 1 was to replicate and extend the work of Ford (1995). It was predicted that two main reasoning strategies would be identified, verbal and spatial, and that written and verbal protocols together with questionnaire responses would provide converging evidence of these strategies. In addition, it was hoped that the WP-only condition would allow for the reliable identification of the strategies without the need for onerous verbal reports. The findings appear to support these predictions. In condition V/WP (the replication), consistent trends in the written and verbal protocols clearly indicated the presence of two major forms of reasoning behaviour and these correspond to those which typify the verbal and spatial strategies described by Ford. In the W-only condition, written protocols show the same general characteristics and participants can be similarly categorised as belonging to one of the two strategy groups on the basis of these alone. In both conditions, questionnaire data supported the strategy categorisation. Participants generally answered in the affirmative to questions which probed *either* verbal or spatial reasoning behaviours and these in turn corresponded to the strategy suggested by their written and verbal reports. However, scores on the reasoning behaviours scale generated from questionnaire responses did indicate some overlap between the two strategy groups, albeit for a minority of reasoners. This may suggest that, although there is clear evidence for qualitatively distinct verbal and spatial strategies, some individuals may be employing a more mixed approach which wasn't immediately apparent from their protocols. The individual differences observed may in fact lie on a strategic continuum, a spectrum of strategies with what appear to be verbal and spatial reasoners positioned at the two extremes.

With the exception of the four same-form, multi-model syllogisms discussed previously, neither the present study, nor that of Ford, found any significant difference in overall performance across different strategy types and both strategies perform least well on DF problems. However, the present data suggests that spatial reasoners do significantly better than verbal on those syllogisms. Moreover, the present study found none of the within-strategy variations in performance which Ford identified as a function of type of substitution rule (verbal reasoners) or constraint of premises (spatial reasoners). In contrast to Ford, reasoners showed a consistent approach for all problems. Verbal reasoners appear to apply naive substitution throughout, performing least well on syllogisms where the mood makes this form of substitution most difficult. There is no evidence that they applied the formal simple substitution rules or sophisticated substitution process which Ford describes. Moreover, on the more difficult problems, verbal reasoners present evidence of precisely those Gricean conversion errors that Ford claims her reasoners eschew. Spatial reasoners produce similar representations for all problems, rarely appearing to represent multiple possibilities even for Ford's less-constrained syllogisms. Where performance variations were observed between syllogisms, they tended to apply similarly to both verbal and spatial reasoners. Even when the nature of fallacious responses is examined, there is little difference between the strategy groups, both verbal and spatial reasoners generate similar types of responses, presenting an outward appearance of universality.

However, such ubiquity may be superficial. Examination of the underlying processes reveals that errors occur for very different reasons. Verbal reasoners rely very literally on premise form which leads them into erroneous conclusions, especially for DF syllogisms. In this aspect, their conclusions give the appearance of those predicted by the matching hypothesis of Wetherick and Gilhooly (1990). Wetherick and Gilhooly asked participants to evaluate presented conclusions and hence responses could be directly matched to one of the premises. In the present study, participants were required to produce their own conclusions and it would appear that verbal reasoners generated responses which appeared to match the form of one of the premises. They appeared to choose the *All* premise to obtain a value for B because it affords the greatest informativeness about the individuals represented by the terms (in matching theory terms, this premise is the least conservative). Hence, the *All* premise is always selected for processing first, even when it is presented as the second premise in the syllogism. The remaining premise (the most conservative) is the one which is in effect "matched". Wetherick and Gilhooly (1995) highlight that such a strategy will work for 14 of the 27 valid syllogistic forms (i.e. the SF problems) but not for the other 13 (the DF). But, as Wetherick and Gilhooly allow, some reasoners do get DF syllogisms right. They state: "subjects who give correct conclusions to syllogisms in these moods cannot be matching, and it seems reasonable to suppose that their correct responses to other moods may not have been obtained by matching either" (1995, pp. 172-3). Wetherick and Gilhooly maintain that matching is a heuristic adopted when subjects are unable or unwilling to reason with logic. But the fact that some verbal reasoners are able to generate correct conclusions to DF syllogisms suggests that they are capable of some form of reasoning, albeit not necessarily one which is compatible with the rules of formal logic. Certainly, it seems that some verbal reasoners are able to modify their strategy to cope with more difficult syllogisms, but there is little evidence of the more formal modifications to premises that Ford describes as sophisticated substitution. The prevalent use of identity as representative of the *All* premise in the present study might suggest that our participants were less sophisticated reasoners then those of Ford. An assumption of identity could lead reasoners into illicit premise conversion and hence to believe that simple substitution was appropriate whenever an *All* quantifier was present.

However, spatial reasoners also find SF syllogisms easier than DF. Mental model theory suggests that the former are in fact easier for everyone because they require the construction of fewer models (Johnson-Laird and Byrne, 1991). The theory predicts errors on DF syllogisms because reasoners often fail to search for counter-examples which may refute their putative model. Certainly few spatial reasoners appear to consider the possibility of alternative models, and when they do, they tend to dismiss it, preferring to retain their original model. Spatial reasoners construct Euler circle type representations, but show no evidence of attempting to construct representations of all possible states of the premises. Erickson (1978) suggests that reasoners can only handle sufficient cognitive load for one diagram per premise, thus, if presented for instance with the premise *All X are Y*, they will encode the relationship as *either* subset or identity. Only one spatial reasoner in the present study adopted both representations. However, nor did they construct separate diagrams for each premise and then combine them, as traditional Euler circle theory would suggest. Rather, every spatial reasoner appeared to follow the same three-stage process:

1. Represent premise one diagrammatically
2. Draw an amendment, or addition, to this initial diagram in order to incorporate the information presented in premise two.
3. Derive a conclusion from the final composite representation.

As such, both interpretation of premise two and combination of the premises occur in one stage. How spatial reasoners then draw their conclusion remains unclear. Neither verbal nor written protocols make explicit how stage 3 above occurs. Conclusions may be based on a literal interpretation of diagram structure, although Johnson-Laird (1989, page 473) has stated that "Mental models are symbolic structures, and the relation of a model to the world *cannot* simply be read off from the model."[my emphasis]. If this is so, either spatial reasoners' diagrams are not manifestations of the use of mental models, or, conclusions may be drawn from a more complex process which may yet involve the application of some as yet unspecified inference rule.

Verbal reasoners can also be accused of failing to search for counter-conclusions, but then they show no sign of generating and/or using mental models. Polk and Newell's Verbal Reasoning hypothesis does not preclude the use of some form of mental model or non-linguistic representation, but unlike mental model theory it does not posit the search for counter-examples as a key element in conclusion generation. Rather, the two premises are encoded individually and then combined to form an initial putative conclusion. If this conclusion is not acceptable, the premises are repeatedly re-encoded until a valid conclusion is suggested. However, the present data present no evidence that verbal reasoners are any more likely than spatial to encode the two premises separately, or that reasoners using either strategy attempt to re-encode or re-formulate premise combinations in any way in an attempt to reconsider their initial conclusion.

Overall, the present findings seem to suggest that individual differences in strategies cannot be wholly accounted for by any one deductive theory. Johnson-Laird and Bara (1984) and others have suggested that manipulation of models may underpin substitution behaviour, and later work by Johnson-Laird (2001) has suggested that, over time, reasoners may learn to construct formal rules and that this is an essential stage in the development of logic. However the present data suggest no more evidence for this than they do for spatial reasoning underpinned by propositional rules. Although the diagrammatic models of spatial reasoners may be likened to Euler circles or mental models, neither theory can fully account for their reasoning. However, Wetherick and Gilhooly (1995) do highlight individual differences in heuristic processing of syllogisms, suggesting that while some participants indeed appear to be matchers, some may employ other non-logical methods. Chater and Oaksford's (1999) model of non-logical heuristic reasoning also predicts that conclusion mood will be the same as that for the least informative premise, but computed probabilistically, rather than obtained by simple matching of quantifiers. For DF syllogisms, initiation of the *p-entailment* heuristic would depend on an ability to identify that a further conclusion is probabilistically entailed by the *min*-conclusion. If participants in the present study were reasoning probabilistically, in these aspects at least, the PHM might account for the fact that both verbal and spatial reasoners performed similarly across different forms of syllogism. However, the fact that W-only participants showed strikingly similar individual differences to those in the V/WP condition appears to refute Chater and Oaksford's suggestion that the verbal-spatial distinction is nothing more than a methodological artefact arising from the use of verbal protocol. This in turn calls into question their claim that the PHM provides an accurate fit for experimental data without the need to account for such differences. Whether verbal and spatial reasoners are indeed similarly probabilistic in their deduction is an area that requires further research.

An important question that remains is why different people prefer to represent premise information in different forms in order to deduce conclusions, especially given that overall performance does not seem to change as a function of strategy. This may reflect individual differences in other cognitive abilities, and such differences may be equally apparent in individual approaches to other forms of reasoning task. Furthermore, participants in both this and Ford's study were drawn from a student population who are (presumably) fairly high in general intelligence. Conversely, Bucciarelli and Johnson-Laird (1999) learned little from verbal reports produced by Italian students who are admitted to university without selection criteria, and hence may be more representative of the general population. Levels of intelligence and education may influence both ability to make inferences and reasoning strategy. Moreover, any association between strategy choice and differences in abilities known to be influential in reasoning (such as working memory capacity) may help to clarify whether protocol data actually reflects differences in underlying cognitive processes, rather than in manner of representation.

Furthermore, some authors (for instance Johnson-Laird et al 2000) have claimed that Euler circle type representations rely on vestigial memories of procedures learned in school. However, item 9 of the questionnaire (see Appendix 3A) asked reasoners:

> "Did you apply a rule or procedure which you already knew about from things you have done in the past, rather than develop a new rule/procedure for this task? If yes, please describe the rule/procedure and how you knew about it".

Of the 51 participants in Experiment 1, only 5 answered in the affirmative to this item, and three of these were verbal reasoners who described algebra and/or equations (Ford describes these as typical of verbal reasoners' methods). Of the other two, one mentioned "spider diagrams" for summarising information and just one mentioned Venn diagrams learned in mathematics lessons at school. Although this indicates that the vestigial memories argument cannot be totally discounted, it does suggest that it applies only to a small minority of individuals and is not limited to spatial reasoners. Such memories may be selective and the nature of recalled material consistent with a predisposition to reasoning in a verbal or spatial way. It is suggested that such predispositions to a particular strategy are inherent factors of the individual, and hence will be robust across different task formats.

Certainly, both strategic processes remain underspecified in a number of aspects. In particular, how spatial reasoners draw conclusions from their representations and how verbal reasoners modify premise information when naive substitution is inappropriate, may prove to involve some form of rule, as yet undetermined. Verbal and spatial reasoning may in fact be examples of the transduction strategies which Polk and Newell overtly reject. The construction of the distinctive spatial representations may arise because of a need to convert premise information into a form which spatial reasoners find more conducive to cognitive processing. Similarly, for syllogisms where the naive substitution heuristic is not effective, verbal reasoners may be required to convert premise information into a form where they can then apply this principle. The mixed strategy group appear to use a combination of verbal and spatial behaviours within or between different syllogisms. The small number of mixed reasoners identified in the present sample precludes drawing any firm conclusions. This strategy may be some kind of compromise because these people are unsure of the best approach to take, or they may possess some cognitive factor which makes this mix of approaches a natural way of working for these individuals. Experiment 2 aimed to discover more about the nature of strategies.

2.3 Experiment 2

2.3.1 Introduction to Experiment 2

Experiment 1 has shown that most individuals adopt one of two reasoning strategies when solving syllogistic problems and that these appear to be verbal-propositional and spatial in character. In this respect Experiment 1 has replicated the findings of Ford (1995) and further shown that such strategies can be identified from written accounts alone, without the need for verbal protocol data. Moreover, a questionnaire has been piloted which seems to support the classification of strategies.

Individuals who adopt the verbal strategy seem to use a form of naive substitution, that is, they gain a value for the B term from a universal affirmative premise and substitute that value for B in the other premise to form a conclusion. Although this process lends itself well to problems where conclusions are of the same-form as one of the premises (SF syllogisms), verbal reasoners have difficulty with those syllogisms where the conclusion is of a different form (DF) to either premise and they perform significantly less well on these problems. Individuals using the spatial strategy use representations of the premises which are similar to traditional Euler circles, or Venn diagrams, though they rarely attempt to represent more than one possible model of the premises. DF syllogisms tend largely to be those problems which both mental models theory and Ford (1995) suggest have more than one possible way to represent the premises. Hence, spatial reasoners also have greatest difficulty with these syllogisms, though they still perform significantly better than verbal reasoners.

Otherwise however, these strategies are a weak indicator of differences in reasoning performance. Overall performance in terms of accuracy of conclusions generated tends not to differentiate between strategy users, which may be why many theories of reasoning have assumed that a single universal process is at work. However, from the pattern of performance observed over different syllogistic forms, it may be possible to make certain predictions about how verbal and spatial reasoners will respond. For instance, if verbal reasoners are indeed substituting terms in the manner described above, they will be likely to produce SF conclusions, even to DF problems. One way of testing such predictions is to examine not only the conclusions that reasoners produce, but also how they choose to evaluate conclusions they are presented with and how they deal with invalid problems. On evaluation tasks, all reasoners tend to perform better than when asked to generate their own conclusions (see for instance Evans et al 1993 for a review). However, substitution behaviour would suggest that verbal reasoners will be likely to endorse SF conclusions as valid, even when logically invalid. Conversely, they may rate valid DF conclusions as invalid. Spatial reasoners rely less on premise form and therefore may make fewer such errors. Experimental evidence in defence of such predictions would add further support to the argument for individual differences in reasoning strategies, as discussed in Experiment 1. Moreover, if the findings of Experiment 1 are robust, they will afford further replication. These were therefore the main aims of Experiment 2.

2.3.2 Aims

The aims of experiment 2 were threefold:

1. To replicate the findings from experiment 1 in terms of identifying verbal and spatial strategies.

2. To discover more about the nature of these strategies by examining their performance across tasks involving the production and evaluation of syllogistic conclusions, and including invalid problems. The following predictions were made:

   a) On a production task (as Experiment 1), verbal reasoners will produce more same-form conclusions than different form, even for invalid/indeterminate syllogisms. They will also be less likely than spatial reasoners to state "no valid conclusion" for invalid/indeterminate syllogisms, instead generating an invalid same-form conclusion.
   b) On an evaluation task, verbal reasoners will be more likely than spatial to endorse same-form conclusions as valid when the conclusion presented is logically invalid. Valid different-form conclusions are more likely to be rated as invalid by verbal reasoners.

3. To further develop the questionnaire as a tool for the identification of strategies

2.3.3 Methods

Participants

73 undergraduate students volunteered to take part in return for course credit. This sample comprised 7 males and 66 females, mean age 20.53 years. None had received formal training in logic.

Materials

Syllogistic Problems

Evans et al (1999) reported percentage endorsement of conclusions under both necessity and possibility instructions, for all 512 possible syllogistic forms. This extensive resource was used to identify syllogisms for use in the present study. As reasoners tend to struggle with DF syllogisms, problems were used which were endorsed as necessary at a high rate in the Evans et al data, in order to avoid a floor effect. Similarly, people tend to find SF problems easiest, so problems were identified where necessity endorsements were high to avoid a ceiling effect. In total, 20 syllogisms were identified from this source, half SF and half DF according to the conclusions presented by Evans et al. All four figures were represented amongst this set. As in Experiment 1, thematic content comprised occupations and persuasions, and this content was varied across the two tasks. In the production task, ten valid problems and five indeterminate ones (not from Evans et al) were presented. For evaluation, the same ten valid problems together with a further ten invalid ones were used. The presentation order was randomised prior to the experiment and then remained constant for all participants. All 25 syllogisms are presented in Appendix 2F, in abstract form, together with the endorsement rates as shown by Evans et al (1999). For the purposes of this experiment, validity was defined as per Johnson-Laird and Byrne (1991) i.e. a problem which has a necessary conclusion, one that must be true given that the premises are true. Some of the problems classed here as invalid may have *possible* conclusions (i.e. a conclusion that can be true for one model of the premises), however this definition of validity is in line with the procedure used in Experiment 1 and the valid syllogisms used by Ford (1995).

Questionnaire Development

Experiment 1 piloted a reasoning behaviours questionnaire designed to elicit information about strategy usage. Items were developed from the descriptions of verbal and spatial reasoning behaviours presented by Ford (1995). These behaviours were clearly replicated in Experiment 1 and strategy groups identified from written and verbal protocols were fully supported by questionnaire responses. Reliability analysis of these responses resulted in a final eight item questionnaire (see Table 2.4 previously) with an Alpha = 0.68. Six of these reliable items were designed to identify verbal reasoning behaviours (in Table 2.4 these are item 1 and items 3 to 7) but only two items appeared to be reliable indicators of spatial reasoning (items 2 and 8 in Table 2.4). Because of this imbalance, the questionnaire was extended for use in Experiment 2, using the eight original reliable items as a basis. Three additional items were developed based on spatial reasoning behaviours identified in Experiment 1 and which appeared to discriminate between them and verbal reasoners (the tendency to group information into sets and subsets, to use diagrams and to deal with premises in the order presented, irrespective of quantifier). A final questionnaire item was also included in the form of a general request for qualitative information about the approach used. Very little useful information was gleaned from the behaviour specific qualitative items in Experiment 1. It appeared that many participants found their specific tactics difficult to articulate and this part of the process laborious, especially coming as it did at the end of the experimental session. However, qualitative data can potentially offer useful supporting evidence and it was anticipated that a single, more general, question may encourage more reasoners to offer some details of their strategy not covered by the other questionnaire items. A copy of this revised questionnaire as used in Experiment 2 is presented in Appendix 2G (the new items added for this experiment are numbers 7,8,9 and 12).

Procedure

Participants were run in small groups of around eight individuals in single sessions lasting around 45 minutes. Each was seated at a separate desk to minimise sight of other participants' work. In a within-subjects design, all participants were presented with both the production and evaluation tasks, with task presentation order counterbalanced to avoid possible order effects. In each case, they were asked to provide written protocols of their reasoning but no verbal reports were collected (as in the WP-only condition of Experiment 1). After the two tasks they completed the questionnaire as in Appendix 2G. All the above were presented on paper in a single workbook. Participants were instructed to work through the booklet in the order presented. Each task was preceded by a set of written instructions and these are shown in Appendix 2H. Within each of the two tasks, syllogism were presented in a different random order for each participant. Each syllogism was presented on a single page with room below for written workings.

2.3.4 Results

2.3.4.1 Identification of Verbal and Spatial Strategies

Strategies were identified from the written protocols presented for the production tasks, according to the same criteria as in Experiment 1. Protocol 2. 18 presents an example of a verbal reasoner who seemed to be using a similar type of naive substitution described in Experiment 1.

Protocol 2.18

Syllogism 2:
Participant 30: All of the freemasons are receptionists
All of the receptionists are joggers

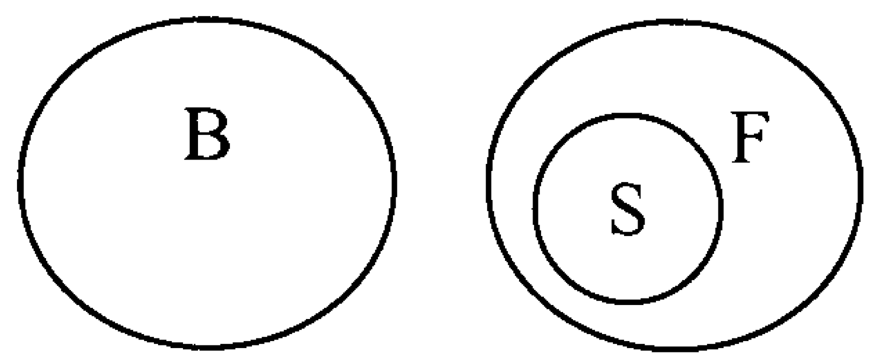
Correct conclusion given: *All of the freemasons are joggers*

Protocol 2.19 below shows a spatial reasoner.

Protocol 2.19

Syllogism 4: None of the builders are flautists
All of the shotputters are flautists

Participant 49:



Correct conclusion given: *None of the builders are shotputters*

Preferred strategy was identified from protocols produced on the conclusion production task as this was comparable to the process used in Experiment 1 and previously by Ford (1995). However, when these protocols were then compared to those produced during the conclusion evaluation task, the latter were found to be remarkably similar. Protocols 2.20 and 2.21 present examples of verbal and spatial reasoners respectively, across the two task types. The latter is interesting because, although using exactly the same approach as the Euler circle type reasoners such as that in protocol 2.19 above, she has chosen to represent the three terms using different shapes, presenting a shape identification key for each item.

Protocol 2.20

Production Task

Syllogism 1: Some of the opticians are shopkeepers
All of the opticians are Buddhists

Participant 17: opticians = Buddhists
some opt = shopkeepers ⇒ some Buddhists

Correct conclusion given: *Some of the Buddhists are shopkeepers*

Evaluation Task

Syllogism 3: All of the biologists are prizewinners
Some of the singers are biologists
Conclusion: Some of the singers are biologists

Participant 17: biologists = prizewinners
some singers = biologists ⇒ prizewinners

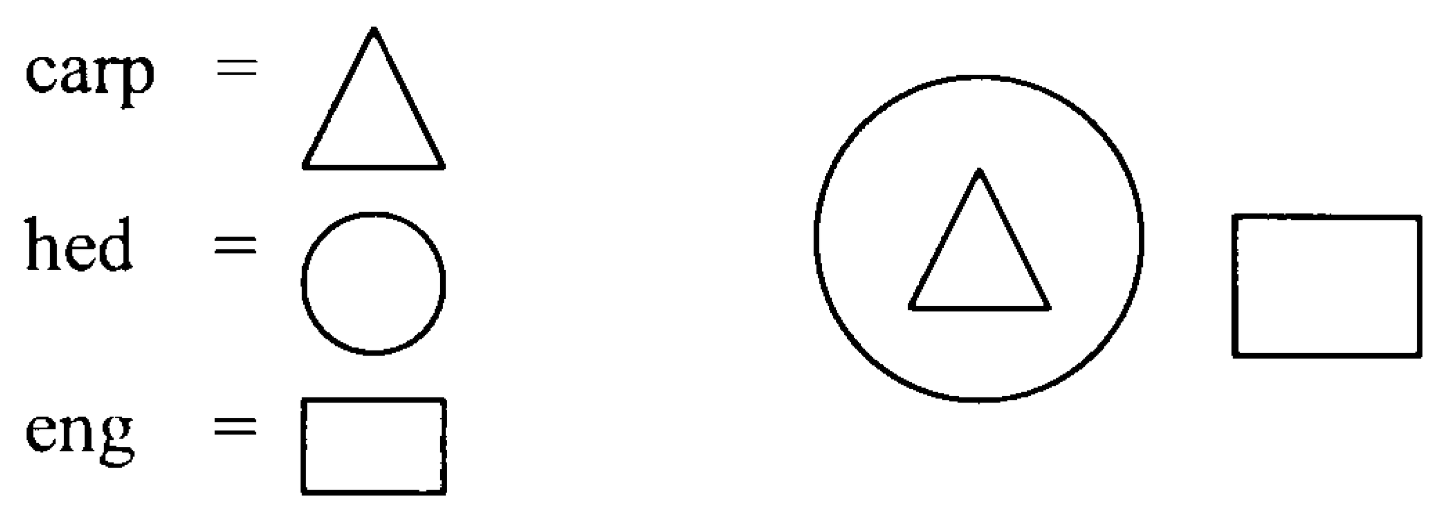
Correct evaluation given: *Conclusion is valid.*

Protocol 2.21

Production Task

Syllogism 5: All of the carpenters are hedonists
None of the engineers are hedonists

Participant 15:

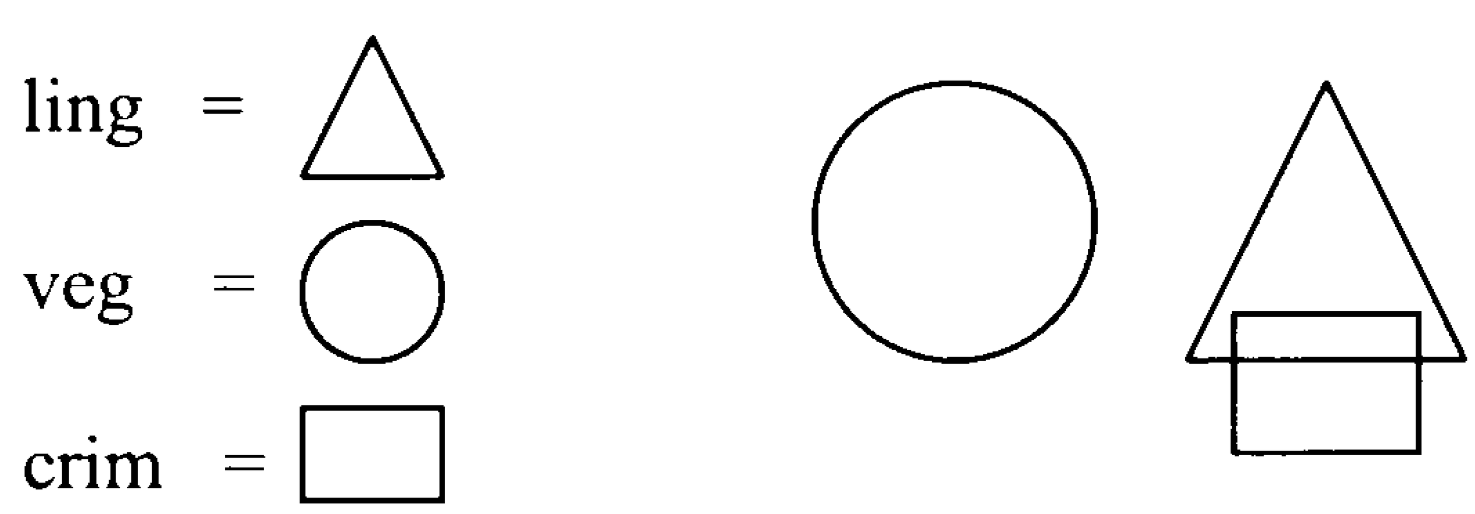


Correct conclusion given: *None of the engineers are carpenters*

Evaluation Task

Syllogism 6: None of the linguists are vegetarians
Some of the criminals are linguists
Conclusion: Some of the criminals are not vegetarians

Participant 15:



Correct evaluation given: *Conclusion is valid*

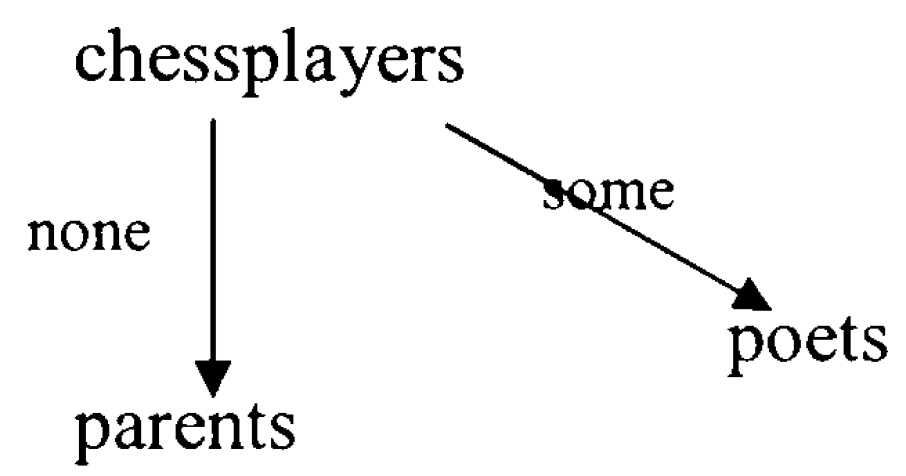
However, the strategy classifications were not as straightforward as in Experiment 1. Whereas many verbal reasoners who perform substitution use arrows to show how terms were moved around within and between premises, others used arrows to represent relationships between terms. They did this on both production and evaluation problems. The protocols typically represent the common, or B, term centrally and then link the other terms to it by arrows. For instance:

Protocol 2.22

Evaluation task

Syllogism 9: Some of the chessplayers are poets
None of the chessplayers are parents
Conclusion: Some of the poets are not parents

Participant 66:



Correct evaluation given: *Conclusion is valid*

A few similar protocols (N=8) were also observed in Experiment 1 and were categorised as evidence of verbal reasoning according to Ford's criteria (see for example Protocol 2.2 previously). The protocols resemble several of those shown by Ford (e.g. example numbers 39 and 43, pages 35 and 37) and the account of verbal reasoning which she gives on page 16 (see also Chapter 1 of this thesis). In the present study, when responding to item 12 of the questionnaire, participants who produced such protocols frequently referred to them as "flow-charts" or "arrow diagrams" and described using the arrow to follow the connection between terms. Participant 66 above is fairly typical in referring to some quantifiers as being "stronger" than others and how this influenced her choice of conclusion.

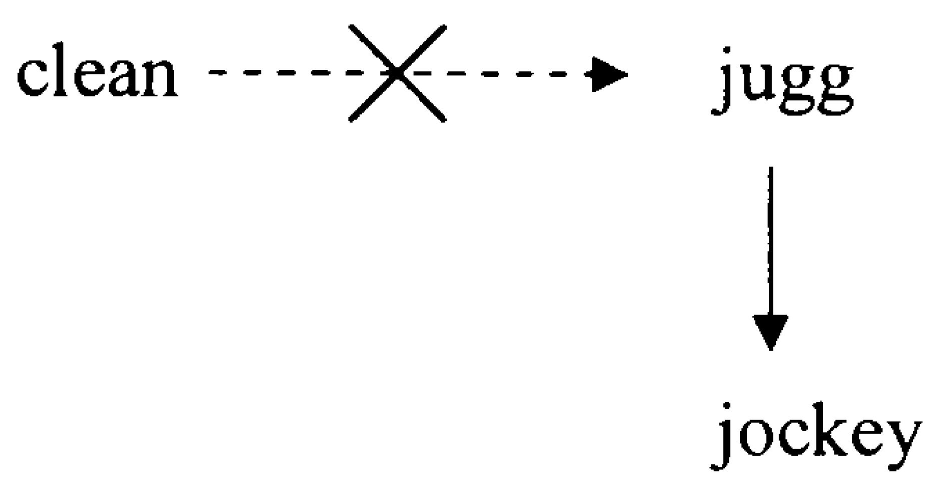
Participant 66; response to Q12: *"I tried to see whether the connection between the occupation which was mentioned twice and the other two occupations was stronger, e.g. All or some. If one of the statements was none I tended to, most of the time, put* *"none of the..." in my conclusion".*

Others made similar comments and indicated the relative quantifiers by thicknesses, or different forms, of arrow as in protocol 2.23.

Protocol 2.23

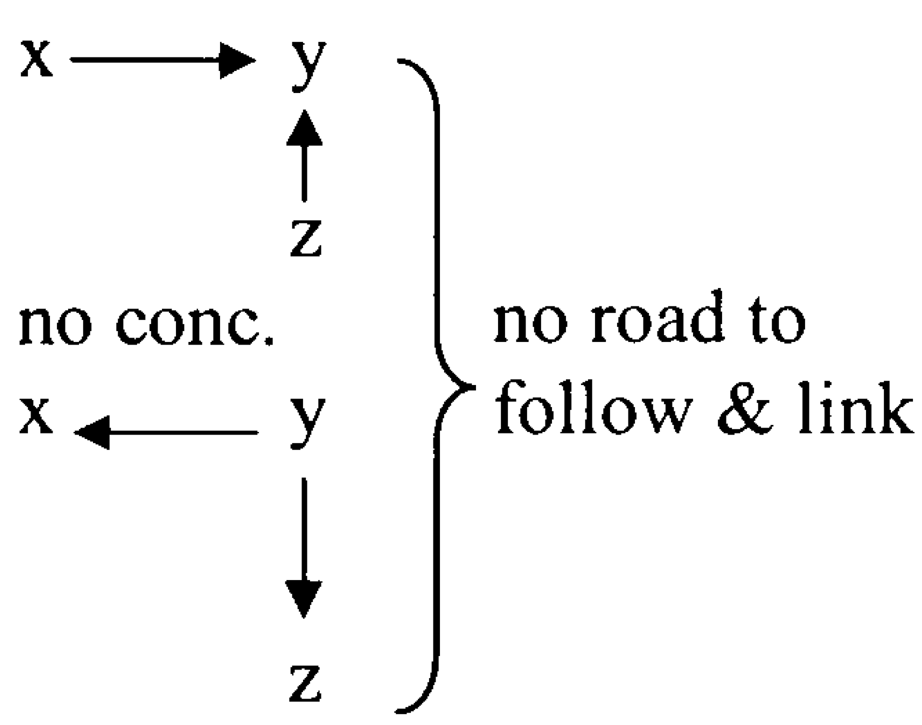
Syllogism 8P:	None of the cleaners are jugglers All of the jugglers are jockeys

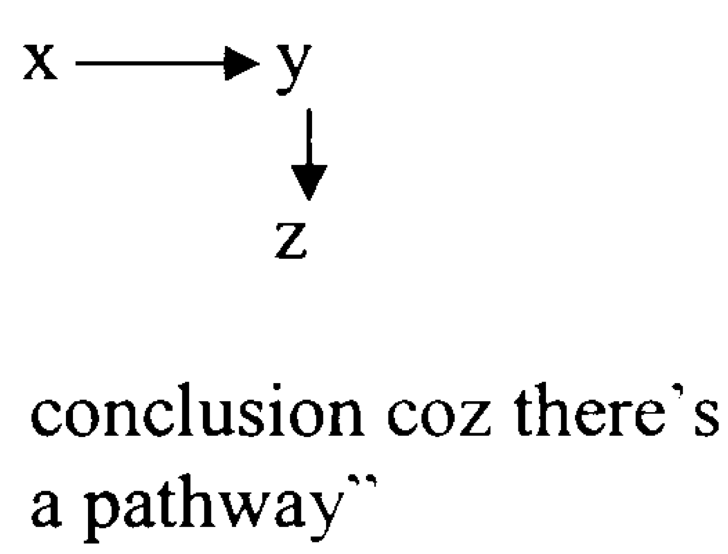
Participant 14:



Incorrect conclusion given: *None of the cleaners are jockeys*

Response to Q12: *"dotted line meant some, solid meant all, cross meant none and the arrow showed the direction of information given. If the arrows did not follow then no conclusion could be made, i.e.*





On this DF problem, the reasoner presents the incorrect conclusion which might be predicted were she using substitution, yet her protocols do not make this explicit. These data suggest that such flow-charts do likely indicate a form of verbal reasoning, but that the strategy does not seem to be substitution, either in terms of the naïve form described in Experiment 1, or the more sophisticated forms described by Ford. Conversely, it could be argued that, although dealing with the information propositionally, this approach seems to have a visuo-spatial element as evidenced by the description of "following pathways".

Altogether 31 (42.3% of total sample) of participants in Experiment 2 showed evidence of using these flow-chart diagrams.

Some variation was also apparent within the spatial strategy group. Of the 7 spatial reasoners identified in Experiment 2, five presented the typical Euler circle type representations. However, two individuals produced protocols of the type shown in protocol 2.24 below.

Protocol 2.24

Syllogism 10P	Some of the managers are boxers None of the boxers are judges	
Participant 65:	M – B M – B M M	J J

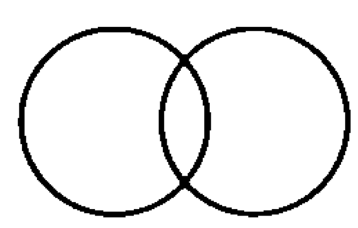
Correct conclusion given: *Some of the managers are not judges.*

These protocols present striking similarities to the classic notation for mental models (see chapter 2) however both participants denied ever having heard of them. Although only two reasoners presented such protocols, they merit attention as they still represent almost a third (28.7%) of the spatial reasoners. However, they present fewer classification problems than the flow-charters as these protocols clearly indicate some form of set based representation, which is an indictor of a spatial strategy.

2.3.4.2 Evidence from Questionnaire Responses

In Experiment 1, a Cronbach's Alpha = 0.68 was observed for eight valid questionnaire items and these formed the basis for the revised form of the questionnaire administered in Experiment 2 (Appendix 2G). In determining reliability of this revised questionnaire, the same procedure was applied as reported previously for Experiment 1. Reliability analysis of the 11 quantitative items suggested the elimination of 8 of these in order to obtain an optimum Cronbach's Alpha of 0.77. SPSS output showing this analysis is presented in Appendix 2I. However, the remaining 3 items (numbers 4, 5 and 6) were all verbal strategy questions. The most reliable spatial item was number 10. As it was desirable to have at least one spatial item included, it was decided to retain item 10 as that had also proved reliable in Experiment 1. These four items produced a final Alpha = 0.70. The list of final 4 valid items is shown in Table 2.8 below.

Table 2.8: The final 4 questionnaire items. Note that items 4, 5 and 6 were presented with a 5 point response scale as shown in original questionnaire, Appendix 2G. V or S in parentheses indictes whther item was intended to identify verbal or spatial reasoning behaviours repectively.

4.	Did you attempt to substitute terms/occupations from one statement to another (i.e. switch the occupations around between statements)? (V)
5.	Did you attempt to combine the two statements into one to form a single, longer verbal description? (V)
6.	To what extent did you attempt to reverse the position of the occupations within the statements? (V)
10.	Did you think about sets/Venn Diagrams (similar to that shown below) either using circles or any other shape? YES/NO  (S)

Questionnaire item responses again appeared to provide converging evidence for the presence of verbal and spatial strategies, with reasoners whose written protocols identified them as verbal, responding positively to verbal questionnaire items and spatial reasoners responding most positively to the spatial items. Participants allocated to the mixed and indeterminate strategy groups generally provided mixed or conflicting responses to the questionnaire items. The low number of spatial reasoners may have reduced the influence of spatial items in questionnaire reliability.

The 31 reasoners who used the "flow-charting" approach presented varying questionnaire response profiles. In order to determine whether these represented a third distinct strategy group, flow-chart users were divided into two groups on the basis of whether or not they answered "Yes" to item 10. Hence it was assumed that one group of flowcharters (N=9) were using some kind of mixed strategy (part verbal and part spatial) whilst others (N=22) were employing a verbal-propositional strategy. This classification is supported by comments made in response to item 12 of the questionnaire where participants were asked to try and describe their overall approach. Participants in the verbal flowchart group tended to describe typical verbal reasoning actions such as:

> "wrote each stage out again and used arrows to form the reinterpreted phrases...inserted some, all or None" (Participant 46)

> "kept sight of words all and some, I tried to make the statements into one sentence" (Participant 63).

Those in the mixed flowchart group tended to use terminology typical of spatial reasoners, for instance:

> "it was easier to look visually at the problem rather than put it in a sentence" (Participant 57).

> "tried to find relationships between sets..." (Participant 4);

Others in this category indicated they had used a mixture of verbal and spatial approaches, for instance, (comments in square brackets added for clarification):

> "...switch the top one for the bottom one" [as verbal substitution]...... "using diagrams to keep the groups distinct from one another" [as spatial reasoning]. (Participant 40).

> "Firstly represent he information diagrammatically [as spatial reasoning]....then put information into a verbal form to move it around"[as verbal substitution] (Participant 18).

Reverse–scaling the questionnaire items relating to the spatial strategy resulted in an overall reasoning behaviour scale with a maximum score of 47, on which a high score represented a high degree of spatial reasoning traits and a low score indicated a high degree of verbal reasoning traits. The mean scores for all five strategy groups (no indeterminate included) are shown in Table 2.9 both when the scale is based on all eleven quantitative questionnaire items, and for when it is based on just the four most reliable items. In both cases, the predicted pattern of scale scores was observed, with the spatial group scoring highest, followed by the mixed groups. The verbal and verbal-flowchart groups present very similar scores and these are also the lowest scores overall.

Table 2.9: Mean reasoning behaviour scores for all strategy groups based on scales derived from questionnaire responses. Shown in terms of scales based on all 11 original items, and on 4 most reliable items (latter as in table 2.8).

Strategy	N	Mean		S. Dev.	
		11 items	4 items	11 items	4 items
Verbal	23	23.57	8.30	5.12	3.39
Spatial	7	33.43	11.14	3.82	3.57
Mixed	2	33.50	10.00	5.00	1.41
Flowchart (verbal)	22	27.00	8.95	4.54	3.42
Flowchart (mixed)	9	30.44	11.67	4.53	3.28

When all 11 questionnaire items are considered, analysis of variance comparing mean scale scores for all five strategies indicated significant differences between the groups; $F(4,58) = 8.47$, MSE = 187.58, $p < 0.001$. Scheffé post-hoc comparisons indicated significant differences in scores between the verbal and spatial groups. Differences are also observed between these and the mixed and verbal flowchart groups respectively, ($p<0.05$ in every case). When this analysis was repeated using a scale comprised of scores on the four most reliable questionnaire items only (as Table 2.8), ANOVA indicated that inter-group differences approached significance: $F(4,58) = 2.19$, MSE = 25.04, $p = 0.08$. All analysis of variance tables pertaining to Experiment 2 are presented in Appendix 2J.

With the indeterminate group again removed from analysis, K-means cluster analysis was applied to the responses to the 4 valid questionnaire items. Two distinct clusters of participants were revealed and the distribution of strategies across these clusters is shown in Table 2.10. However, because several cells have an expected count of less than 5, Pearson's Chi square was not computed in this case.

Table 2.10: K-means cluster analysis of questionnaire responses: cluster membership by strategy. Figures represent within strategy group numbers and percentages.

Strategy	Cluster 1 (N=34)	Cluster 2 (N=29)	Totals (100%)
Verbal	16 (69.6%)	7 (30.4%)	23
Spatial	3 (42.9%)	4 (57.1%)	7
Mixed	1(50%)	1 (50%0	2
Flow (verbal)	13 (59.1%)	9 (40.9%)	22
Flow (mixed)	1 (11.1%)	8 (88.9%)	9

2.3.4.3 Reasoning Performance

Overall

The performance of all five strategy groups was compared across the two tasks. All reasoners performed best on evaluation as might be expected and this difference was significant in every case ($p < 0.01$). Table 2.11 presents the descriptive statistics. Groups are compared in terms of percentages correct as there are unequal numbers of items in the two tasks. A 5 (strategy) x 2 (task) mixed analysis of variance duly indicated a significant main effect of task type, $F(1,58) = 225.21$, MSE = 21477.53, $p < 0.001$; but no significant interaction between task and strategy, nor main effect of strategy ($p > 0.05$).

Table 2.11: Overall performance in terms of percentage of correct items compared across strategy group and task type

Strategy	N	Mean		Std. Dev.	
		Production	Evaluation	Production	Evaluation
Verbal	23	25.22	61.52	15.43	12.74
Spatial	7	34.28	62.14	24.17	9.06
Mixed	2	13.33	67.50	9.43	12.12
Flow (verbal)	22	22.72	57.27	10.82	10.20
Flow (mixed)	9	25.93	63.33	9.69	10.61

Performance on the Conclusion Production Task

Table 2.12 compares accuracy in terms of mean numbers of correct conclusions to valid problems and "no valid conclusion" (NVC) responses to invalid problems. The mean numbers of same-form conclusions generated to all three types of syllogism are also presented. The accuracy values in Table 2.12 show that, in line with Experiment 1, spatial reasoners perform slightly better than verbal.

Table 2.12: Conclusion production task: Descriptive statistics comparing all strategy groups in terms of accuracy and number of SF conclusions produced.

Strategy	Valid Same-form (Item N = 5)		Valid Different-form (Item N = 5)		Invalid (Item N = 5)	
	Mean	S. dev.	Mean	S. dev.	Mean	S. dev.
Accuracy						
Verbal	2.17	1.40	0.91	0.95	0.70	1.11
Spatial	3.14	0.38	0.71	0.50	1.29	1.98
Mixed	2.00	0.41	.00	.00	.00	.00
Flow (verbal)	2.45	1.18	0.55	1.01	0.41	0.80
Flow (mixed)	2.00	1.41	.00	.00	0.67	1.12
SF conclusions						
Verbal	3.65	1.70	2.39	1.41	1.91	0.79
Spatial	4.86	0.38	3.71	1.38	1.43	0.98
Mixed	4.50	0.71	4.00	1.41	2.00	1.41
Flow (verbal)	3.91	1.60	3.59	1.65	2.23	0.75
Flow (mixed)	4.67	0.71	3.44	1.51	2.11	1.41

This holds across all three syllogism types, however the differences were not significant ($p > 0.05$ in every case). For invalid syllogisms, accuracy entailed producing a NVC response and, in line with prediction, verbal reasoners produced fewer of these than spatial, though not significantly so ($p > 0.05$).

Again in line with the findings from Experiment 1, all reasoners performed less well on valid DF problems, compared to valid SF. A mixed analysis of variance shows a significant main effect of form; $F(1,58) = 59.03$, $MSE = 61.16$, $p < 0.001$. But again there is no significant interaction between form and strategy or main effect of strategy. The à priori predictions for invalid syllogisms seem to be supported as verbal reasoners indeed produce more invalid SF conclusions than spatial reasoners, and this difference is marginally significant; $t(51) = 1.97$, $p = 0.06$. Furthermore, when verbal reasoners' responses to invalid syllogisms are compared, they are found to produce significantly more invalid SF conclusions than they do correct NVC responses; $t(45) = 6.41$, $p < 0.001$.

Performance on Conclusion Evaluation Task

For all items on this task there were only two possible responses, the conclusion presented could be rated as valid or invalid. Table 2.13 compares performance between reasoners in terms of how many items they correctly evaluated across the four types of syllogism presented.

Table 2.13: Conclusion evaluation task: Descriptive statistics comparing all strategy groups in terms of numbers of conclusions correctly evaluated.

Strategy	Valid Same-form (N = 5)		Valid Different-form (N = 5)		Invalid Same-form (N = 5)		Invalid Different-form (N = 5)	
	Mean	S. dev.	Mean	S. dev.	Mean	S. dev.	Mean	S. dev.
Verbal	4.26	0.96	3.43	1.34	2.47	1.08	2.13	1.35
Spatial	5.00	0.00	3.57	1.62	2.00	1.41	1.86	1.07
Mixed	5.00	0.00	4.00	0.00	2.00	0.00	2.50	2.12
Flow (verbal)	4.23	0.92	3.23	1.34	1.82	1.33	2.18	1.33
Flow (mixed)	4.56	0.73	2.89	1.17	2.56	1.33	2.67	1.32

On the valid SF conclusions, all participants perform well, and spatial reasoners are significantly better than verbal. All spatial reasoners got all these items correct and hence endorsed significantly more SF conclusions as valid, $t(45) = 5.15$, $p < 0.001$. There was little difference between the performance of verbal and spatial reasoners on the other three forms of syllogism in this task ($p > 0.1$ in every case). A 5x2x2 (strategy x form x validity) mixed analysis of variance indicated significant main effects for syllogism form; $F(1,58) = 44.68$, $MSE = 95.86$, $p < 0.001$; and for validity; $F(1, 58) = 11.77$, $MSE = 8.77$, $p = 0.001$; but not for strategy ($p > 0.05$). A significant interaction between syllogism form and validity; $F(1,58) = 8.24$, $MSE = 12.17$, $p < 0.05$ was also apparent. On valid syllogisms, reasoners evaluated more conclusions correctly when they were the same form as one of the premises, compared to when the conclusion was of a different form. On invalid syllogisms, however, they performed similarly on all problems, irrespective of conclusion form.

Given the diversity of strategies identified in Experiment 2, and the small numbers in some groups, it was difficult to draw any firm conclusions based on the data, hence analysis was concluded at this point. The nature of the verbal and mixed flowchart strategies, as evidenced by the protocols and questionnaire responses discussed previously, provides a rationale for combining these with the verbal and mixed strategy groups respectively. When the above analyses were replicated with these new, larger, strategy groupings, the results were virtually identical.

2.3.4 Discussion

The aims of Experiment 2 were to further support the evidence for individual differences in strategies identified in Experiment 1. The first aim was to replicate the verbal and spatial strategies identified previously. However, although clear evidence of these strategies was again observed, for some participants the classification was not straightforward. These individuals presented diagrams which they tended to describe as flowcharts, and which seemed to present both verbal and spatial reasoning characteristics. They used verbal tools such as arrows and propositional terms, but laid out the diagrams in a spatial fashion. Moreover, when asked to describe their approach in the questionnaire, some participants wrote of representing sets/groups (a typically spatial reasoning trait), whilst others described typically verbal reasoning behaviours such as rewriting the premises into one sentence or swapping terms between premises. Examination of the quantitative questionnaire responses for these individuals suggested that some presented a mixed profile whilst others appeared to use a verbal strategy. Without further evidence however, it may be inappropriate to simply amalgamate these participants with the verbal and mixed strategy users.

There are a couple of possible reasons why this may have occurred. One major procedural difference to Experiment 1 was that participants were run in groups. Whilst every attempt was made to arrange seating so as to minimise the sight of others work, it was impossible to eliminate this completely. It is therefore possible that some participants sought inspiration from their neighbour's work and saw flowcharts as a recognisable and understandable approach which they could emulate. In Experiment 1, some participants were unsure about the written protocols aspect and sought reassurance from the researcher that they were "doing it right". In a group situation, subjects are reluctant to do this, or to ask for clarification at all. This may also explain the relatively high number of indeterminate strategies, where, despite instructions, no protocols were produced.

Another explanation may simply be that there are variations within the verbal and spatial strategies – with not all spatial reasoners using Euler circles and not all verbal reasoners substitution. They may still however be reasoning with information in a spatial or propositional form respectively. This requires further investigation. Certainly the above data suggest some justification for classifying verbal-flowchart and mixed-flowchart reasoners within the verbal and mixed strategy groups respectively. Furthermore, Experiment 1 showed that even when the vast majority of reasoners seem to adopt either a verbal or spatial approach, some overlap in their reasoning scale scores is observed. If a strategic continuum is envisaged, then it may be that the "flowcharters" lie somewhat towards the centre of the spectrum, with some showing strategies which incline towards the verbal end of the continuum, with others more allied to spatial strategies.

The second aim of Experiment 2 was to test predictions regarding the nature of verbal and spatial strategies, based on observations from Experiment 1. Two reasoning tasks were presented, involving the production and evaluation of syllogistic conclusions, and including invalid problems. It was expected that naive substitution would lead verbal reasoners towards a tendency to produce SF conclusions even for indeterminate syllogisms. Similarly, they were expected to be more likely to endorse SF conclusions as valid, than DF, even when the conclusion was logically invalid. However, the diversity of approaches identified in Experiment 2 meant that it was difficult to analyse performance meaningfully in line with these predictions. The figures presented in Tables 2.12 and 2.13 above, however, suggest little differences between strategy groups.

Little difference in accuracy was observed between strategy groups in terms of the predictions regarding performance across conclusion and evaluation tasks. Indeed, on invalid syllogisms where the greatest differentiation between verbal and spatial reasoners was expected, both groups seemed similarly reluctant to suggest a *No Valid Conclusion* response, preferring to produce an incorrect SF conclusion. For instance, syllogism 21 (*Some A are B; Some B are C*), is a problem where previous studies have indicated that reasoners often do produce the correct response (for instance, Dickstein, 1978, found that 58% of subjects correctly presented a *NVC* response to this syllogism). However, in Experiment 2, only 9 reasoners (12.5%) of the entire sample produced this response, of whom only 1 fell into the verbal strategy group and 2 into the spatial group, the remaining 6 exhibiting other strategies.

The third aim was to further develop the questionnaire as tool for identifying reasoning strategy. Reliability analysis again produced a satisfactory alpha value, though only for a minority of items, and only one of these measured spatial reasoning behaviours. Given that this questionnaire was an extension of the reliable eight item measure from Experiment 1, it might be expected that these items at least would prove to be good predictors of strategy in Experiment 2. One possible explanation that the sample population differed in some way. Certainly a more diverse array of representational approaches was observed in the later study. However, the majority of these do seem to be closely related to the verbal and spatial strategies observed previously. The issue of group administration mentioned above may be a factor here also, with some participants possibly producing a style of protocol which was not representative of their own reasoning processes. However, the four most reliable items, (numbers 4, 5, 6 and 10) were also identified as reliable in Experiment 1, suggesting that they may indeed provide useful indicators of these aspects of strategy. Moreover, the addition of a final qualitative item provided some useful information which assisted with the understanding of the "flowchart" strategies. However, the predominance of such non-typical approaches raises further questions about the nature of strategies and possible intra-strategy variations which Ford's work (and hence also the replication in

Experiment 1) does not address.

2.4 Overall Summary of Chapter 2

Chapter 2 presented two initial experiments which investigated the existence and nature of individual differences in strategies adopted for syllogistic reasoning. Experiment 1 successfully replicated the work of Ford (1995) which identified two main types of strategy. Like Ford, the findings of Experiment 1 suggest that these strategies involve reasoning with information in either a verbal-propositional or visuo-spatial form. This is clearly supported by the evidence for written and verbal protocols presented in Protocols 2.1 to 2.17. However, whereas the spatial reasoners adopted Euler circle type representations like those presented by Ford, the present data showed little evidence of the substitution rules which Ford described for verbal reasoners. Rather, these reasoners use a naive substitution which relies little on quantifiers, and heavily on premise form, in drawing a conclusion. Also contrary to Ford's findings, the only significant differences in conclusion accuracy observed as a function of strategy were between SF and DF syllogisms, with spatial reasoners performing significantly better than verbal on the latter.

These findings led to the rationale for Experiment 2. Firstly it was desirable to further replicate the findings regarding the presence of verbal and spatial strategies. Also, if verbal strategy users are indeed reasoning by means of naive substitution and relying on premise form, it is possible to predict how they will perform over different types of syllogism. However, the strategy data from Experiment 2 were not as clear as expected. Although verbal and spatial reasoners were again identified (as evidenced in Protocols 2.18 to 2.21), a fair proportion of individuals in Experiment 2 presented written protocols which resembled flowchart diagrams which appeared to incorporate elements of both verbal and spatial representations (see protocols 2.22 and 2.23). Qualitative questionnaire data suggests that some of these reasoners are adopting an alternative verbal strategy whilst others seem to use a mixed approach. Moreover, some spatial reasoners appeared to use a process which, although clearly spatial, did not involve the characteristic Euler circle type representations (See Protocol 2.24).

A questionnaire was developed with the aim of identifying verbal and spatial reasoning behaviours, but only 4 items (3 verbal and one spatial) have shown consistent reliability. As such, the questionnaire seems better at identifying verbal than spatial behaviours, though the latter remain seriously underspecified anyway, especially in terms of how the conclusions are actually drawn from the spatial diagrams.

Despite the data from Experiment 2 being less clean than expected, the fact that it was possible to replicate Ford (1995) strategies in Experiment 1, and that clear evidence for verbal and spatial strategies was again observed in Experiment 2, does suggest that the strategies are present and robust. This raises the question as to what factors give rise to one strategy rather than another. Moreover, the strategies have thus far been identified only for syllogistic reasoning. If there exists some inherent difference between individuals that leads to the employment of either a verbal or spatial strategy, those strategies should be apparent across other reasoning tasks also. Experiment 3 began an investigation into this possibility. Moreover, Experiment 3 returned to the original methodology of running subjects individually, in order to minimise any group effects and try to ensure that protocols produced are a reflection of an individual's own work, and not inspired by that of others around them.

CHAPTER 3

STRATEGIES IN OTHER REASONING DOMAINS

3.1 General Introduction to Experiments 3 and 4

Experiments 1 and 2 have revealed clear evidence of individual differences in strategies for syllogistic reasoning. Verbal and written protocol evidence has suggested that, for most people, reasoning with syllogistic premises involves using information in *either* a verbal-propositional or visuo-spatial form. A smaller number of individuals have also presented mixed profiles whereby they seem to use a combination of the two approaches, either across different problems or at different stages within the generation/evaluation of a conclusion. Now chapter 3 presents two experiments which extend these findings to a wider range of reasoning tasks. If the tendency to adopt a verbal or spatial strategy on syllogistic tasks is due to inherent individual differences in preference for how information is cognitively processed and manipulated, then it would be expected that these individuals will also adopt related strategies in other reasoning domains. Experiments 3 and 4 will investigate this hypothesis with regard to two other reasoning paradigms, transitive inference and sentence-picture verification. In both cases, previous research has indicated possible individual differences in strategies.

3.2 Experiment 3: Strategies in Transitive Inference

3.2.1 Introduction

Many everyday decisions are based on inferences about the relationship between objects, for instance, evaluating the relative merits of two products before deciding which to purchase. Many such relations are transitive, that is, when the relationship holds for two entities *a* and *b*, and also for *b* and another entity, *c*. Then it can be inferred that the relationship also holds between *a* and *c*. For instance, if *a* is greater than *b*, and *b* is greater than *c*, then it follows necessarily that *a* is greater than *c*. Relationships such as "is less than," "is greater than," "is subsequent to," "is parallel to" are all transitive and can link together three terms in a unique fashion. Conversely, other everyday relationships such as "is next to" is not transitive.

Such inferences have usually been studied in the form of three-term series problems, i.e. linear syllogisms. For instance, if three terms can be arranged in a linear sequence, ABC according to their relative properties, and given that property to be height, the information can be expressed as two premises in the linear syllogism:

A is taller than B	or alternatively,	B is not as tall as A
B is taller than C		C is shorter than B

Both describe the same linear array of objects unambiguously. By introducing inverse relational adjectives and negations, and given two possible premise orders in which these can be presented, overall there are 32 possible pairs of premises which describe a single given three-term array, such as that above, and which give rise to valid transitive inferences. Experimental tasks which present such linear syllogisms typically ask participants to determine the relationship between the initial and end terms (A and C in the above example), or to decide which is the taller/shorter etc of the two terms. Although most people tend to get the correct answers to such simple linear syllogisms, the time taken to reach those conclusions varies between problems and such variations in latencies suggests variations in difficulty (see Evans, Newstead and Byrne, 1993 for review). In general, problems which contain negation or inverse relational adjectives tend to require longer response times.

The mental models versus logical rules debate has also been contested over linear syllogisms. Johnson-Laird and Byrne (1991) have proposed that reasoners construct a mental array of the entities described in terms of their spatial layout, relative properties etc. As with categorical syllogisms discussed previously, in some cases the layout is ambiguous and more than one possible model can be constructed. Hence, errors on such problems can be attributed to failure to consider all possible models of the premises. In contrast, rule theorists (e.g. Rips, 1994; Hagert, 1985) suggest that humans possess an implicit logical rule for transitive inference. For simpler problems, the required relationship can easily be inferred from such a rule. On more complex problems however, for instance those with larger number of terms, which operate in more than one spatial dimension, or which involve negation, additional rules are required. Again as for categorical syllogisms, more rules equals more errors. However, the problem complexity suggested by rules theories does not map cleanly onto that predicted by model theory and some one-model problems require more rules than do multi-model problems (Byrne and Johnson-Laird (1989). Some studies of relational inference, where the adjectives describe a spatial relationship between terms (i.e. to the left/right of, behind, above etc) or temporal reasoning (before, after) have suggested that such relations encourage a spatial strategy and present support for a models approach (Vandierendonck and De Vooght, 1996). However, Roberts (2000) has shown that although a mental-model based approach seems to be the prevalent strategy, some individuals (around 10%) will spontaneously use a verbal strategy and their data seem to match the predictions of the rule theories. Johnson-Laird and Byrne (1991) conclude that the model theory assumes an understanding of the semantic meaning of the relational adjective, rather than of its logical properties. They argue, for instance, that both "on the right of" and "on the left of" are transitive. Being able to make an appropriate inference based on such information relies on semantic knowledge. An abstract rule for transitivity would not be able to distinguish between the two meanings.

Theories specific to linear syllogisms have also tended to favour either a models or rules perspective. Spatial array theories such as those proposed by DeSoto, London and Handel (1965) and Huttenlocher (1968) were precursors of the early mental model theories (e.g. Johnson-Laird, 1983). They suggest that information from the two premises is integrated and then represented in a spatial array. To return again to the previous example:

A is taller than B
B is taller than C

the spatial array would be:

A
B
C

Such theories propose that some linguistic descriptions more readily lead to the construction of models than others. Descriptions such as "taller" typically lead to a vertical array (as above), whereas others, e.g. "faster", typically lead to a horizontal array. Moreover, whilst an array such as that above will tend to be constructed in a top down-order (i.e. A, then B, then C). Alternative descriptions, which contain negations and/or inverses, such as:

B is not as tall as A
C is shorter than B

would lend for construction of the same array but from the bottom up (i.e. C, then B, then A). Such bottom-up construction is more difficult and hence leads to longer response times for such problems.

An alternative view is that people use a linguistic representation which allows them to understand the relational meaning of the premises (Clark, 1969). Information from the two premises is not integrated, rather it represented by a set of linguistic propositions, for instance in the first example above, as a dimension of tallness, (A is tall, B is less tall, C is even less tall). Premises which describe negation are constructed on an inverse dimension, for instance in the second example above, a dimension of badness (C is bad, B is less bad, A is even less bad). In this latter case, longer response times result from the increased difficulty of constructing negated propositions.

In addition, the linguistic theories attempt to explain differences in difficulty between problems which have the same adjectives (e.g. A is better than B, B is better than C). It suggests that the information is compressed in working memory, i.e. the representation is reduced to information about the most extreme relationship only (i.e. A is best) effectively eliminating the pivot (common) term B. This must then be recalled from memory in order for comparison with premise two. For other forms of this problem however, (such as, B is better than C, A is better than B) the common term is retained in the compressed representation (B is best) allowing for an immediate comparison to be made, B being then discarded in favour of A.

A related model was proposed by Quinton and Fellows (1975) who claim that only a superficial linguistic representation of premises is sufficient to solve most linear syllogisms. This is facilitated by a simple algorithm which removes the need for more sophisticated (and hence cognitively intensive) representations proposed by the above linguistic theories. In addition, Sternberg (1980) proposed a mixed model whereby premise information was first decoded into a linguistic format and then recoded into a spatial format. The final response is elicited from scanning the spatial array, supplemented by a rechecking the linguistic propositions where required on more difficult problems.

A further source of difficulty concerns marked and unmarked adjectives. Dimensions often have a marked and unmarked end. The unmarked form of an adjective is usually that used to name the scale (e.g. taller, faster), whereas the marked adjective for is the negative form, usually used in a contrastive sense (e.g. shorter, slower). Research has shown that inferences involving marked adjectives are more difficult than those involving unmarked ones - though again the differences are measured in milliseconds. The spatial array view (e.g. DeSoto, et al 1965) maintains that individuals prefer to construct top-down arrays, so problems which contain premises which allow for this will be processed fastest. These premises are also those which contain unmarked adjectives. Those with marked adjectives in contrast, are those which take longer to process because their model requires bottom-up construction (e.g. "A is worse than B" requires a bottom-up array). The linguistic rule view (e.g. Clark, 1969) suggests that a marked adjective provides information about both entities involved (e.g. the taller and the shorter) and hence requires more information about their position in the array to be stored in working memory, than is required for unmarked adjectives.

All of the above theories, as with those concerning syllogistic reasoning, tend to assume universality of processes across all individuals. However, using the theories outlined above as a basis, Sternberg and Weil (1980) have shown that individuals may adopt different representational strategies for linear syllogisms, both spontaneously and under the influence of instruction. They found that many subjects spontaneously adopted a visuo-spatial strategy, but that others, who initially used a more linguistic approach, could be prompted to do so through training. The algorithmic model (Quinton and Fellows, 1975) proved the most efficient in terms of latencies, but, for logically untrained individuals, the mixed model advocated by Sternberg (1980) and involving a combination of verbal and spatial processes was most likely used.

Egan and Grimes-Farrow (1982) also found evidence of individual differences in strategy for this task, identifying two strategy types from written and verbal protocols. They presented linear syllogisms which they claimed encouraged the use of spatial representations: problems involving relational adjectives (e.g. above-below, left-right) or visual comparatives (e.g. rougher-smoother, fatter-thinner). The three terms were always the geometric figures circle, square and triangle, and premises were presented in both auditory and visual modalities over two experiments. Retrospective verbal and written reports indicated that some participants established a scale or continuum onto which the three terms were placed and then compared according to their relationship. This same process occurred whatever the relational adjective in the problem and the appearance of the terms remained standardised throughout (i.e. the three shapes appeared uniform throughout). These people were termed *abstract directional* thinkers. Other subjects attributed physical properties to the terms according to the relational adjectives and represented the objects as having these properties (i.e. the shapes were variously represented as large, small, light, dark etc) in their protocols. The relative properties were then simply compared to reach a conclusion. These individuals were termed *concrete properties* thinkers. Egan and Grimes-Farrow (1982, page 301) describe the two strategies as "verbal and non-verbal modes of describing problem representations". Although no diagrammatic data is presented in their paper, Egan and Grimes-Farrow do present examples of written retrospective reports produced by participants when asked to describe their reasoning. Some examples are shown in Table 3.1.

Table 3.1: Examples of written reports of reasoning on linear syllogisms for the two strategies identified by Egan and Grimes-Farrow (1982, page 301).

Abstract directional thinkers	Concrete properties thinkers
Problem relation: Rougher-smoother	
"Rather than imagining a rough/smooth figure, I put the figures in a horizontal line, in my mind, in the order of left/right rather than rough/smooth"	"I drew a picture, and if something was rough – I would put craters in it in my mind-smooth was just plain white".
Problem relation: Darker-Lighter	
"I set up a scale with the lightest on the far right and darkest on the far left and [placed the figures on their appropriate spots".	""In my mind, I coloured in the object that was darkest"
Problem relation: Fatter-Thinner	
"Put shapes in order from thinner to fatter"	""Made the figures fatter and thinner in my head".

A content analysis of the protocols was conducted, and found that the two strategy groups differed in the key words they used when describing their reasoning. Concrete properties (CP) thinkers used more words suggesting an image (e.g. picture) whilst Abstract directional (AD) thinkers tended to use words which suggested a scale (e.g. line, order). The words used to describe the process of creating such arrays also differed, with CP thinkers preferring words such as "draw" whilst AD thinkers used functional words such as "put".

Importantly, Egan and Grimes-Farrow also found differences in the errors made by the two groups of reasoners. AD thinkers made significantly fewer errors than CP which was attributed to the greater efficiency of generating a linear mental array over a concrete pictorial representation. Moreover, the two groups also seemed to differ in the types of problem they found most difficult. Spatial model theories (e.g. DeSoto et al, 1965) suggest an end-anchoring effect which leads reasoners using a directional representation to have least difficulty constructing their mental array if it is "end anchored", that is, conducted from the ends towards the middle. The implication for AD thinkers is that the difficulty of solving a problem should be related to the number of premises which have the common, or pivot, term first. This premise form leads reasoners to construct an array from the middle (starting with the pivot) to the ends (non-pivot terms then placed in appropriate positions in relation to the pivot). Egan and Grimes-Farrow's data suggested that this was indeed the case, AD thinkers had greatest difficulty with problems where the pivot term was presented first in at least one premise, their error rate increasing monotonically with number of premises where this was the case.

CP thinkers showed no such effect. Egan and Grimes-Farrow claim that the qualitative differences in the spatial arrays constructed by those reasoners suggest that end-anchoring is not an issue for them. However, they make greater use of inverse adjectives than do AD thinkers which leads to an alternative source of difficulty. Problems for CP thinkers occur where a change or alternation in relations is presented (for example, where one premise contains the adjective larger and the other smaller). If a relation and its inverse require differing concrete representations, then both will require storage and retrieval from working memory. CP error rate increased monotonically with the number of changes or alternations in relational adjectives contained in problems. AD thinkers showed no such effect.

According to the model proposed by Egan and Grimes-Farrow, AD reasoners encode premise one, establish a mental scale to encompass those two terms, the grammatical subject being placed first. In encoding the second premise, the subject searches for the third, missing term, which once located, is placed appropriately on the scale. The easiest case is when this position occurs immediately following the two terms already established (e.g. if the first two are placed smooth > rough and term three is the roughest of the three). Conversely, the more difficult cases arise when the third term must be placed at the other end of the scale associated with the inverse (e.g. if the first two terms are placed smooth >

rough and then term three turns out to be the smoothest of the three). CP thinkers on the other hand, are assumed to encode the premises and then create an image pairing, in which the subject of the premise takes on a property (e.g. rough) and the object remains neutral. Two such pairs are generated, one for each premise. When the question is encoded, the two image pairs are scanned for the response. CP thinkers find difficulty in generating pairs requiring an inverse relation, and where two pairs either present alternate relations or are inconclusive. Egan and Grimes-Farrow presented a process model of their data which accounted for the strategic differences and errors. Tested with multiple regression, the model was shown to account for 90% of variance in performance for AD and 80% in CP.

Both of Egan and Grimes-Farrow's strategies are assumed to possess a spatial element, indeed they set up their experiment in such way as to encourage this. However, clear qualitative differences are apparent between the two strategies, one (CP) is clearly visuo-spatial, the other (AD) comprising a more abstract representation which remained constant irrespective of the relational properties. An obvious parallel with the verbal and spatial strategies identified for categorical syllogisms in Experiments 1 and 2 suggest itself, more so given also that Sternberg and Weil (1980) have identified spatial, verbal and mixed strategies in solving linear syllogisms. Hence, Experiment 3 attempted to investigate a possible relationship between strategies used for syllogistic reasoning and for transitive inference involving linear three-term series problems. To avoid possible confusion between the two forms of syllogism, henceforth, linear syllogisms will be referred to as three-term series problems. When the term syllogism is used, it will refer to categorical syllogisms as presented in Experiments 1 and 2.

3.2.2 Aims

For syllogisms, individuals differ in whether they use visuo-spatial or verbal-propositional representations. The review above suggests that there may also be individual differences in strategies used for solving three-term series problems. A further question to investigate is whether there is cross-task consistency in strategy usage. A positive finding would indicate that strategy usage is a stable and consistent factor in reasoning, rather than merely an artefact of a given reasoning task. Hence Experiment 3 aimed to investigate the relationship between strategies on syllogistic reasoning and three term series problems. Specifically it was predicted that:

1. Verbal and spatial strategies would be identified on a syllogistic task, as for Experiment 1,

2. Verbal reasoners would be most likely to adopt an abstract directional (AD) strategy on three term series problems, whilst spatial reasoners would be most likely to adopt a concrete properties (CP) strategy.

3.2.3 Methods

Participants

66 undergraduate students from University of Plymouth volunteered to take part in return for course credit. The sample comprised 10 males and 56 females with a mean age of 22.52 years. None had received formal training in logic.

Materials

Categorical Syllogisms

Ten categorical syllogisms were presented as shown in Appendix 3A. Five were same-form problems and five different-form, with all four syllogistic figures represented. As in Experiments 1 and 2, the three terms comprised hobbies and occupations. Written instructions were presented exactly as for the syllogistic conclusion production task in Experiment 2, and shown in Appendix 2H.

Three-term Series Problems

Sixteen problems were presented as shown Appendix 3B. These were chosen to represent a range of premise forms and hence difficulty. Four problems contained two positive premises, four had two negated premises, and eight problems comprised one of each, four in each order. The positions of the pivot term varied, with 4 problems having neither pivot occurring first, 8 having one pivot first and four having both pivots first. The task was preceded by written instructions and two practice items, similar to those presented for the syllogistic task. These are reproduced in Appendix 3C. In order to avoid jargon for the participants' benefit, the above tasks were referred to as a "Conclusion Production Task" and an "Object Relationship Task" respectively.

Questionnaires

A reasoning behaviour questionnaire was administered following each task. For syllogisms, reliability analysis on the questionnaire administered in Experiment 1 has established a Cronbach's Alpha of 0.86 for 8 reliable items. Experiment 2 presented an 11 item questionnaire based on these items which proved less reliable due to the wide variety of strategies which became apparent in that experiment (Alpha = 0.72 but for just four reliable items, see Table 3.9). However, it also included an additional qualitative question which did yield some useful additional information about how people reason. For Experiment 3, an 9 item questionnaire was compiled which comprised the 8 most reliable spatial and verbal items Experiment 1 (which also included the four reliable items from Experiment 2) together with the qualitative item which asked participants to describe their approach. The questionnaire items are presented in Appendix 3D.

For the three-term series problems, a questionnaire was devised along similar lines, again with 8 items followed by the same final qualitative enquiry into the approach used. As it was intended to compare strategies across the two tasks, the questions are similar to those for syllogisms, but tailored to the task and tied into some of the key tenets of transitive inference described above. Four items attempted to capture the CP reasoning strategy by reference to the use of images and physical properties whilst four aimed at identifying the AD strategy, referring to representations of series, scales, continua etc. The 9 items are shown in Appendix 3E and include a final qualitative item as for syllogistic reasoning questionnaire.

Procedures

This experiment returned to the original procedure of running each participant individually, in sessions lasting around 45 minutes. Each participant completed both tasks, presentation order counterbalanced to control for possible order effects. Items were presented in a different randomised order for each person. The two tasks were presented together in a booklet, each task preceded by the relevant instructions and followed by the appropriate questionnaire. For both tasks, participants were instructed to write down their working out as in the previous two experiments and hence each item was again presented on a separate page, with room beneath to present the written protocol and conclusion. Participants were instructed, both in writing on the front cover of the booklet, and verbally by the experimenter, to work through the booklet in order presented.

3.2.4 Results

The data from 4 participants was excluded from analysis as they were found to have attended a course where logic had been taught. Hence for all the following analyses, total N = 62.

3.2.4.1 Strategies for Syllogistic Reasoning

The vast majority of subjects were classified as either verbal (N = 38) or spatial reasoners (N = 17) based on their written protocols. Where necessary for clarification, responses to qualitative item 9 of the questionnaire were also taken into account. A few mixed reasoners were also identified (N=5) with protocols and questionnaire responses clearly suggesting a combination of verbal and spatial approaches. Only a small minority (N = 2) were using a totally indeterminate strategy, and interestingly did so on both syllogisms and the three-term series task.

Questionnaire Responses

As in Experiment 1, strategy classifications based on protocols were fully supported by questionnaire responses. Reliability analysis conducted on the 8 quantitative items in the syllogistic reasoning behaviours questionnaire (Appendix 3D) suggested a Cronbach's Alpha of 0.83 when all 62 participants are included. SPSS output showing this analysis is presented in Appendix 3F. When just the verbal and spatial reasoners are included (N = 55), a repeat of this analysis produced an Alpha of 0.84. Reverse scoring the verbal questionnaire items produced a reasoning behaviour scale on which a high score indicated a high level of spatial reasoning behaviour and a low score indicated a high level of verbal reasoning behaviour. When the verbal and spatial strategy groups (as determined from their protocols and qualitative questionnaire item 9) are compared in terms of their scores on this scale, spatial reasoners (M = 31.24, sd = 3.4) scored significantly higher than verbal reasoners (M = 17.26, sd = 4.4); t (53) = 11.57, $p < 0.001$. As in Experiment 1, the mixed strategy group score lay in between (M = 24.6, sd = 4). K-means cluster analysis (see chapter 3, Experiment 1 for further explanation) on the 8 questionnaire items revealed 2 distinct clusters of participants. The distribution of verbal and spatial reasoners across these clusters is illustrated in Table 4.2. As so few mixed reasoners were identified, they have not been included in this analysis. The association between strategy grouping and cluster membership proved highly significant; $\chi^2 = 43.1$, df = 1, $p < 0.001$.

Table 3.2 K-means cluster analysis of syllogistic reasoning questionnaire responses: cluster membership by strategy. Figures represent within strategy group numbers and percentages.

Strategy	Cluster 1 (N=35)	Cluster 2 (N=20)	Totals
Verbal	35 (92.1%)	3 (7.9%)	38 (100%)
Spatial	0	17 (100%)	17 (100%)

Characteristics of the strategies identified in Experiment 3 will now be considered in more detail.

Verbal Reasoners (N = 38)

Of the verbal reasoners identified, 11 showed clear evidence of classic verbal substitution behaviour, for example:

Protocol 3.1

Participant 54
Syllogism 6

All singers are athletes
Some plumbers are singers

Correct conclusion given: Some plumbers are athletes

Some people even managed this on different form (DF) syllogisms, for instance:

Protocol 3.2

Participant 12
Syllogism 3

No biologists are musicians
Some vegetarians are musicians

Correct conclusion given: Some vegetarians are not biologists

However, others made the typical error for substitution on DF syllogisms. The correct conclusion in Protocol 3.3 should be *Some bellringers are not gardeners*.

Protocol 3.3

Participant 17
Syllogism 5

No gardeners are dentists
All dentists are bellringers

Incorrect SF conclusion given: No gardeners are bellringers

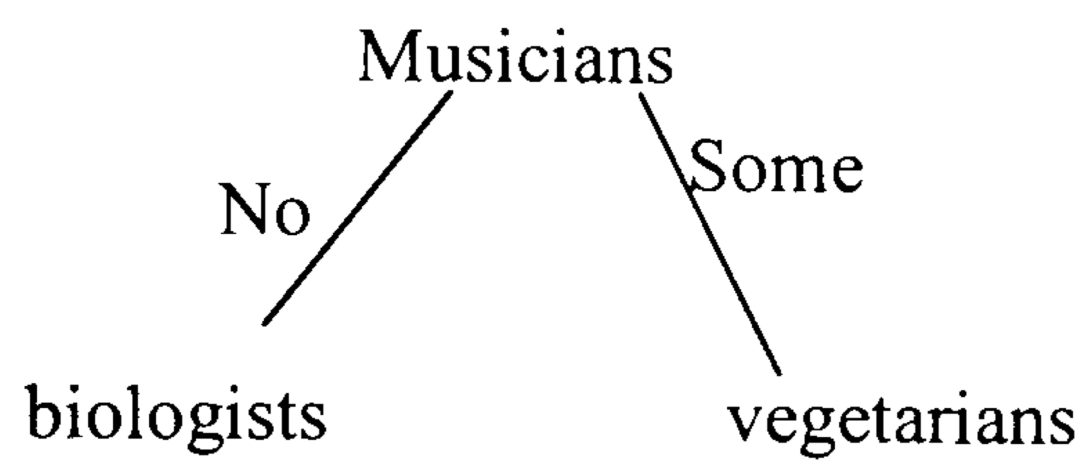
The remainder of the verbal reasoners produced protocols along the lines of those observed in Experiment 2, with the flowchart type diagram being again prevalent. These people typically showed a classic verbal questionnaire profile (answering in the affirmative to verbal items and negative to spatial items). However, on this occasion, the inclusion of questionnaire item 9 revealed some insight into this reasoning. By far the most prevalent form of protocol typically showed the middle term as central, with the other terms linked to it by lines or arrows, with the quantifiers indicated by words. as typified by the following:

Protocol 3.4

Participant 11
Syllogism 3

No biologists are musicians
Some vegetarians are musicians

Protocol:



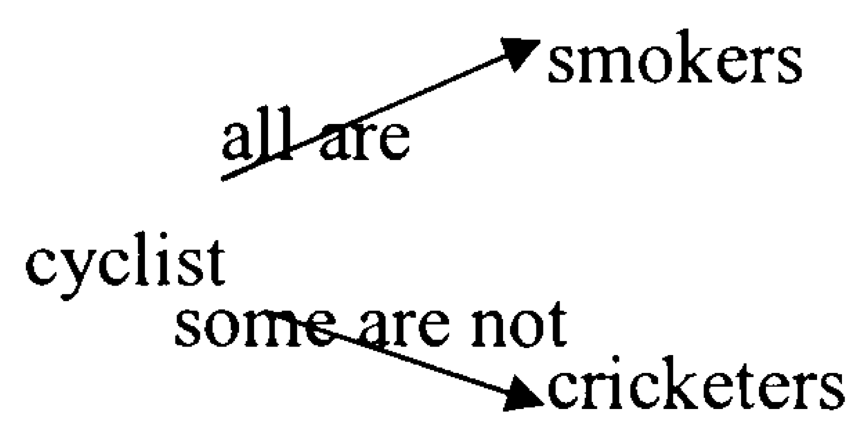
Correct conclusion given: Some of the vegetarians are not biologists

Protocol 3.5

Participant 34
Syllogism 2

All smokers are cyclists
Some cricketers are not cyclists

Protocol:



Correct conclusion given: Some cricketers are not smokers

For these people, the diagram seems to be a way of simplifying and integrating the premises in propositional form. After drawing the diagram, they frequently cancel out the middle term (literally on paper or mentally) and link the other two terms together according to their quantifiers to form a conclusion. This latter part of the process is frequently difficult to articulate and remains under-specified in some cases though it is probable that some kind of informal rule is applied. Participants frequently spoke of linking or joining the two end terms and their quantifiers to get a conclusion, frequently ignoring the middle terms, for instance;

Participant 34:

"...mentally cancelled out the common link and then joined the other parts together to get the conclusion."

Participant 30

..."attempted to combine the parts of the sentences not concerning the repeated one"

Participant 66

"...conclusion drawn by looking at the two occupations and whether they were some, all etc and then linking them together according to the arrows..."

Participant 50 was one who was able to describe this process in more detail. She described a rule for linking quantifiers to produce a conclusion along the lines of:

all + some = some

all + no = no

all + some-not = some-not

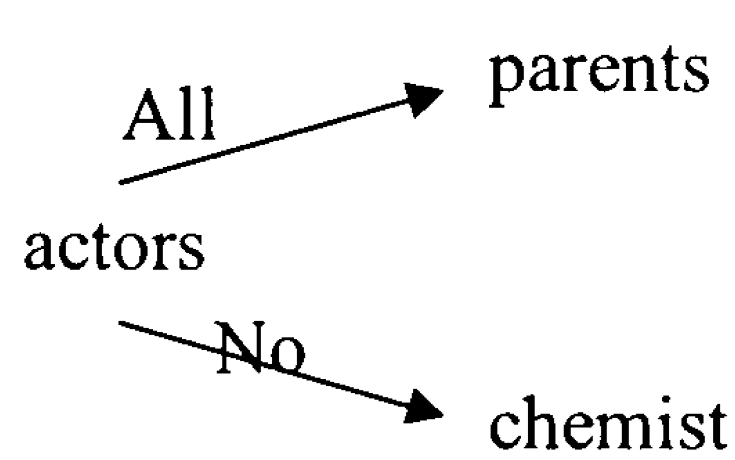
no + some = no

This would produce a typical verbal outcome with end quantifiers in line with those predicted by matching theory (Wetherick and Gilhooly, 1990; 1995, described earlier) and by verbal substitution. As the "linking-rule" shown above clearly results in a conclusion containing a quantifier corresponding to that from one of the premises, its application leads to errors for different-form syllogisms. Many of these participants seem to have a feel for the substitution process but have not quite cottoned on to the fact that they can just swap terms around to get the answer. Instead they try to work the problem out (i.e. actually attempt to reason) but are often unsure how to. Protocols 3.4 and 3.5 above show that the strategy is sufficient to obtain the correct conclusion to both SF and DF syllogisms, however, many of these people present a similar error profile to those using substitution. The use of a rule along the lines of that outlined above may lead to particular problems with DF syllogisms as Protocol 3.6 shows. In this case, by following the *all + no = no* rule, the incorrect conclusion produced is that which would generally be predicted by substitution. The correct conclusion should be: *Some parents are not chemists*:

Protocol 3.6

Participant 34
Syllogism 7 — All actors are parents
No chemists are actors

Protocol:



Incorrect conclusion given: No parents are chemists

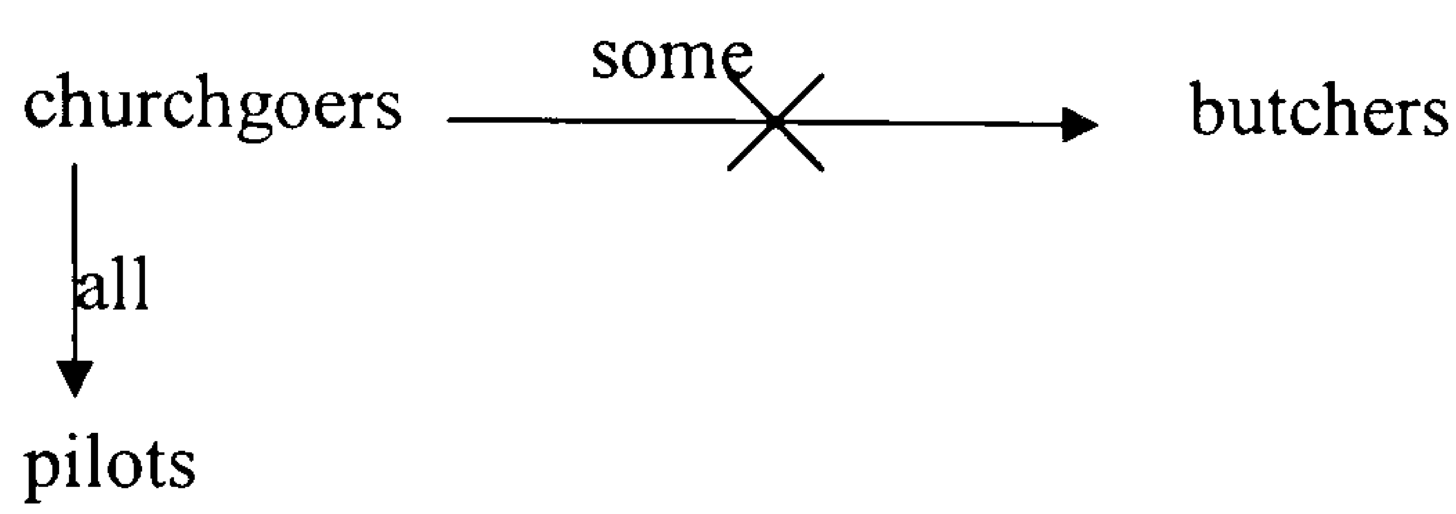
However, in some cases, incorrect conclusions to DF syllogisms may differ to those predicted by substitution. Whereas substitution often leads to an incorrect end quantifier (as in Protocol 3.6 above), Protocol 3.7 shows that when these linking rules are used, the end quantifier may be correct but term order is incorrectly reversed. The correct conclusion here should be *Some pilots are not butchers*. This may reflect the order in which reasoners have simply read the terms off their diagram. This in turn, may be a reflection of presentation order in the syllogism.

Protocol 3.7

Participant 57
Syllogism 1

Some churchgoers are not butchers
All churchgoers are pilots

Protocol:



Incorrect conclusion given: Some butchers are not pilots

This "linking rule" approach is clearly verbal-propositional in nature and related to substitution, though not as effective: Comparison of performance revealed that people who were clearly using substitution (N = 11, sd = 13.5) performed significantly better (65% correct) than those who employed the linking rule strategy (N = 27, 35.7% correct, sd = 15.5); $t(36) = 5.2, p < .01$.

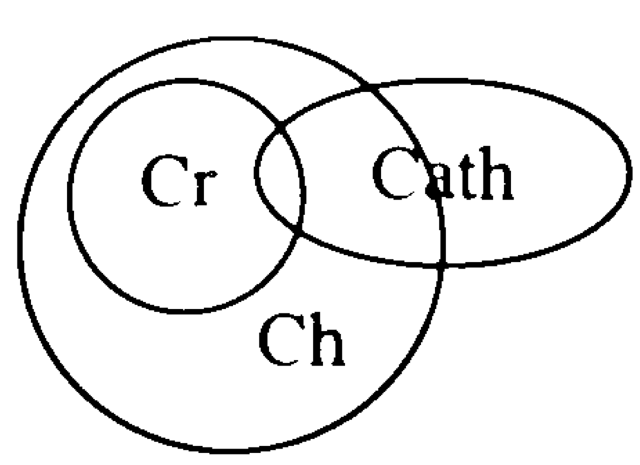
Spatial Reasoners (N = 17)

There was far less variation in tactics amongst the spatial reasoners with all but one producing the typical Euler circle type diagrams observed in Experiments 1 and 2, and by Ford (1995). Protocols 3.8 and 3.9 below illustrate cases for a SF and a DF syllogism respectively.

Protocol 3.8

Participant 48
Syllogism 4

All criminals are chessplayers
Some criminals are not Catholics

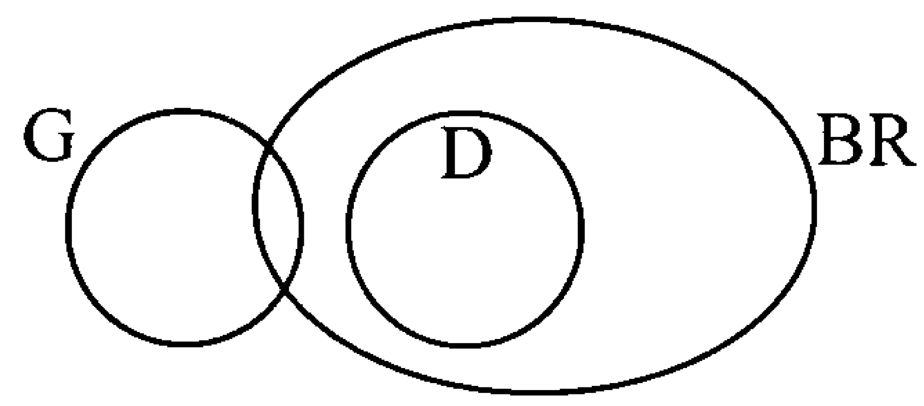


Correct conclusion given: Some chessplayers are not catholics

Protocol 3.9

Participant 13
Syllogism 5

No gardeners are dentists
All dentists are bellringers



Correct conclusion given: Some bellringers are not gardeners

The one remaining spatial reasoner presented the "mental models" type approach as shown by participants in Experiment 2, Protocol 3.24 previously.

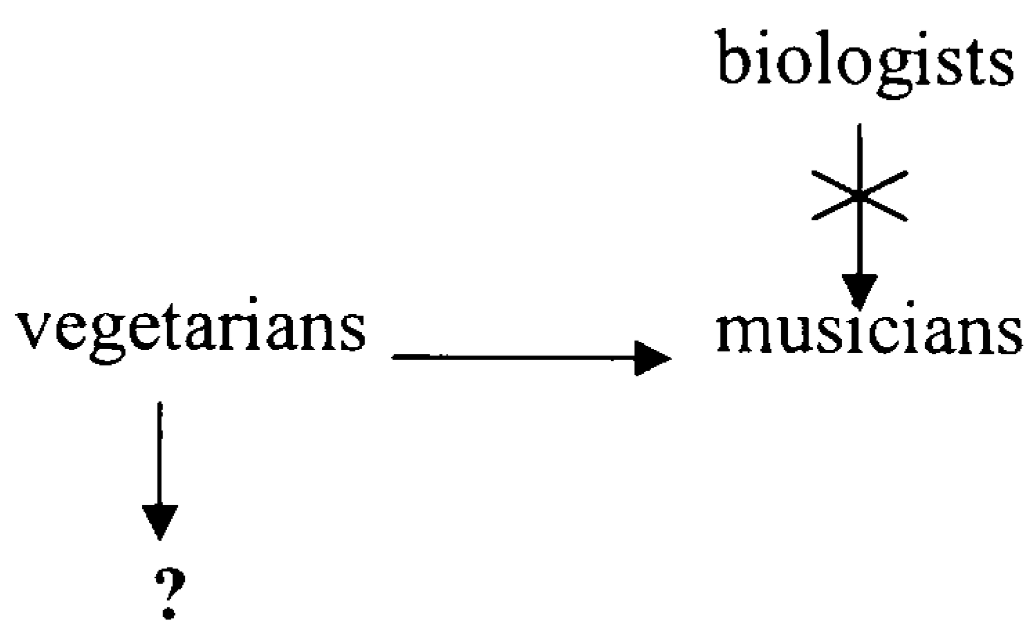
Mixed Reasoners (N = 5)

The 5 mixed reasoners all presented flow-chart type protocols very similar to those shown above for verbal reasoners. These participants tended to first draw the diagram which they described (questionnaire item 9) as a spatial layout which helped them to clarify information and organise the occupation groups in their mind. Then, secondly, they applied some kind of verbal rule to link the A and C parts of the model and draw a conclusion according to the quantifiers – in a very similar way to the verbal flow-charters described earlier. Participant 41 gave a very good description of her rule which was like that described earlier but with the addition of *some* + *no* = *some-not*. She used this principle successfully on the DF syllogism in Protocol 3.10.

Protocol 3.10

Participant 41
Syllogism 3

No biologists are musicians
Some vegetarians are musicians



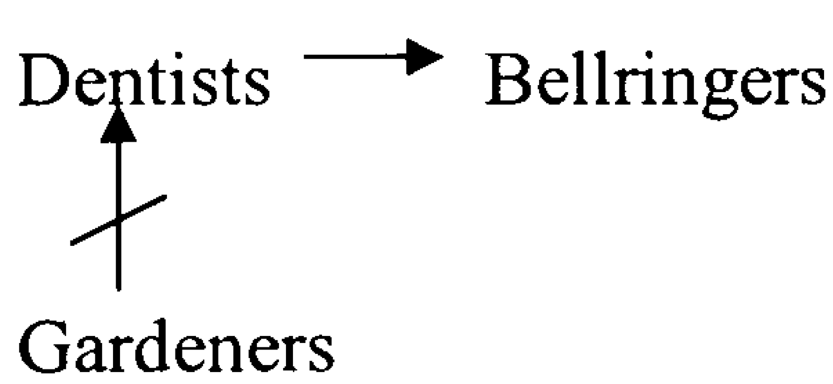
Correct conclusion given: Some vegetarians are not biologists

Other mixed reasoners applied substitution after drawing their diagram, though not always successfully. In Protocol 3.11, the participant makes a typical substitution error on this DF syllogism, even with the additional aid of his diagram. The correct conclusion is *Some bellringers are not gardeners.*

Protocol 3.11

Participant 40
Syllogism 5

None of the gardeners are dentists
All of the dentists are bellringers



No gardeners are dentists
All dentists are bellringers
create new statement

Incorrect conclusion given: No gardeners are bellringers

The mixed reasoners typically gave mixed questionnaire responses and emphasised that they needed to use *both* spatial and linguistic processes in the ways described above. As such, this is a two stage strategy, involving the spatial organisation of material, followed by application of a verbal rule. However, it remains unclear whether these people really are actually generating their conclusions verbally, or whether they are first forming a putative conclusion from the spatial representation, then verifying it propositionally.

Performance on Syllogisms

Measuring inter-strategic performance was not the main aim of Experiment 3 but some basic analysis indicated very similar trends to Experiment 1. Again spatial reasoners performed best overall (52.9% of syllogisms correct, sd = 15.3, as opposed to 43.4%, sd = 19.8, for verbal reasoners) though this difference was not significant ($p > 0.05$). Table 3.3 summarises performance across same-form and different-form syllogisms. On same-form problems the two strategy groups performed comparably and both verbal and spatial reasoners performed better on these than on different-form problems. On the DF problems alone, spatial reasoners did noticeably better than verbal.

Table 3.3: Percentage correct on same-form and different-form syllogisms, by strategy.

Strategy	Same form	S. Dev.	Different form	S. Dev.
Verbal	95.8	79.8	43.2	96.9
Spatial	85.8	75.4	52.9	62.4

A 2 (strategy) x 2 (form) mixed Analysis of Variance indicated a significant main effect of syllogism form; $F(1,53) = 5.35$, $MSE = 43004.20$, $p < 0.05$ though not of strategy ($p > 0.9$). The interaction between strategy and syllogism form suggested by the data in Table 3.3 was not found to be significant; $F(1,53) = 0.28$, $MSE = 2276.93$, $p > 0.1$.

3.2.4.2 Strategies for Transitive Inference

To recap, Egan and Grimes-Farrow (1982) suggested two strategy types for this task which they identified from written and verbal protocols. Some of their subjects established a scale or continuum onto which the three terms were placed and then compared according to their relationship. These people were termed *abstract directional* thinkers. Other subjects attributed physical properties to the terms and simply compared the relative properties to reach a conclusion. These individuals were termed *concrete properties* thinkers. The characteristics described by Egan and Grimes-Farrow were clearly evident in the present protocol data with abstract directional (N = 39) and concrete properties (N = 22) strategy groups identified. However the two strategies did not seem to be as distinct as Egan and Grimes-Farrow suggested, as the more detailed descriptions to follow will illustrate. Just one participant presented indeterminate protocols. For this task, the strategy classification could be made from protocols alone without recourse even to qualitative item 9 of the questionnaire.

Questionnaire Responses

The transitive inference questionnaire responses were subject to the same analyses as the syllogistic reasoning questionnaire, described above. Reliability analysis on all 8 items (see Appendix 3E) initially revealed a disappointing Cronbach's Alpha of 0.38. By removing 6 items stepwise it was possible to obtain a maximal Alpha = 0.82, but this was for just 2 questionnaire items. Perhaps surprisingly, all of the items aimed at capturing the AD strategy (such as items 2 and 6 which deal with putting objects onto a scale) were eliminated in early stages of this analysis. The two remaining invalid items (numbers 1 and 3) both involved the use of physical properties. SPSS output from this analysis is presented in Appendix 3G. As only two reliable items were identified, no further analysis was conducted on the questionnaire. The nature of the two strategies identified for three-term series reasoning will now be discussed in more detail.

Abstract Directional Reasoners (N = 39)

39 participants (62.9% of sample) presented written protocols which clearly showed the characteristics of AD reasoning as described above. Protocols 3.12 and 3.13 are typical.

Protocol 3.12

Participant 18			
Problem 10	E is darker than V V is not as light as T What is the relationship between T and E?		
Protocol	E dark	V light	T lighter

Correct conclusion given: T is lighter than E

Protocol 3.13

Participant 34			
Problem 4	G is happier than M G is sadder than V What is the relationship between V and M?		
Protocol:	V	G	M
	←———————————	———————————	——→
	happy		sad

Correct conclusion given: V is happier than M

Others chose to represent terms in a vertical array, for instance:

Protocol 3.14

Participant 30	
Problem 7	H is not as rough as Z O is not as smooth as Z

What is the relationship between O and H?

Protocol

O roughest
Z rough
H smoothest

Correct conclusion given: O is rougher than H.

However, people who chose to do this tended to do so on all problems, rather than just on those which theories suggest induce a vertical representation. In addition to the above, five participants presented protocols which, although clearly abstract and propositional (and hence warranted inclusion within the abstract strategy group), did not include an explicit scale. Protocol 3.15 is typical .

Protocol 3.15

Participant 26
Problem 15

N is not as rough as D
J is rougher than D
What is the relationship between J and N?

Protocol:

N not as rough
J rougher

Correct conclusion given: J is rougher than N

For AD reasoners, questionnaire responses typically indicated strong positive responses to items 2 and 6, and negative responses to items 1 and 3. Comments at item 9 typically describe putting the letters on a scale, for instance:

Participant 12:

"Picked out the stronger relationship (i.e. rougher). Wrote two letters down in order with a symbol (>) to show which was greater than other. Fitted the third letter into sequence".

Participant 46:

"... I placed the letters in order from big to small, from thick to thin, from rough to smooth."

Concrete Properties Reasoners (N = 22)

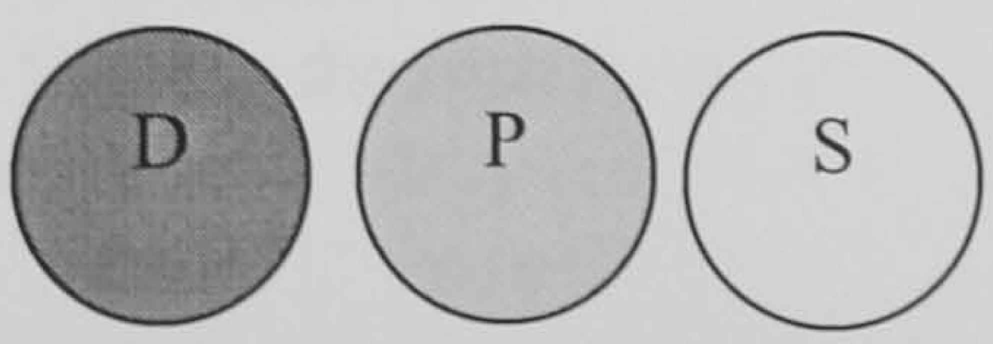
22 participants (35.5%) showed evidence of using the physical properties of the terms. Their protocols clearly showed representations of these properties but, unlike the subjects in the Egan and Grimes-Farrow study, they also tended to put the letters on a scale or continuum. Only 2 of the people who used a CP strategy seemed to consider the properties only, without any directional element to their reasoning. Protocols 3.16 and 3.17 present examples typical of the CP reasoners:

Protocol 3.16

Participant 36
Problem 1

S is lighter than P
P is lighter than D
What is the relationship between S and D?

Protocol:



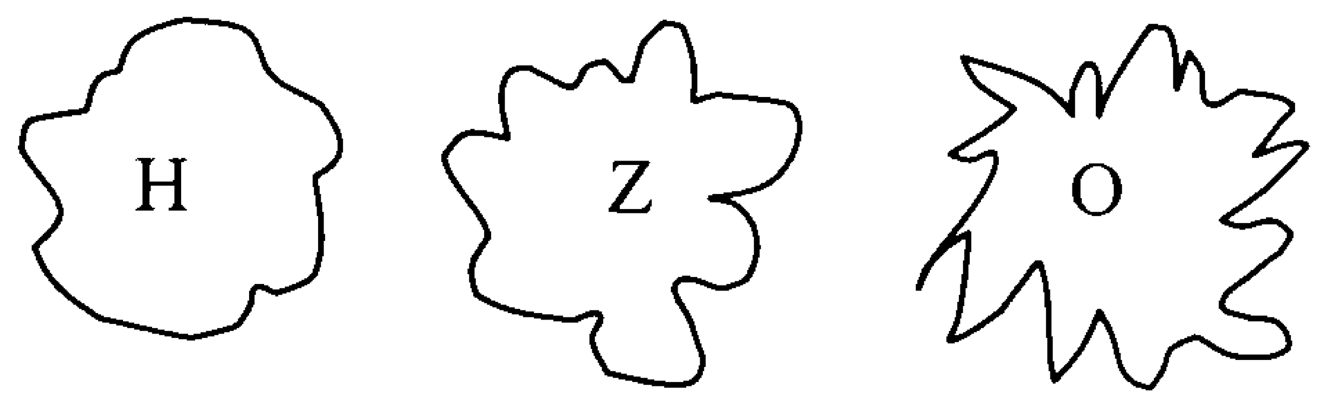
Correct conclusion given: S is lighter than D

Protocol 3.17

Participant 37
Problem 7

H is not as smooth as Z
O is not as smooth as Z
What is the relationship between O and H?

Protocol:



Correct conclusion given: H is not as rough as O

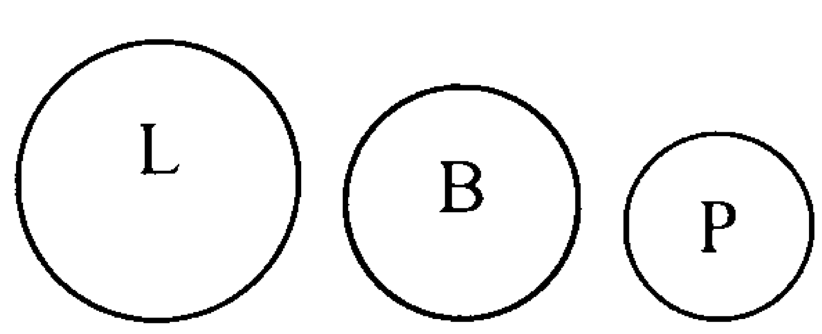
These examples are fairly typical of the representations presented by CP reasoners. However, some protocols were not as detailed as these. For instance, the participant in Protocol 3.18 gave questionnaire responses which indicated that physical properties were central to his reasoning. However, rather than attempting to represent all the properties literally, many of his protocols were of the form below, with relationships shown by relative size, simply because he found this easiest to draw. In this example, a larger object represents strength, and a smaller one relative weakness. Other CP reasoners used height in a similar way.

Protocol 3.18

Participant 49
Problem 16

B is not as weak as P
B is weaker than L
What is the relationship between L and P?

Protocol:



Correct conclusion given: L is not as weak as P

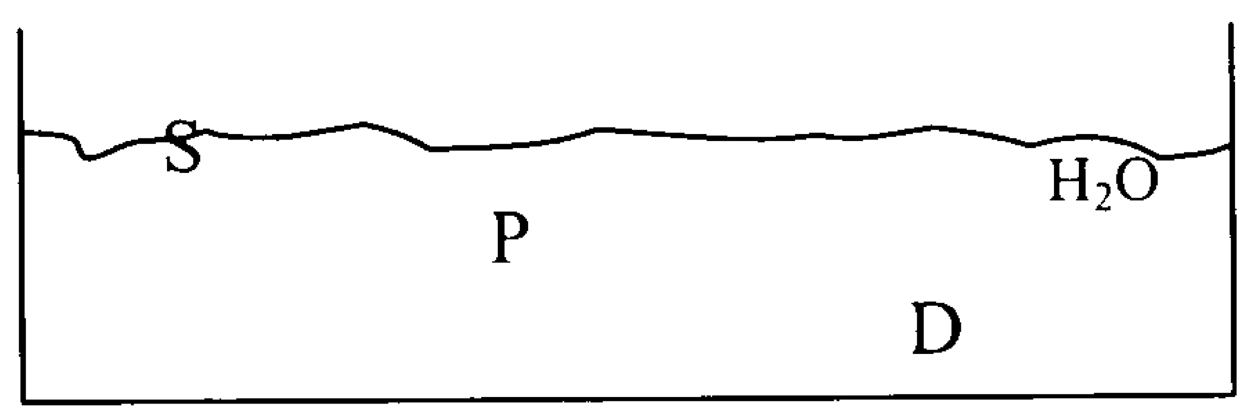
Conversely, some participants embellished their protocols enormously, drawing arms with muscles for strength, floating balloons for lightness and faces with changing expressions for happiness and sadness. Participant 52 produced some lovely examples, one of which is shown in Protocol 3.19.

Protocol 3.19

Participant 52
Problem 1

S is lighter than D
P is lighter than D
What is the relationship between S and D?

Protocol



Correct conclusion given: S is lighter than D

CP reasoners typically answered affirmatively not only to the items referring to use of physical properties (particularly item 1) but also to those which referred to use of a scale (especially item 2) and which had been intended to capture the AD approach. These responses are fully in line with their protocols as the above examples show.

Performance on Transitive Inference

Latencies were not measured in this experiment as the aim was to identify strategies rather than measure performance. As would be expected, accuracy rate was high with CP (94.6% correct, sd = 2.5) and AD (94.2% correct, sd = 1.5) reasoners performing almost identically overall.

3.2.4.3 Comparing Strategies across the Two Tasks

Table 3.4 below shows the relationship between strategies adopted for syllogistic reasoning and that adopted for three-term series problems. For this analysis, total N = 55 as all participants who used either a mixed or indeterminate strategy on either/both tasks have been removed.

Table 3.4: Crosstabulation showing the association between strategies adopted for syllogistic reasoning and transitive inference.

Syllogistic Strategy	Transitive inference strategy		
	Abstract directional	Concrete properties	Total
Verbal	30	8	38
Spatial	5	12	17
Total	35	20	55

Table 3.4 clearly suggests an association between the strategies that people adopt across the two tasks. The majority of people who adopt a verbal strategy for categorical syllogisms use the AD strategy for three-term series problems. The spatial reasoners on syllogisms however, tend to adopt the CP approach to three-term problems described above, using concrete properties but placing the terms on a scale. These findings are fully in line with predictions. Pearson Chi square suggests that this association between strategies on the two tasks is highly significant, $\chi^2 = 12.45$, df = 1, $p < 0.001$. Strategy choice cannot be attributed to presentation order effect. Presentation order of the two tasks was counterbalanced and there seems to be no association between this and strategy used for either task; $\chi^2 = 1.87$, df = 1, $p > 0.05$ for syllogisms; $\chi^2 = 1.03$, df = 1, $p > 0.05$ for transitive inference

3.2.5 Discussion

The aims of Experiment 3 were twofold: Firstly, and especially in the light of the somewhat fragmented strategies identified in Experiment 2, to further replicate the verbal and spatial strategies identified by Ford (1995) and in Experiment 1. The second aim was to investigate whether such strategies were associated with those identified on a transitive inference task involving three-term series problems. Such an association would support the idea that strategies are an inherent individual difference. Egan and Grimes-Farrow (1982) identified two strategies for transitive inference, a *concrete properties* (CP) strategy (where reasoners use detailed representations of the actual physical properties of the terms as described by the relational adjectives) and an *abstract directional* (AD) strategy (where reasoners simply place terms on a directional scale according to their relative properties). Although Egan and Grimes-Farrow claim a spatial element for both of these approaches, they do seem to differ in the level of semantic representation required, as do the spatial and verbal strategies identified for syllogistic reasoning in Experiments 1 and 2. Hence, for Experiment 3, it was predicted that people who adopted a verbal strategy for syllogisms would use an AD strategy for transitive inference, whilst spatial syllogistic reasoners would adopt a CP strategy. The above results strongly suggest that both aims were fully realised.

On the syllogistic task, both verbal and spatial reasoners were easily identified from protocols and, as in Experiment 1, the classification was supported by questionnaire responses. Moreover, qualitative data provided further insight into the strategy behind the "flowchart" representations identified in Experiment 2. The vast majority of these are clearly verbal strategies (as Ford's work also suggested) but they do not rely purely on substitution (either naive or sophisticated) - although this is an element for some individuals. In many cases, reasoning involves the application of a simple rule for obtaining a conclusion based on linking of the two quantifiers. Such rules present similarities to the principles highlighted in the matching theory (Wetherick and Gilhooly, 1990; 1995), and make similar predictions, the linking-rule also resulting in a conclusion containing a quantifier corresponding to that from one of the premises. Hence, like substitution, its application leads to errors for different-form syllogisms.

It could be argued that there is also a spatial element to the use of flowchart diagrams and a minority of subjects who presented such data appeared to use a mixed reasoning strategy. For them, the strategy appeared to involve two stages, first spatial *then* verbal, the inference process therefore being propositional, even though the premise information is initially represented spatially. A similar description of the relationship between spatial representations and rule-based reasoning has been suggested by Rips (1986), who purports that reasoning with mental models also depends on what he terms "non-standard" rules (i.e. not formal logic) in the manipulation of models and evaluation of conclusions.

On the transitive inference task, protocols clearly suggested the presence of both the concrete properties (CP) and abstract directional (AD) strategies, however the two were not as distinct as Egan and Grimes-Farrow suggested. 39 participants presented protocols which identified them as using the AD strategy. Just as Egan and Grimes-Farrow described, they placed the terms on a scale according to their physical properties and drew their conclusion from their relative positions on that scale. 20 subjects showed evidence of using a CP strategy. Their protocols clearly showed representations of the actual physical properties of terms, for instance drawing shapes which were taller or shorter, with darker or lighter shading, with rougher or smoother edges etc, but, unlike the subjects in the Egan and Grimes-Farrow study, they *also* tended to put the letters on a continuum or in order of their linear relationship. Only 2 of the people who used the concrete properties approach included no directional element in their protocol.

In terms of the theories of transitive inference outlined earlier, the vast majority of subjects used some form of dimensional spatial array, either with (CP reasoners) or without (AD reasoners) explicit representation of physical properties of the terms. Only a small minority of subjects overall (N=5, 8% of sample) used a purely propositional strategy (such as that advocated by Clark, 1969), a similar percentage to that observed by Roberts (2000) in studies of relational inference. These individuals presented evidence of reformulating premises to simplify difficult relationships (e.g. rewriting *X is not as tall as Y* as *Y is taller than X*) and also tended to present protocols which effectively eliminated the pivot term, as linguistic theories suggest occurs with problems which contain two matching relational adjectives. However, in the present data, these few individuals tended to adopt this approach for all problems, see for instance Protocol 3.15 above. If the continua used by AD reasoners is classed as a form of spatial array (as it is by Egan and Grimes-Farrow) than it would seem that spatial strategies dominate in the present study. However, AD reasoners may also be incorporating a linguistic element in their strategy. The mixed model advocated by Sternberg (1980) suggested that premises are first encoded linguistically and then represented spatially. This may be what is occurring here with the AD reasoners. Certainly these make up the majority of subjects (63% of sample) and Sternberg and Weil (1980) have suggested Sternberg's mixed model to be the preferred approach for the majority of untrained reasoners.

In the light of this, responses to the questionnaire completed after the transitive inference task are interesting. The questionnaire was designed to identify strategy use based on the criteria suggested by Egan and Grimes-Farrow – i.e. discriminate between two distinct strategies, one based on properties, one on an abstract scale. However, the only reliable questionnaire items were those two pertaining specifically to the use of physical properties (items 1 and 3). Those items referring to use of a scale/continuum were found to be some of the least reliable. Initially, this was a surprising result given the predominance of the AD strategy. However, the protocols offer some explanation. The explicit representation of physical properties seems to be the only feature which discriminates between the two strategies. Both use a directional aspect, and as such, both could be said to possess a spatial element. The directional arrays are similar to those proposed by spatial theorists such as Huttenlocher, however they do not vary with adjective, or with markedness, participants preferring to use either a vertical or horizontal array throughout. However, where the two strategy groups *differ* is in the degree to which the physical properties of the terms are represented. Abstract directional thinkers can simply use letters to denote the terms, but concrete properties thinkers seem to require a more explicit visual comparison – rather as do spatial reasoners for syllogisms. Johnson-Laird and Byrne (1991) have claimed that the use of mental models relies on semantic knowledge of premises, for CP reasoners, this would seem to be an important factor, whilst for AD reasoners, beyond an initial recognition of the transitivity of relationships, semantics are immaterial, just as they are for syllogistic reasoners using substitution and allied verbal strategies.

In line with the above, a highly significant association was observed between strategies adopted across the two tasks, verbal syllogistic reasoners tending to opt for the AD strategy for transitive inference and spatial syllogistic reasoners for the CP strategy. This is fully in line with *a priori* predictions and clearly suggests the presence of an inherent reasoning preference which differs across individual. The verbal strategy for syllogisms and abstract directional strategy for transitive inference are alike in that semantic meaning is not the major factor in determining a conclusion. In the former, the quantifiers are largely ignored during substitution and the relevant term simply replaced, whilst with the linking rule method, quantifiers are used, but only in so far as identifying the most appropriate heuristic rule to apply. Their actual logical meaning is irrelevant. In the AD strategy, reasoners need to understand the transitive nature of the relationships depicted in order to place the terms appropriately upon a scale. Over and above this basic function however, the actual nature of terms is immaterial and they are represented in abstract form. However, the actual nature is and the terms simply rearranged and represented in an abstract form. Consequently, in both tasks, identification of the conclusion is straightforward and the reasoning process is identical for all problems.

Both transitive inference strategies are similarly effective in terms of accuracy, however, informal observation of subjects performing the AD strategy suggests that it is highly efficient in terms of processing speed (although actual latencies were not directly measured). Although CP reasoners also place terms in a linear order, they need to visualise the material more clearly in order to understand the semantics of the problem, hence they supplement their representations with visual images which make clear the physical properties as suggested by the relational adjectives in the problems. This must involve additional processing stages, and hence it can be predicted that their response times would be greater than those of AD reasoners.

Overall, the findings of Experiment 3 further support the existence of individual strategies for reasoning and suggest that the tendency to employ either an abstract/syntactic or a concrete/semantic strategy is likely to be a stable and robust feature of individual reasoning preference. Experiment 4 will extend these findings to a further task, sentence-picture verification, and also allow for the comparison of problem solution latencies between strategy groups.

3.3 Experiment 4: Do Strategic Differences also Extend to Sentence-Picture Verification?

3.3.1 Introduction

Experiment 3 has shown that strategic preference is not limited to syllogistic reasoning tasks and that strategies which appear to be propositional and spatial in nature are also observed in transitive inference. Moreover, the strategies on the two tasks are strongly related, with almost all individuals consistently preferring one or the other. However, whilst there is great variation in difficulty in syllogistic problems and inter-strategic performance can be measured in terms of accuracy across problem type (e.g. as function of conclusion form or model count), participants almost always provide correct responses to transitive inference problems. In this case, differences are more easily measured by latencies, the time taken to solve the problems reflecting relative level of difficulty. This aspect was not measured in Experiment 2 which aimed primarily to establish the existence of individual differences in strategy. Having fulfilled that aim, Experiment 4 aimed to extend this line of enquiry in two ways, firstly by examining the performance of the two transitive inference strategies in terms of latencies, and also by comparing those with strategies and latencies for a further reasoning task, sentence-picture verification.

Deciding whether a linguistic statement accurately reflects an observation of the world is a key task in everyday language comprehension. Representations of both media have to be formed and compared, and the sentence-picture verification (SPV) task has been a primary research tool in investigating how this process occurs. In the SPV task (first devised by Clark and Chase, 1972) participants are presented with a simple sentence such as:

A is above B

and then a picture, for instance either

A	or	B
B		A

The task is to indicate whether the sentence is a true representation of the picture. Latencies can be measured from the onset of the picture, the chief independent variable being the linguistic complexity of the sentence. Within this paradigm, four trial types can be produced by combining affirmative or negative sentences (e.g. A is above B, or, A is not above B) with pictures for which the sentences are either true or false representations. Carpenter and Just (1975) proposed a three stage model of sentence-picture verification:

1. Sentence representation: Mental representations are generated in the form of logical propositions analogical to the sentences.

2. Picture representation: mental representations of the pictures are generated which also take the form of logical propositions analogous to the affirmative statement which describes the picture.

3. Comparison: The two representations are compared, one component at a time. When a mismatch is identified, those components are tagged as "resolved" and the comparison process begun again for another component. When all constituents are either found to be resolved or to match another, a response is generated. The nature of the responses depends on whether an odd or even number of attempts at completing a comparison was required.

This model predicts that more difficult trial types result in longer response times, and this increases monotonically in the order TA > FA > FN > TN. The increase in latencies occurs as function of the number of constituent comparisons which need to be made. Table 3.5 illustrates the SPV sentence-pairs as a function of trial type, hypothetical representation and number of constituent comparisons suggested by Carpenter and Just's model. This illustration is based on one presented by MacLeod, Hunt and Mathews (1978) and assumes that each constituent comparison process takes a constant amount of time (*K*).

Table 3.5: The sentence-picture stimulus pairs by trial type, hypothetical representation and number of constituent comparisons (c.f. MacLeod et al, 1978, Table 1, Page 494).

Trial type	Sentence	Picture	Sentence representation	Picture represent--ation	No. of constituent comparisons
True Affirmative (TA)	A is above B B is below A	A B	[AFF (A, TOP)]	(A, TOP)	K
False Affirmative (FA)	B is above A A is below B	A B	[AFF (B, TOP)]	(A, TOP)	$K + 1$
True Negative (TN)	A is not above B B is not below A	A B	{NEG [AFF (B, TOP]}	(A, TOP)	$K + 5$
False Negative (FN)	B is not above A A is not below B	A B	{NEG [AFF (A, TOP)]}	(A, TOP)	$K + 4$

In effect, the constituent comparison model places each trial type at a unique point on an interval scale, and predicts that observed response times will be a linear function of this scale. As Table 3.5 suggests, data typically present an interaction in response times between truth and polarity of trials (TA > FA; FN > TN). Carpenter and Just (1975, Tables 4 and 5, pages 59-60) reviewed 15 earlier studies of SPV and show that their linear model effectively accounts for over 95% of the variance in response time across the trial types. Moreover, although Carpenter and Just refer to the internal representations as "propositions" they suggest that these may not necessarily always be linguistic in nature, but may take the form of abstract tokens. The SPV task clearly presents both a linguistic and a visuo-spatial element and so, given the nature of the two main reasoning strategies identified in Experiments 1 to 3, would seem a useful topic of study where both spatial and verbal strategies might be observed, either as general strategic preferences, or, at different stages of the SPV process.

MacLeod, Hunt and Mathews (1978) tested the Carpenter and Just model against SPV latency data provided by 70 participants. Two response time measures were taken, firstly for sentence comprehension (CRT) (from sentence onset to initial key press) and verification (VRT) (from onset of picture to press of true or false key). They identified two groups of participants based on how well their data fitted the model. The "well-fit group" they suggested used a linguistic strategy compatible with Carpenter and Just's model. This strategy consisted of representing the sentence linguistically, without recoding any negation as affirmation, followed by a linguistic recoding of the picture, after which the two linguistic representations were compared. These reasoners presented patterns of verification RTs in line with those suggested in Table 3.5 in terms of the truth-polarity interaction. The "poorly-fit group" did not present this pattern of latencies and were assumed to use a strategy involving an iconic recoding of the sentence into a pictorial form before the onset of the picture. This was then compared to the image. As such, the linguistic group took longest over the verification stage, whilst the pictorial group required a longer comprehension time. The effect of markedness was significant in the linguistic group, increasing both comprehension and verification times. No such effect was observed for the pictorial group. Table 3.6 presents the pattern of response times (RTs) for both groups across the two phases of the task.

Table 3.6: Response times for the comprehension and verification stages of SPV (in msecs) across two strategy groups as recorded by MacLeod et al (1978)

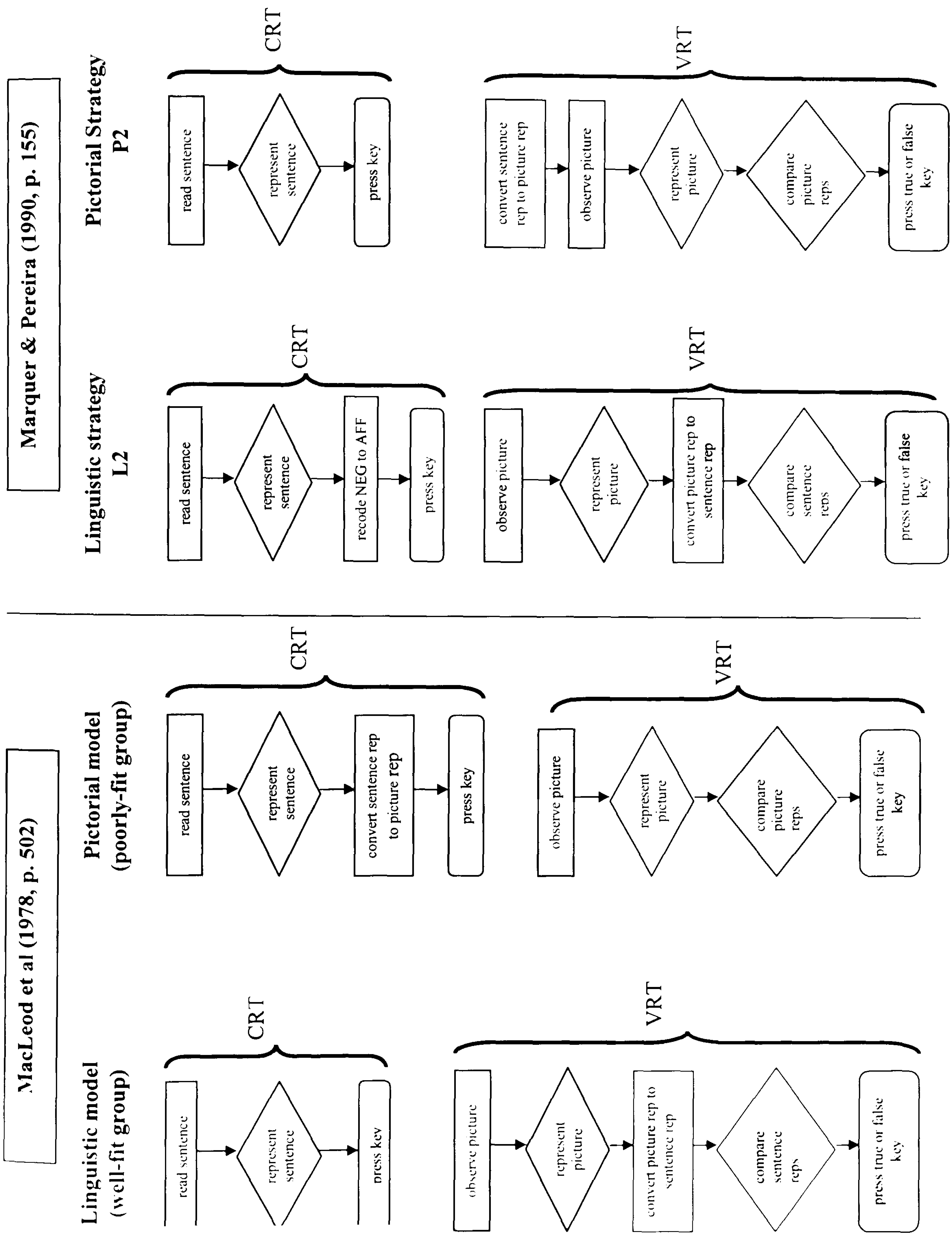
	CRT	VRT
Well-fit group (linguistic strategy)	1652	1210
Poorly-fit group (pictorial strategy)	2579	651

MacLeod et al also examined verbal and spatial ability scores from the relevant components of the Washington Pre-college Test, a test similar to SAT taken by Washington state high school students considering higher education. Subjects also completed a reading comprehension test. In general terms, none of these measures seemed to predict CRT very well, whilst *both* verbal and spatial ability significantly predicted VRT, and to a similar magnitude (r = between 0.46 and 0.6 across the four trial types). This aspect will be discussed further in Chapter 5.

However, MacLeod et al also observed considerable inter-subject variability and claim that a single model of the type proposed by Carpenter and Just (1975) was inadequate for capturing individual differences in RT, along with the underlying strategic differences. In replicating this work, Marquer and Pereira (1985) and Richards and French (1987) emphasise that MacLeod et al imposed strategies on their data, and that these may differ from spontaneous strategies. Their replications showed that whilst verification response times supported the strategy groups identified by MacLeod et al in terms of the patterns of RTs across trial type (one group presenting the truth x polarity interaction, whilst the other did not), verbal protocols presented a different picture. The well-fitting group reported a diverse range of strategies, including the pictorial one thought to be used by the poorly fitting group. Conversely, in the poorly-fitting group, only one person reported such a strategy, most having used a linguistic approach. However, what these reasoners had done, was recode negations as affirmations *prior* to picture presentation, resulting in a RT profile more typical of the pictorial strategy, hence their inclusion in the poorly fitting group. Marquer (1990) also showed that individual differences may be masked by RTs and suggested that RT data be supplemented with that from other sources, such as verbal reports.

Marquer and Pereira (1990) also commented on the lack of homogeneity in MacLeod et al's two strategy groups, especially the "poorly–fitting" pictorial group. Marquer and Pereira compared latencies across four strategies, two linguistic and two pictorial. Those strategies they termed L1 and P1 reflect the linguistic and pictorial strategies respectively described by MacLeod et al. Participants were instructed to use those strategies and RT data was expected to present the well and poorly fitting profiles described by MacLeod et al. In linguistic strategy 2 (L2) participants were instructed to reason as in L1 except that they were told to recode negative sentences into affirmatives during the comprehension stage. Hence only the true false decision should affect VRT, as with P1. Similarly, for pictorial strategy 2 (P2) they were instructed to proceed as P1 except that the sentence must be revised iconically only *after* pressing a computer key to record the end of comprehension time. As such, both affirmative-negative and true-false factors should influence VRT as with L1. In addition, subjects provided both retrospective and think aloud protocols. Similar RT profiles were observed for strategies L1 and P2, and for strategies L2 and P1. Hence Marquer and Pereira (1990) succeeded in showing that similar patterns of RT latencies can in fact reflect the use of either a pictorial or linguistic strategy, whereas differing RT patterns may result from the use of a single strategy. Figure 3.1 presents a comparison of their models (L2 and P2) and those of MacLeod et al (L1 and P1).

Fig. 3.1: SPV strategy models proposed by MacLeod et al (1978) and Marquer & Pereira (1990)



The work of MacLeod et al (1978) and Marquer and Pereira (1990), together with the findings from Experiment 3, informed the rationale for Experiment 4. Experiments 1-3 of the present thesis have shown two distinct strategies for syllogistic reasoning and for transitive inference. For the latter however, accuracy ceiling effects mean that no differences in performance between strategies could be identified. However, informal observation suggests that the CP strategy is more time consuming than AD and hence overall latencies may discriminate between the strategies, offering further converging evidence for the distinction. Also, if the strategies are processing semantic information differentially, latencies across the different three-term series trial types may further differentiate between them. As the introduction to Experiment 3 described, AD reasoners show an effect of end-anchoring whilst CP reasoners find problems containing inverse adjectives most problematic. Linguistic and spatial strategies have also been identified in SPV, although Marquer and Pereira's work casts some doubt on whether latencies alone can effectively discriminate between strategy types. However, a convergence between RT data on this task and strategy grouping from syllogistic reasoning would suggest that different strategies are indeed present. The purpose of Experiment 4 was to investigate these issues.

3.3.2 Aims

Experiment 4 had two main aims:

1. To replicate and extend Experiment 3 in two ways:
   a) firstly to test the robustness of the findings regarding strategies for transitive inference and their relationship with those identified for syllogistic reasoning,

b) secondly, to investigate whether the inter-strategic differences observed for transitive inference are also reflected in differences in latencies across trial types in that task.

2. To compare the findings from 1 above with latencies for SPV and specifically with the models and patterns of latencies suggested by MacLeod et al (1978) and Marquer and Pereira (1990). In terms of the strategies identified for syllogistic reasoning, it was predicted that verbal reasoners would present a pattern which suggests the use of a linguistic SPV strategy, whilst spatial reasoners would present latencies suggestive of a pictorial SPV strategy.

3.3.3 Methods

Participants

70 undergraduates from University of Plymouth volunteered to take part in return for course credit. This sample comprised 21 males and 49 females, mean age 25.20 years. None had participated in Experiment 3.

Materials and Procedures

Participants were run individually in sessions lasting around 1 hour. Each experimental session comprised two phases:

Phase 1: Pen and paper tasks to elicit evidence of strategies

Each participant was issued with a booklet containing a shortened version of Experiment 3 with 5 each of the syllogistic and three-term series problems. Again participants were asked to write down their working out and complete a questionnaire following each part of the task. As in Experiment 3, task presentation order was counterbalanced. Format of problems, instructions and questions were exactly as in Experiment 3 and as presented in Appendices 3A to 3C. The questionnaire presented after the syllogistic reasoning task was identical to that which proved reliable in Experiment 3 ($\alpha = 0.80$) and which was presented in Appendix 3D. For the three-term series problems, the questionnaire administered in Experiment 3 yielded only 2 reliable items, numbers 1 and 3 ($\alpha = 0.82$). Thus, for Experiment 4 a revised questionnaire was presented. This comprised the 6 most reliable items from Experiment 3, plus a qualitative item. It is presented in Appendix 3H.

Phase 2: Computer based tasks to measure response times

Secondly, each participant completed the following two tasks which were presented on an RM Accelerator computer with Intel Pentium IV processor and a 17 inch screen. Again, task presentation order was counterbalanced.

Task 1: Sentence-Picture Verification: Procedures were based on those used by MacLeod et al (1978). 64 sentence-picture verification items were presented, which comprised 4 sets of 16 basic item types as shown in Appendix 3I. Items were presented in a different random order for each participant. Premise terms were single letters (as in the example presented earlier in section 3.3.1 and Table 3.6) and these letters varied with repeated presentation of trial. Written instructions as shown in Appendix 3J were first presented on the computer screen, followed by 16 practice items. Each trial was preceded by a red warning dot in the centre of the screen which appeared for 500 msecs. This was followed by the stimulus sentence presented horizontally across the centre of the screen, together with a button marked "ready". This stimulus remained on the screen until the participant clicked the ready button. The picture then appeared after 500 msecs, together with two buttons marked "true" and "false". Participants were instructed to click on either button as appropriate to indicate their response. The computer recorded the comprehension response time (CRT) from the point of sentence onset to when the ready button was clicked. The verification response time (VRT) was measured from the onset of the picture to when either the true or false button was clicked. After an interstimulus interval of 1000 msecs another warning dot then appeared to herald presentation of the next trial.

Task 2: Transitive Inference: This task comprised 64 three-term series problems, four of each of the 16 problem types (see Appendix 3B). As in Experiment 3 and the pen and paper task above, terms were represented by letters. Instructions presented on the computer screen prior to the experiment replicated those from Experiment 3, as shown in Appendix 3C. Following these, 16 practice trials were presented. Each trial was preceded by the red warning dot for 500 msecs followed by the premise pair and a "ready" button. The premises remained on screen until the button was clicked. After 500 msecs, a conclusion appeared beneath the premises together with "true" and "false" buttons. The task required the participants to click the appropriate button to indicate whether they thought the conclusion was true or false. Again two sets of response times were collected, CRT was measured from onset of the premises to when the ready button was clicked. VRT was measured from onset of conclusion to when either the true or false response button was selected. After 500 msecs, another red warning dot appeared. This remained on screen for 1000 msecs, followed by the next trial.

As such, although the conclusion appears alongside the premises, this form of presentation cannot be considered truly parallel (i.e. when premises and question are presented together from the outset) nor serial (when the premises disappear after the comprehension stage). Clark (1969) concluded that verbal process tended to be employed with parallel presentation of material whilst Huttenlocher (1968) concluded that with serial presentation, spatial strategies were employed because verbal strategies were too onerous on working memory unless all material was presented simultaneously. Potts and Scholz (1975) and Ormrod (1979) have presented evidence which further supports these. Roberts (2000) has described how parallel presentation of material is associated with the use of task-specific heuristic strategies and how serial presentation may be employed in an attempt to overcome this. However, in both cases, the predominant strategy for unpractised reasoners tends to be spatial, with possibly only around 10% of individual attempting a verbal approach (ibid.). However, in the present study, the method of presentation was chosen in order to as nearly as possible replicate the conditions of Experiment 3 and of the paper version of this task in the present study. These are those tasks which have presented evidence for the strategies which Experiment 4 predicts will be reflected in latencies measured during the computer version of the task. Uniformity of this part of the procedure was intended to ensure compatibility of strategies adopted across the two forms of the task, whilst hopefully not unduly encouraging either a verbal or spatial strategy.

3.3.4 Results

3.3.4.1 Strategies for Syllogistic Reasoning

The two short pen and paper tasks were designed to elicit evidence of strategies for syllogistic and three-term series reasoning. The findings of Experiment 3 were easily replicated even with only 5 items per task. For syllogistic reasoning, 38 verbal, 27 spatial and 5 mixed reasoners were identified. Written protocols for verbal and spatial reasoners were of the type identified previously, as the examples in Protocols 3.20 and 3. 21 respectively illustrate.

Protocol 3.20

Participant 65
Syllogism 4

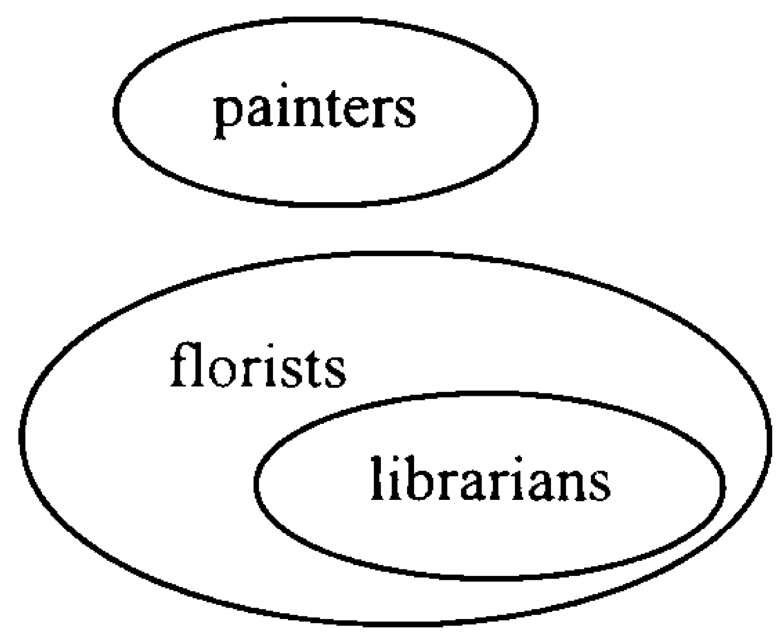
Some of the bankers are golfers
All of the doctors are golfers

Correct conclusion given: Some of the bankers are golfers

Protocol 3.21

Participant 17
Syllogism 3 — All of the librarians are florists
None of the painters are florists

Protocol:



Correct conclusion given: None of the painters are florists

Reasoning Behaviours Questionnaire (Syllogisms)

Experiment 3 established a reliable 8 item questionnaire for discrimination between syllogistic reasoning strategies (α = 0.82). This same questionnaire (see Appendix 3D) was administered in Experiment 4 and the responses analysed by the same methods as previously. Again this proved a reliable 8 item questionnaire with a Cronbach's Alpha = 0.7. This analysis is presented in Appendix 3K. Reverse scoring the items referring to the spatial strategy (numbers, 1, 6 and 7) resulted in a reasoning behaviour scale on which a high score indicted a high level of spatial reasoning behaviour and a low score a high level of verbal reasoning behaviour. As previously, the spatial strategy group ($M = 27.52$) scored significantly higher than the verbal group ($M= 18.58$); $t\ (63) = 8.02$, $p < 0.001$. With mixed reasoners removed from analysis, K-Means cluster analysis revealed two clusters of participants and their relationship with the verbal and spatial strategies is shown in Table 3.7.

Table 3.7 K-means cluster analysis of syllogistic reasoning questionnaire responses: cluster membership by strategy. Figures represent within strategy group numbers and percentages.

Strategy	Cluster 1 (N=37)	Cluster 2 (N=28)	Totals
Verbal	10 (26.3%)	28 (73.7%)	38 (100%)
Spatial	27 (100%)	0	27 (100%)

A highly significant association is observed between strategy grouping and cluster membership; $\chi^2 = 34.95$, df = 1, $p < 0.001$. Once again, as in Experiments 1 and 3, it would appear that questionnaire responses strongly support strategy categorisation by protocol.

3.3.4.2 Strategies for Transitive Inference

For transitive inference, protocols produced while solving the five three-term series problems indicated 41 abstract directional reasoners, 24 concrete properties reasoners and 2 reasoners who presented a mixed approach (using AD on some problems and CP on others). Three participants produced indeterminate protocols as all they had done was to rewrite the premises. Protocols 3.22. and 3.23 show typical examples of the AD and CP strategies respectively:

Protocol 3.22

Participant 13
3-TS Problem 1 — W is fatter than A
Q is thinner than A

Protocol: W A Q
fat ←——→ thin

Correct relationship given: Q is thinner than W

Protocol 3.23

Participant 54
3-TS Problem 4 — D is taller than X
L is not as tall as X

Protocol: D X L

Correct relationship given: D is taller than L

Reasoning Behaviour Questionnaire (Transitive inference)

For the three-term series questionnaire (Appendix 3E) a similar analysis was carried out. In Experiment 3, the questionnaire for this task proved less reliable than that for syllogisms (initial $\alpha = 0.38$, final $\alpha = 0.82$ for just 2 items). The revised questionnaire used in Experiment 4 proved even less reliable (Cronbach's $\alpha = 0.21$ initially with all 8 items included, rising to just $\alpha = 0.39$ for two items 4 and 5). The SPSS output for this analysis is presented in Appendix 3L. Interestingly, neither of these items relate directly to either AD or CP reasoning, rather they are intended to capture alternative strategies, allied to the verbal strategies in reasoning and which include behaviours suggested by verbal inference theories such as that of Clark (1969). However, protocols clearly indicate the presence of AD and CP strategies (as in Experiment 3) and very few people showed any indication in their protocols of either reversing terms within the premises or combining them into one sentence, as the two reliable questionnaire items would suggest. As so few items again proved reliable, as in Experiment 3, questionnaire analysis was concluded at this point. Overall, it would seem that although written protocols present clear evidence for CP and AD strategies, this questionnaire is not a particularly reliable predictor of these.

3.3.4.3 Comparison of Strategies across the Two Tasks

When the mixed and indeterminate reasoners from both/either task were removed from the sample, strategies on the two tasks were compared as in Experiment 3, with the same outcome. As in Experiment 3, verbal syllogistic reasoners were most likely to adopt an abstract directional strategy for solving three-term series problems, whilst spatial syllogistic reasoners were most likely to adopt a concrete properties strategy for three-term series problems. This association was highly significant: $\chi^2 = 20.95$, df = 1, $p < 0.001$. The relationship is illustrated in Table 3.8.

Table 3.8: Crosstabulation showing the association between strategies adopted for syllogistic reasoning and transitive inference

Syllogistic Strategy	Transitive inference strategy		
	Abstract directional	Concrete properties	Total
Verbal	30	4	34
Spatial	8	18	26
Total	38	22	60

3.3.4.4 Response Times for Transitive Inference

These latencies were analysed in terms of the two strategies identified for this task as above. Table 3.9 shows the mean RTs across the two response stages (comprehension and verification) for the two groups:

Table 3.9: Mean response times in msecs for the two strategy groups identified for transitive inference

Strategy	CRT (msecs)		VRT (msecs)	
	Mean	**S. Dev.**	**Mean**	**S. Dev.**
Abstract directional	9384.84	3465.11	4008.17	3057.84
Concrete properties	8726.36	3660.40	4416.38	3315.24

As Table 3.9 shows, although CP reasoners were quickest of the two groups on CRT and AD quickest on VRT, overall both groups took significantly longer for the comprehension stage of the task. A 2 (strategy) x 2(response time) mixed ANOVA indicated a significant main effect of RT; $F(1, 63) = 38.01$, $p < 0.001$, however no significant main effect of strategy or interaction between strategy and response stage was apparent ($p > 0.1$ in both cases).

Egan and Grimes-Farrow (1982) have shown that AD reasoners perform least well on problems which are not end anchored (i.e. where the pivot term is presented first in one or both premises). CP thinkers have most difficulty with problems where the two premises contain inverse adjectives. Latencies presented by the two strategy groups were therefore compared in terms of these two factors. Table 3.10 presents the mean response times.

Table 3.10: Mean response times (in msecs) for problems with and without end-anchoring and with and without inverse adjectives, by strategy group.

Strategy		End anchoring				Inverse adjectives			
		Yes (pivot is first term in at least 1 premise)		No (pivot is second term in both premises)		Yes (premises have inverse adjectives)		No (adjective same in both premises)	
		CRT	VRT	CRT	VRT	CRT	VRT	CRT	VRT
Abstract directional (N = 41)	Mean	9175.09	4042.22	9812.61	3864.96	9198.87	3944.26	9099.74	4227.84
	S.Dev.	3362.58	2988.04	3872.01	3090.54	3629.11	2938.41	8852.02	2931.21
Concrete properties (N = 24)	Mean	8549.93	4481.57	9252.15	4255.80	8405.43	4419.92	8852.02	4614.63
	S.Dev.	3601.23	3128.22	4132.67	2905.11	3459.05	3219.52	3862.75	3454.02

As Table 3.10 indicates, the RT trends were fairly similar across strategy groups. Firstly to consider end-anchoring, a 2 (RT) x 2 (end-anchoring status, yes or no) X 2 (strategy) mixed Analysis of Variance presented a significant main effect of RT; $F(1,63) = 41.16$, $MSE = 1535910852.7$, $p < 0.001$, and interaction between RT and anchoring status; $F(1, 63) = 11.51$, $MSE = 11494750.44$, $p = 0.001$. The main effect of end-anchoring approximated to significance ($p = 0.06$). No main effect of strategy, or significant interactions involving strategy were observed ($p > 0.8$ in every case). When possible effects of inverse adjectives are examined, a 2 (RT) x 2 (inverse status, yes or no) x 2 (strategy) Analysis of Variance showed a significant main effect of RT only; $F(1, 63) = 35.86$, $MSE = 1274282932.2$, $p < 0.001$. No other main effects or significant interactions were observed ($p > 0.1$ in every case).

3.3.4.5 Response Times for Sentence-Picture Verification

Comprehension response time

Previous research, reviewed earlier, has shown that during comprehension time (CRT), affirmative sentences are processed more quickly than negative sentences. The present data indicate this to be the case for participants in Experiment 4 (N = 70); affirmative (M = 1611.12 msecs, sd = 717.08), negative (M = 2365.89 msecs, sd = 1054.24), t (69) = 9.02, p < 0.001.

Verification response time

For verification response time (VRT), an interaction between truth and polarity of items is expected. A 2 (truth) x 2 (polarity) Analysis of Variance indicated a significant main effects of truth; F (1,69) = 74.80, MSE = 2736910.37, , p < 0.001, and of polarity; F (1, 69) = 32.41, MSE = 6096809.48, p < 0.001. A significant interaction between the two was also observed and is further illustrated in the chart in Figure 3.2; F (1, 69) = 6.86, MSE = 342439.41, p = 0.01. The relevant mean response times are presented in Table 3.11.

Figure 3.2: SPV: Verification response times (in msecs) as a function of truth and polarity (N = 70).

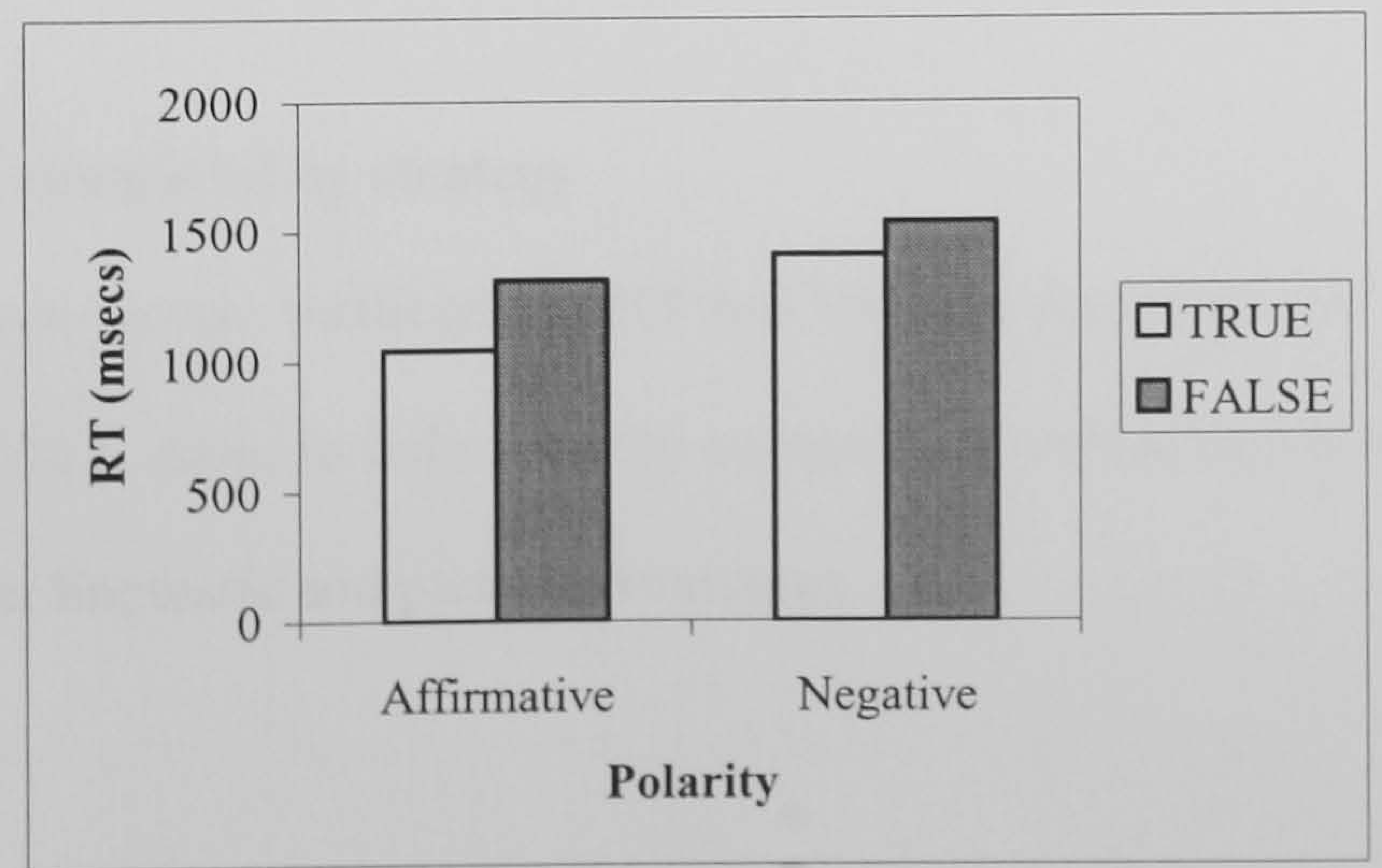


Table 3.11: SPV Verification response times (in msecs), by trial type as a function of truth and polarity.

True affirmative		True negative		False affirmative		False negative	
Mean	**S. Dev.**	**Mean**	**S. Dev.**	**Mean**	**S. Dev.**	**Mean**	**S. Dev.**
1043.69	258.87	1408.76	652.59	1311.37	339.25	1536.55	613.90

Overall therefore, it would appear that data in Experiment 4 replicates the usual patterns observed for sentence-picture verification latencies.

3.3.4.6 Strategies for Sentence-Picture Verification

Aim 2 predicted that participants using differing strategies for SPV would present distinct patterns of latencies on that task. Throughout Experiments 1 to 3, consistent individual differences in syllogistic reasoning strategies have been identified, and these are thought to be verbal and spatial in nature. Section 3.3.4.1 above has shown that these strategies have again been observed in Experiment 4. In Experiment 3, and during the analysis in Section 3.3.4.2 above, these strategies have been successfully mapped onto data provided for a transitive inference task. Hence this procedure will also be followed for SPV. The syllogistic strategy groupings identified in Section 3.3.4.2 were used to define potential strategy groups for SPV, but in terms of latencies. All mixed strategy syllogistic reasoners (N = 5) have been removed from sample prior to the following analyses.

Response times compared by strategy

The mean sentence-picture verification CRT and VRT for the verbal and spatial groups are shown in Table 3.12. Also, in italic, are the comparative values presented by MacLeod et al (1978) for their linguistic and pictorial strategies.

Table 3.12: Mean comprehension and verification response times (in msecs) as a function of strategy group as defined by written protocols during syllogistic reasoning, and by total sample.

Strategy	CRT		VRT	
	Mean	S. Dev.	Mean	S. Dev.
Verbal (N = 38)	1898.06	603.98	1269.91	381.23
Linguistic	*1652*	-	*1210*	-
Spatial (N = 27)	2064.15	963.15	1402.62	471.30
Pictorial	*2579*	-	*651*	-
Total (N = 65)	1967.05	771.08	1325.03	422.61
Total (MacLeod et al)	*4231*	-	*1831*	-

Comparison of data with that of MacLeod et al shows a clear discrepancy with regard to VRT for the spatial group. Whilst in MacLeod et al's study, the pictorial group present very fast latencies in comparison with the linguistic (mean difference, – 559), the spatial group identified here in Experiment 4 show an RT higher than that of the verbal group (although the mean difference in this case is only 132.71 msecs). A 2 (strategy) x 2 (RT) Analysis of Variance presented a significant main effect of RT as expected; $F(1, 63) = 36.97$, $MSE = 13127068.04$, $p < 0.001$. However, no main effect of strategy or interaction between strategy and RT were observed ($p > 0.1$ in both cases).

Verification response times compared by strategy

Another potential way to differentiate between the strategy groups is by examining the truth-polarity interaction which is typical of verification response times on which the Carpenter and Just model was based (and which was observed in the overall sample in the present study). If the verbal and spatial strategies do map onto the linguistic and pictorial ones identified by MacLeod et al, it would be expected that the verbal strategy data will fit the Carpenter and Just model well (i.e. the interaction will be observed) whilst the spatial group will fit the model less well (the interaction will not be observed). To test this prediction, a 2 (strategy) x 2(truth) x 2 (polarity) mixed Analysis of Variance was conducted. The expected significant main effects of both truth; $F(1, 63) = 60.28$, $MSE = 2357492.32$, $p < 0.001$, and polarity; $F(1, 63) = 30.64$, $MSE = 23989.34$, $p < 0.001$, were both observed, together with a significant interaction between the two; $F(1, 63) = 7.74$, $MSE = 371480.46$, $p < 0.01$. However, no main effect of strategy or 2-way interaction involving strategy was apparent ($p > 0.2$ in every case). Most important for the present purpose however was the absence of a significant 3-way interaction between strategy, truth and polarity ($p > 0.1$). Hence, it would seem that patterns of latencies on SPV cannot discriminate between the verbal and spatial strategies, at least, not according to the MacLeod et al criteria. The means for each trial type are shown in Table 3.13.

Table 3.13: SPV: VRT (in msecs) as a function of strategy group, truth and polarity.

Strategy	TA	TN	FA	FN
Verbal				
Mean	1021.55	1312.57	1267.94	1491.63
S. Dev.	262.81	576.94	292.78	594.61
Spatial				
Mean	1083.33	1564.12	1376.81	1618.12
S. Dev.	273.98	762.37	396.08	678.41

3.3.5 Discussion

Experiment 4 had two aims. The first of these was to replicate and extend the findings of Experiment 3. To recap, Experiment 3 replicated earlier findings which suggested that a verbal and a spatial strategy for syllogistic reasoning could be identified from written protocols. Moreover, Experiment 3 found that these strategies respectively could be mapped directly onto abstract directional and concrete properties strategies for transitive inference. These latter strategies had previously been identified by Egan and Grimes-Farrow (1982). In phase 1 of Experiment 4, participants completed a pen and paper task which comprised a mini version of Experiment 3. Again, written protocols clearly indicated verbal and spatial strategies for syllogistic reasoning, and abstract directional and concrete properties strategies for transitive inference. The nature of these strategies was just as observed in Experiment 3. However, and again in line with Experiment 3, the transitive inference strategies were not as distinct as previous research by Egan and Grimes-Farrow (1982) had suggested, with both strategy groups tending to incorporate a directional element in their protocols. Responses to reasoning behaviour questionnaires strongly supported strategy group membership for syllogistic reasoning, though not so for transitive inference. This was also the case in Experiment 3 and may again be attributable to the lack of distinction between the two transitive inference strategies, the only differentiation between them being the degree to which physical properties were used. Furthermore, when strategies across the two tasks were examined, a highly significant association was observed between use of the verbal syllogistic strategy and abstract transitive inference strategy, and between the spatial syllogistic strategy and the concrete properties transitive inference strategy. Altogether these are important findings. They indicate that the effects are robust and replicable across new samples. Moreover, they add substantial weight to the argument that forms a key tenet of this thesis, namely that preference for a visuo-spatial or abstract-propositional strategy is a stable and inherent cognitive factor which differs between individuals.

In phase two of the investigation, latencies for both transitive inference and SPV problems were measured. In both cases, two response time measures were recorded, a comprehension response time (CRT, the time from onset of premises to participant indicating they were ready for next stage) and a verification response time (VRT, from onset of conclusion/picture to when a true or false response was given). For transitive inference, and in line with previous research by Egan and Grimes-Farrow, it was predicted that abstract directional reasoners would present longer response times for those problems which were not end-anchored (i.e. those which presented the pivot term first in at least one premise). Concrete properties reasoners, on the other hand, would show no such effect. Instead, they would present longer times on problems which contained inverse adjectives. No such effects were observed, both strategies presenting similar patterns of latencies across both sets of problems. Overall, any differences observed were consistent across both strategy groups. Both groups presented significantly longer latencies at the comprehension stage, compared with verification. Informal discussion with participants suggested that almost all were using a similar approach: the question was presented simultaneously to the premises (as in the pen and paper task). Participants tended to work out the relationship in question from the premises during the comprehension stage, and then simply compared this with the conclusion they were later presented with. If it matched, they clicked true, if not, they clicked false. Hence, the majority of subjects in Experiment 4 were apparently using a heuristic approach which may have masked any potential discriminatory trends in the data. Overall, the complete lack of an effect of strategy on transitive inference latencies suggests that these data alone are not an accurate predictor of strategy for this task.

Marquer and Pereira (1990) have made a similar observation about latencies for sentence-picture verification (SPV), and phase two of Experiment 4 also measured latencies for this task. It was predicted that the strategies identified for syllogistic reasoning would map onto those for SPV identified by MacLeod et al (1978). Specifically, verbal reasoners would present the latency profiles for SPV suggested by the linguistic SPV strategy whilst the spatial reasoning strategy would map onto the latency profile typical of the pictorial SPV strategy. Previous research reviewed earlier has shown that typically, during the comprehension stage of SPV, affirmative sentences are processed more quickly than negative sentences. This was also the case in Experiment 4. Also, on verification latencies, an interaction between truth and polarity is typically observed and again this was the case in the present data. Having established that the overall data matched the expected patterns of latencies, response times were examined in terms of the strategy groups identified for syllogistic reasoning. However, no differences between the two groups were observed. Both presented longer comprehension time than verification. The model proposed by Carpenter and Just (1975) and which was found to fit the latencies presented by the linguistic strategy group identified by MacLeod et al (178) describes a latency profile which presents the typical interaction between truth and polarity described previously. It was therefore expected that if the verbal syllogistic group were tending to use a linguisitc strategy for SPV, the latencies produced by that group would present the interaction. No such effect was predicted for the spatial group. However, this interaction was in fact observed in the data from *both* verbal and spatial reasoners.

Overall, an association between strategies identified on syllogisms and those for SPV had been predicted, and this was clearly not the case. The former strategy groupings are now established and replicable phenomena. However, they do not seem to map onto SPV latencies, and SPV latencies in turn cannot seem to predict strategies for that task. There is a possibility that participants may not have been performing the task according to instructions. However, the findings from phase 1 are well in keeping with those from Experiments 1 and 3, and suggest that participants *were* attempting to perform the tasks they had been set. There is no reason to doubt whether this motivation continued into phase 2. There are two further possible explanations for the SPV findings. Firstly, individual differences in strategies for SPV may simply not exist. This is unlikely, given that such differences are so clear in the other reasoning tasks. However, SPV is in some ways a very different kind of task to either syllogisms or transitive inference. SPV bears resemblance to other reasoning tasks in that it involves the comparison of representations in order to conclude whether or not they match. As such, it has been regarded as a simple reasoning task, however, response times are extremely fast compared to those for say syllogisms or transitive inference. Consider, for instance, the overall latencies shown in Tables 3.9 (transitive inference) and 3.11 (SPV). Whereas for the former, VRT varies between about 4000-4400 msecs, for SPV, RTs are less than half this magnitude (around 1900-1500 msecs for VRT). This raises the question regarding the degree to which a conscious strategy may be employed during such a rapid process, and reservations about how much actual reasoning takes place during solution of an SPV item. A similar point has been made by Evans (1982) who suggested that reasoners on SPV do not need to understand any logical principles (for example, that the truth of the premises necessitates the truth of the conclusion) they merely have to decide whether or not a description is accurate. Hence he concludes that such tasks "do not actually require deductive reasoning" (page 24).

A further possibility is that distinct strategies for SPV are present, but that latency information alone is insufficiently sensitive to discriminate between them. Marquer and Pereira suggested that such data be supplemented by verbal protocols in order for strategies to be identified. They claimed that different strategies may present similar latency profiles and this certainly seemed to be the case in the present study. Experiment 4 collected no verbal reports which, in retrospect, might have proved useful. However, these in turn may have raised other concerns. Given the nature of the SPV task, the only protocols which can be collected are retrospective verbal accounts and, although Marquer and Pereira offer some persuasive arguments in their favour, such accounts are likely to consist mainly of post-hoc rationalisation, rather than an accurate record of reasoning processes. This may be further exacerbated with SPV given that responses are so rapid. Participants must be consciously aware of their strategies if they are to articulate them (be it concurrently or retrospectively). There is some doubt whether this could ever be the case with SPV.

Given that no effects of strategy on RT were found, it would seem that for the vast majority of individuals, RTs remain fairly consistent throughout the task. Those who find the task most difficult will produce longer RTs at both comprehension and verification stages, whereas those who find the task easier, will produce faster RTs. Whatever strategy they may be using, reasoners tend to produce the typical truth-polarity interaction as has been observed in previous research and which, it was hoped given MacLeod et al's claims, would discriminate between verbal (linguistic) and spatial (pictorial) reasoners in the present study. A tendency towards either verbal or spatial reasoning may however be present in SPV and, with the above reservations in mind, future studies may wish to collect verbal accounts in addition to latencies in an attempt to identify such differences. However, the above data suggest that even then, strategies may not map cleanly onto those for other tasks. It may be that the combined verbal and pictorial demands of SPV mean that, for many reasoners, a mixed strategy involving both approaches is the most efficient. Given that individuals are likely to experience some difficulty in articulating their SPV reasoning processes in detail, identifying any strategic differences among them will probably require converging evidence form a range of measures.

3.4 Summary of Chapter 3

Chapter 3 has presented two further experiments which have supported the existence of individual differences in reasoning strategies. Following up Experiments 1 and 2, which presented evidence for verbal and spatial strategies in syllogistic reasoning, Experiment 3 was able to replicate and extend these finding to a further task, transitive inference. For this task, although both of the strategies identified possessed a directional element, a strong association was observed between these and the strategies for syllogistic reasoning. The verbal syllogistic reasoning strategy mapped onto the abstract strategy for transitive inference, whilst the spatial syllogistic strategy mapped onto the concrete strategy. The main discriminating feature between the two transitive inference strategies was the degree to which the properties of the terms were necessarily represented in order for a conclusion to be reached, concrete properties reasoners requiring the most explicit representation. Comparisons can be drawn between this and the differing semantic requirements of the two respective syllogistic strategies, where spatial reasoners present explicit representations of relationships between terms, whereas most verbal reasoners simply manipulate the linguistic form of the premises. Experiment 4 introduced a third task, sentence-picture verification with the aim that differing patterns of latencies for this task would reveal corresponding strategic differences. The findings from Experiment 3 were replicated but unfortunately when SPV latency profiles across the groups of reasoners identified during syllogistic reasoning were compared, no differences were observed. Overall, however, it remains that evidence for a verbal-spatial distinction in strategies has now been replicated four times for syllogistic reasoning (what Evans, 1982, would call a *bona fide* deductive reasoning task) and these strategies map strongly onto those identified for transitive inference. Moreover, the findings for SPV do not necessarily refute the possibility of individual differences in strategies for that task. Latency data alone may be simply not sensitive enough to capture such differences. However, the question as to what determines verbal or spatial strategic preferences still remains, and Experiments 5 and 6 investigate this issue. A number of researchers have associated strategies with individual differences in verbal and spatial cognitive ability (including MacLeod et al's, 1978, work on SPV discussed in Experiment 4), or with working memory (both mental models and rules theories suggest that working memory capacity is a key factor in reasoning performance). Experiment 5 will investigate whether the reasoning strategies identified thus far are associated with individual differences in verbal and spatial abilities, and/or in other aspects of cognitive style. Experiment 6 will then extend this line of research to consider whether individual differences in verbal and spatial working memory capacity influences whether a verbal or spatial reasoning strategy is adopted.

CHAPTER 4

Reasoning Strategies, Abilities and Cognitive Style

4.1 Introduction to Experiment 5

Chapters 2 and 3 have presented a series of studies which show that individual differences in strategy use for reasoning do exist and that they are robust across different reasoning domains. This strongly suggests that the strategies are somehow inherent to the nature of the individual, rather than artefacts of experimental reasoning tasks. One interpretation of these findings is that the two main strategies involve the representation and manipulation of information in *either* a verbal-propositional, *or* a spatial, manner. However, a major question which remains is why individuals differ in the sorts of strategies they adopt. Experiment 5 aimed to address this question by means of an investigation into a possible association between reasoning strategy preference and two aspects of cognitive function where individual differences have consistently been observed; spatial-verbal ability and cognitive style.

In discussing the concept of individual differences research generally, Chapter 1 presented a definition of cognitive abilities offered by Carroll (1983). Carroll described abilities as traits which determine how an individual processes mental information when performing a cognitive task such as reasoning. Such abilities necessarily involve not only performance itself, but the potential for performance, across a range of tasks. Such abilities are suggested to be stable characteristics, possessed to a greater or lesser extent by all individuals. Section 1.5.3 presented a general discussion of research on verbal and spatial abilities, hence this review will limit itself to some issues which are of particular relevance to Experiment 5. As Chapter 1 described, researchers have frequently relied on measures such as Scholastic Aptitude Test (SAT) scores in evaluating verbal and spatial ability, and it has been suggested that these can discriminate between individuals who use different strategies for reasoning. A good example is the work of MacLeod, Hunt and Mathews (1978) reviewed in Chapter 3. To recap, MacLeod et al suggested that patterns of latencies across the comprehension and verification stages of a sentence-picture verification task were indicative of two distinct strategies, a pictorial strategy and a linguistic strategy. These strategies were also found to be related to spatial and verbal ability respectively. Abilities were assessed from verbal and spatial components of the Washington pre-college test (a test similar to SAT taken by Washington state high school students considering further education) together with an additional verbal measure of reading comprehension. The correlations between verification response times and ability scores were highly significant for both strategy groups (linguistic group, $r = -.47$ and $-.52$ with the two verbal ability measures; pictorial group $r = -.68$ with the spatial measure). The negative values indicate that higher ability individuals produced faster latencies. No significant relationships were observed between linguistic group latencies and spatial ability, or between pictorial group latencies and verbal ability.

Other measures specifically designed to assess verbal and spatial ability have also been associated with reasoning strategy. Cooper and Mumaw (1985) presented a series of studies which suggested that individual differences in strategies for spatial problem solving were related to differences in spatial ability. They assessed ability by means of a mental rotation test developed by Shepard and Metzler (1971) and also a visual comparison task which involved matching a three-dimensional figure to a corresponding (or not) orthographic representation. Two strategies were first identified from verbal protocols; a constructive strategy (thought to involve the construction of a three-dimensional internal representation) and an analytic strategy (involving the sequential comparison of individual features). It was predicted that high spatial ability individuals would adopt the former strategy and low ability individuals the latter. However, when ability test scores were examined, the opposite trend was observed, high spatial ability was associated with a tendency to use the analytic strategy, whilst low spatial ability individuals tended to opt for the constructive method. However, comparison of latencies indicated that for the minority of high ability individuals who did adopt the constructive strategy (thought to be most compatible with their ability) reasoning was both faster and more accurate. Low spatial ability individuals using an analytic strategy however, performed similarly, suggesting that this approach was more suited to their level of aptitude. Kyllonen, Woltz and Lohman (1981) also showed that even when aptitude is not reflected in initial strategy choice, it can influence accuracy and efficiency in appropriate tasks (e.g. high spatial ability individuals performing a spatial task). On the reasoning tasks examined in Experiments 1 to 4, it might be predicted that individuals with high spatial ability will prefer a spatial strategy, whilst those with high verbal ability will prefer a verbal approach. However, it should be noted that in both the above cases, all strategies are assumed to be spatial, no mention is made of the possibility of a propositional strategy being employed for the spatial task.

A further related factor may be individual differences in cognitive or thinking style. Research in this area has been extensive, as is the number of dichotomies in the literature (see for instance Roberts and Newton, 2001, for a discussion), hence, this short review will limit itself to some of the major work involving human reasoning. Sternberg (e.g. 1997) describes how individuals may possess preferred ways of thinking and Sternberg and Grigorenko (1997) state that these preferences are but one aspect of what has been termed "cognitive styles", characteristic, stable and typically preferred modes of processing information. Such styles have also been variously termed "thinking dispositions" (e.g. Stanovich and West, 1998; Perkins, Farady and Bushey, 1991), "intellectual styles" (e.g. Sternberg, 1988) and "information processing styles" (e.g. Klaczynski, Gordon and Fauth , 1997), these terms often being used interchangeably. For the present purposes, and to avoid confusion, the term cognitive style will be used throughout, assuming the useful definition of "an individual's preferred and habitual approach to organising and representing information" (Riding and Rayner, 1998, page 8). According to Sternberg (1997), these styles are not in themselves cognitive abilities, but reflect how abilities are used. Hence, individuals may have similar abilities, but their style of thinking may differ considerably.

A large amount of research into individual differences in thinking dispositions, cognitive ability and reasoning has been carried out by Keith Stanovich and colleagues (e.g. Stanovich, West and Sá, 1999; Stanovich and West, 1998). These studies examined a range of reasoning tasks, including syllogistic reasoning and the Wason Selection Task, and also took into consideration cognitive abilities as determined by SAT scores. They found evidence indicating there was a systematic variance in performance on reasoning tasks which was not accounted for by the computational limitations imposed by cognitive ability. They proposed that abilities and styles are constructs which exist at different levels of analysis within cognitive theory and differ in the degree of malleability that they afford. Abilities may be relatively stable psychological mechanisms, whereas styles may be influenced by training, prior knowledge or intentional factors such as personal goals or motivation (Stanovich et al, 1999). Others (e.g. Baron, 1988) have made a similar point, highlighting also that cognitive abilities have traditionally tended to be measured using psychometric instruments such as IQ tests, which do not tap into differences in cognitive style.

Stanovich and West (1998) used a composite measure of cognitive style which assessed individuals on four subscales, epistemological absolutism, willingness to switch perspective, to decontextualise and to consider counter-arguments. Their data suggest that these factors did offer predictable and systematic explanation for some of the outstanding variance in reasoning performance after ability had been accounted for. However, examining their data reveals that the actual amount of variance explained was quite small, SAT scores accounted for 19% whilst cognitive style accounted for just 7%. The fact that these values were found to be highly significant (($p < 0.001$ in both cases) may be, at least in part, attributable to the large sample size in Stanovich and West's study (N = 546).

However, similar findings have been reported by Klaczynski, Gordon and Fauth (1997) and Klaczynski, Fauth and Swanger (1998). Verbal and inductive ability were measured by the Shipley Institute of Living Scale. This scale provides scores for both vocabulary and verbal abstract reasoning. These skills are considered to be key elements in critical reasoning and indices of fluid and crystallised intelligence respectively (Cattell, 1967). In basic terms these "intelligences" can be considered as verbal IQ and domain general reasoning ability (see for instance, Kline, 1991, for further explanation). However, the data presented by Klaczynski et al suggested that although some aspects of critical reasoning were predicted by ability, reasoning biases (and hence performance influenced by them) were more likely to be associated with cognitive style. This finding is consistent with earlier work on informal reasoning (Perkins, 1985; Perkins, Jay and Tishman, 1993) which has suggested that factors other than those associated with intelligence may be necessary to explain why some intelligent and well-educated people seem incapable of reasoning logically when logic conflicts with their beliefs, whilst others seem able to override and disregard their prior knowledge. Klaczynski et al (1997; 1998) used a measure of cognitive style developed initially by Epstein, Pacini, Denes-Raj and Heier (1996) and based around their Cognitive Experiential Self Theory (CEST). This theory suggests that individuals use two parallel but interacting information processing systems, one rational, the other experiential. Epstein et al (1996, page 391) describe these in the following terms:

"the rational system operates primarily at the conscious level and is intentional, analytic, primarily verbal and relatively affect free. The experiential system is assumed to be automatic, preconscious, holistic, associationistic, primarily non-verbal and intimately associated with affect"

Although behaviour is influenced by both systems, one may dominate according to situational, expectational or emotional factors. Importantly for the present purpose, individual differences have been observed whereby some people are more inclined towards rational approaches whilst others seem to prefer experiential. Moreover, given the apparent nature of the reasoning strategies identified thus far, the two systems are also of interest in view of the verbal – non verbal distinction made by Epstein et al. The Rational Experiential Inventory (REI) is a psychometric tool for measurement of individual differences in rational-experiential cognitive styles. In the original version (Epstein et al, 1996), the rational thinking element was based on the concept of "need for cognition" defined as the tendency to enjoy and engage in thinking (Cacioppo and Petty, 1982). The experiential element evolved from a new scale "Faith in intuition" developed by Epstein and his colleagues.

Recent work which has investigated the REI with respect to syllogistic reasoning has been conducted by Handley, Harley, Wright and Farrelly (in press). This research found that although intellectual ability (as measured by a psychometric test of intelligence) was a good predictor of logical syllogistic reasoning where belief conflicts were present, cognitive motivation (i.e. need for cognition) consistently predicted performance on non-conflict problems. Need for cognition has consistently been shown to predict willingness to search for counter-examples inconsistent with prior belief (e.g. Furlong (1993) and hence to be a predictor of syllogistic reasoning performance (e.g. Handley, Dennis, Evans and Capon, 2000; Newstead, Handley and Buck, 1999). Torrens, Thompson and Cramer (1999) present similar findings and also data to show that abstract reasoning ability (again as measured by the Shipley Institute of Living Scale) was also associated with logical reasoning ability.

Riding and Cheema (1991) have presented an alternative explanation of cognitive style which also proposes that thinking behaviours may be grouped into two principal dimensions:

a) The wholist-analytic style dimension of whether an individual tends to *organise* information into wholes or parts.
b) The verbal-imagery style dimension of whether an individual is inclined to *represent* information verbally or in mental pictures.

The Cognitive Styles Analysis (CSA, Riding, 1991) is a computer presented tool for the assessment of thinking style on these two constructs, i.e. it measures how individuals both organise *and* represent information. In contrast to measures of either spatial or verbal ability (i.e. levels of performance), the CSA claims to measure manner of performance. It presents three subtests comprised of simple tasks which claim to reflect underlying processing style and how individuals habitually organise and represent material during thinking. A ratio of subtest scores determines an individuals position on the two style dimensions which allows them to be allocated to a given cognitive style category, as illustrated in Figure 4.1.

Figure 4.1: The dimensions of cognitive style (reproduced from Riding, 1998)

Wholist- Analytic Dimension	Verbal-Imagery Dimension: Ratio < 0.98	Ratio > 0.98 and < 1.09	Ratio > 1.09
Ratio > 1.35	ANALYTIC VERBALISER	ANALYTIC BIMODAL	ANALYTIC IMAGER
Ratio >1.02 and <= 1.35	INTERMEDIATE VERBALISER	INTERMEDIATE BIMODAL	INTERMEDIATE IMAGER
Ratio < 1.02	WHOLIST VERBALISER	WHOLIST BIMODAL	WHOLIST IMAGER

Numerous studies have claimed to support the notion of style according to the two dimensions (i.e. Riding, 1997; Rayner and Riding, 1997) and neurological evidence has shown associated individual differences in recorded EEG during information processing (for instance, Riding, Glass, Butler and Pleydell-Pearce, 1997). The dimensions have been shown to differentiate between behaviours as diverse as learning performance in school (e.g. Riding and Douglas, 1993) and social harmony among flat-sharers (Riding and Wright, 1995). Riding and Rayner (1998) present a comprehensive review of research. The style dimensions appear to evaluate individuals on the basis of their tendency towards thinking in either a verbal or a imagistic way. However, to this author's knowledge, the CSA has not been investigated as a possible predictor of performance or strategy formal reasoning tasks.

Roberts and Newton (2001) however, are sceptical of the cognitive style approach to understanding strategies. They criticise advocates of cognitive style for a reluctance to integrate the diverse array of cognitive style constructs and personality variables which have been proposed to influence reasoning. They accuse cognitive styles of simply redescribing behaviour, rather than explaining it, as Entwistle (1979) has pointed out, fuzzy definitions may lead to confusion between the nature of cognitive styles and of strategies themselves. Moreover, connotations of the word style imply a flexibility under differing task demands which is frequently not observed. For instance, Newton and Roberts (2000) have highlighted that cognitive style may lead to some individuals sabotaging their performance on a compass-point direction task because of an inability to discover alternative methods.

Overall, despite these reservations, it remains that a wealth of research into cognitive styles has suggested that individual differences do exist. The research reviewed above has indicated associations between styles, cognitive abilities and reasoning performance, even if the association with reasoning strategies is rather more tenuous. Individual differences in verbal and spatial ability, on the other hand, have been shown to predict strategies on various reasoning tasks. Experiments 1 to 4 previously do suggest the presence of some cognitive factor which is stable within individuals and which facilitates strategy choice. Given the apparent verbal and spatial nature of reasoning strategies so far identified, a possible association with levels of verbal and spatial ability seems reasonable. Similarly, with regard to cognitive style, the nature of style dimensions on the CSA and REI particularly do seem to suggest some level of differentiation between analytic-propositional and imagistic-spatial thinking. Hence, Experiment 5 aimed to investigate the possible relationship between these factors and reasoning strategies. As Roberts and Newton concede, even if cognitive styles cannot themselves explain strategies, they may act as a useful starting point for the greater understanding of individual differences in strategy use.

4.2 Aims

The overall aim of Experiment 5 was to investigate the relationship between verbal and spatial reasoning strategies, verbal-spatial ability and cognitive style. More specifically:

1. Abilities: It was predicted that individual differences in verbal and spatial ability would in turn predict whether an individual adopted a verbal or spatial syllogistic reasoning strategy respectively.

2. Cognitive style: Two measures, the REI and CSA were administered. These were exploratory measures and no specific predictions were made as to how cognitive style and reasoning strategy might be associated.

4.3 Methods

Participants

Participants were 110 undergraduate students at University of Plymouth who participated in return for their choice of course credit or £5 cash. The sample comprised 24 males and 86 females with total mean age of 21.7 years. All were native English speakers and none had received formal training in logic.

Procedures

Participants were run in groups of up to 4 people seated at separate desks, each with a computer terminal. Each participant was presented with six experimental tasks in a session lasting around 1 hour. Two syllogistic reasoning tasks were first administered, one to identify strategy and one to assess reasoning performance. Two measures of ability were used, the Vandenberg Mental Rotation task (spatial ability) and The Shipley Institute of

Living Scale (verbal ability). As the experiment would be employing student participants (who are all presumably of fairly high IQ and have all reached university level of education), a measure of general intelligence or scholastic ability (such as that employed in the studies reported by Stanovich, 1999) may not provide much insight into differences in abilities. Hence, ability measures were selected which would feasibly tap into distinct cognitive capacities which may differentiate between verbal and spatial reasoners. Both measures are also established psychometric tools with good levels of internal reliability. Two cognitive style measures were also included, the Rational Experimental Inventory and the Cognitive Styles Analysis. These were selected because both have been subject to a substantial amount of research and been shown to be reliable. Moreover, both measure style in terms of subscales which suggest a possible differentiation between propositional and spatial ways of thinking. Further details of all tasks are presented below. The six tasks are described to follow in the order presented to every participant. All are pen and paper tasks except syllogistic performance and Cognitive Style Analysis, which were presented on a computer as described.

1. Reasoning Strategy

This comprised exactly the same procedure as used in Experiments 3 and 4. Each participant was presented with five syllogistic problems and asked to deduce the conclusion whilst writing down their working out. These were the same five problems used in Experiment 4 in terms of figure and mood, however the occupation names (terms) were changed for this study. The problems were presented in a booklet, one to a page, preceded by written instructions identical to those used previously and shown in Appendix 2H. No reasoning behaviour questionnaire was completed on this occasion.

2. Syllogistic Performance

Participants were asked to generate conclusions to 16 syllogistic problems presented individually on a computer screen. Conclusions were typed into a box provided and the computer recorded these responses. No written working out was allowed. The 16 problems, together with their logical conclusions, are shown in Appendix 4A in the order presented to every participant. All possible syllogistic figures and mental model counts are represented within this list, affording a range of problem difficulty. Prior to beginning the task, participants were presented with written instructions and a practice item as shown in Appendix 4B. This task was one previously employed by Capon, Handley and Dennis (2003) and shown by them to possess a high level of internal reliability (Cronbach's alpha = 0.82).

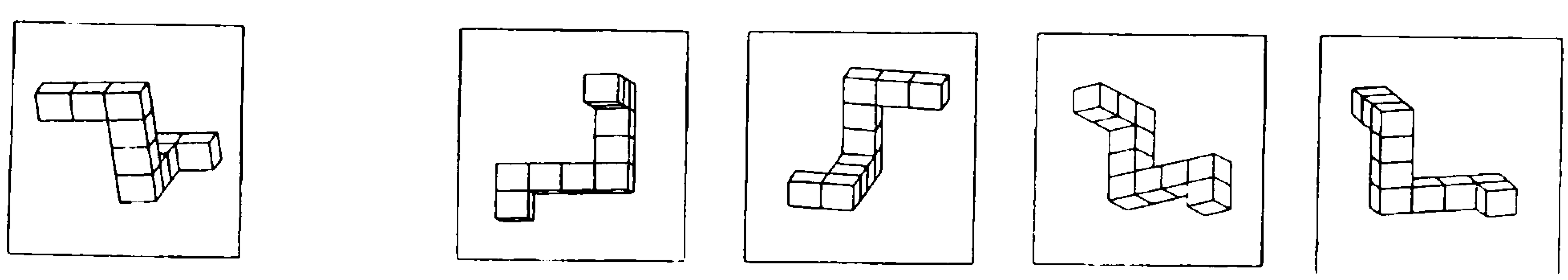
3. Rational Experiential Inventory (REI)

The REI is a self-report inventory which claims to measure cognitive style on two subscales, a rational scale (verbal, analytic thinking) and an experiential scale (non-verbal, associationistic thinking). The latest version (Pacini and Epstein, 1999) contains 40 items, 20 on each scale. Within each, there are two further subscales, corresponding to engagement (the enjoyment of using either rational or intuitive thinking) and ability (the confidence in carrying out such thinking). Pacini and Epstein report internal reliabilities of $\alpha = 0.9$ and $\alpha = 0.87$ for the rationality and experiential scales respectively. The two were shown to be independent factors and the subscales presented a robust distinction in each case. Handley, Newstead and Wright (2000) have replicated these findings in the UK with both undergraduate students and a sample from the general population, and also present test-retest reliabilities of between 0.6 and 0.88 across the four subscales. The 40 REI items, together with the 5 point response scale used, are presented in Appendix 4C. Scoring according to the scale shown results in overall scores for rational and experiential thinking and for each of the two subscales.

4. Vandenberg Mental Rotations Test (VMRT)

The VMRT (Vandenberg and Kuse, 1978) is a pen and paper test of spatial visualisation ability developed from the classic Shepard and Metzler (1971) mental rotation task, as used by Cooper and Mumaw (1985) described earlier. Vandenberg and Kuse (1978) presented significant correlations between the VMRT and 20 other established and validated tests of spatial ability (r = between 0.32 and 68). Internal reliability of the test was high (r = 0.88) and test-retest reliability has been shown to be between 0.7 and 0.83. The test has been associated with spatial working memory ability on various cognitive tasks (e.g. Hamilton, 1999). The VMRT comprises 24 items, each consisting of a criterion figure, presented with two correct alternative representations and two incorrect distracters. The figures are three-dimensional drawings of cubes stacked in various arrays In each case, the task is to identify the two alternative representations which are identical to the criterion object, but presented rotated about the vertical axis to varying degrees. For instance, in Figure 4.2 below, the correct response would be to highlight the second and fourth alternatives as these are rotated representations of the criterion figure, shown far left.

Figure 4.2: Example item from the Vandenberg Mental Rotation Task (Vandenberg and Kuse, 1978).



A 5 minute time limit was imposed for the task. Participants scored 1 point if they identified both correct alternatives for an item. Performance on the VMRT was treated as an indicator of spatial ability. It was predicted that if a spatial reasoning strategy is indeed a reflection of enhanced spatial abilities or preference for working with spatial representations, then spatial reasoners would outperform verbal on this task.

5. Shipley Institute of Living Scale (SILS)

The SILS is a self-administered pen and paper task comprising 2 subtests, a 40 item vocabulary test and a 20 item test of abstract reasoning. The SILS has also been widely used as a measure of general intellectual ability, and the well-established Wechsler Adult Intelligence Scale - Revised (WAIS-R) (Wechsler, 1989) has served as one of the basic criterion measures for validating SILS in predicting individual IQ scores. The two tasks correlate highly (r = 0.8) and WAIS-R full-scale IQ scores can be estimated from SILS data (Zachary, 2000). SILS can also be used as a measure of cognitive impairment in which case a 10-minute time limit is imposed for each subtest. However, for a student population, this time proved excessive and no time limits were imposed in the present study, all participants finishing both SILS tests in well under 10 minutes.

Although the SILS yields six major summary scores, only four were used for the present purposes; a) vocabulary score, b) abstraction score, c) combined total score and d) estimated WAIS-R IQ score. The other two scores (the conceptual and abstraction quotients) are derived from the above but are generally utilised in assessing cognitive impairment in clinical settings and hence were not appropriate for current purposes. For more information on these, see Zachary (2000).

The Vocabulary Test: 40 items are presented in order of increasing difficulty. Multi-choice format in which the respondent is asked to select which of four possible words is closest in meaning to a target word. For instance, in Figure 4.3 below, the correct response would be to circle the word "rashness":

Figure 4.3: An example item from the SILS vocabulary subtest.

TEMERITY	rashness	timidity	desire	kindness

One point was awarded for every correct response. 0.25 of a point is added for each item which was not attempted, to give a total vocabulary score. This task relies on verbal skills including reading ability, comprehension, acquired knowledge and long-term memory.

The Abstraction Test: 20 items presented in order of increasing difficulty. Each comprises a sequence of numbers, letters or words with the final element in each sequence omitted. The task is to complete the sequence. In Figure 4.4 below, the correct response would be to finish the sequence with the word "at".

Figure 4.4: Example item from the SILS abstraction subtest.

knit in	spud up	both to	stay - -

One point is awarded for each correct item, this total is then multiplied by 2 to give a final abstraction score. This task also relies on long-term memory, but also on attentional abilities, abstract concept formulation and cognitive flexibility.

The combined total score is obtained by summing the vocabulary and abstraction scores. The estimated WAIS-R IQ score is obtained from conversion table D2 provided within the SILS test manual (Zachary, 2000, pp. 87-88). In the present study, SILS performance was treated as an indicator of verbal/abstract reasoning ability, and the estimated WAIS-R score as a measure of a more general intellectual ability.

It was predicted that verbal reasoners may score more highly than spatial on SILS, especially on the abstract reasoning subtest. Vocabulary may be more concerned with educational attainment and would be assumed to be similar for all of the undergraduate population.

6. Cognitive Styles Analysis (CSA)

The CSA (Riding, 1991) is a computer presented tool for the assessment of thinking style on the wholist-analytic and verbal-imagistic dimensions, i.e. it measure of how individuals both organise *and* represent information. Whereas the VMRT and SILS claim to assess spatial and verbal abilities respectively (i.e. levels of performance), the CSA claims to measure manner of performance. It presents three subtests comprised of simple tasks which claim to reflect underlying processing style and how individuals habitually organise and represent material during thinking.

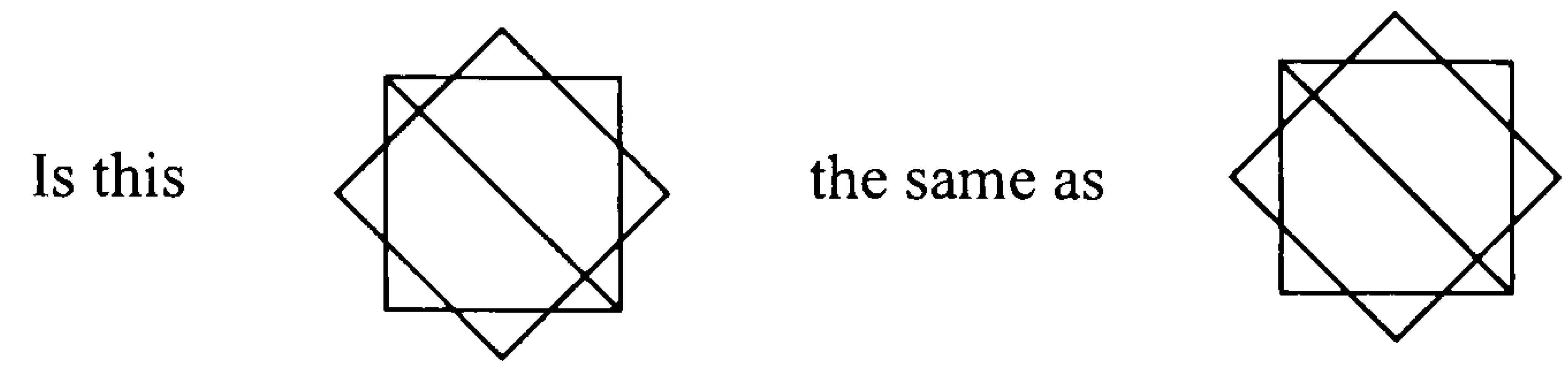
The Verbal - Imagery Subtest: 48 items each consisting of a statement about two objects, either that they are the same type (conceptual category) or colour (appearance). Half of each type of item are true. It is assumed that imagers will respond more quickly to the appearance statements because the objects can be readily represented as mental pictures and a comparison rapidly made from these images. Verbalisers are expected to present shorter response times for the conceptual category items because the semantic conceptual category membership is verbally abstract and cannot be represented visually (Riding and Rayner, 1998). The task is to indicate whether the statement is right or wrong by pressing one of two marked keys. For instance, in Figure 4.5, the correct response would be "Right" because both cricket and skiing are sports.

Figure 4.5: Example items from the CSA verbal-imagery subtest.

SKIING and CRICKET are the same TYPE

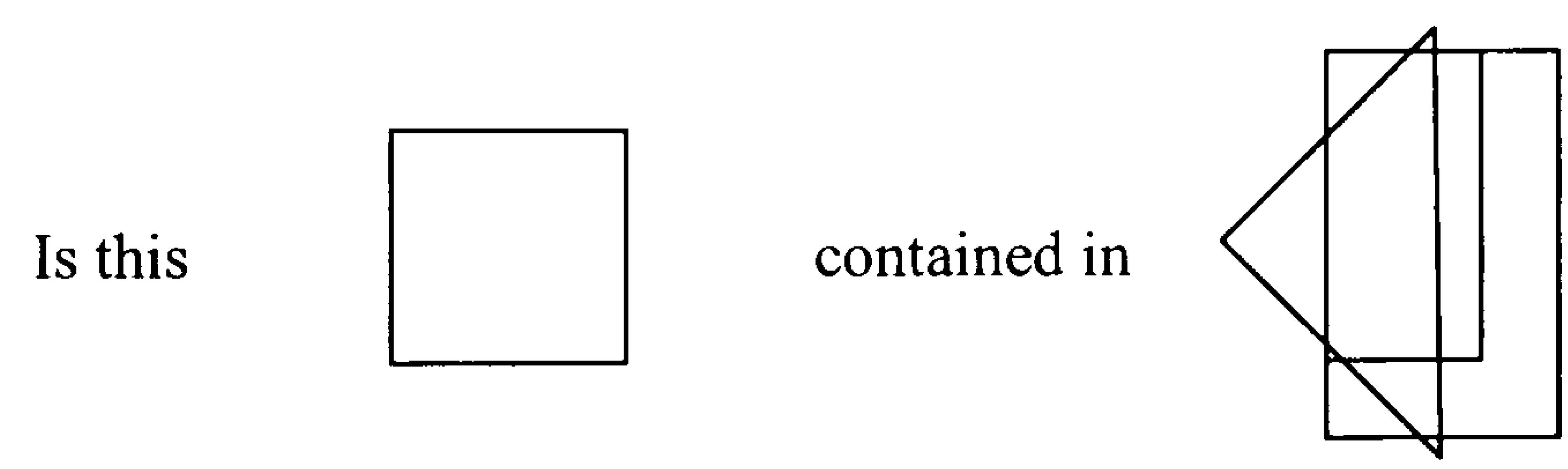
The Two Wholist - Analytic Subtests: The first wholist-analytic subtest presents 20 items each comprising two complex geometrical figures. Participants are asked to respond "Yes" or "No" by pressing marked keys to indicate whether the two shapes are the same. As this task involves judgement of similarity, it is assumed that faster responses will be attained by wholists (ibid.). In Figure 4.6, the correct response is "Yes".

Figure 4.6: Example item from the first CSA wholist-analytic subtest.



In the second wholist-analytic subtest, again 20 items of two shapes are presented, a simple geometric shape and a complex figure. This time the yes/no response is made according to whether the simple shape is contained within the other. This task requires a degree of disembedding and analytic thinkers are assumed to be quicker at this (ibid.). In Figure 4.7, the correct response is "No".

Figure 4.7: Example item from the second CSA wholist-analytic subtest.



The computer records the number of correct items and an index of the overall speed for completion of the items on each dimension. Based on this data, a ratio value is calculated for each of the two style dimensions and this is used to place each participant in one of nine cognitive style categories, as illustrated previously in Figure 4.1. Central positions on the above two dimensions are referred to as bimodal or intermediate styles respectively. For Experiment 5, it was predicted that the verbal-imagery dimension would be particularly useful in differentiating between verbal reasoners (verbal representation) and spatial (imagistic representation).

4.4 Results

4.4 .1 Strategies for Syllogistic Reasoning

Once again reasoning strategies were easily identifiable from written reports produced when generating conclusions to the five syllogisms in the strategy identification task. Table 4.1 below shows the proportion of participants who presented reasoning behaviours indicative of each of the strategy types.

Table 4.1: Syllogistic reasoning strategies identified in Experiment 5

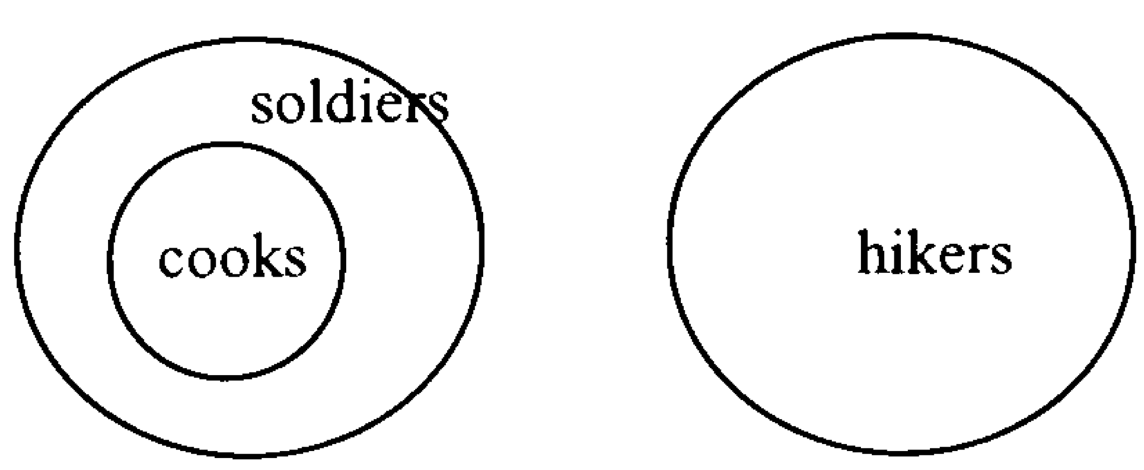
Strategy	N
Verbal	69
Spatial	34
Mixed	6
Indeterminate	1
Total	110

Protocols were highly alike to those observed in Experiments 1 to 4. Protocols 4.1 and 4.2 present examples for spatial and verbal reasoners respectively.

Protocol 4.1:

Syllogism 2 — All of the cooks are soldiers
None of the hikers are cooks

Participant 55



Correct conclusion given: None of the hikers are soldiers

Protocol 4.2

Syllogism 3 — All of the bookbinders are drivers
Some of the bookbinders are not poets

Participant 62 — BOOKBINDERS = DRIVERS
(instead of bookbinders)

SOME DRIVERS ARE NOT POETS

Correct conclusion given: Some of the drivers are not poets

4.4.2 Syllogistic Reasoning Performance

The descriptive statistics for syllogistic performance are in line with those found in Experiments 1 to 4. Again spatial reasoners (M= 8.03 correct out of possible 16, sd = 2.79) performed slightly better than verbal (M= 7.68, sd = 2.34) but the difference was far from significant (p= 0.5).

4.4.3 The Ability Measures (VMRT and SILS)

As the descriptive statistics in Table 4.2 show, the two ability measures indicated little difference in verbal or spatial ability between the verbal and spatial reasoners.

Table 4.2: Descriptive statistics showing verbal and spatial reasoners' performance on the two ability measures

Measure	Means		Std. Dev.	
	Verbal (N=69)	Spatial (N=34)	Verbal (N=69)	Spatial (N=34)
VMRT (Spatial ability)				
No. items correct (max. 24)	7.3	6.94	3.9	3.7
No. items attempted in 5 mins.	14.4	15.7	4.6	4.7
SILS (Verbal ability)				
Vocabulary (max. 40)	29.3	30.1	3.5	3.6
Abstraction (max. 40)	31.3	30.4	4.6	3.8
Total score (max. 80)	60.6	60.4	6.7	5.7
Est. WAIS-R IQ score	104.9	104.5	6.9	5.7

Although the above data do suggest a slight trend towards spatial reasoners performing best on the VMRT and verbal reasoners on SILS abstract reasoning, in no instance were these differences significant ($p > 0.05$ in every case). Estimated WAIS-R IQ scores were also virtually identical.

4.4.4 The Cognitive Style Measures (REI and CSA)

No significant differences were observed between verbal and spatial reasoners on any of the REI or CSA constructs, as Table 4.3 shows, ($p > 0.05$ in every case).

Table 4.3: Descriptive statistics showing performance of verbal and spatial reasoners on the two thinking style measures

Measure	Means		Std. Dev.	
	Verbal (N=69)	Spatial (N= 34)	Verbal (N=69)	Spatial (N=34)
REI				
Rational ability	31.4	31.7	3.2	2.4
Rational engagement	28.1	28.6	3.3	3.3
Rational total	59.4	60.3	4.8	3.5
Experiential ability	32.0	31.9	3.1	3.4
Experiential engagement	30.2	30.5	2.7	3.3
Experiential total	62.1	62.4	4.4	5.5
CSA				
Wholist-analytic ratio	1.7	1.1	0.4	0.3
Wholist-analytic speed	6.5	7.2	1.8	1.6
Wholist analytic correct	92.0	92.9	4.5	4.5
Verbal-imagery ratio	1.0	1.0	0.1	0.1
Verbal-imagery speed	4.2	4.3	1.1	1.1
Verbal-imagery correct	90.9	90.2	6.2	7.3

The CSA also categorises participants according to their position on the wholist-analytic and verbal-imagery dimensions. Table 4.4 presents a crosstabulation of these 9 possible categories with the verbal and spatial strategy groups. No significant relationships are observed between strategy group membership and CSA cognitive style classification; $\chi^2 = 8.04$, df = 8, $p > 0.05$. For definitions of the abbreviated CSA categories, cf. Figure 4.1 above.

Table 4.4: Relationship between reasoning strategy and cognitive style as measured by CSA. Figures represent numbers in category and within strategy group percentages.

Strategy	Cognitive style classification (from CSA)									Total
	A -V	I - V	W - V	A - B	I-B	W - B	A - I	I - I	W - I	
Verbal	4	7	12	9	7	5	2	8	15	69
%	5.8	10.1	17.4	13.0	10.1	7.2	2.9	11.6	21.7	100
Spatial	2	4	4	2	5	8	1	4	4	34
%	5.9	11.8	11.8	5.9	14.7	23.5	2.9	11.8	11.8	100
Total N.	6	11	16	11	12	13	3	12	19	103

4.4.5 Correlational Analyses

Although no inter-strategic differences were observed in either abilities or cognitive style, associations between the measures may yet differentiate between strategy groups and/or differentially predict syllogistic reasoning performance. Hence a series of correlational analyses was also conducted, initially using the whole sample (N = 110) and then examining verbal and spatial reasoners separately.

A matrix showing the relationships between all measures, for the entire sample, is presented in Appendix 4D. Table 4.5 present the significant correlations with syllogistic performance.

Table 4.5: Significant correlations between ability and cognitive style measures and syllogistic reasoning.

Measure	Correlation with syllogistic reasoning (r)
Verbal Ability	
Vocabulary	0.27**
Abstraction	0.24 *
WAIS IQ	0.31**
Spatial Ability	
VMRT correct	0.23 *
Cognitive Style: REI	
Experiential ability	-0.25**
Cognitive style: CSA	
Verbal-imagery speed	0.23 *

* sig. at 0.05 level
** sig. at 0.01 level

All correlations are of a similar magnitude and suggest that spatial ability, verbal ability in terms of vocabulary, and also IQ all predict syllogistic performance. Few aspects of the cognitive style measures have any impact on reasoning. Spatial and verbal abilities were correlated in terms of verbal abstraction (r = 0.31), but not vocabulary (r = 0.09). IQ necessarily predicted both aspects of verbal ability, given that the IQ estimate was derived from these scores (r = 0.72 for vocabulary and r = 0.86 for abstraction). However, it also predicted spatial ability, though not as strongly (r = 0.29). No relationships were observed between the various subscales of the cognitive style measures (refer to Appendix 4D for r values). Table 4.6 presents the inter-correlations between ability measures and syllogistic performance, by strategy group.

Table 4.6: Correlations between syllogistic reasoning performance and ability measures

Measure	Strategy	Syllogisms correct	SILS vocabulary	SILS abstraction	SILS total	WAIS-R
SILS vocabulary						
	V	0.22	-	-	-	-
	S	0.37*	-	-	-	-
SILS abstraction						
	V	0.27*	0.40*	-	-	-
	S	0.27	0.19	-	-	-
SILS total						
	V	0.30*	0.78**	0.88**	-	-
	S	0.41*	0.76**	0.79**	-	-
WAIS-R IQ score						
	V	0.30*	0.74**	0.88**	0.98**	-
	S	0.40*	0.69**	0.82**	0.98**	-
VMRT correct						
	V	0.17	0.20	0.41**	0.39**	0.43**
	S	0.35*	0.05	0.06	-0.01	0.00

* Correlation is significant at the 0.05 level (2-tailed)
** Correlation is significant at the 0.01 level (2-tailed)

As the table shows, the main predictors of syllogistic performance for verbal reasoners appear to be the SILS scores (verbal ability), and the WAIS-R IQ score devolved from this. WAIS and SILS (especially the abstract reasoning subtask) also correlated significantly with performance on the VMRT (spatial reasoning task). However, for spatial reasoners, performance in the VMRT is *also* a significant predictor of syllogistic reasoning, whereas it seems to have no relationship with SILS at all. The tendency for spatial ability to predict performance for spatial reasoners, but not verbal, is fully in line with prediction. However, when the correlations are compared across strategy group, the difference does not reach significance ($z = 0.89$, $p > 0.3$). Table 4.6 indicates that many of the correlations are of similar magnitude across strategy and none present significant differences between verbal and spatial reasoners (z between 0.03 and 1.43, $p > 0.1$ in every case).

When a similar analysis was conducted for the cognitive style measures, very few significant associations were observed. In fact, for many of the relationships observed between variables, the r value is less than 0.1. The correlations are all presented in Appendix 4E. In summary, verbal reasoners presented a similar profile as the overall sample shown in Table 4.5. Syllogistic performance was predicted by the REI experiential ability subscale ($r = -0.37$) and by speed of completion of the CSA verbal-imager trials ($r = 0.49$) only. For spatial reasoners, these two correlations were very small ($r = -0.17$ and $r = -0.07$ respectively) and no other measures predicted syllogistic reasoning accuracy either.

Overall, cognitive style, at least as measured by these particular tests, does not seem to be a factor in syllogistic reasoning performance. Moreover, almost all correlations are of similar magnitude when compared across strategy groups. No significant differences were observed between r values ($p > 0.1$ in every case).

4.5 Discussion

The aim of Experiment 5 was to investigate whether strategies employed for syllogistic reasoning could be predicted by verbal-spatial abilities and/or by cognitive style. Firstly, written protocols produced during a syllogistic reasoning task again identified verbal and spatial reasoners, as in Experiments 1 to 4. All participants also completed two ability measures; the VMRT (spatial ability) and the SILS (verbal ability). It was predicted that verbal reasoners would score more highly on the SILS (especially abstract reasoning which is less a measure of educational level than vocabulary) and spatial reasoners would do best on the VMRT. However, this was not the case and the scores attained by the two strategy groups were virtually identical. However, correlational analysis indicated some slight differences. Although the relationship between SILS performance and syllogistic reasoning was similar for both strategies, there was a tendency for spatial ability (as measured by VMRT) to predict reasoning performance for spatial reasoners, but not for verbal. This was in line with prediction, but the difference between the groups did not reach significance. Estimated WAIS-R IQ scores did significantly predict performance on the VMRT for verbal reasoners, perhaps surprisingly given that the SILS scores (from which the IQ estimate was derived) did not. Furthermore, those verbal reasoners who did well on spatial reasoning, also did so on abstract reasoning.

There are two possible explanations. Firstly, those verbal reasoners who are good abstract reasoners may have better spatial ability. Secondly, these verbal reasoners may be using an effective, but propositional, strategy for the VMRT. Roberts and Newton (2001) argue that individual differences in strategies have implications for the interpretation of the very ability measures with which styles may be experimentally associated. The psychometric properties of such measures are based on statistical analyses which assume that the same cognitive factor loadings apply to everyone, i.e. that everyone employs the same strategy. They particularly highlight spatial ability tasks which involve cube folding. These tasks tend to assume that high spatial ability individuals will use an imagistic strategy to imagine folding the cube. In fact, many use an analytical strategy whereby they make inferences from the appearance of the flattened cube presented in test items. When the results of such a test are factor analysed, the two strategic users typically present different patterns of factor loadings. Hence, even when test results correlate with performance at another task, it cannot be inferred which strategy was being used. Lohman and Kyllonen (1983) and Kyllonen, Lohman and Snow (1994) have identified numerous instances of strategic impurity on spatial tasks such as mental rotation and found that strongest performers had the ability to be flexible in their strategy choice. Rather than using complex spatial or imagistic representations, better performers simplified their strategies to reduce cognitive load. Cooper and Mumaw (1985) found that high spatial ability people were actually least likely to use a visuo-spatial representative strategy for such tasks, preferring a less demanding analytic scanning strategy.

However, such results might equally point to strategic flexibility, facilitated in some individuals and not others through contrasting cognitive styles. An association between cognitive style and reasoning strategy was also predicted but the data again indicate no such association. One explanation could be that the tests have not measured the constructs they claim to describe. However, a considerable body of research using both the CSA (see Riding and Rayner, 1998 for review) and REI (see Handley et al, 2000) have found them to be valid measures. The present study cannot really draw this conclusion without further evidence. The second possibility is that there really is no relationship between cognitive style and reasoning strategy. With regard to the CSA, given the description of styles presented by Riding and Rayner et al, it was expected that spatial reasoners' scores would place them highly on the imagery and wholistic ends of the style dimensions, whilst verbal reasoners would tend towards analytic and verbal styles. The present results show no evidence of this. The most noticeable between-strategy differences occur in the analytic-bimodal and wholist-bimodal categories where a higher percentage of verbal and spatial reasoners fall respectively. This is a slight trend towards the *a priori* prediction, but the fact that both categories are described as bimodal suggests that on the verbal-imagery scale (that where we might have expected the most differentiation), both verbal and spatial reasoners fall in the middle of the dimension, i.e. they *both* present evidence of using *both* verbal and imagery based styles. Furthermore, Table 4.3 above shows that accuracy and processing speed were almost identical for the two strategy groups, across both dimensions.

On the REI, again no differences were observed. To recap, the REI measures style according to two subscales concerned with rational and experiential thinking. No specific predictions were made as to how such measures would associate with reasoning strategy, and as Table 4.3 shows, mean scores on all four scales were remarkably similar for both strategy groups. However, when describing the REI, Epstein et al (1996) refer to the rational and experiential dimensions as conscious/intentional and automatic/preconscious respectively. Both reasoning strategies identified are clearly intentional and conscious even if their users do not always possess the flexibility to switch at will from verbal to spatial and vice versa. This is evident from the fact that the vast majority of subjects are able to articulate, at least to some extent, the tactics they have used in solving the syllogisms (cf. Experiment 1). Moreover, as discussed in the previous chapters, many verbal reasoners are not using logical reasoning – even to the extent of the substitution rules suggested by Ford (1995). Rather they use a heuristic process of naive substitution, or, if rules are applied, they tend to take the form allied to those suggested by matching theory (c.f. Experiment 2). According to CEST (Epstein, 1994) heuristic or "illogical" thinking originates from the experiential system - no differences were observed between strategies according to this subscale. Moreover, the work of Klaczynski et al (1997) suggested that the REI will predict illogical, heuristic reasoning performance (i.e. that influenced by belief bias) but not logical, rational thought. It may be that both verbal and spatial strategies represent different heuristic processes. Furthermore, neither the CSA nor the REI were found to predict reasoning accuracy. Newstead et al (in press) also found no association between

REI scores and reasoning accuracy which seems to be in line with the current findings. However, to this author's knowledge, such comparisons have not previously been made between reasoning accuracy and the cognitive style as measured by the CSA.

All of the above factors may contribute to the lack of a difference observed between verbal and spatial reasoners in terms of cognitive styles and abilities, at least as measured by these tasks. The key may lie in individual flexibility, the ability to change or modify a preferred strategy to effectively meet differing task demands or reduce cognitive workload. In other words, individuals may possess a preference for either a verbal or spatial strategy, but, some may have the capacity to change or adapt their usual approach to meet differing task demands. The composite cognitive style measure which Stanovich and West (1998) suggest may predict some of the variance in reasoning performance, included measures which claim to indicate an ability, or willingness, to engage in cognitive flexibility. Cary and Reder (2003) suggest that strategy selection, within and between individuals, may be regulated by metacognitive processes which respond to dynamic environmental features of which the reasoner may not even be aware. In terms of the REI, this seems in line with the claims of Epstein (1994) and Klaczynski et al (1997), that the experiential system is influenced by affective factors which may lead to the development of heuristic strategies in otherwise educated (and presumably rational) individuals.

Developmental work has suggested that strategy development is associated with cognitive mechanisms which allow for strategic flexibility in response to environmental changes (for instance, Crowley, Shrager and Siegler, 1997) and Crowley and Siegler (1999) have shown that children create novel goal structures which allow them to persist in the use of a given strategy, even when other demands compete, and suggest that such problem solving tactics persist into the adult cognitive repertoire. So it is possible that whilst some reasoners possess strategic flexibility, others persist with their usual strategy, or a variation of it, even when new tasks seem to be demanding something else. Furthermore, Roberts and Newton (2001) have argued that individual differences in strategies have implications for the interpretation of the very ability measures with which they may be experimentally associated. The psychometric properties of such measures are based on statistical analyses which assume that the same cognitive factor loadings apply to everyone, i.e. that everyone employs the same strategy. They particularly highlight spatial ability tasks which involve cube folding. These tasks tend to assume that high spatial ability individuals will use an imagistic strategy to imagine folding the cube. In fact, many use an analytical strategy whereby they make inferences from the appearance of the flattened cube presented in test items. When the results of such a test are factor analysed, users of the two different strategies typically present different patterns of factor loadings. Hence, even when test results do correlate with performance at another task, it cannot be inferred which strategy was being used. However, such results might equally point to strategic flexibility, facilitated in some individuals and not others through contrasting cognitive styles. The most efficient verbal reasoners are also likely to possess generally higher intellectual ability, explaining why WAIS-R IQ seems to predict VMRT performance for the verbal group. Spatial reasoners on the other hand, have a natural inclination towards a strategy which is compatible with the demands of VMRT, and so intellectual ability is not necessarily a predictor of performance. On the SILS, spatial reasoners seem to possess the skills, or perhaps flexibility, to perform comparably with verbal reasoners, in fact cognitive flexibility has been cited as a predictor of success on the abstract reasoning subtest (Zachary, 2000). However, given that this was a student sample, general educational attainment might explain the lack of variance on the SILS, especially for the vocabulary subtest, which it could be argued relies on knowledge and memory rather than reasoning ability.

In conclusion, Experiments 1 to 4 presented evidence for two distinct reasoning strategies which appear to be verbal-propositional and visuo-spatial in nature. This effect was replicated yet again in Experiment 5, and it was predicted that individual differences in verbal and spatial ability and/or cognitive style would underpin such strategic differences. However, neither style nor ability measures suggested any distinction between the two groups of reasoners. Overall, it would appear that these two factors (at least as measured by the four tasks used here) may not be base factors in determining individual differences in reasoning strategy. This leaves us with a number of outstanding questions regarding the nature of the two strategies. The distinct verbal and spatial representations which are presented in written protocols may in fact be external manifestations of the same, or similar, underlying process, as has been suggested by Stenning and Yule (1997, discussed in Chapter 1). Alternatively, strategies may be distinct but unrelated to abilities. In this case, even if individuals were resilient to changing task demands and used their preferred strategy throughout, no differences would be observed as function of strategy.

However, some individuals may be able to draw effectively on both verbal and spatial resources. This type of cognitive flexibility may mask any underlying differences in strategy preference resulting in similar results on verbal and spatial ability measures, regardless of syllogistic reasoning strategy preference. Such flexibility has been associated with individual differences in verbal and spatial working memory systems (Schunn, Lovett and Siegler, 2001). Furthermore, individual differences in working memory capacity has also been associated with verbal and spatial abilities, *and* with reasoning strategies. Handley et al (2002) presented evidence to show that distinct working memory systems for dealing with spatial and verbal representations were involved in both conditional reasoning and spatial problem solving, whilst Capon, Handley and Dennis (in press) have suggested that syllogistic reasoning performance is predicted by both spatial and verbal working memory span. Hence, maybe it is not verbal and spatial abilities *per se* that underpin strategy choice, but rather the ability to draw appropriately on verbal and spatial resources, including those within the working memory system. Experiment 6 addressed this question by means of an investigation into the relationship between verbal and spatial abilities, working memory capacity and syllogistic reasoning strategies.

CHAPTER 5

Reasoning Strategies and Working Memory

5.1 Introduction

As we have seen, individuals have a strong tendency to use either visuo-spatial or verbal/propositional representations when reasoning syllogistically. However, it remains unclear what leads these individuals to prefer one strategy over another. Part of the answer may relate to their working memory capacity – a factor long associated with reasoning performance (see for instance Gilhooly, 1998; Johnson-Laird, 2001). Reasoning and memory are inextricably linked (Gilhooly and Logie, 1998). Memory involves the encoding, retention and retrieval of information, and storage may either be temporary or form part of a long-term knowledge base. Reasoning relies on both of these storage mechanisms. Whereas information in long-term memory (for instance previous experience with the task, or prior knowledge of the problem content) may be drawn upon, reasoning also necessarily involves the use of some kind of mental workspace where information is stored temporarily during intermediate stages of the task (Gilhooly, 1998). A short-term memory store can hold details of the situation in-hand and maintain a record of the processing stages involved in working through the problem. Similarly, information about the current task can be passed back to this short-term memory store. That information is retained just long enough for it to be used for the task. Hence this workspace is referred to as *working memory* (Baddeley, 1986). Hence working memory (WM) refers to aspects of on-line cognition, the moment-to-moment processing, monitoring and maintenance of information. Baddeley and Logie (1999, page 29) define it as comprising

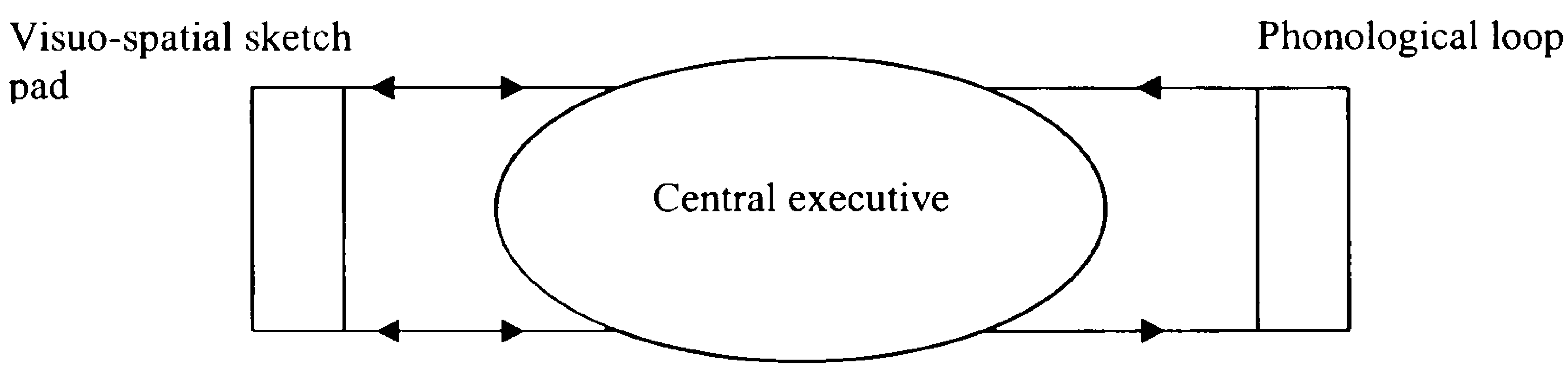
> "those functional components which allow humans to comprehend and mentally represent their immediate environment, to retain information about

their immediate past experience, to support the acquisition of new knowledge, to solve problems, and to formulate, relate, and act on current goals".

5.1.1 The Working Memory Model

Thinking is constrained by the ease with which stored knowledge can be retrieved, maintained and processed. Working memory is in turn constrained by it's architecture and the efficiency of operations within that architecture. Baddeley (1986, 1997) described a tripartite model of working memory comprising a controlling attentional system, termed the *central executive* which oversees and co-ordinates the operation of two slave systems: the *phonological loop* is concerned with processing linguistic material and comprises a phonological store served by an articulatory control process which allows for the maintenance of memory traces by means of subvocal rehearsal. The memory trace in the phonological loop (PL) is thought to decay after around one and a half seconds, but can be refreshed by the subvocal rehearsal. The other system, the *visuo-spatial sketchpad* is responsible for the construction and manipulation of visuo-spatial images. It can be fed directly through visual perception or indirectly through the generation of a visual image. Hence, information presented in non-visual modalities can be converted into a spatial code through imagery. Storage is maintained by a spatial control process thought to involve eye-tracking (what Baddeley and Logie, 1999, refer to as the *inner scribe*). Considerable experimental and neuropsychological evidence supports these distinctions within WM (see Baddeley, 1986; 1997 for extensive discussion). However, the slave systems are thought to play a rather more passive role in higher level cognitive tasks which involve processing as well as storage of information. For instance, simple measures of phonological loop capacity, such as digit recall span, do not predict accuracy in reading comprehension as do more complex measures which are assumed to also reflect the role of the central executive (e.g., Daneman & Carpenter, 1980; Gilhooly, 1998; Shah & Miyake, 1996). Jonides and Smith (1997) suggest that WM also comprises a further conceptual resource, over and above the verbal and spatial ones, which stores information such as semantic categories and facilitates the extraction of meaning from propositional material represented in the PL during semantically meaningful operations such as reasoning. For a discussion of recent developments in the working memory model see Baddeley and Logie (1999). Figure 1 presents a simplified representation of the WM model.

Figure 5.1: A simplified representation of the working memory model (from Baddeley, 1997, Figure 4.2, page 71).



Smith and Jonides (1997) present evidence from positron-emission tomography (PET) studies which clearly support the existence of different WM systems for spatial, object and verbal information. Within the verbal and spatial systems, separable components seem to be responsible for the passive storage and active maintenance of information. Moreover, separate executive functions processing the contents of WM (rather than simply storing or retrieving) were suggested, mediated by distinct brain areas in the frontal cortex.

A number of studies have investigated the involvement of the three components of working memory in reasoning, primarily through the use of dual-task methodology, whereby secondary tasks are used to reduce available capacity for particular subcomponents. The phonological loop is typically loaded by secondary tasks which suppress subvocal articulation (such as concurrent repetition of a given word or number) and the visuo-spatial sketchpad by concurrent spatial tasks, such as moving the non-preferred hand in a set spatial pattern. Gilhooly, Logie, Wetherick and Wynn (1993) found that such secondary tasks did not significantly reduce syllogistic reasoning performance, although a small but consistent interference by articulatory suppression was observed. Another secondary task, random number generation, did significantly affect performance and latencies for premise processing. This task was intended to interfere with the working of the central executive and the findings suggest that the central executive plays a major role in syllogistic reasoning with some minor involvement of the phonological loop (see also Gilhooly, Logie, & Wynn, 2002, in which overall working memory load of the task was raised).

However, no evidence was found for a visuo-spatial sketchpad contribution. Similar findings have also been reported in relation to propositional reasoning (Klauer, Stegmaier and Meiser, 1997) and conditional reasoning (Klauer et al 1997; Toms, Morris and Ward, 1993). Gilhooly et al. (1993) note that this goes against some model-oriented theories such as Euler circles, which have suggested the use of mental imagery. However, many subjects in their studies used superficial heuristics which place little demand on WM. Gilhooly et al. suggest that greater involvement of the visuo-spatial sketchpad might be observed if subjects were induced to employ other strategies. In more recent work, Gilhooly, Logie and Wynn (1998) explored performance of individuals who had been trained in syllogistic reasoning and found that those who produced the highest scores following training were most susceptible to the disruptive effects of random number generation. The training seemed to have resulted in better performance, but at the expense of greater load on WM.

5.1.2 Verbal and Spatial Strategies and Working Memory

Both the mental models and rules theories of syllogistic reasoning reviewed in Chapter 1 suggest that working memory capacity is an important factor in reasoning performance. According to model theorists, working memory plays a key role in the storage of initial putative models and mental footnotes, and in searching for counterexamples (e.g. Gilhooly, 1998; Johnson-Laird and Byrne, 1991). Rule based accounts claim its importance in the retention of temporary reasoning subgoals (e.g. Rips, 1994) or the retrieval and application of reasoning schema (e.g. O'Brien, 1998). See also Section 1.5.2 previously. If the working memory model is considered in conjunction with the strategic differences identified thus far, the propositional nature of the verbal strategy suggests involvement of the phonological subsystem of WM. The research reviewed above suggests minimal (if any) involvement of the visuo-spatial sketchpad in reasoning. However, some of the reasoners in the previous five experiments clearly use a visuo-spatial approach which suggests that the visuo-spatial component of WM may be important for these people.

In discussing the working memory model and strategies, Gilhooly (1998) referred only to "verbal" strategies. For instance, he describes an implementation of the "atmosphere" strategy, whereby rules are retrieved from long-term memory and applied to the premises. These rules are remarkably similar to the "quantifier linking rules" used by some verbal and mixed reasoners and described in Experiment 2. Gilhooly suggests this approach may be adopted because it places relatively little load on WM. He suggested also that, during reasoning, it is the central executive which stores information relevant to immediate processing and very recent processing, whilst the actual processing is carried out unconsciously by rules held in long-term memory that respond to the information in CE. Hence the CE is a key storage site during on-line processing. The slave systems act as temporary caches which back up the CE to avoid overload (Logie, 1995). The information in the CE is held in an abstract propositional code, whilst that in the slave systems is in more literal codes (Toms et al 1993). Premise information in either a articulatory (PL) or visual image (VSSP) is used together with intermediate processing results (Gilhooly, 1998). In terms of the verbal syllogistic strategies, the latter may include temporary premise rearrangements in substitution, or the two quantifiers in the linking rule approach.

However, for spatial reasoners, such on-line information might feasibly include an intermediate spatial arrangement of the A and B terms as presented in premise one of a syllogism. Although the research reviewed above and by Gilhooly (1998) has suggested little involvement for VSSP, tasks which are thought to have a very strong spatial component, such as linear syllogisms containing spatial adjectives, *are* affected by secondary tasks which load the visuo-spatial sketchpad, however, the central executive also seems to pay a major role (e.g. Vandierendonck & De Vooght, 1997).

Shah and Miyake (1996) postulated a fractionated central executive with resources for spatial and linguistic information. To demonstrate this, they employed a series of simple and complex measures of verbal and spatial working memory span, the simple measures drawing on one or other of the slave systems and the complex measures drawing on either of these plus the central executive also. Using similar methodology, Handley, Capon, Copp and Harper (2002) presented evidence that two systems for dealing with spatial and verbal representations were involved in both conditional reasoning and spatial problem solving. However, although these appear to be distinct systems, factor analysis did suggest some common processing requirements.

Capon, Handley and Dennis (2003) extended this research to syllogistic reasoning. They argued that if Shah and Miyake's thesis was correct, and verbal and spatial central executive resources are dissociable, then established theories of reasoning would offer differing predictions regarding the role of working memory in reasoning. Rule-based theories, which emphasise the role of propositional or language based representations (hence presumably a verbal strategy), would suggest that syllogistic tasks would draw preferentially on verbal resources. In contrast, model theories. (which encompass the spatial strategy), would predict a more important role for spatial working memory. Subjects completed two syllogistic reasoning tasks (with visual and verbal presentation of the premises) plus a series of working memory span measures. Correlational analysis indicated that syllogistic reasoning performance was predicted by both spatial and verbal working memory span. Furthermore, a confirmatory factor analysis showed that an orthogonal three factor model, comprising a verbal, spatial and general factor, fitted the data well. Interestingly, Capon et al. have since tested this three factor model on Shah and Miyake's data and found it to be an excellent fit. Overall, syllogistic reasoning performance (irrespective of presentation modality) loaded significantly, and to a similar degree, on *both* verbal and spatial working memory resources, and also on a third, general factor.

Capon et al. offer two possible explanations; either that syllogistic reasoning involves both verbal and spatial forms of representation or, and perhaps more importantly for the present discussion, that individual differences exist: loadings on the two factors may in fact reflect different groups of individuals. In terms of the verbal and spatial strategies described earlier, the groups may differ in terms of the verbal and spatial working memory capacity they have available. They may also differ in the extent to which they draw on working memory resources. For visually presented material, we would expect a degree of phonological loop involvement for all reasoners. However, spatial reasoners may also draw on visuo-spatial sketchpad resources. Experiment 6 aimed to investigate these questions.

5.2 Experiment 6

5.2.1 Aims

Experiment 6 aimed to investigate the relative role of the three working memory components for verbal and spatial reasoners. Given that individuals seem to possess a predilection towards one or other strategy, there were two possible predictions:

1. either that the strategy groups differ in their verbal and spatial working memory capacities, or

2. that their capacities are similar but they draw differentially on verbal and spatial working memory resources.

The procedures followed were similar to those of Handley et al (2002) and Capon et al (2003).

5.2.2 Methods

Participants

Participants were 155 undergraduate and postgraduate students from University of Plymouth who volunteered to take part in return for their choice of course credit or £7.50 cash. The sample comprised 128 females and 27 males with an overall mean age of 22.08 years. All were native English speakers and none had received formal training in logic.

Materials and procedure

Participants were run in pairs, each seated at a separate desk with a computer terminal. Each participant was presented with six experimental tasks in a session lasting around one and a half hours. The six tasks are described below in the order presented to every participant. The computerised tasks were all developed in Visual Basic 6.0 and presented on a Pentium IV PC. Tasks 3 to 6 below were based on ones used by Shah and Miyake

(1996). These were also the four WM measures employed by Handley et al (2002) and were among those used by Capon et al (2003). In summary, the simple span tasks are designed to measure PL and VSSP storage capacity whilst the complex span tasks also include a processing component, and hence assume involvement of the central executive.

1. Reasoning Strategy

Exactly the same procedure as employed in Experiments 1 to 5. Again, the five syllogism task employed in Experiments 4 and 5 was used with only the occupation names changed.

2. Syllogistic Reasoning Performance

Exactly the same procedure was used as in Experiment 5. The 16 syllogisms presented were shown in Appendix 4A. For Experiment 6 the figure, mood remained as shown, only the occupation names were changed. Task instructions remained as shown in Appendix 4B.

3. Simple Verbal Word Span (words)

This task was designed to measure passive storage capacity for verbal information, with no explicit processing requirement. The words used in this task were all two syllable concrete nouns. They were selected by Capon et al from the Oxford Psycholinguistic Database (Quinlan, 1992) and their usage frequency rating ranged from 5 to 216 according to the Kucera and Francis (1967) corpus of over 1 million words.

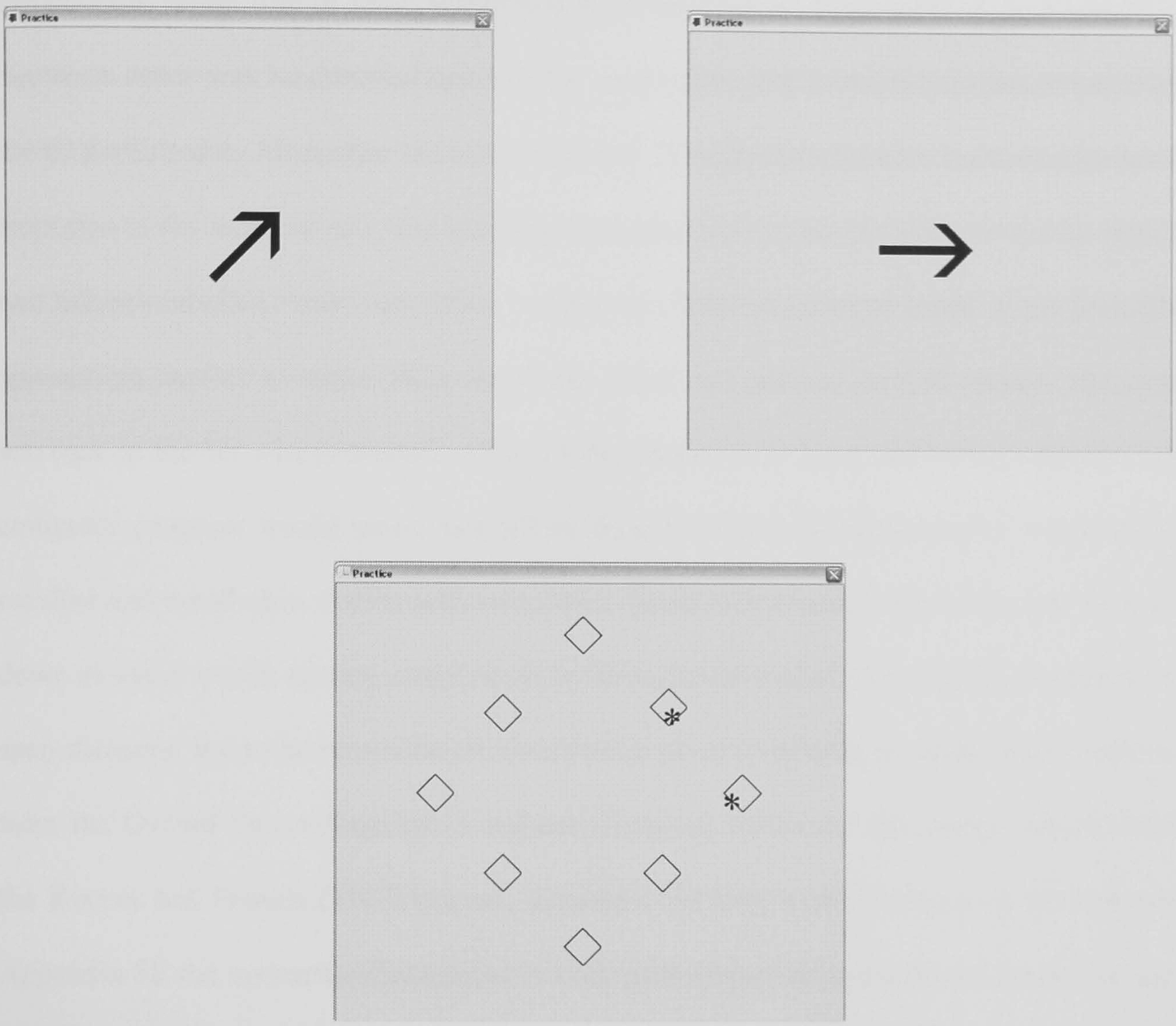
To-be-remembered words were presented in sets of increasing size from two to seven words, with five sets of each size. Each word appeared on the computer screen individually for 800 msecs, with an interstimulus interval of 50 msecs. Written instructions were presented to participants at the outset, followed by two practice sets of two words. The instructions are shown in Appendix 5A and all the word sets in Appendix 5B. At the end of each word set, a lightbulb symbol appeared on the screen and the program paused to allow the participant time to write down the words presented on a simple paper response sheet. They were instructed to write down as many words as they could recall in the order presented on computer. Hence difficulty increased with set size.

The simple verbal word span score was calculated by awarding one point for each correctly recalled word, in each correctly recalled set, regardless of level. For instance, if a participant completed four 2 word sets correctly and two three word sets correctly; score = (4x2) + (2x3) = 14. The maximum possible score was 135.

4. Simple Spatial Arrow Span (arrows)

Again this task originated in the work of Shah and Miyake and developed in the present form by Capon et al. The task was designed to measure passive storage capacity for spatial information, with no explicit processing component. Again the stimuli were presented in sets, with three sets at each level of two to six items, 15 sets in total. However this time each item comprised an arrow pointing in one of eight possible orientations. Each arrow remained on the screen for 1000 msecs with an interstimulus interval of 250 msecs. Following each set of arrows, a grid was presented, showing all eight possible orientations. The task was to indicate, through a mouse click, the directions of each of the arrows in the set, in the order presented. The written instructions presented to each participant are presented in Appendix 5C and were followed by two practice sets of two arrows each. Figure 5.2 below shows a two arrow set and the response grid which followed all sets. On the grid shown, the correct response would be to click on the grid points indicated with an asterisk (*).

Figure 5.2: A typical two arrow stimulus set and response grid as presented in the Simple Spatial Span task (arrows).



The Simple Spatial Arrow Span score was calculated by a similar procedure to that for the simple word span described previously. One point was awarded for each correct arrow, in each fully correct set; i.e. if a participant got three two arrow sets correct and one three arrow set, (3x2) + (1x3) = 9. The maximum total score for this task was 60.

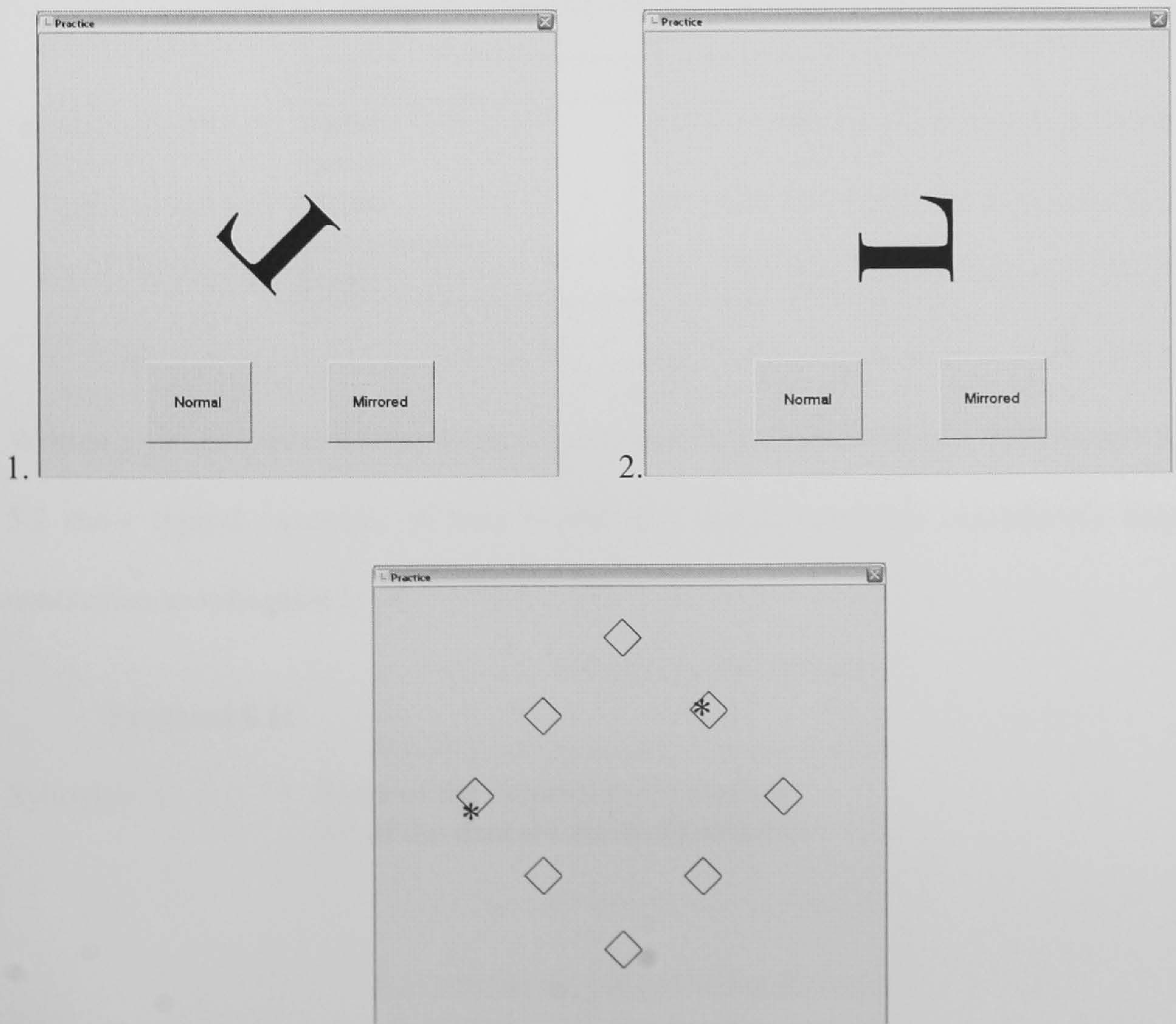
5. Complex Verbal Sentence Span (sents)

This task was based on the reading span test devised by Daneman and Carpenter (1980) and was also used by Capon et al (2003). It is intended to be a span measure of functional working memory capacity for verbal material and also requires the simultaneous maintenance and processing of information. The task required participant to verify sets of sentences as either true or false, whilst also remembering the final word of each sentence. Sentence order was randomised prior to the experiment and remained in that same order for all participants. Altogether the test comprised 25 sentence sets, with 5 sets at each level from two to six sentence sets. Each sentence remained on screen for 800 msecs after which two buttons labelled "true" and "false" appeared. Participants were asked to click on the appropriate button to make their response. After, an interval of 250 msecs, the next sentence in the set was presented. After all the sentences in a set had been presented, the computer program would pause and allow time for the to-be-remembered words to be recalled and noted on a written response sheet. Again, participants were instructed to write down as many words as they could recall in the order presented. As with the Simple word span measure, the to-be-remembered words were all two syllable concrete nouns selected from the Oxford Psycholinguistic Database (Quinlan, 1992) and frequency verified with the Kucera and Francis (1967) corpus. Appendix 5D shows all sentences in the task and Appendix 5E the instructions presented to each participant and example test screen format. Scoring was by the same principle as for the previous two tasks. Hence, the maximum score was 100. The number of errors made in the verification component of the task (true/false judgement) was also recorded.

6. Complex Spatial Letter Span (letters)

This task was a span measure of functional working memory capacity for spatial information which also required the simultaneous maintenance and processing of spatial material. Again, the test items were presented in sets, this time of 2 to 5 items, with five sets at each level. Each item comprised one of five letters (either F, J, L, P or R) presented in one of eight possible orientations and either as a normal or mirror image. Presentation was constrained so that opposing orientations were not presented successively within a set and that each orientation appeared only once per set. The letters remained on the screen for a maximum of 5000 msec. and participants were required to respond whether the letter was a normal or mirror image by clicking on an appropriate button. After an inter-stimulus interval of 250 msecs the next letter would appear. After each set, participants were presented with a grid identical to that used in the simple arrow span, and shown in Example 5.1 previously. They used the mouse to click and indicate the direction in which the top of each letter was oriented, in the order the letters were presented. The instructions presented to each participant are presented in Appendix 5F and an example of a two item set in Figure 5.3 below. In the example, the correct response to the letter 1 would be to click the button labelled "mirrored" and for letter 2 the button labelled "normal". On the grid shown after this set, the correct boxes are again marked with an asterisk.

Figure 5.3: A typical two letter stimulus set and response grid as presented in the Complex Spatial Span task (letters)



Scoring was calculated exactly as for the complex verbal sentence span. If a set was correct, regardless of level, one point was recorded per letter. The total possible score was therefore 80. The number of errors on the verification component (normal/mirrored judgement) was also recorded.

5.2.3 Results

5.2.3.1 Strategies for Syllogistic Reasoning

Verbal and spatial strategies were again identifiable from written protocols according to the same criteria as used in Experiments 1 to 5. In addition, 9 participants seemed to be using a mixture of the two approaches and 5 either produced protocols which were indeterminate, or no protocols at all despite instructions. Table 5.1 shows the proportion of participants within each strategy group.

Table 5.1: Syllogistic reasoning strategies identified in Experiment 6.

Strategy	N
Verbal	93
Spatial	48
Mixed	9
Indeterminate	5
Total	155

Written protocols were similar to those presented in Experiments 1 to 5. Protocols 5.1 and 5.2 show typical examples of how verbal and spatial reasoners respectively reached a conclusion to syllogism 1.

Protocol 5.1:

Syllogism 1 — Some of the carpenters are dancers
All of the dancers are birdwatchers

Participant 7 Some of the carpenters are dancers

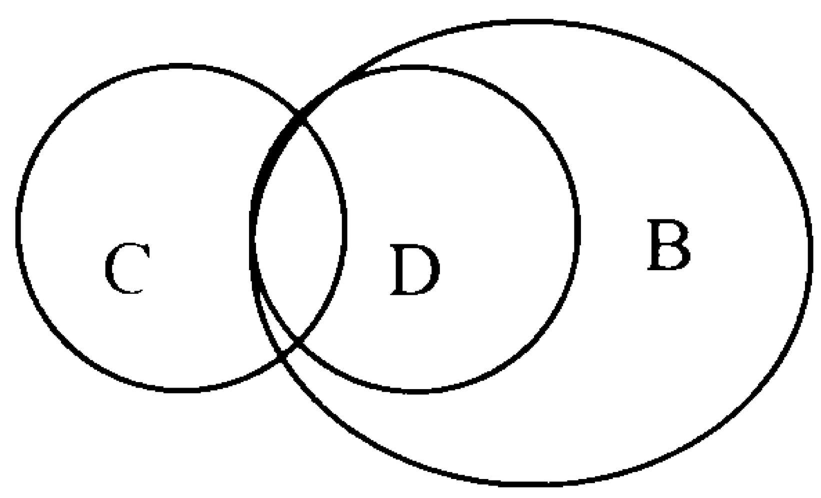
swap so start with a larger group of people. All of the dancers are birdwatchers

Correct conclusion given: Some of the carpenters are birdwatchers

Protocol 5.2

Syllogism 1 as above

Participant 20



Correct conclusion given: Some of the carpenters are birdwatchers

5.2.3.2 Performance on all Measures

Table 5.2 below shows descriptive statistics for the overall sample (N = 155) and for verbal and spatial reasoners on syllogistic reasoning and on all six working memory measures.

Table 5.2: Descriptive statistics showing performance of whole sample (N = 155), verbal (N = 93) and spatial (N = 48) reasoners on syllogistic reasoning and all working memory measures.

Measure	Mean			Std. Dev.			Reliability **
	All	Verbal	Spatial	All	Verbal	Spatial	
Syllogisms (max 16)	7.59	7.46	8.02	3.12	2.70	3.63	0.82
Words (max. 135)	47.22	47.26	47.71	15.12	12.72	19.15	0.87
Sents (max. 100)	27.60	27.76	26.33	16.98	15.68	17.77	0.93
Sents error (max. 100)	8.59	8.16	9.10	5.56	5.63	5.95	-
Arrows (max. 80)	22.84	21.87	23.73	9.31	8.78	9.93	0.89
Letters (max. 100)	18.70	18.12	18.23	14.64	12.86	15.94	0.50
Letters error (max. 100)	11.41	12.61	9.52	8.88	8.69	8.67	-

* * Reliability estimates (split-half/Spearman-Brown) from those presented for identical tasks by Capon et al (2003, Experiment 3, page 227).

As Table 5.2 shows, performance was remarkably similar across all three sets of participants (all, verbal and spatial) on all measures. On syllogistic reasoning, in line with the performance observed in Experiments 1-5, spatial reasoners (M = 8.02 correct out of

16) performed slightly better than verbal (M= 7.46), but the difference was not significant: t(74.6) = 0.94, p > 0.05. With regard to the working memory measures, of some interest are the processing error rates. Overall, participants made between 8-9% errors on the processing element of the complex verbal span task (sents. error) and 9-12% on the equivalent components of the complex spatial task (letter error). This is well below the chance rate of 50% which would be expected if they were simply guessing and suggests that participants were actively attempting the processing element of these tasks. Comparing the strategy groups on these measures, spatial reasoners made significantly fewer errors on the processing component of the complex spatial task (letters error) than verbal reasoners; t(139) = 2, p = 0.05. The opposite trend was observed with regard to the verbal processing errors, but this effect did not reach significance (p > 0.3). Across all other measures, verbal and spatial performance was comparable (p > 0.05 in every case). There was slightly more variance in the spatial group, especially on the verbal working memory measures (words and sents) which was caused by a few exceptionally good all round performers. Overall however, the values shown in Table 5.2 were very similar to those found in the three experiments presented by Capon et al (2000).

5.2.3.3 Correlational Analyses

Working memory span measures

Table 5.3 presents correlations between the six working memory measures for the whole sample (N = 155). Table 5.3 indicates that the strongest correlation is between the two verbal span measures (r = 0.54, p < 0.001) and the second strongest between the two spatial span measures (r = 0.47, p < 0.001). These results are to be expected if the verbal and spatial measures are tapping different working memory resources. However, both verbal spans also correlated significantly with the spatial span measures, the relationship being particularly strong with the complex spatial span ($r = 0.42$, $p < 0.001$) and the two complex span measures are also correlated significantly ($r = 0.42$, $p < 0.001$). The measure of verbal processing error correlated negatively, significantly and to a similar magnitude with all four span measures, whilst spatial error correlated significantly and negatively only with the spatial span measures.

Table 5.3: Correlations between all working memory measures, across entire sample (N = 155).

Measure	Simple verbal span (words)	Complex verbal span (sents)	Verbal processing error (sents error)	Simple spatial span (arrows)	Complex spatial span (letters)
Complex verbal span (sents)	0.54**	-	-	-	-
Verbal processing error (sents error)	-0.19*	-0.21*	-	-	-
Simple spatial span (arrows)	0.13	0.28**	-0.18*	-	-
Complex spatial span (letters)	0.25**	0.42**	-0.29**	0.47**	-
Spatial processing error (letter error)	-0.05	-0.7	-0.00	-0.28**	-0.24**

* Correlation is significant at the 0.05 level (2-tailed)
** Correlation is significant at the 0.01 level (2-tailed)

Overall, these findings are closely in line with those of Capon et al (2003). They suggest that the span measures share some common variance (possibly due to shared central executive resource on complex tasks) but, as the highest correlations tend to be between same-modality measures, spatial and verbal resources can be considered to possess a fairly high degree of independence.

Working memory span and syllogistic reasoning

Table 5.4 shows correlations between the span measures and syllogistic reasoning performance both for the overall sample, and for the verbal and spatial strategy groups.

Overall, reasoning correlates significantly with all six measures, negatively with the two error measures. Again this is in line with Capon et al's findings. For the individual strategy groups, the correlations show that simple verbal span predicted syllogistic performance for verbal ($r = 0.36$, $p < 0.05$) but not spatial reasoners ($r = -0.05$, $p > 0.05$); $z = 2.27$, $p < 0.05$. This finding is fully in line with prediction. The other WM measures show almost identical correlations with syllogistic reasoning, across the two strategy groups. Not surprisingly given the coefficients in Table 5.4, none of these inter-strategic differences reached significance ($p>0.05$ in every case).

Table 5.4: Correlation between syllogistic reasoning performance and working memory span, for overall sample and for verbal and spatial reasoners.

Measure / Strategy	words	sents	sents error	arrows	letters	letter (error)
Syllogisms						
Overall	0.18*	0.38**	-0.22**	0.34**	0.25**	-0.21*
Verbal	0.36*	0.37**	-0.27**	0.30**	0.20	-0.18
Spatial	-0.05	0.37*	-0.13	0.37**	0.19	-0.15

* Correlation is significant at the 0.05 level (2-tailed)
** Correlation is significant at the 0.01 level (2-tailed)

Table 5.5 presents a similar analysis which allows for examination of correlations between the working memory measures and verbal and spatial strategy. Again, many of the correlations are remarkably similar across strategy groups. In some of these cases (for instance, the relationship between sents error and words) correlations of similar magnitude are observed to be significant for the verbal group, but not for spatial. This can be attributed to the differing sample sizes within the two groups (N = 93 and 48 respectively).

Again, the largest correlations are between related span measures for both strategy groups; for words*sents, $r = 0.57$ and $r = 0.53$ for verbal and spatial reasoners respectively, $p < 0.01$ in both cases; for arrows*letters, $r = 0.33$, $p < 0.05$ and $r = 0.59$, $p < 0.01$, for verbal and spatial reasoners respectively. The latter data suggest some difference between strategy groups, and this difference approximates to significance; $z = 1.86$, $p = 0.06$.

Table 5.5: Correlations between all six working memory measures, by strategy group.

Measure / Strategy		words	sents	sents error	arrows	letters
sents	V	0.57**	-	-	-	-
	S	0.53**	-	-	-	-
sents. error	V	-0.27*	-0.23*	-	-	-
	S	-0.15	-0.22	-	-	-
arrows	V	0.26*	0.22*	-0.24*	-	-
	S	-0.02	0.26	-0.20	-	-
letters	V	0.36**	0.37**	-0.33**	0.33*	-
	S	0.14	0.30*	-0.33*	0.59**	-
letters error	V	-0.05	-0.04	0.03	-0.28**	-0.24*
	S	-0.09	-0.06	-0.10	-0.21	-0.13

Some inter-strategic differences are apparent in the relationship between the simple word span (words) and the two spatial spans. For verbal reasoners, the former seems to predict the latter in both cases ($r = 0.26$, $p < 0.05$ for arrows, $r = 0.36$, $p < 0.001$ for letters). For spatial reasoners, the correlations are much smaller ($r = -0.02$ and 0.14 respectively) and non-significant. However, when the correlations were compared across strategy groups, no significant differences were observed; for words*arrows, $z = 1.58$, $p > 0.1$; for words*letters, $z = 1.31$, $p > 0.1$.

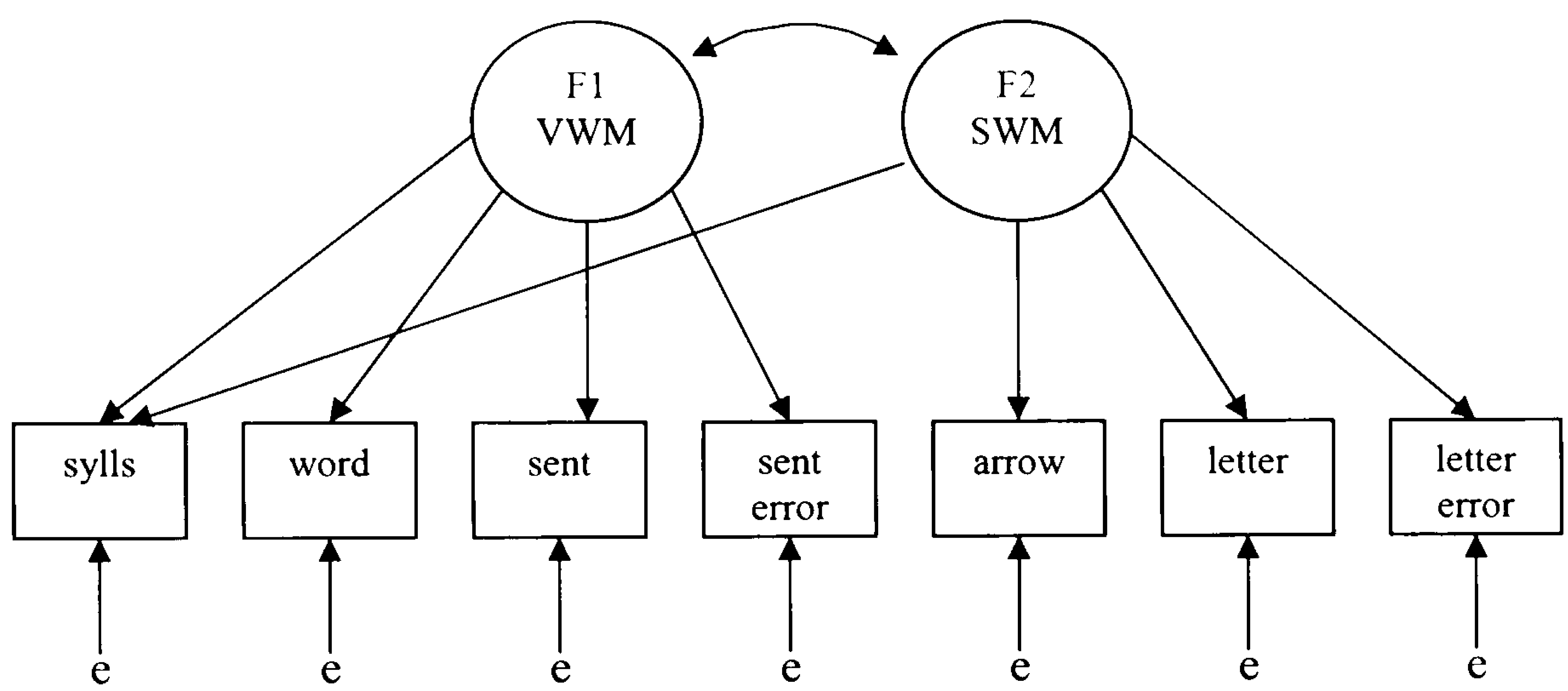
Summary of findings so far

Overall, the above analyses present evidence that verbal and spatial working memory are fairly distinct pools of resource. The simple and complex span measures are highly inter-correlated within modality, both for the whole sample and within each of the strategy groups. The dissociation seems to be especially strong in the spatial group where the spatial spans are particularly highly correlated, compared to the verbal strategy. However, some overlap is apparent, which is to be expected if a shared central executive resource operates in both complex spans. Moreover, there is tendency for an association between simple verbal span and both spatial measures for verbal reasoners, but not for spatial. This suggests that verbal reasoners were in fact drawing partly on their verbal resource during the spatial span tasks, whilst spatial reasoners had no need to do so. In terms of syllogistic reasoning (the task by which strategy groups were determined), the simple verbal span predicted performance for the verbal reasoners, but not for spatial, suggesting the latter group do not rely to such an extent on verbal working memory ability when reasoning. However, the complex verbal span was a predictor for both groups, and to the same extent. This was also the case with the simple spatial span, although, perhaps surprisingly, the complex spatial span predicted syllogistic performance for neither group. Overall, when conducting syllogistic reasoning, it would seem that whilst spatial reasoners draw almost exclusively on their spatial working memory capacity, verbal reasoners may need to draw on both verbal and spatial resources. Both groups also, of course, presumably draw upon central executive resources, as necessary, during the processing of information. To test these assumptions, a confirmatory factor analysis was conducted. This process allows for the testing of theoretically motivated models against the above data.

5.2.3.4 Confirmatory Factor Analysis

Motivated by the findings of Handley et al (2002) described above, a correlated two factor model was tested. This assumes some common processing resources and is hence consistent with the WM framework (Baddeley, 1996;1997). The sample sizes of N=93 (verbal) and N=48 (spatial) are quite small for this type of analysis, especially in the latter case. However, this is a fairly simple model with 16 parameters to be estimated. A minimum of five subjects per free parameter is generally recommended (Bentler, 1995) and in this respect the present sample size is sufficient overall, though if taken alone, the spatial group is a little small. The two factors were assumed to correspond to verbal WM (F1) and spatial WM (F2). Syllogistic reasoning was allowed to load freely on either factor, the verbal measures (words, sents and sents errors) on the VWM factor and the spatial capacity measures (arrows, letters and letter errors) on the SWM factor. Syllogistic reasoning remained unconstrained in order to examine whether its relationship with the working memory components differed across strategy groups. Figure 5.4 below presents a path diagram illustrating the model.

Figure 5.4: Path diagram of 2 correlated factor model



A multi-population analysis, using the Maximum Likelihood method, was conducted to determine whether the same model was appropriate for both verbal and spatial strategy groups. The six working memory measures were constrained to equality across the two groups, whilst syllogistic reasoning performance remained free to vary across groups. The full EQS program output for the model is presented in Appendix 5G. The Goodness-of-fit summary Chi-square suggested a fairly good fit; $\chi^2 = 34.96$ df = 28, p = 0.17. This was further supported by a comparative fit index (CFI) of 0.95 and a Root Mean Square Error of Approximation (RMSEA) of 0.04. The 95% confidence interval for the RMSEA was 0 - 0.081. The higher end of this interval would correspond to a model with less than satisfactory fit, although at this small sample size (especially in the spatial group) it is not possible to test the model more rigorously. Overall, the two factor model was likely to be appropriate for both strategy groups. The standardised factor loadings are shown in Table 5.6.

Table 5.6: Standardised factor loadings for Model 1 (all measured variables constrained). CFI = 0.96; RMSEA = 0.04.

Variable	Strategy	Factor 1 (VWM)	Factor 2 (SWM)	R^2
Syllogisms				
	Verbal	0.40*	0.15	0.26
	Spatial	0.25#	0.30*	0.21
Words				
	Verbal	0.73*		0.53
	Spatial	0.60*		0.36
Sents.				
	Verbal	0.77*		0.59
	Spatial	0.90*		0.82
Sents. error				
	Verbal	-0.31*		0.10
	Spatial	-0.36*		0.13
Arrows				
	Verbal		0.60*	0.36
	Spatial		0.80*	0.64
Letters				
	Verbal		0.63*	0.40
	Spatial		0.73*	0.53
Letters error				
	Verbal		-0.25*	0.06
	Spatial		-0.34*	0.12

* Loading significant at the 0.05 level
\# Loading approximates to sig. (p = 0.07, 1-tailed)

As these data show, some differences between groups were apparent. For verbal reasoners, syllogistic reasoning loaded significantly on Factor 1 (verbal resource) but hardly at all on Factor 2 (spatial resource). Spatial reasoners on the other hand, seem to draw similarly on verbal and spatial resources. This seems to suggest a pattern of loadings which would be in line with prediction 2. The correlation between Factors were also disparate (r = 0.67 for verbal reasoners and r = 0.37 for spatial) suggesting that the relationship between verbal and spatial memory resource may be stronger for verbal reasoners. Otherwise, little difference between the two groups was observed and some low R^2 values suggest that a fair amount of variance remains unexplained. Overall, it must be acknowledged that the differences observed between patterns of loadings across the two groups have not been directly shown to be significant, and, whilst they are fully in line with prediction 2, given the limited sample size, this finding should be seen as suggestive, rather than definitive.

Another possibility is that the working memory measures may load differentially on the factors, across the two strategies. The pattern of inter-correlations shown previously in Table 5.5 indicated different patterns of relationships between verbal and spatial measures for the two groups. Similarly, correlations presented in Experiment 5 (see Table 4.6) showed that IQ scores correlated significantly with spatial ability for verbal reasoners ($r = 0.43$, $p < 0.01$) but not at all for spatial reasoners ($r = 0$) suggesting that verbal reasoners may have been drawing, to an extent, on the same resource for both the verbal ability and mental rotation tasks. Therefore, a second model was tested, identical to the first, except that all constraints on the working memory measures were lifted, in effect fitting two separate models for the two strategy groups. The full EQS output from this model is presented in Appendix 5H.

For Model 2, the fit indices again suggest a fairly well fitting model; $\chi^2 = 29.82$, df = 24, $p = 0.19$, CFI = 0.96, RMSEA = 0.04; RMSEA CI = 0 - 0.08. A chi square change test was conducted to compare the fit of the two models. This resulted in a χ^2 change = 5.14, df = 4. For 4 df, the critical value of χ^2 when $p = 0.05$ is 9.49, which indicates that Model 2 does not fit the data significantly differently to Model 1. The standardised factor loadings for Model 2 are shown in Table 5.7. A pattern of loadings is observed which is again in line with the second á priori prediction for Experiment 6, namely that the two strategies draw differentially on working memory resources. However, this trend seems to be in the opposite direction to that assumed following the correlational analyses of task performance data reported earlier. Like Model 1 previously, Model 2 indicates that, for verbal reasoners, syllogistic performance draws significantly on verbal working memory (Factor 1) but not on Factor 2. Spatial reasoners on the other hand, appear to draw significantly, and similarly, on *both* spatial and verbal factors. All the WM measures loaded significantly on their respective factors. The correlation between the two factors was moderate, though significant, in the spatial group ($r = 0.30$, $p < 0.01$) and strong and significant ($r = 0.68$, $p < 0.01$) for the verbal group.

Table 5.7: Standardised factor loadings for Model 2 (WM variables constrained). CFI = 0.93; RMSEA = 0.05.

Variable	Strategy	Factor 1 (VWM)		Factor 2 (SWM)		R^2	
Syllogisms							
	Verbal	0.43*		0.10		0.26	
	Spatial		0.28*		0.31*		0.22
Words							
	Verbal	0.75*				0.56	
	Spatial		0.53*				0.28
Sents.							
	Verbal	0.73*				0.53	
	Spatial		1.00*				1.00
Sents. error							
	Verbal	-0.39*				0.15	
	Spatial		-0.22#				0.05
Arrows							
	Verbal			0.55*		0.29	
	Spatial				0.92*		0.85
Letters							
	Verbal			0.68*		0.45	
	Spatial				0.64*		0.41
Letters error							
	Verbal			-0.34*		0.11	
	Spatial				-0.22~		0.05

* Significant loading at 0.05 level
# Loading approximates to sig. (p = 0.06, 1-tailed)
~ Loading approximates to significance (p = 0.08, 1-tailed)

5.3 Discussion

Much research has related reasoning accuracy to working memory capacity, and the two most influential theories of reasoning (mental models and rules) both suggest that it plays a key role in reasoning processes (see Chapter 1). However, studies of working memory and reasoning have not tended to consider the possibility of individual differences in how people reason, and whether this may lead them to use their working memory resource in different ways. Experiment 6 aimed to investigate the relationship between working memory capacity and verbal and spatial strategies for syllogistic reasoning. Given that Experiment 5 has already shown that, whatever their strategic preference, individuals do not seem to differ in verbal or spatial ability, two possible predictions regarding working memory were suggested; either that strategy preference reflects differences in verbal-spatial working memory capacity, or, that individuals using different strategies draw differentially on those particular working memory resources.

The first of these possibilities was investigated by examining the performance of the two groups of reasoners across the measures of working memory capacity. On the complex span tasks (those which involve both a storage and processing element) spatial reasoners made significantly fewer processing errors on the spatial task, whilst on the verbal task, the opposite trend was observed, although there was no overall difference in span scores. This suggests that indeed, the two groups may differ in their ability to process information within spatial and verbal domains. Nevertheless, the overall span score is unaffected, suggesting that spatial processing ability is not a key factor contributing to performance on the task.

The second prediction was that strategy groups would draw differentially on working memory resources during syllogistic reasoning, spatial reasoners tending to draw on the spatial resource and verbal on the verbal resource. Correlational analysis suggested two main inter-strategic differences: firstly, that the simple verbal span (i.e. verbal storage capacity) predicted syllogistic reasoning performance for verbal, but not spatial, reasoners; secondly, this measure correlated significantly with both spatial spans for verbal reasoners, but not for spatial. This might suggest that verbal reasoners, to an extent, draw upon the same resources for verbal and spatial tasks, whilst spatial reasoners draw almost exclusively on spatial resources for spatial tasks and verbal resources for verbal tasks.

Confirmatory factor analysis (CFA) allowed for the testing of a correlated two factor model against the data from Experiment 6, the factors corresponding to verbal and spatial working memory. This model was motivated by that used by Handley et al (2002), who have demonstrated that both verbal and spatial resource is implicated in reasoning with conditional premises. Capon et al (2003) have presented similar findings regarding syllogistic reasoning, their model also comprising a third general factor. However, Experiment 6 had too few participants (especially in the spatial group) for a three factor model to be technically viable.

As previous work has suggested little involvement of the visuo-spatial scratchpad in reasoning (e.g. Gilhooly et al, 1993), all reasoners might be expected to draw on verbal resources to some extent, and the CFA indicted that this was indeed the case in Experiment 6. However, spatial reasoners *also* load significantly onto the spatial factor. In reviewing a number of studies, Gilhooly (1998) suggested that the general lack of evidence for VSSP involvement goes against model based theories which posit the involvement of mental imagery. One theory he cites particularly is that of Euler circles which, as all six Experiments in this thesis have shown, approximate to the protocol representations consistently presented by spatial reasoners. Gilhooly's participants however, used mainly heuristic strategies and he concedes that more VSSP involvement may be observed if reasoners were induced to use other strategies. Certainly tasks which are known to possess a strong spatial component have been shown to require VSSP involvement. Experiments 1-6 here have not required any inducement for around a third of reasoners to spontaneously produce representations allied to Euler circles, and to describe reasoning processes which seem distinctly spatial in nature. These are the individuals who, in Experiment 6, have been shown to draw on spatial working memory resources during reasoning. The verbal reasoners identified in Experiments 1-6, on the other hand, present strategies involving the substitution of terms and/or linking of quantifiers. In Gilhooly's terms, such processes would be classed as heuristic (certainly by comparison to the complex logical rules proposed by Rips, 1994, or even the substitution rules of Ford, 1995). For these individuals, syllogistic reasoning loads almost entirely on the verbal factor, totally as the research that Gilhooly describes would predict.

Capon et al, in common with other research reviewed in this chapter, assumed strategic universality. Although they investigated individual differences in terms of the relative extent to which reasoners draw on the three working memory components, they did not take into account possible differences in strategy – rather, they assumed that all reasoners were using basically the same functional approach and drawing on both resources. However, in discussing their findings, Capon et al offer an alternative, post-hoc, explanation - that individual differences in strategies for syllogistic reasoning may exist, and that some reasoners draw primarily on verbal resource whilst others draw on spatial. This latter explanation is fully supported by the findings of Experiment 6.

However, although inter-strategic differences are indicated, these are preliminary findings only and need to be treated with some caution. A fair amount of variance in performance is not explained by the two factor model presented here and hence, there may be other determinants of performance which are as yet unaccounted for, and which may be related to strategy choice. In addition, the sample size was quite small for this type of analysis (especially in the spatial strategy group). Overall though, it remains that Experiment 6 has indicated inter-strategic differences in how individuals draw on the verbal and spatial components of working memory. This in turn further supports the hypothesis which this thesis set out to investigate, i.e. that the two groups of reasoners, which have been repeatedly identified throughout Experiments 1-6, are indeed employing functionally distinct reasoning strategies. Chapter 6, to follow, will summarise the findings presented in this thesis, the evidence they present for individual differences in strategies for syllogistic reasoning and the implications of such differences, across both theoretical and applied situations.

CHAPTER 6

GENERAL DISCUSSION

6.1 Introduction

The primary aim of the experimental programme presented in this thesis was to investigate the nature and veracity of individual differences in strategies for human reasoning. The six studies described have presented clear evidence for the existence of such strategies. The aim of this final chapter is to discuss the findings of Experiments 1 to 6 and their implications. It will begin with a brief summary of the previous 4 experimental chapters in terms of the aims and results of the studies in question. The implications of these results will then be discussed in terms of the theories of reasoning presented in Chapter 1. Possible future directions for the continuation of this line of research will then be considered before the final concluding comments.

6.2 Summary of Experimental Findings

The initial experiment was inspired by the work of Ford (1995). From verbal and written protocols produced during syllogistic reasoning, Ford concluded that two main types of reasoning behaviour could be observed and she termed these verbal and spatial strategies. Those individuals which Ford termed spatial reasoners used shapes such as circles or squares placed in different spatial configurations to represent the relationships between the terms in the premises. They provided written protocols illustrating such procedures diagrammatically, and their verbal reports frequently described these relationships in terms of group membership, sets and subsets. The verbal reasoners, on the other hand, displayed various types of substitution behaviour. To reach a conclusion, they replaced the middle term from one premise with the end term from the other, as if solving an algebraic problem. Ford presented some fairly complex substitution rules which she claimed were being used by the majority of verbal reasoners in her study.

However, Ford's work has been criticised on methodological grounds, mainly due to her small sample size (just N = 8 in each strategy group) and for her reliance on verbal protocol. Furthermore, although some other researchers have found similar evidence to Ford in written protocols, verbal reports have proved inconclusive (e.g. Bucciarelli and Johnson-Laird, 1999). Hence, Experiment 1 attempted to replicate and extend Ford's study by using a larger sample and by investigating whether strategies may be identified without the need for verbal protocols. In pursuit of these aims, participants were allocated to one of two conditions, in one condition the procedure replicated that of Ford, in the other only written protocols were collected. Furthermore, all participants completed a questionnaire designed to elicit evidence of the reasoning behaviours described by Ford.

Data from Experiment 1 clearly replicated the findings of Ford (1995). In the replication condition, both written and verbal protocols presented evidence of verbal and spatial strategies with characteristics as described in Ford's work. Moreover, in the other condition, written accounts of verbal and spatial strategies were identical in nature, clearly indicating that they were not just artefacts generated as result of asking participants to verbally articulate their reasoning. A small number of reasoners appeared to be using a mixed approach. The questionnaire also proved reliable for 8 items and this provided useful converging evidence for the strategic distinction made from written reports. The fact that this questionnaire had not previously been piloted, and yet proved reliable, is further evidence that Ford's findings and those of Experiment 1 were compatible.

However, Experiment 1 did not fully replicate some of the more detailed performance data presented by Ford. Firstly, although verbal reasoners in Experiment 1 present clear evidence of using substitution behaviour, it only tended to be of the form Ford describes as "naive substitution", i.e. it occurred in consistent way across all problems, irrespective of mood or figure. There was no evidence of reasoners attempting to use the more sophisticated substitution rules that Ford describes for more difficult problems. For instance, Ford makes a distinction between problems which require the modus ponens and modus tollens forms of her rules, and shows that her verbal reasoners found the latter most difficult. However, in no instance did the data from Experiment 1 suggest any similar difference. Rather than attempting the more formal or sophisticated substitution process that Ford describes for more difficult syllogisms, verbal reasoners appeared to use naïve substitution throughout. On more difficult problems which contained an *All* premise, they frequently made precisely the type of conversion errors that Ford claims her reasoners eschew (assuming that *All A are B* is equivalent to *All B are A*). Similarly, Ford describes how syllogisms where premises are less constrained (in essence, she means that they can be represented in multiple ways) are more intractable for spatial reasoners. Experiment 1 found no such effect. For all problems, spatial reasoners presented a single model of the premises and showed little evidence of considering alternative representations on less-constrained problems.

Moreover, when individual syllogisms are considered, more disparity with Ford's findings became apparent. For verbal strategy users, naive substitution can be successfully applied to all same-form syllogisms and hence these are easy for them. But, because the verbal strategy relies on premise quantifiers in generating conclusions, it can lead to difficulty for different-form problems, and verbal reasoners duly tended to produce conclusions which were incorrect by virtue of containing the same quantifier as one of the premises. However, in Experiment 1, a similar trend was also observed for spatial reasoners. In mental-model terms, the different-form syllogisms are three-model and hence the most difficult of all (Johnson-Laird & Byrne, 1991), and these are also described by Ford as less-constrained. However, unlike Ford, we found very little evidence of spatial reasoners attempting to represent multiple models of the premises, and this seemed to be a major factor in their difficulty with different-form syllogisms.

Experiment 1 presented an important, and overdue, replication. The findings presented some initial evidence that the two strategic behaviours are reliable and replicable and hence possibly a robust feature of how individuals reason. However, Experiment 1 (in line with Ford, 1995) only considered syllogistic reasoning in terms of generating conclusions to valid problems. Experiment 2 was based on the premise that a further replication of the verbal and spatial strategies generally would present more credence to the proposal that they are inherent individual differences, and more so if they were also found to be employed during a conclusion evaluation task. Moreover, the nature of the verbal strategy would suggest that, when presented with invalid syllogisms, they would be likely to generate or evaluate more incorrect SF conclusions, and produce fewer NVC responses, than spatial reasoners. Further development of the reasoning behaviour questionnaire was also desirable.

Experiment 2 presented all participants with both a syllogistic conclusion generation and evaluation task, followed by a revised version of the questionnaire. They were asked to provide written protocols whilst completing both reasoning tasks, but no verbal accounts were collected. Although verbal and spatial strategies were clearly identified once again, much more variation in the nature of the protocols was observed than in Experiment 1. A significant number of participants produced protocols which they tended to describe as "flowcharts", and which showed terms linked together by arrows. In this respect they resembled several of the verbal reasoners' protocols presented by Ford and indeed, the few of these that were observed in Experiment 1 were allocated to the verbal group. However, they were far more prevalent in Experiment 2, more so than substitution. Moreover, their protocols often suggested a spatial element to their representations and some participants did answer affirmatively to spatial questionnaire items. Most however, simply answered negatively to almost all items, seemingly showing that they used neither substitution, nor spatial strategies. Not surprisingly given this, the reasoning behaviour questionnaire proved less reliable than in Experiment 1. Experiment 1 suggested that verbal reasoners were using a strategy which relies heavily on premise form, but that this was less important for spatial reasoners. As such, it was predicted that in Experiment 2, verbal reasoners would generate more invalid SF conclusions than they correct NVC responses, whilst spatial reasoners would be more successful at identifying these problems as invalid. A trend in line with this prediction was observed, however none of the differences between strategies reached statistical significance.

Overall, Experiment 2 was successful at replicating the verbal and spatial strategies observed in Experiment 1, but also suggested that that they may not be as homogeneous as Ford (1995) and Experiment 1 had suggested. Experiment 3 aimed to further investigate both inter- and intra-strategic differences as well as extending the research beyond syllogistic reasoning, to transitive inference and sentence-picture verification. Firstly, and most importantly, evidence for verbal and spatial reasoners was once again observed through verbal protocols. Some of the flowchart behaviour was again present, but it was not as prevalent as in Experiment 2 and, in this experiment, the questionnaire once more proved reliable. Moreover, the inclusion of a qualitative questionnaire item which asked participants to say more about their overall strategy revealed some very useful information about how "flowcharters" reasoned. The vast majority it appeared were reasoning propositionally, though not by substitution. Rather they applied an informal rule to link together the two end terms and form a conclusion, the form of the rule varying according to the quantifiers present. These linking rules are described in more detail in Chapter 2. These participants were therefore designated as verbal reasoners. A minority of people used a combination of verbal and spatial approaches, some involving linking rules, others substitution, together with a spatial representation of premises. These were allocated to the mixed strategy group.

A second important finding of Experiment 3 was that, for the most part, the two main strategy groups identified for syllogistic reasoning mapped onto two strategy groups identified for transitive inference. These latter groups were also identified from written protocols which suggested them to be the abstract directional and concrete properties strategies previously identified by Egan and Grimes-Farrow (1982), the former mapping onto the verbal syllogistic strategy group and the latter onto the spatial group. Only a small minority of participants did not present this correspondence of strategies. The nature of these transitive inference strategies lend further support to the claim that the two groups of reasoners are using either a abstract propositional or a visuo-spatial approach. For transitive inference they differed in terms of the degree of explicit visual representation of the relationship between terms which was required, just as they did for syllogistic reasoning. The findings regarding sentence-picture verification were less clear, and it would seem that latency data alone is insufficient to identify strategies on this task, and other supporting data (such as verbal protocol) may be required, as Marquer and Perieira (1990) have suggested. However trends in the data suggest that similar strategies cannot be wholly discounted.

Having repeatedly replicated the verbal and spatial strategies for syllogistic reasoning, Experiments 5 and 6 began an investigation into possible factors which may underpin such strategic differences. The rationale was that if strategies are an inherent feature of the individual, then they may be reflected in other aspects of their cognitive function. Experiment 5 examined verbal and spatial abilities and cognitive style. Two ability measures were used. For verbal ability, the Shipley Institute of Living Scale produces two measures, vocabulary and abstract reasoning, together with an estimation of IQ in line with that assessed by the well recognised WAIS. The spatial ability measure was the Vandenberg Mental Rotation Task, which involves the mental manipulation of a series of novel objects. Two cognitive style measures were also used, the Rational Experiential Inventory , and the Cognitive Style Analysis which measures style along two dimensions verbal-analytic style versus wholistic-imagistic which it was hoped would reflect verbal and spatial reasoners. All measures have been extensively researched previously and found reliable. Again, on a short syllogistic reasoning task, the distinctive verbal and spatial written protocols were observed indicating individual difference in strategy. Perhaps surprisingly, given the apparent nature of these strategies, no significant association between these and differences in verbal and spatial ability was apparent. The main predictors of syllogistic performance for verbal reasoners were the SILS verbal ability and estimated IQ score, with this and SILS abstract reasoning correlating significantly with performance on the VMRT. However, for spatial reasoners, performance on the VMRT was *also* a significant predictor of syllogistic reasoning, whereas it had no relationship with SILS at all. The tendency for spatial ability to predict performance for spatial reasoners, but not verbal, was to be expected, however the difference in correlation between strategies was not significant. The relationships between the cognitive style measures were strikingly similar in both verbal and spatial strategy groups, correlations tended to be very weak and no style measure predicted strategic preference.

Experiment 6 continued this line of enquiry, this time focussing on working memory. Previous work has indicated individual differences in terms of how syllogistic reasoning loads on the verbal and spatial components of working memory. Experiment 6 aimed to discover if this was the case between verbal and spatial reasoners. Six measures of working memory span were used, three each for verbal and spatial resource. Firstly, the verbal and spatial strategies were again replicated for syllogistic reasoning. However, no differences in verbal or spatial working memory capacity were observed between the groups. However, a confirmatory factor analysis was also conducted and a two correlated factor model was found to fit the data well, the factors relating to verbal and spatial working memory. It was found that whereas verbal reasoners drew primarily on the verbal resource, spatial reasoners drew significantly, and similarly, on both this and the spatial resource. Given the small sample size (just N = 48 in the spatial group) these results need to be treated with some caution, but nevertheless do present some preliminary evidence that there is a fundamental difference in how verbal and spatial strategy users draw on working memory resource in syllogistic reasoning.

Overall, the six experiments presented in this thesis have strongly suggested that individual differences in reasoning strategies do exist and have presented evidence suggesting that most individuals prefer to represent information in either a verbal-propositional or a spatial form when reasoning. Moreover, this extends beyond syllogistic reasoning suggesting that strategic preference is a domain general phenomena, not simply a function of syllogistic reasoning processes, or an artefact of verbal protocol. Moreover, the final experiment has presented some preliminary evidence to suggest an association between strategic preference and differences in how individuals draw on the verbal and spatial components of working memory. In all, this evidence suggests that strategic preferences are an intrinsic cognitive feature within individuals, and hence may influence how people reason not only in experimental situations, but in their everyday dealings with the world. The two following sections of this chapter will expand on this to discuss the implications of such differences, firstly for theories of human reasoning and then for everyday and applied situations.

6.3 Implications for Theories of Reasoning

As Chapter 1 has described, theories of reasoning have tended to assume cognitive universality, that is, that all individuals reason in basically the same way. This applies not only to general reasoning theories, such as those based on logical rules (e.g. Rips, 1994; Braine and O'Brien, 1998) and mental models (e.g. Johnson-Laird and Byrne, 1991)which have been applied across reasoning domains, including syllogisms, but also to theories which have been developed specifically to explain syllogistic reasoning. The evidence presented in thesis clearly refutes such assumptions. The key finding presented in this thesis is concerned with the identification of verbal and spatial reasoning strategies. This section offers a summary of the overall findings as they relate to the theories of reasoning described in Chapter 1, with particular reference to syllogistic reasoning which has remained a consistent theme throughout this thesis. The reader is referred to Chapter 1 for a fuller explanation of the theories discussed below.

Mental models accounts (e.g. Johnson-Laird and Byrne, 1991; Bara and Bucciarelli, 2000) suggest that individuals construct analogous mental representations of possible states of affairs suggested by syllogistic premises. Spatial reasoners can be said to construct and manipulate spatial mental models of the premises, and a few presented protocols which were uncannily like mental model notation, although they denied having knowledge of the theory (e.g. Protocol 2.24). However, the vast majority of spatial reasoners presented models that were more like Euler circle representations (more on this later). Mental models theory does not make explicit exactly what form a model takes, indeed this has been one of the ongoing criticisms levelled especially by rules theorists (e.g. Rips, 1986, 1994). Johnson-Laird (1983, page 125) describes mental models as "an internal tableau containing elements that stand for members of sets" and later descriptions add little to this definition. Certainly spatial reasoners represent sets, but they show no sign of representing the "finite sets of tokens mapped onto other finite sets of tokens" that Johnson-Laird and colleagues propose. Rather their circles represent the class or set itself, and not members of it. Ford (1995) made exactly this observation about her spatial reasoners, indeed one of the initial aims of her study was to show not only that some reasoners did not reason with models, but also that those who did use that approach, did not use the form of model proposed by the theory. However, mental models theory does make some predictions which seem to be borne out by spatial reasoners, especially with regard to the source of errors. Model theory predicts that errors occur due to the failure to flesh out an initial model, or to search for counterexamples. As the data in Chapter 2 show, very few spatial reasoners attempted to do either of these things, rather they constructed an initial model, representing premise 1 first, then adding the information from premise 2 according to their initial interpretation of the relationships involved. Hence, their errors on more difficult problems, especially different form syllogisms or those which contained *some...not* premises (which they tended to represent as *some*) arose because alternative representations were not considered.

Conversely, theories based on mental logic assume reasoning takes the form of a logical proof which uses inherent formal inference rules. A range of different rules are proposed for dealing with different quantifiers and connectives and representations are assumed to be propositional, or linguistic, in nature (e.g. Rips, 1994; Braine and O'Brien, 1998). Although some verbal reasoners in the present series of studies did employ rules of sorts (for instance to link quantifiers, see Chapter 3) these were of a superficial and heuristic nature, and certainly do not suggest an inherent understanding of logical principles. At no stage were reasoners observed to be using rules even approaching the levels of complexity suggested by Rips et al. Ford (1995) proposed that, on more complex problems, verbal reasoners employed what she termed sophisticated substitution rules, and claimed that these were similar to some of the formal logical rules advocated by Rips (1994). However, although the verbal reasoners in the present series of studies present many of the naive substitution behaviours suggested by Ford, they did not appear to employ these more sophisticated substitution rules. The verbal protocols in Experiment 1 present very clear indication of what verbal reasoners did, namely, obtain a value for the common/B term from the universal premise and then simply substitute this for B in the other premise to reach a conclusion (Protocol 2.1 is an excellent example). In support of her claim that verbal reasoners were able to use more formal reasoning, Ford cites evidence from verbal protocols, or states, simply, that they were able to generate conclusions to more difficult problems. However, in the present studies, although those verbal reasoners who were more successful were able to recognise that naive substitution would not give the correct answer, they simply modified this approach to suit the problem (as in Protocol 2.14), rather than adopted more sophisticated rules. However, mental logic theories do acknowledge that some reasoners use propositional representations, as opposed to spatial mental models. Moreover, the version offered by Braine and O'Brien (e.g. 1998) also acknowledges the role of natural language understanding of quantifiers. Almost all reasoners here tended to make natural language based assumptions (such as assuming *all* to mean literally all, and not *some*, as formal logic would claim).

However, Polk and Newell (1995) claim that the use of both mental models and logical rules are a form of transduction paradigm, processes designed to transform presented stimuli into an internal representation upon which specialised reasoning mechanisms can then operate. The results of such processes are then transduced back into a form suitable for presentation as a conclusion. Polk and Newell's verbal reasoning hypothesis (VRH) describes this process. The useful aspect here, given the nature of the strategies observed, is that the VRH allows for individual differences in representations, they may be either spatial/model based or propositional/language based. However, from the evidence presented in this thesis, it is not possible to draw any firm conclusions about the degree to which the strategies observed reflect (in Polk and Newell's terms) reasoning or transduction processes. The spatial strategy remains underspecified especially in terms of how the conclusion is actually drawn once the representation has been constructed. Both verbal and written protocols were inconclusive on this point, indeed, when questioned informally, spatial reasoners were unable to articulate exactly how they had arrived at their conclusion, whether they had simply read it off the diagram, or applied some form of rule. Hence, it may be that the construction of such spatial representations is a form of transduction, with further, as yet unspecified, reasoning processes then taking place. However, for verbal reasoners, there is very little evidence that they formed an internal representation upon which they then operated. Rather, their substitution, or linking rules, allow them to reason heuristically with little need to pay more detailed attention to the premise content. However, Polk and Newell do state that, because in the VRH reasoning processes are based in knowledge from natural language, no mental logic is required, even for individuals who represent information propositionally.

One theory which does seem to offer some explanation, for the spatial strategy at least, is that of Euler circles (e.g. Erickson, 1974). These reasoners present diagrammatic representation of premises alike to those shown in Table 1.8, although they tended not to attempt more than one possible representation of a premise. Even those few individuals who considered alternative representations only did so in terms of the combination of two premises, not in terms of single premises. Hence, for an *All* premise, either an identity or a subset-set relationship would be used (see also Experiment 1, Chapter 2 for further discussion of this point). A *some* premise was always expressed as an overlap, as was *some...not*, many reasoners assuming that one implied the other. No reasoners attempted to use the alternative set-subset representations for *some...not*. However, contrary to Euler circles theory, rather than constructing two models, one for each premise, and then combining them, spatial reasoners constructed a single model, initially from premise one, then augmented with information from premise two.

Conversely, Stenning and Oberlander (1995) have suggested that only one Euler circle type representation per premise is actually required, and that such an approach reduces abstraction and allows individuals to examine relationships between terms more simply. This idea certainly seems to fit with the approach of the spatial reasoners, who frequently claimed that their diagrams helped them to visualise the problem. Their equivalent representations for transitive inference further support the idea that these people dislike abstraction and require an explicit representation of the situation in order to reason. Hence they need to represent explicit relationships between sets in syllogistic reasoning, and between relative properties in transitive inference.

However, although both of the above implementations of Euler circles offer some explanation regarding the behaviour of spatial reasoners, neither can offer an account of verbal syllogistic reasoning. However, Sternberg and Guyote's (1981) transitive chain theory, although based on Euler circles, exhibits no resemblance to the processes which seem to be used by spatial reasoners. Rather, it seems applicable to some of the processes observed in verbal reasoners' protocols. Several reasoners described a process of cancelling out the common/B term and linking the other terms together, in effect forming a transitive chain. However, while Sternberg and Guyote assume that reasoners are using some form of set based representation, there was no evidence that the vast majority of such individuals were reasoning anything other than propositionally. A small minority of individuals did claim to represent terms within an spatial layout before any cancellation or linking rules were applied. These were classed as mixed reasoners, but the numbers were too small to draw any firm conclusions about them (see discussion in Chapter 3).

A very different explanation of syllogistic reasoning was offered by Geurts (2003) monotonicity theory. Geurts does offer an explanation of substitution behaviour in syllogistic reasoning, and this seems to predict the behaviour of verbal reasoners in terms of the conclusions they produce. However, whether verbal reasoning is based on an implicit knowledge of monotonicity is difficult to say. Certainly, according to their verbal and written protocols, their reasoning appears rather less sophisticated than the approach proposed by Geurts. Monotonicity theory is a very recent addition to the literature, and, although it presents some interesting and novel concepts, remains relatively under developed. Furthermore, it cannot account at all for spatial reasoners behaviour.

A number of heuristic accounts of syllogistic reasoning have also been proposed, and three of these (atmosphere, matching and the probability heuristic model) will be considered here. All make similar predictions regarding the type of conclusion reasoners are likely to generate, although they attribute them to differing processes. The atmosphere theory (e.g. Woodworth and Sells, 1936) claims that the mood of a conclusion arises from the atmosphere created by the mood of the premises. For instance, a syllogism with the quantifiers *all; none*, will create a negative atmosphere leading to a negative (*None*) conclusion. Matching theory (e.g. Wetherick and Gilhooly, 1990; 1995) suggests that reasoners will select the most conservative premise quantifier and match their conclusion to that mood. Hence for the same *all; none* syllogism, again a *none* conclusion would be predicted. Both these approaches suggest that reasoning is carried out in a non-logical heuristic fashion and both clearly predict conclusions typically presented by verbal reasoners, although they generally arrive at them by performing substitution. Interestingly however in view of the individual differences in strategies observed here, Wetherick and Gilhooly have pointed out that although some reasoners use matching, many do not. Some have an awareness of more logical principles whilst other may use different processes altogether. They do not make suggestions as to what these alternatives may be, but do acknowledge that individual differences in strategy are likely. Verbal reasoners in these six experiments do not seem to be matching in quite the way described by Wetherick and Gilhooly, but their approach has some similarities. Their end quantifier does tend to be the most conservative and linking rules bear comparison to the idea that certain quantifier combinations necessarily (at least to those reasoners) suggest which should be chosen for the conclusion.

Similar predictions are also made by the probability heuristics model (PHM, Chater and Oaksford, 1999) although in this case, heuristics are used to decide a conclusion based on the relative informativeness of the premises, as suggested by probability. Hence, to return to the above example, for an *all; none* syllogism, the quantifier *none* will be selected for the conclusion. Again, although the types of conclusions drawn by verbal reasoners can be accounted for, neither the PHM nor matching can fully explain the process of verbal reasoning by substitution. There is little evidence that relative conservativeness (or informativeness) is a factor in the choice of end-quantifier for those people. However, for some other verbal reasoners (such as those using linking rule approach) some participants did describe thinking of some quantifiers as "stronger" than others, with *all* being the strongest (see for instance Protocol 2.22) which may be an indicator that some kind of quantifier evaluation process was occurring for those people. However, this was mentioned by only a few participants. Also, although spatial reasoners often produce similar conclusions to verbal (especially to same-form problems), their strategy is simply not accounted for in the above heuristic theories.

At this point, it is pertinent to briefly mention theories of transitive inference, a task which was employed in both Experiments 3 and 4, and for which strategies were found to map closely onto those for syllogistic reasoning. For this task, neither the abstract directional not the concrete properties strategy can be fully accounted for by any one theory. Only a small proportion of participants used a purely propositional strategy (Clark, 1969), and those who did so, used the same approach for all problems, irrespective of the relational adjectives present, contrary to the theory. Moreover although the AD reasoners certainly produced a spatial array (as per DeSoto et al, 1965; Huttenlocher, 1968) almost all produced a horizontal array to all problems, even those with adjectives such as "taller" which the theories suggest tend to prompt a vertical array. The few reasoners who presented vertical arrays did so as a matter of course, again to all problems. The same can be said for concrete properties reasoners who, although showing explicit representations of relational properties, still presented the terms in a directional array, either vertically or (most commonly) horizontally. Hence the tendency towards using one or other form appears to be a function of individual preference, rather than problem type. The findings of Experiments 3 and 4 and their implications for mainstream theories of transitive inference are discussed in more detail in Chapter 3.

Overall, it is apparent that although some of the theories discussed above can account for some aspects of reasoning behaviour, in one or other strategy, none of them can account, even partially, for *both* verbal and spatial reasoning strategies. This is a clear reflection of the dichotomous nature of reasoning theories, as discussed previously in Section 1.7. Consequently, the findings presented in this thesis have some important implications for theories of reasoning as they stand at present. The data present evidence of individual differences in reasoning strategies, but indicate few distinctions in perfomance as a consequence of these. Verbal and spatial reasoners do not differ substantially in their overall patterns of accuracy, even though their errors arise for different reasons. Hence traditional behavioural measures, based on accuracy and types of errors (on which many theories are based), may not tap into these strategies. A number of accounts claim to have successfully modelled syllogistic reasoning data from differing theoretical perspectives (e.g. Bucciarelli and Johnson-Laird, 1999; Chater and Oaksford, 1999; Polk and Newell, 1995; Rips, 1994) but these models may not capture the subtleties which underlie the production of that data. Any comprehensive theoretical account of human reasoning needs to present an explanation of the differing nature of strategies and what underlies their development in logically untrained individuals.

6.4 Directions for Future Research

Although the observation of verbal and spatial strategies seems to be a fairly robust phenomena (replicated six times across these experiments) a number of issues remain unresolved. Firstly, the spatial strategy especially remains underspecified. Further work needs to investigate particularly how such reasoners draw their conclusions once a spatial representation been constructed. Moreover, further evidence is needed to clarify the extent to which the strategies really are verbal and spatial in nature.

The findings presented in this thesis suggest that spatial reasoners require a more explicit visual image in order to reason. Hence, it could be predicted that these people will have more difficulty with material which does not lend itself to such representation. Comparisons between reasoning performance on material presented in visual and verbal modalities, as a function of strategy, could prove informative. Using problems which conflict with belief may also support an inter-strategic distinction, with spatial reasoners more prone to belief bias as results of their greater semantic involvement with problem content. It is also still unclear whether other individual differences in cognitive function underlie strategic differences. In Experiment 6, confirmatory factor analysis has suggested some differences in how the strategies draw on the verbal and spatial components of working memory. However, these were preliminary findings only and require further investigation with a larger sample.

Another outstanding issue concerns how and why logically untrained individuals develop different strategies for reasoning - what Johnson-Laird, Savary and Bucciarelli (2000, page 238) have called "the Holy Grail" of reasoning research. Given that some strong evidence for strategic distinctions has been observed, an investigation into whether the verbal-spatial preferences are reflected in the selective remembering and subsequent application of learned material is suggested. Training studies, possibly incorporating a test-retest methodology with control for inherent strategic preference, may indicate effects on reasoning performance as a function of whether the strategy participants are made to use, after training, matches their natural preference.

Further to the above, and in pursuit of the same aim, developmental research might investigate the stage at which strategies develop during cognitive maturation and the extent to which any inherent approach to reasoning reflects educational attainment. An interaction between verbal-spatial strategies and learning has implications for teaching methods, for instance, number lines are often used to teach young children mathematics and these apply a spatial dimension to the solving of abstract numerical problems. Reasoning is integral to the education process (e.g. Perkins, Farady and Bushey, 1991) and if individuals do indeed represent and manipulate information in different ways, as the present research suggests, then this has obvious implications for learning and problem solving in the real world, both in formal educational settings, and as part of the general lifelong learning process.

Although a fair amount of research has already been conducted into individual differences in thinking process during learning, much of this work has been concerned primarily with differences in cognitive styles and/or associated learning styles. For instance, the wealth of studies discussed by Riding and Rayner (1998, Chapter 7), have all used learning style measures evolved from the Cognitive Style Analysis used in Experiment 5. Here, no association was found between cognitive style , as measured by the CSA, and reasoning strategy. Learning styles were not examined in Experiment 5, but these have often been developed in line with associated theories of cognitive style (see for instance Riding and Rayner, 1998, who present a comprehensive summary of both their own (CSA based) learning style model and others). Given that Experiment 5 found absolutely no effects of cognitive style on strategy use, it may also be the case that inter-strategic differences in learning style would equally not be observed. Similarly, much research examining individual differences in factors such as motivation for learning (e.g. the work of Klaczynski and colleagues) has also used thinking style measures such as the REI, which again was used in Experiment 5 and found not be associated with reasoning strategy. Perkins, Jay and Tishman (1993) have argued that such cognitive styles have implications not only for reasoning generally, but also for issues of culture and education. However, given the findings of Experiment 5, for the present purpose, alternative forms of investigation, which examine differences in thinking and reasoning strategy on educational tasks such as mathematics, are advocated, rather than reliance on psychometric style measures.

Finally, what about other forms of everyday reasoning? If individual differences in strategies for reasoning are observable under experimental conditions, it is reasonable to suppose that they may also be present in other settings. This is even more likely given the present findings which strongly suggest that such strategies are an intrinsic feature of how individuals reason, rather than an experimental artefact. Galotti (1989) has stated that the justification for studying syllogistic reasoning must be grounded in its relationship to everyday reasoning, and that the results need to extend to reasoning which is typical of that in ordinary life. Although everyday reasoning seems to requires a rather different form of "logic" to formal tasks (see for instance Galotti, 1989, page 335 for a comparison), Perkins et al (1991) have highlighted how the two forms of reasoning may challenge individuals in very similar ways, errors often arising from incomplete consideration of alternatives and belief bias. The lack of formal structure and containment in many everyday reasoning situations may mean that verbal and spatial strategies per se are inappropriate. However, if individual differences in approach were observed, strategy classifications may prove to map onto those for syllogistic reasoning, just as they did with transitive inference strategies in Experiments 3 and 4.

Overall, the above suggestions would allow for the line of research presented in this thesis to be extended, and provide further supporting evidence for the strategic distinctions highlighted in Experiments 1 – 6. Although the terms verbal and spatial have been used to describe the strategies throughout this thesis (after Ford, 1995, and in line with the propositional and spatial representations they produce in protocols) further evidence is required before a definitive statement regarding their nature can be made. Moreover, a greater understanding of strategic acquisition and development, together with associated effects on learning processes, can be translated into more effective teaching methods and the acquisition of new and flexible skills to meet the changing needs of society.

6.5 Concluding Comments

To recap, the overall purpose of this series of experiments was to establish evidence for the existence of individual differences in reasoning strategies. The programme aimed to investigate the nature of such differences, whether they generalised across differing reasoning domains, and some of the underlying cognitive factors which may give rise to them. From an accumulation of evidence across the six experiments, four particularly important findings have emerged:

1. Experiments 1-6 have presented some strong and consistent evidence for individual differences in strategies for human reasoning. Verbal and written protocols produced during Experiment 1 showed clear differences in how individuals chose to represent syllogistic premise information, and have suggested of some key qualitative differences between two main strategies. Those individuals that have been termed verbal reasoners seem to reason propositionally, manipulating the written form of the premises in order to arrive at a conclusion. Conversely, the spatial reasoners prefer to represent premises in the form of sets and subsets, generating a diagram to show the relationships between terms spatially. This finding was replicated a further *five* times during the course of this programme of study, and without the need for verbal reports. Moreover, a questionnaire has been developed which has proven to be a reliable indicator of the presence of these strategic differences.

2. No significant differences are observed between reasoning performance of verbal and spatial reasoners. Not only do they tend to perform similarly overall, they present similar patterns of errors, both performing least well on problems where the conclusion is of different form to that of either premise. However, such errors occur for different reasons. The verbal reasoning strategy (be it substitution or linking of quantifiers) uses premise form quite literally in generating a conclusion where the quantifier tends to be the same as that for one or other premise. Spatial reasoners on the other hand, only tend to consider one possible representation of the premises and their failure to incorporate alternatives into their representations can lead to the correct conclusion being overlooked. Hence individuals can present an outward appearance if universality, despite using very different reasoning strategies. These are important findings given the ubiquity assumed by many theories of reasoning.

3. Strategic preferences have been shown to extend beyond syllogistic reasoning. Two strategies have also been identified for transitive inference, an abstract directional and a concrete properties strategy. These differ mainly in the degree to which the relative qualities of the terms are represented, the former being based on a simple abstract scale and the latter on a more complex and explicit representation of the qualitative nature of the terms. Individuals using verbal and spatial syllogistic strategies also tended to be those who adopted an abstract or a concrete transitive inference strategy respectively. This clearly suggests that individual differences in strategies may be an intrinsic cognitive factor, one which remains consistent between individuals across reasoning domains.

4. Verbal and spatial strategies draw differentially on working memory resources. Confirmatory factor analysis has indicated that whilst verbal reasoners draw on verbal working memory resources during syllogistic reasoning, spatial reasoners draw on both verbal and spatial resources. This further supports the idea of some form of basic cognitive distinction between the two groups of individuals.

Overall, and in conclusion, the experiments described herein seem to have fulfilled their overall aims. Having demonstrated consistent differences in syllogistic reasoning strategies, they also present some preliminary indications that the preferences are domain-general and that they may be underpinned by fundamental cognitive differences. These findings have potentially important and far reaching implications for the theories of human reasoning which assume strategic universality. It is suggested that future research concentrate on consolidating these findings with regard to the distinctive qualitative natures of strategies and to establishing their cognitive antecedents. With these features firmly documented, comes the task of getting the concept of individual differences in strategies for reasoning acknowledged in the wider academic community. As fresh insights and data come to the fore, theoretical accounts are constantly being updated, and new models proposed. These cannot hope to offer a comprehensive explanation of human reasoning behaviour unless the possibility of individual differences is taken into account.

References

Armstrong, S.J. (2000). Individual differences in cognitive style and their potential effects on organisational behaviour: a summary of recent empirical studies. In R. Riding & S. Rayner (Eds.). *International Perspectives on Individual Differences: Volume 1, Cognitive Styles* (pp. 215-237). Stamford: Ablex.

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

Baddeley, A.D. (1997). *Human Memory: Theory and Practice* (Revised Ed.) Hove: Psychology Press.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multi-component model. In A. Miyake, & P. Shah (Eds.), *Models Of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 28-61). Cambridge: Cambridge University Press.

Bara, B.G. & Bucciarelli, M. (2000). Deduction and induction: reasoning through mental models. *Mind & Society, **1(1)***, 95-107.

Bara, B.G., Bucciarelli, M. and Johnson-Laird, P.N. (1995). Development of syllogistic reasoning. *American Journal of Psychology, **108(1)***, 157-193.

Baron, J. (1988). *Thinking and Deciding*. Cambridge: Cambridge University Press.

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

Bentler, P.M. (1995). *EQS: Structured Equation Program Manual*. Encino, CA.: Multivariate Software.

Borg, M.G. & Riding, R. (1993). Teacher stress and cognitive style. *British Journal of Educational Psychology, **63(1)***, 2.

Bradon, P., Capon, A., Dennis, I., Evans, J. St.B.T., Handley, S. & Newstead, S.E. (2002). *Difficulty modelling of analytical reasoning problems*. Unpublished report delivered to Educational Testing Services, Princeton, USA.

Braine, M.D. (1990). The "natural logic" approach to reasoning. In W.F. Overton (Ed.). *Reasoning, Necessity and Logic: Developmental Perspectives* (pp. 135-157). Hillsdale, NJ: Lawrence Erlbaum Associates.

Braine, M.D. (1993). Mental models cannot exclude mental logic, and make little sense without it. *The Behavioural and Brain Sciences, 16.*

Braine, M.D., Reiser, B.J. & Rumain, B. (1984). Some empirical justification for a theory of mental propositional logic. In G. Bower (Ed.). *The Psychology of Learning and Motivation: Advances in Research and Theory, Volume 18* (pp. jhg). New York: Academic Press.

Braine, M.D. and O'Brien, D.P. (1998). *Mental Logic.* Mahwah, NJ.: Lawrence Erlbaum Associates.

Bucciarelli, M. and Johnson-Laird, P.N. (1999). Strategies in syllogistic reasoning. *Cognitive Science, 23(3)*, 247-303.

Byrne, R.M. and Johnson-Laird, P.N. (1989). Spatial reasoning. *Journal of Memory and Language, **28***, 564-575.

Cacioppo, J.T. & Petty, R.E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116-131.

Capon, A., Handley, S.J., & Dennis, I. (2003). Working memory and reasoning: An individual differences perspective. *Thinking and Reasoning, **9(3)***, 203-244.

Carpenter, P.A. & Just, M.A. (1975). Sentence comprehension: a psycholinguistic processing model of verification. *Psychological Review, **82(1)***, 45-73.

Carroll, J.B. (1983). Studying individual differences in cognitive abilities: through and beyond factor analysis. In R.F. Dillon & R.R. Schmeck (Eds.). *Individual Differences in Cognition, Volume 1* (pp. 1-33.). New York Academic Press.

Carroll, J.B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. New York: Cambridge University Press.

Cary, M. & Reder, L.M. (2003). Metacognition in strategy selection: giving consciousness too much credit. In Izaute, M., Chambres, P. & Merescaux, P.J. (Eds.). *Metacognition: Process, Function and Use* (pp. 63-77). New York: Kluwen Press.

Chater, N. and Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology, **38***, 191-258.

Cheng, P.W. & Holyoak, K.J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, **17(2)***, 285-313.

Cherubini, P., Garnham, A., Oakhill, J. and Morley, E. (1998). Can any ostrich fly? some new data on belief bias in syllogistic reasoning. *Cognition, **69(2)***, 179-218.

Clark, H.H. (1969). Linguistic processes in deductive reasoning. *Psychological Review, 76*, 387-404.

Clark, H.H. & Chase, W.G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, **3(3)***, 472-517.

Cooper, L.A. and Mumaw, R.J. (1985). Spatial aptitude. In R.F. Dillon (Ed.). *Individual Differences in Cognition, Volume 2* (pp. 67-94). New York: Academic Press.

Copi, I.M. & Cohen, C. (1998). *Introduction to Logic* (10th ed.). Prentice-Hall.

Crowley, K. & Siegler, R.S. (1999). Explanation and generalisation in young children's strategy learning. *Child Development, **70(2)***, 304-316.

Crowley, K., Shrager, J. & Siegler, R.S. (1997). Strategy discovery as a competitive negotiation between metacognitive and associative mechanisms. *Developmental Review, **17 (2)***, 462-489.

Daneman, M. & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal learning and Verbal Behaviour, **19(3)***, 450-466.

Daneman, M. & Merikle, P.M. (1996). Working memory and language comprehension: a meta-analysis. *Psychonomic Bulletin & Review, **3(3)***, 422-433.

Dickstein, L.S. (1978). The effect of figure on syllogistic reasoning. *Memory and Cognition, **6***, 76-83.

Donaldson, M. (1978). *Children's Minds*. London: Penguin.

Egan, D.E. (1983). Retrospective reports reveal differences in people's reasoning. *Bell Systems Technical Journal, 62(6)*, 1675-1697.

Egan, D.E. and Grimes-Farrow, D.D. (1982). Differences in mental representations spontaneously adopted for reasoning. *Memory and Cognition, 10(4)*, 297-307.

Engle, R.W., Kane, M.J. & Tuholski, S.W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In A. Miyake, & P. Shah (Eds.), *Models Of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 102-134). Cambridge: Cambridge University Press.

Entwistle, N.J. (1979). Stages, levels, styles or strategies: dilemmas in the description of thinking. *Educational Review, 31(1)*, 123-132.

Epstein, S. (1994). An integration of the cognitive and psychodynamic unconscious. *American Psychologist, 49(5)*, 709-724.

Epstein, S., Pacini, R., Denes-Raj, V. & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology, 71(2)*, 390-405.

Erickson, J.R. (1978). Research on syllogistic reasoning. In R. Revlin and R.E. Mayer (Eds.). *Human Reasoning* (pp. 39-50). New York: Wiley.

Ericsson, E.A. and Simon, H.A. (1980). Verbal reports as data. *Psychological Review, 87(3)*, 215-251.

Ericsson, E.A. and Simon, H.A. (1993). *Protocol Analysis: Verbal Reports as Data* (Rev. ed.). Cambridge, Mass: MIT Press.

Evans, J. St. B. T. (1982). *The Psychology of Deductive Reasoning*. London: Routledge.

Evans, J. St.B. T. (1989). *Bias in Human Reasoning: Causes and Consequences*. Hove: Lawrence Erlbaum Associates.

Evans, J. St. B. T. (2000). What could and could not be a strategy in reasoning. In W. Schaeken, G. De Vooght, A. Vandierendonck and G. d'Ydewalle. (Eds.). *Deductive Reasoning and Strategies* (pp. 1-22). Mahwah, NJ.: Lawrence Erlbaum Associates.

Evans, J.St.B.T. (2002). Logic and human reasoning: an assessment of the deduction paradigm. *Psychological Bulletin, **128(6)***, 978-996.

Evans, J. St.B. T., Newstead, S.E. and Byrne, R.M.J. (1993). *Human Reasoning: The Psychology of Deduction*. Hove: Lawrence Erlbaum Associates.

Evans, J. St.B. T. & Over, D.E. (1996). *Rationality and Reasoning*. Hove: Psychology Press.

Evans, J.St.B.T., Handley, S.J., Harper, C.N. & Johnson-Laird, P.N. (1999). Reasoning about necessity and possibility: a test of the mental model theory of deduction. *Journal of Experimental Psychology: Learning Memory & Cognition, **25(6)***, 1495-1513.

Evans, J.St.B.T., Handley, S.J., & Over, D.E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory and Cognition, **29(2)***, 321-335.

Evans, J.St.B.T., Over, D.E. & Handley, S.J. (in press). Rethinking the model theory of conditionals. In W. Schaeken, A. Vandierendonck, W. Schroyens & G. d' Ydewalle (Eds.). *The Mental Models Theory of Reasoning. Refinements and Extensions*. Hove: Lawrence Erlbaum.

Ford, M. (1995). Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition, **54(1)***, 1-71.

Furlong, P.R. (1993). Personal factors influencing informal reasoning of economic issues and the effect of specific instructions. *Journal of Educational Psychology, **85(1)***, 171-181.

Galotti, K.M. (1989). Approaches to studying formal and everyday reasoning. *Psychological Bulletin, **105(3)***, 331-351.

Garnham, A. and Oakhill, J. (1994). *Thinking and Reasoning*. Oxford: Blackwell.

Gattis, M. and Dupeyrat, C. (2000). Spatial strategies in reasoning. In W. Schaeken, G. De Vooght, A. Vandierendonck and G. d'Ydewalle. (Eds.). *Deductive Reasoning and Strategies* (pp. 153-175). Mahwah, NJ.: Lawrence Erlbaum Associates.

Geurts, B. (2003). Reasoning with quantifiers. *Cognition, **86(3)***, 223-251.

Gilhooly, K.J. (1996). *Thinking*. (3rd ed.). London: Academic Press.

Gilhooly, K. J. (1998). Working memory and reasoning. In R. H. Logie, & K. J. Gilhooly (Eds.). *Working Memory and Thinking* (pp. 7-22). Hove: Psychology Press.

Gilhooly, K.J. & Logie, R.H. (1998). Thinking in working memory. In R. H. Logie, & K. J. Gilhooly (Eds.). *Working Memory and Thinking* (pp. 1-5). Hove: Psychology Press.

Gitomer, D.H. (1988). Individual differences in technical troubleshooting. *Human Performance, **1(2)***, 111-131.

Gustafsson, J.E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence **8(1)***, 189-203.

Guyote, M.J. & Sternberg, R.J. (1981). A transitive chain theory of syllogistic reasoning. *Cognitive Psychology*, *13 (4)*, 461-525.

Hagert, G. (1985). Modelling mental models: experiments in cognitive modelling of spatial reasoning. In T. O'Shea (Ed.). *Advances in Artificial Intelligence: Proceedings of the Sixth European Conference on Artificial Intelligence* (389-398). Amsterdam: North Holland.

Handley, S.J., Capon, A., Copp, C. and Harper, C. (2002). Conditional reasoning and the Tower of Hanoi: the role of spatial and verbal working memory. *British Journal of Psychology, 93*, 501-518.

Handley, S.J., Dennis, I., Evans, J. St. B.T. & Capon, A.J. (2000). Individual differences in the search for counterexamples in syllogistic reasoning. In W. Schaeken, G. De Vooght, A. Vandierendonck and G. d'Ydewalle. (Eds.). *Deductive Reasoning and Strategies* (pp. 241-265). Mahwah, NJ.: Lawrence Erlbaum Associates.

Handley, S.J., Newstead, S.E. & Wright, H. (2000). Rational and experiential thinking: a study of the REI. In R. Riding & S. Rayner (Eds.). *International Perspectives on Individual Differences: Volume 1, Cognitive Styles* (pp. 97-113). Stamford: Ablex.

Hartigan, J.A. and Wong, M.A. (1979). A K-means clustering algorithm: algorithm AS136. *Applied Statistics, 28(1)*, 126-130.

Horn, J.L. & Cattell, R.B. (1967). Age differences in fluid and crystallised intelligence. *Acta Psychologica, **26(1)***, 107-129.

Hunt, E., Lunneborg, C. & Lewis, J. (1975). What does it mean to be high verbal? *Cognitive Psychology, **7(1)***, 194-227.

Huttenlocher, J. (1968). Constructing spatial images. *Psychological Review, **75(4)***, 550-560.

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

Johnson-Laird, P.N. (1989). Mental models. In M.I. Posner, (Ed.). *Foundations of Cognitive Science*. Cambridge, Mass: MIT Press.

Johnson-Laird (1995). Inference and mental models. In S.E. Newstead & J. St. B. T. Evans (Eds.). Perspectives on Thinking and Reasoning: Essays in Honour of Peter Wason (pp. 115-146). Hove: Lawrence Erlbaum Associates.

Johnson-Laird, P.N. (1998). Imagery, visualisation and thinking. In J. Hochberg (Ed.). *Perception and Cognition at Century's End* (pp. 441-467). San Diego: Academic Press.

Johnson-Laird, P.N. (2001). Mental models and deduction. *Trends in Cognitive Sciences, **5(10)***, 434-442.

Johnson-Laird, P.N. and Steedman, M.J. (1978). The psychology of syllogisms. *Cognitive Psychology, **10(1)***, 64-99.

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

Johnson-Laird, P.N. and Byrne, R.M.J. (1991). *Deduction*. Hove: Lawrence Erlbaum Associates.

Johnson-Laird, P. N. & Byrne, R.M. (2000). Mental models and pragmatics. *Behavioral and Brain Sciences, **23(2)***, 284-286.

Johnson-Laird, P.N., Savary, F. and Bucciarelli, M. (2000). Strategies and tactics in reasoning. In W. Schaeken, G. De Vooght, A. Vandierendonck and G. d'Ydewalle. (Eds.). *Deductive Reasoning and Strategies* (pp. 209-240). Mahwah, NJ.: Lawrence Erlbaum Associates.

Johnson-Laird, P. N. & Byrne, R.M. (2002). Conditionals: a theory of meaning, pragmatics and inference. *Psychological Review, **109(4)***, 646-678.

Jonides, J. & Smith, E. E. (1997). The architecture of working memory. In M.D. Rugg, (Ed.). *Cognitive Neuroscience* (pp. 243-276). Hove: Psychology Press.

Just, M.A. & Carpenter, P.A. (1985). Cognitive co-ordinate systems: accounts of mental rotation and individual differences in spatial ability. *Psychological Review, **92(1)***, 137-172.

Kagan, J. (1966). Developmental studies in reflection and analysis. In A.H. Kidd & J.L. Rivoire (Eds.). *Perceptual Development in Children*. New York: International University Press.

Kirton, M.J. (1994). *Adaptors and Innovators* (2nd. ed.). London: Routledge.

Klaczynski, P.A. (1997). Bias in adolescent's everyday reasoning and its relationship with intellectual ability, personal theories and self-serving motivation. *Developmental Psychology, 33(2), 273-283.*

Klaczynski, P.A., Gordon, D.H. and Fauth, J. (1997). Goal-oriented critical reasoning and individual differences in critical reasoning biases. *Journal of Educational Psychology, 89(3)*, 470-485.

Klaczynski, P.A., Fauth, J. and Swanger, A. (1998). Adolescent identity: Rational vs. experiential processing, formal operations, and critical thinking beliefs. *Journal of Youth and Adolescence, 27(2)*, 185-207.

Klauer, K.C., Stegmaier, R. and Meiser, T. (1997). Working memory involvement in propositional and spatial reasoning. *Thinking & Reasoning, 3 (1),* 9-47.

Kline, P. (1991). *Intelligence: The Psychometric View.* London: Routledge.

Kucera, H. & Francis, W.N. (1967). *Computational Analysis of Present-Day American English.* Priovidence, RhI: Browns University Press.

Kyllonen, P.C. (1996). Is working memory capacity Spearman's g? In I. Dennis & P. Tapsfield (Eds.). Human abilities: Their nature and measurement (pp. 49-75). Mahwah, NJ: Lawrence Erlbaum Associates.

Kyllonen, P.C. & Christal, R.E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14(3)*, 389-433.

Kyllonen, P.C. & Stephens, D.L. (1990). Cognitive abilities as determinants of success in acquiring logic skill. *Learning and Individual Differences, 2(1)*, 129-160.

Larkin, J.H. & Simon, H.A. (1987). Why a diagram is (sometimes) worth more than ten thousand words. *Cognitive Science 11(1)*, 65-99.

Logie, R.H. (1995). *Visuo-spatial Working Memory.* Hove: Lawrence Erlbaum Associates.

Lohman, D.F. and Kyllonen, P.C. (1983). Individual differences in solution strategy on spatial tasks. In R.F. Dillon & R.R. Schmeck (Eds.). *Individual Differences in Cognition, Volume 1* (pp. 105-135). New York Academic Press.

MacLeod, C.M., Hunt, E.B. and Mathews, N.N. (1978). Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning and Verbal Behaviour.*, ***17***, 493-507.

Manktelow, K. (1999). *Reasoning and Thinking*. Hove: Psychology Press.

Markovits, H. and Nantel, G. (1989). The belief bias effect in the production and evaluation of logical conclusions. *Memory and Cognition*, ***17(1)***, 11-17.

Marquer, J. (1990). Reaction times and verbal reports in the study of cognitive strategies: a reply to Evans. *Quarterly Journal of Experimental Psychology,* ***42A(1)***, 171-172.

Marquer, J. & Pereira, M. (1990). Reaction times in the study of strategies in sentence-picture verification: a reconsideration. *Quarterly Journal of Experimental Psychology,* ***42A(1)***, 147-168.

Martinsen, Ø. & Kaufmann, G. (2000). The assimilator-explorer cognitive styles and their relationship to affective-motivational orientations and cognitive performance. In R. Riding & S. Rayner (Eds.). *International Perspectives on Individual Differences: Volume 1, Cognitive Styles* (pp. 3-39). Stamford: Ablex.

Matarazzo, J.D. (1992). Psychological testing and assessment in the 21st century. *American Psychologist,* ***47 (5)***, 1007-1018.

Maybery, M.T., Bain, J.D. & Halford, G.S. (1986). Information processing demands of transitive inference. *Journal of Experimental Psychology: Learning, Memory & Cognition,* ***12(4)***, 600-613.

Newstead, S.E. (in press). Can natural language semantics explain syllogistic reasoning? *Cognition.*

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

Newstead, S.E., Handley, S.J. & Buck, E. (1999). Falsifying mental models: testing the predictions of theories of syllogistic reasoning. *Memory and Cognition, 27(2)*, 344-354.

Newstead, S.E., Handley, S.J., Harley, C., Wright, H. and Farrelly, D. (in press). Individual differences in deductive reasoning. *Quarterly Journal of Psychology.*

Newton, E.J. & Roberts, M.J. (2000). An experimental study of strategy development. *Memory and Cognition, **28(4)***, 565-573.

Oakhill, J.V., Yuill, N. & Parkin, A.J. (1986). On the nature of the difference between skilled and less-skilled comprehenders. *Journal of Research in Reading, **9(1)***, 80-91.

Oakhill, J.V., Johnson-Laird, P.N. and Garnham A. (1989) Believability and syllogistic reasoning. *Cognition, **31(2)**, 117-140.*

Oaksford, M. & Chater, N. (1995). Theories of reasoning and the computational explanation of everyday intelligence. *Thinking and Reasoning, **1(2)***, 121-152.

O'Brien, D.P. (1995). Finding logic in human reasoning requires looking in the right places. In S.E. Newstead & J. St. B.T. Evans (Eds.). *Perspectives on Thinking and Reasoning: Essays in Honour of Peter Wason* (pp. 189-216). Hove: Lawrence Erlbaum Associates.

Pacini, R. & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs and the ratio-bias phenomenon. *Journal of Personality and Social Psychology, **76(6)***, 972-987.

Paivio, A. (1971). Styles and strategies of learning. *British Journal of Educational Psychology, **46(1)***, 146-148.

Pellegrino, J.W. & Goldman, S.R. (1983). Developmental and individual differences in verbal and spatial reasoning. In R.F. Dillon & R.R. Schmeck (Eds.). *Individual Differences in Cognition, Volume 1* (pp. 137-180). New York Academic Press.

Perkins, D.N. (1985). Post-primary education has little impact on informal reasoning. *Journal of Educational Psychology, **37(5)***, 562-571.

Perkins, D.N., Farady, M. and Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In J.F. Voss, D.N. Perkins and J.W. Segal. (Eds.). *Informal Reasoning and Education* (pp. 83-105). Hillsdale, NJ.: Lawrence Erlbaum Associates.

Perkins, D.N., Jay, E. & Tishman, S. (1993). Beyond abilities: towards a dispositional theory of thinking. *Merrill-Palmer Quarterly, **39(1)***, 1-21.

Polk, T.A. and Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review, **102(3)***, 533-566.

Quinlan, P.T. (1992). *The Oxford Psycholinguistic Database*. Oxford: Oxford University Press.

Quinton, G. & Fellows, B. (1975). "Perceptual" strategies in the solving of three-term series problems. *British Journal of Psychology, **66(1)***, 69-78.

Rayner, S. and Riding, R. (1997). Towards a categorisation of cognitive styles and learning styles. *Educational Psychology, **17(1)***, 5-27.

Reichle, E.D., Carpenter, P.A. and Just, M.A. (2000). The neural bases of strategy and skill in sentence-picture verification. *Cognitive Psychology, **40(4)***, 261-295.

Richards, A. & French, C.C. (1987). The effects of independently validated strategies on visual hemifield asymmetry. *British Journal of Psychology, **78(1)***, 163-181.

Riding, R. (1991) *Cognitive Styles Analysis*. Birmingham: Learning and Training Technology.

Riding, R. (1997). On the nature of cognitive style. *Educational Psychology, **17(1)***, 29-49.

Riding, R. (1998). *Cognitive Styles Analysis – Research Administration*. Birmingham: Learning & Training Technology.

Riding, R. & Cheema, I. (1991). Cognitive styles – an overview and integration. *Educational Psychology, **11(3)***, 193-215.

Riding, R. and Douglas, G. (1993). The effect of cognitive style and mode of presentation on learning performance. *British Journal of Educational Psychology, **63(1)***, 297-307.

Riding, R., Burton, D., Rees, G. and Sharratt, M. (1995). Cognitive style and personality in 12 year old children. *British Journal of Educational Psychology, **65(1)***, 113-124.

Riding, R. and Wright, M. (1995). Cognitive style, personal characteristics and harmony in student flats. *Educational Psychology, **15(2)***, 337-349.

Riding, R., Glass, A., Butler, S. and Pleydell-Pearce, C. (1997). Cognitive Style and individual differences in EEG alpha during information processing. *Educational Psychology, **17(1)***, 219-234.

Riding, R. & Craig, O. (1997). Cognitive style and problem behaviour in boys referred to residential special schools.

Riding, R. & Rayner, S. (1998). *Cognitive Styles and Learning Strategies*. London: David Fulton.

Rips, L. J. (1986). Mental muddles. In M. Brand & R. Harnish (Eds.). *The Representation of Knowledge and Belief* (pp. 258-286). Tucson: University of Arizona Press.

Rips, L. J. (1994a). *The Psychology of Proof.* Cambridge, MA.: MIT Press.

Rips, L. J. (1994b). Deductive reasoning and its cognitive basis. In R.J. Sternberg (Ed.). *Thinking and Problem Solving* (2nd. ed.), (pp. 1510-178). San Diego: Academic Press.

Roberts, M.J. (1993). Human reasoning: deduction rules, mental models or both? *Quarterly Journal of Experimental Psychology, **46A(4)***, 569-589.

Roberts, M.J. (2000). Individual differences in reasoning strategies: a problem to solve or an opportunity to seize? In Schaeken et al. (pp. 23-48).

Roberts, M.J., Gilmore, D.J. and Wood, D.J. (1997). Individual differences and strategy selection in reasoning. *British Journal of Psychology*, ***88***, 473-492.

Roberts, M.J. and Newton, E.J. (2001). Understanding strategy selection. *International Journal of Human-Computer Studies, **54(1)***, 137-154.

Sá, W.C., West, R.F. & Stanovich, K.E. (1999). The domain specificity and generality of belief bias: searching for a generalizable critical thinking skill. *Journal of Educational Psychology*, ***91(3)***, 497-510.

Salmon, M.H. (1991). Informal reasoning and informal logic. In J.F. Voss, D.N. Perkins and J.W. Segal. (Eds.). *Informal Reasoning and Education* (pp. 153-168). Hillsdale, NJ.: Lawrence Erlbaum Associates.

Schunn, C.D., Lovett, M.C. & Reder, L.M. (2001). Awareness and working memory in strategy adaptivity. *Memory and Cognition, **29(2)***, 254-266.

Schribner, S. (1986). Thinking in action: some characteristics of practical thought. In R. J. Sternberg and R. Wagner. (Eds.). *Practical Intelligence: Nature and Origins of Competence in the Everyday World* (pp. 13-30). Cambridge: Cambridge University Press.

Sells, S.B. (1936). The atmosphere effect: an experimental study of reasoning. *Archives of Psychology, **200***, 1-72.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, **125(1)***, 4-27.

Shepard, R.N. & Metzler, J. (1971). Mental rotation of three dimensional objects. *Science, 171*, 952-954.

Siegel, H. (1993). Not by skill alone: the centrality of character to critical thinking. *Informal Logic, XV:3, Fall*, 163-177.

Siegler, R.S. (1996). *Emerging Minds: How Children Discover New Strategies*. New York: Oxford University Press.

Siegler, R.S. and Stern, E. (1998). Conscious and unconscious strategy discoveries: a microgenetic analysis. *Journal of Experimental Psychology: General, 127(4)*, 377-397.

Smith, E.E. & Jonides, J. (1997). Working memory: a view form neuroimaging. *Cognitive Psychology, 33(1)*, 5-42.

Snow, R.E. & Yallow, E. (1982). Education and intelligence. In R. Sternberg (Ed.). *Handbook of Human Intelligence* (pp. wkuvhb). Cambridge: Cambridge University Press.

Spearman, C. (1904). "General intelligence": objectively determined and measured. In Kline, P. (1991).

Stanovich, K.E. (1999). *Who is Rational? Studies of Individual Differences in Reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Stanovich, K.E. & West, R.F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology, 89(3)*, 342-357.

Stanovich, K.E. & West, R.F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127(2)*, 161-188.

Stanovich, K.E., West, R.F. and Sá, W.C. (1999). Thinking dispositions and decontextualised reasoning. In Stanovich, K.E. *Who is Rational? Studies of Individual Differences in Reasoning* (pp. 153-189). Mahwah, NJ: Lawrence Erlbaum Associates.

Stenning, K. & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: logic and implementation. *Cognitive Science, 19(1)*. 97-140.

Stenning, K. & Yule, P. (1997). Image and language in human reasoning: a syllogistic illustration. *Cognitive Psychology, 34(1)*, 109-159.

Sternberg, R.J. (1980). Representation and process in linear syllogistic reasoning. *Journal of Experimental Psychology, General, **109(2)***, 119-159.

Sternberg, R.J. (1982). Reasoning, problem solving and intelligence. In R. Sternberg (Ed.). *Handbook of Human Intelligence* (pp. 225-307), Cambridge: Cambridge University Press.

Sternberg, R.J. (1988). Mental self-government: a theory of intellectual styles and their development. *Human Development, **31(1)***, 197-224.

Sternberg, R.J. (1997). *Thinking Styles*. Cambridge: Cambridge University Press.

Sternberg, R.J. & Weil, E.M. (1980). An aptitude x strategy interaction in linear syllogistic reasoning. *Journal of Educational Psychology, **72(2)***, 226-239.

Sternberg, R.J. & Grigorenko, E.L. (1997). Are cognitive styles still in style? *American Psychologist, **52(7)***, 700-712.

Toms, M., Morris, N. & Ward, D. (1993). Working memory and conditional reasoning. *Quarterly Journal of Experimental Psychology, **46A***, 679-699.

Torrens, D., Thompson, V.A. & Cramer, K.M. (1999). Individual differences and the belief bias effect: mental models, logical necessity and abstract reasoning. *Thinking and Reasoning, **5(1)***, 1-28.

Van der Henst, J-B, Mental model theory and pragmatics. *Behavioral and Brain Sciences, **23(2)***, 283-284.

Vandenberg, S.G. & Kuse, A.R. (1978). Mental rotation: a group test of three-dimensional spatial visualisation. *Perceptual and Motor Skills, 47*, 599-601.

Vandierendonck, A., & De Vooght, G. (1997). Working memory constraints on linear reasoning with spatial and temporal contents. *Quarterly Journal of Experimental Psychology, **50A***, 803-820.

Wechsler, D. (1981). ***WAIS-R Manual: Wechsler Adult Intelligence Scale – Revised.*** San Antonio: Psychological Corporation.

Wetherick, N.E. (1989). Psychology and syllogistic reasoning. *Philosophical Psychology, 2(1)*, 111-124.

Wetherick, N.E. and Gilhooly, K.J. (1990). Syllogistic reasoning: effects of premise order. In K.J. Gilhooly, M.T. Keane, R.H. Logie, and G. Erdos (Eds.). *Lines of Thinking: Reflections on the Psychology of Thought, Vol. 1*. (pp. 99-108). Chichester: Wiley.

Wetherick, N.E. and Gilhooly, K.J. (1995). "Atmosphere", matching and logic in syllogistic reasoning. *Current Psychology: Developmental, Learning, Personality and Social, 14 (3)*, 169-178.

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

Zachary, R.A. (2000). *Shipley Institute of Living Scale: Revised Manual.* Los Angeles: Western Psychological Services.

APPENDICES

All appendices are numbered in line with the chapter to which their contents refer.

APPENDIX 2A

Experiment 1: The 27 valid syllogistic problems presented to all participants. Shown with their valid conclusions.

1	Some of the archaeologists are atheists None of the archaeologists are smokers Some of the atheists are not smokers	15	Some of the painters are linguists All of the linguists are birdwatchers Some of the painters are birdwatchers
2	All of the parents are teachers Some of the teetotallers are not teachers Some of the teetotallers are not parents	16	All of the greengrocers are bellringers None of the greengrocers are hedonists Some of the bellringers are not hedonists
3	All of the athletes are lawyers All of the lawyers are comedians All of the athletes are comedians	17	Some of the doctors are not singers All of the doctors are intellectuals Some of the intellectuals are not singers
4	None of the pianists are mechanics Some of the experts are mechanics Some of the experts are not pianists	18	None of the criminals are florists Some of the florists are trainspotters Some of the trainspotters are not criminals
5	All of the zookeepers are surfers Some of the zookeepers are not homeowners Some of the surfers are not homeowners	19	Some of the gymnasts are nurses None of the nurses are alcoholics Some of the gymnasts are not alcoholics
6	None of the philosophers are saxophonists Some of the boxers are philosophers Some of the boxers are not saxophonists	20	Some of the hikers are not politicians All of the cyclists are politicians Some of the hikers are not cyclists
7	None of the winedrinkers are biologists All of the biologists are potters Some of the potters are not winedrinkers	21	Some of the farmers are freemasons All of the farmers are prizewinners Some of the freemasons are prizewinners
8	None of the bankers are buddhists All of the jugglers are bankers None of the jugglers are buddhists	22	None of the astronomers are naturists Some of the astronomers are flautists Some of the flautists are not naturists
9	All of the cooks are bookworms None of the poets are bookworms None of the cooks are poets	23	Some of the dentists are guitarists None of the skydivers are dentists Some of the guitarists are not skydivers
10	All of the librarians are skaters Some of the sculptors are librarians Some of the sculptors are skaters	24	None of the psychologists are catholics All of the psychologists cricketers Some of the cricketers are not catholics
11	All of the churchgoers are chemists None of the chemists are footballers None of the churchgoers are footballers	25	All of the weavers are gardeners All of the vegetarians are weavers All of the vegetarians are gardeners
12	All of the historians are philatelists None of the beekeepers are historians Some of the philatelists are not beekeepers	26	All of the students are golfers All of the students are drivers Some of the golfers are drivers
13	Some of the clubbers are pilots None of the rockclimbers are pilots Some of the clubbers are not rockclimbers	27	None of the chessplayers are bookbinders All of the dancers are bookbinders None of the chessplayers are dancers
14	All of the plumbers are gamblers Some of the plumbers are snowboarders Some of the gamblers are snowboarders		

APPENDIX 2B

Experiment 1: The Reasoning behaviours questionnaire administered to all participants. V or S in parentheses indicates whether item was intended to identify a verbal or spatial strategy behaviour, and were not shown on original document when presented to participants.

This short questionnaire aims to gather some further information about how you personally approached the 27 reasoning problems. If none of the options in a particular question seem appropriate to you, just write a short note alongside explaining your approach. If you have any difficulties or need to clarify anything, please ask the researcher at any time.

1. Please indicate the extent to which you thought about each of the following during the task.

	A lot				Not at all
a) Groups or sets of people (S)	1	2	3	4	5
b) The words used in the statements (V)	1	2	3	4	5
c) The structure of the statements (all, some etc) (V)	1	2	3	4	5
d) The order in which the occupations were presented (either)	1	2	3	4	5

2. Please describe the kind of mental images/representations which came to mind while you were solving the problems. Mention anything you thought of, and how you used/manipulated the image to help you solve the problems. (S)

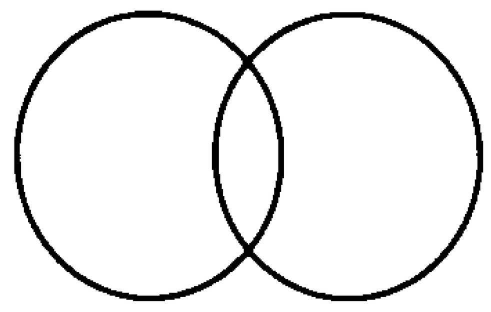
....

2a. Which of the following best sums up the mental image/representations which were *most helpful to you*. Indicate the extent to which this representations were helpful in deciding on a conclusion

	A lot				Not at all
a) Pictorial (of the people described) (S)	1	2	3	4	5
b) Shapes (circles/other shapes in spatial relationships) (S)	1	2	3	4	5
c) Linguistic (e.g. sentences/words) (V)	1	2	3	4	5
d) Numeric (e.g. numbers/equations) (V)	1	2	3	4	5

3. Did you think about sets/Venn diagrams (similar diagrams to that shown below) either using circles or any other shape? (S)

YES/NO



3a. If yes, what symbols did you imagine within the shapes? (S)

a) People of different occupations
b) letters
c) numbers
d) other symbols (please describe)..

3b. If you answered letters, numbers or other to question 3a above, describe why you chose to do this to help you solve the problems. ....

4. Did you attempt to substitute terms/occupations from one statement to another (i.e. switch the occupations around between statements)? YES/NO. (V)

	A lot				Not at all
4a. If yes, roughly how often did you do this?	1............	2............	3............	4............	5

5. Did you attempt to combine the two statements into one to form a single, longer, verbal description? YES/NO (V)

	A lot				Not at all
5a. If yes, how often did you do this?	1............	2............	3............	4............	5

6. Did you attempt to reverse the position of the occupations within the statements? YES/NO (V)

	A lot				Not at all
6a. If yes, how often did you do this?	1............	2............	3............	4............	5

7. **Did you develop a rule or set of rules to help you, which you then re-applied to each subsequent problem?** YES/NO (V)

7a. If yes, at what point in the overall task did you develop the rule(s)?

a) Near the start (i.e. during the practice items or first one or two test statements)
b) Fairly early on (i.e. around statements 6/7)
c) About half way through (i.e. around statements 13/14)
d) Towards the end (i.e. after statement 20)

7b. Please try to describe the rule(s)..

8. If you did not develop any rules, did you approach each problem in an entirely individual way? YES/NO (Either)

9. Did you apply a rule or rules which you already knew about from things you have done in the past, rather than develop a new rule for this task? YES/NO (Either)

9a. If yes, please describe the rule(s) and how you knew about it..

10. Most people find some of these problems harder than others. Describe the type of problem you found hardest. (either)...

10a. What made these problems harder than others in the test? ....

11. Please try to sum up your overall approach (either)..

12. Did you change your approach entirely at any point during the task? YES/NO (either)

12a. If yes, roughly at what point?

a) Fairly early on (i.e. around statement 6/7)
b) About half way through (i.e. around statement 13/14)
c) Towards the end (i.e. after statement 20)

13. Did you adapt or develop your initial approach (rather than change it completely) as the task went on and you became accustomed to the problems? YES NO (either)

13a. How did you do this? ....

APPENDIX 2C

Experiment 1: SPSS output from reliability analysis of reasoning behaviours questionnaire completed by all participants.

Analysis 1: all qualitative items

RELIABILITY ANALYSIS - SCALE (ALPHA)
Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Alpha if Item Deleted
Q1A	29.4118	40.0871	.1127	.6421
Q1B	31.0588	38.0565	.3569	.5971
Q1C	31.6863	42.4596	.1688	.6261
Q1D	30.0000	40.3200	.1551	.6298
Q2B	29.6471	34.7929	.4111	.5806
Q2C	30.7843	36.1725	.4211	.5824
Q2D	29.2745	39.9231	.1456	.6338
Q3	32.4510	40.6925	.4654	.6049
Q4A	29.7059	33.7318	.4777	.5651
Q5A	29.4314	36.2102	.2890	.6086
Q6A	30.0784	35.1937	.4419	.5760
Q7	32.7647	42.5835	.2183	.6244
Q2A	29.4706	41.4941	.0419	.6532

Reliability Coefficients

N of Cases = 51.0 N of Items = 13

Alpha = .6301

Analysis 2: Item 2A removed

RELIABILITY ANALYSIS - SCALE (ALPHA)
Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Alpha if Item Deleted
Q1A	25.9020	36.8902	.1625	.6600
Q1B	27.5490	36.4525	.2998	.6325
Q1C	28.1765	39.8682	.1956	.6487
Q1D	26.4902	37.3749	.1964	.6495

Q2B	26.1373	32.4008	.4228	.6064
Q2C	27.2745	34.0031	.4151	.6107
Q2D	25.7647	37.4235	.1538	.6592
Q3	28.9412	38.4165	.4518	.6306
Q4A	26.1961	31.2808	.4964	.5893
Q5A	25.9216	34.7137	.2428	.6477
Q6A	26.5686	32.6502	.4652	.5988
Q7	29.2549	40.1137	.2296	.6482

Reliability Coefficients

N of Cases = 51.0 N of Items = 12

Alpha = .6532

Analysis 3: Item 1A removed

RELIABILITY ANALYSIS - SCALE (ALPHA)
Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Alpha if Item Deleted
Q1B	23.9804	32.2596	.2850	.6426
Q1C	24.6078	35.4431	.1796	.6576
Q1D	22.9216	32.6737	.2174	.6548
Q2B	22.5686	28.3302	.4177	.6150
Q2C	23.7059	29.3718	.4487	.6104
Q2D	22.1961	33.6408	.1072	.6783
Q3	25.3725	33.8784	.4699	.6355
Q4A	22.6275	26.6784	.5374	.5849
Q5A	22.3529	29.9129	.2733	.6511
Q6A	23.0000	29.2400	.4086	.6176
Q7	25.6863	35.3396	.2790	.6526

Reliability Coefficients

N of Cases = 51.0 N of Items = 11

Alpha = .6600

APPENDIX 2C cont.

Analysis 4: Item 2D removed

RELIABILITY ANALYSIS - SCALE (ALPHA)
Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Alpha if Item Deleted
Q1B	20.2745	28.8831	.3122	.6595
Q1C	20.9020	32.2902	.1725	.6786
Q1D	19.2157	29.5725	.2168	.6774
Q2B	18.8627	25.4008	.4192	.6379
Q2C	20.0000	25.7600	.5081	.6191
Q3	21.6667	31.1067	.4055	.6602
Q4A	18.9216	24.3137	.5013	.6169
Q5A	18.6471	26.4729	.3025	.6689
Q6A	19.2941	26.3718	.4024	.6415
Q7	21.9804	32.0596	.2995	.6712

Reliability Coefficients

N of Cases = 51.0 N of Items = 10

Alpha = .6783

Analysis 5: Item 1C removed

RELIABILITY ANALYSIS - SCALE (ALPHA)
Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Alpha if Item Deleted
Q1B	18.9804	27.7796	.2965	.6632
Q1D	17.9216	28.3137	.2141	.6797
Q2B	17.5686	24.0902	.4277	.6349
Q2C	18.7059	24.8118	.4849	.6227
Q3	20.3725	29.8784	.3915	.6625
Q4A	17.6275	23.0384	.5096	.6126
Q5A	17.3529	24.8729	.3283	.6631
Q6A	18.0000	25.3600	.3851	.6452
Q7	20.6863	30.6996	.3081	.6723

Reliability Coefficients
N of Cases = 51.0 N of Items = 9
Alpha = .6786

Experiment 1: Summary of conclusions produced by verbal and spatial reasoners for all 27 valid syllogisms. Presented by syllogism form and mental model count together with Chi square statistics indicating association between strategy group and number of correct and incorrect conclusions. Correct conclusions and significant associations in bold type. Numbers in parentheses indicate table no. of syllogism in Ford (1995) analysis.

Same-form; single model syllogisms

Syll. no. Strategy	Conclusion produced (within strategy %)									χ^2 *and sig.* (*df* = 1)
	All A - C	**All C - A**	**Some A - C**	**Some C - A**	**No A - C**	**No C - A**	**Some A not C**	**Some C not A**	*NVC/other*	
3										
Verbal (5)	**92.9**	3.6	-	-	-	-	-	-	3.6	$\chi^2 = 0.05$
Spatial (10)	**94.4**	5.6	-	-	-	-	-	-	-	p = 0.83
8										
Verbal (5)	-	3.6	-	3.6	**14.3**	**67.9**	-	-	10.7	$\chi^2 = 0.13$
Spatial (10)	-	11.1	-	-	**33.3**	**44.4**	-	5.6	5.6	p = 0.72
9										
Verbal (5)	3.6	-	-	7.1	**25.0**	**60.7**	-	-	3.6	$\chi^2 = 3.65$
Spatial (10)	-	-	-	5.6	**22.2**	**38.9**	11.1	11.1	11.1	p = 0.06
10										
Verbal (5)	-	-	**35.7**	**53.6**	-	-	-	-	10.7	$\chi^2 = 0.00$
Spatial (11)	-	-	**33.3**	**55.6**	-	-	-	5.6	5.6	p = 0.97
11										
Verbal (none)	-	-	-	-	**71.4**	**25.0**	-	-	3.6	$\chi^2 = 1.02$
Spatial (10)	-	-	-	-	**77.8**	**11.1**	-	-	5.6	p = 0.31
14										
Verbal (8)	3.6	3.6	**57.1**	**32.1**	-	-	3.6	-	-	$\chi^2 = 2.06$
Spatial (11)	-	-	**61.1**	**38.9**	-	-	-	-	-	p = 0.15
15										
Verbal (5)	3.6	-	**71.4**	**21.4**	-	-	-	-	3.6	$\chi^2 = 0.05$
Spatial (11)	-	-	**83.3**	**11.1**	-	-	-	-	5.6	p = 0.83
21										
Verbal (8)	3.6	-	**536.**	**42.9**	-	-	-	-	-	$\chi^2 = 0.10$
Spatial (11)	5.6	-	**66.7**	**27.8**	-	-	-	-	-	p = 0.75
25										
Verbal (5)	53.6	**35.7**	3.6	-	3.6	-	-	-	3.6	$\chi^2 = 2.85$
Spatial (10)	38.9	**61.1**	-	-	-	-	-	-	-	p = 0.09
27										
Verbal (5)	-	-	-	-	**53.6**	**28.6**	-	3.6	14.3	$\chi^2 = 0.39$
Spatial (10)	-	-	5.6	-	**44.4**	**44.4**	-	-	5.6	p = 0.53

Same-form; multi-model syllogisms (2 model)

Syll. no. Strategy	Conclusion produced (within strategy %)									χ^2 *and sig.* *(df = 1)*
	All A - C	**All C - A**	**Some A - C**	**Some C - A**	**No A - C**	**No C - A**	**Some A not C**	**Some C not A**	*NVC/other*	
2										
Verbal (9)	-	-	17.9	14.3	7.1	3.6	21.4	**21.4**	14.2	$\chi^2 = 2.74$
Spatial (14)	-	-	11.1	11.1	11.1	5.6	11.1	**44.4**	5.6	$p = 0.10$
5										
Verbal (9)	-	-	17.9	3.6	-	-	**71.4**	3.6	3.6	$\chi^2 = 2.16$
Spatial (14)	-	-	22.2	16.7	-	-	**50.0**	5.6	5.6	$p = 0.14$
17										
Verbal (9)	-	-	14.3	7.1	-	-	-	**67.9**	10.7	$\boldsymbol{\chi^2 = 7.05}$
Spatial (14)	-	-	22.2	16.7	-	-	33.3	**27.8**	-	$\mathbf{p = 0.01}$
20										
Verbal (9)	-	-	46.4	10.7	3.6	3.6	**17.9**	7.1	10.7	$\chi^2 = 1.44$
Spatial (14)	-	-	44.4	-	-	5.6	**33.3**	5.6	11.1	$p = 0.23$

Different-form; multi-model problems (3 model)

Syll. no. Strategy	Conclusion produced (within strategy %)									χ^2 *and sig.* *(df = 1)*
	All A - C	**All C - A**	**Some A - C**	**Some C - A**	**No A - C**	**No C - A**	**Some A not C**	**Some C not A**	*NVC/other*	
1										
Verbal (7)	3.6	-	28.6	-	32.1	14.3	**14.3**	-	7.2	$\boldsymbol{\chi^2 = 5.17}$
Spatial (12)	-	-	5.6	11.1	27.8	5.6	**44.4**	-	5.6	$\mathbf{p = 0.02}$
4										
Verbal (6)	-	-	10.7	10.7	25.0	25.0	7.1	**17.9**	3.6	$\chi^2 = 0.13$
Spatial (12)	-	-	5.6	11.1	16.7	16.7	11.1	**22.2**	5.6	$p = 0.72$
6										
Verbal (6)	-	-	3.6	14.3	14.3	25.0	3.6	**28.6**	10.7	$\chi^2 = 0.53$
Spatial (12)	-	-	5.6	11.1	16.7	22.2	5.6	**38.9**	-	$p = 0.47$
7										
Verbal (7)	3.6	-	3.6	-	42.9	35.7	-	-	14.3	$\boldsymbol{\chi^2 = 4.99}$
Spatial (13)	-	-	5.6	-	38.9	27.8	5.6	**16.7**	5.6	$\mathbf{p = 0.03}$
12										
Verbal (7)	-	-	-	3.6	25.0	57.1	-	3.6	10.7	$\chi^2 = 3.25$
Spatial (13)	-	-	-	-	22.2	50.0	**11.1**	11.1	5.6	$p = 0.07$

Different-form; multi-model syllogisms cont.

Syll. no. Strategy	**Conclusion produced (within strategy %)**									χ^2 *and sig. (df = 1)*
	All A - C	**All C - A**	**Some A - C**	**Some C - A**	**No A - C**	**No C - A**	**Some A not C**	**Some C not A**	*NVC/other*	
13										
Verbal (6)	-	-	10.7	10.7	14.3	32.1	**25.0**	3.6	3.6	$\chi^2 = 1.00$
Spatial (12)	-	-	11.1	5.6	27.8	11.1	**38.9**	-	5.6	p = 0.32
16										
Verbal (7)	-	-	-	3.6	57.1	25.0	**7.1**	-	7.1	$\chi^2 = 3.32$
Spatial (13)	-	-	5.6	-	44.4	16.7	**27.8**	-	5.6	p = 0.06
18										
Verbal (7)	-	-	21.4	3.6	39.3	3.6	7.1	**14.3**	10.7	$\chi^2 = 0.05$
Spatial (12)	-	-	16.7	-	50.0	5.6	5.6	**16.7**	5.6	p = 0.83
19										
Verbal (6)	-	-	17.9	-	32.1	3.6	**32.1**	3.6	10.7	$\chi^2 = 0.22$
Spatial (12)	-	-	11.1	5.6	22.2	5.6	**38.9**	11.1	5.6	p = 0.64
22										
Verbal (7)	-	-	10.7	14.3	17.9	17.9	10.7	**10.7**	17.9	$\chi^2 = 3.56$
Spatial (12)	-	-	-	5.6	38.9	11.1	11.1	**33.3**	-	p = 0.06
23										
Verbal (7)	-	-	10.7	14.3	17.9	31.2	**14.3**	7.1	3.6	$\chi^2 = 0.48$
Spatial (12)	-	-	-	5.6	16.7	33.3	**22.2**	22.2	-	p = 0.48
24										
Verbal (7)	-	-	3.6	-	17.9	57.1	-	**10.7**	10.7	$\chi^2 = 1.13$
Spatial (13)	-	-	-	-	27.8	44.4	5.6	**22.2**	-	p = 0.29
26										
Verbal (7)	71.4	17.9	**-**	**3.6**	-	-	-	-	7.1	$\chi^2 = 3.93$
Spatial (lone)	50.0	16.7	**16.7**	**5.6**	-	-	-	-	11.1	**p = 0.05**

APPENDIX 2E

Experiment 1: Analysis of Variance (ANOVA) tables.

Table 2E.1: Comparison of mean reasoning behaviour scale scores across three strategy groups (c.f. section 3.1.1.2) .

	Sum of squares	df	Mean square	F	Sig.
Between groups	486.49	2	243.25	10.35	.00
Within groups	1128.02	48	23.50		
Total	1614.51	50			

Table 2E.2: Comparison of verbal and spatial reasoners' performance on same and different form syllogisms: Within-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Syllogism Form	58909.12	1	58909.12	144.05	.00
Syll form *Strategy	1191.53	1	1191.53	2.91	0.09
Error (Form)	17993.55	44	408.94		

Table 2E.3: Comparison of verbal and spatial reasoner' performance on same and different form syllogisms: Between-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Intercept	189263.67	1	189263.67	449.82	0.00
Strategy	853.29	1	853.29	2.03	0.16
Error	18513.35	44	420.76		

Table 2E.4: Simple main effects analysis of interaction between syllogism form and strategy: Same-form problems

	Type 111 Sum of squares	df	Mean square	F	Sig.
Corrected model	14.08	1	14.08	0.07	0.80
Intercept	229676.91	1	229676.91	1055.50	.00
Strategy	14.08	1	14.08	0.07	0.85
Error	9574.43	44	408.94		
Total	251479.59	46			
Corrected Total	9588.51	45			

Table 2E.5: Simple main effects analysis of interaction between syllogism form and strategy: Different -form problems

	Type 111 Sum of squares	df	Mean square	F	Sig.
Corrected model	2030.74	1	2030.74	3.31	0.08
Intercept	18495.88	1	18495.88	30.22	0.00
Strategy	2030.74	1	2030.74	4.97	0.03
Error	26932.47	44	408.94		
Total	45680.47	46			
Corrected Total	28963.21	45			

Table 2E.6: Comparison of verbal and spatial reasoners performance on the four same-form, multi-model syllogisms: Within subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Syllogism Pair	13695.65	1	13695.65	10.96	.002
Syll pair *Strategy	13695.65	1	13695.65	10.96	.002
Error (pair)	55000.00	44	1250.00		

Table 2E. 7: Comparison of verbal and spatial reasoners performance on the four same-form, multi-model syllogisms: Between subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Intercept	152899.41	1	152899.41	85.22	.000
Strategy	725.50	1	725.50	0.40	.528
Error	78948.41	44	1794.28		

APPENDIX 2F

Experiment 2: Syllogisms presented to all participants. N-P% refers to percentages of Necessary and Probable conclusions endorsed by subjects in Evans et al (1999).

Valid Same-form (both conditions)				Valid Different-form (both conditions)			
No.	**Figure**	**Syllogism**	**N-P %**	**No.**	**Figure**	**Syllogism**	**N-P %**
1	4	Some B are A All B are C Some C are A	73-87	6	2	No B are A Some C are B Some C are not A	73-80
2	1	All A are B All Bare C All A are C	73-80	7	2	All B are A All C are B Some A are C	43-80
3	2	All B are A Some C are B Some C are A	83-87	8	1	No A are B All B are C Some C are not A	47-57
4	3	No A are B All C are B No A are C	80-93	9	4	Some B are A No B are C Some A are not C	70-83
5	3	All A are B No C are B No C are A	80-80	10	1	Some A are B No B are C Some A are not C	83-93

Invalid Same-form (evaluation only)				Invalid Different-form (evaluation only)			
No.	**Figure**	**Syllogism**	**N-P %**	**No.**	**Figure**	**Syllogism**	**N-P %**
11.	1	Some A are B No B are C Some C are A	50-47	16.	1	No A are B No B are C Some C are not A	40-70
12.	4	Some B are A Some B are C Some C are A	57-90	17.	2	Some B are A All C are B Some C are not A	63-77
13.	3	All A are B Some C are B Some A are C	67-70	18.	3	All A are B All C are B Some C are A	47-60
14.	3	No A are B No C are B No A are C	43-77	19.	2	Some B are A Some C are B Some A are not C	83-87
15.	3	All A are B All C are B All A are C	60-83	20.	2	Some B are A No C are B Some C are not A	67-77

APPENDIX 2`F cont.

Experiment 2: Syllogisms presented to all participants.

Indeterminate (production only)

No.	Figure	Syllogism	N-P %
21.	1	Some A are B Some B are C	-
22.	3	All A are B All C are B	-
23.	3	Some A are B All C are B	-
24.	2	Some B are A All C are B	-
25.	3	All A are B Some C are B	-

APPENDIX 2G

Experiment 2: Revised questionnaire presented to all participants after the two reasoning tasks. V or S in parentheses indicates whether item intended to identify verbal or spatial reasoning behaviours respectively. Items 1 to 10 were presented with a 5 point response scale with possible responses from "a lot" to "not at all".

Reasoning Style Questionnaire

This short questionnaire aims to gather some further information about how you personally approached the reasoning problems you have just completed. Please answer every question. For questions 1 to 11 indicate your answer by circling the appropriate response. For question 12 please try to describe your overall approach as clearly as you can. If you have any difficulties or need to clarify anything, please ask the researcher at any time.

1. To what extent did you think about the words used in the statements? (V)

2. To what extent did you use mental images of shapes (e.g. circles/other shapes in spatial relationships) in deciding on a conclusion? (S)

3. To what extent did you represent the information in linguistic form (e.g. sentences/words) in deciding on a conclusion? (V)

4. Did you attempt to substitute terms/occupations from one statement to another (i.e. switch the occupations around between statements)? (V)

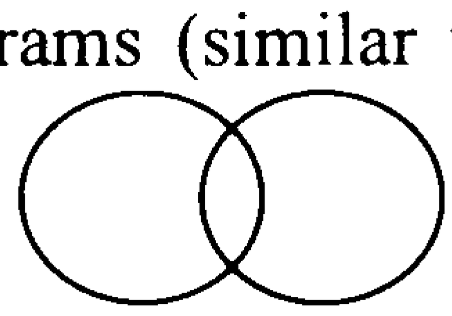
5. Did you attempt to combine the two statements into one to form a single, longer, verbal description? (V)

6. To what extent did you attempt to reverse the position of the occupations within the statements? (V)

7. Did you attempt to organise the occupations into groups and sub-groups/sets and sub-sets? (S)

8. Did you represent the relationship between the occupations diagrammatically? (S)

9. Did you work with the statements in the order they were presented (i.e. dealt with information in the first statement first, then went on to the second statement)? (S)

10. Did you think about sets/Venn diagrams (similar to that shown below) either using circles or any other shape? YES/NO  (S)

11. Did you develop a rule, or set of rules, to help you which you then re-applied to each subsequent problem? YES/NO If yes, please try to describe to describe the rule(s). (V)

12. Please try to sum up your overall approach. (Both)

APPENDIX 2H

Experiment 2: Written instructions presented to participant prior to the two syllogistic reasoning tasks. In each case, two practice items were also presented following the instructions.

1. Instructions for the conclusion production task.

Please read the instructions below before you begin.

Conclusion Production Task

This task consists of 10 pairs of statements about groups of people, similar to those presented in examples 1 and 2 below:

Example 1: All of the beauticians are actors
Some of the chemists are beauticians

Example 2: None of the plumbers are naturalists
Some of the naturalists are tapdancers

Your task is to read the statements and in each case work out what conclusion you can draw, if any, about the relationship between the people. Your conclusion should be based solely on the information given and not on any prior knowledge.

You will notice that the pairs of statements have an occupation in common (in the above example 1 it is beauticians and in example 2 it is naturalists). Your conclusion should be a statement which describes the relationship between the two other occupations, i.e. those which appear only once. For instance, possible conclusions to example 1 above might be:

All of the chemists are actors
All of the actors are chemists
None of the chemists are actors
None of the actors are chemists
Some of the actors are chemists
Some of the chemists are actors
Some of the actors are not chemists
Some of the chemists are not actors
No valid conclusion is possible

Each pair of statements is presented on a separate page. Because we are trying to find out how people solve these problems it is vital that we have a written record of your "working out". Therefore, we would like you to show how you have worked out the conclusion in the space provided beneath each problem. This can include diagrams, words, arrows, symbols or any other written form that helps you to use the information in the statements to find a conclusion. When you have decided on a single conclusion, state that conclusion clearly in writing and move on to the next page. You must only state one conclusion for each problem. If you think there is no possible valid conclusion to a particular problem, state that instead of a conclusion. On the following page the above examples are presented again. Try and work out the correct conclusion to each example and practice representing your work on paper. You can write in the space beneath each example.

APPENDIX 2H cont.

2. Instructions for the conclusion evaluation task.

Please read the instructions below before you begin.

Conclusion Evaluation Task

This task consists of 20 pairs of statements about groups of people. In each case, the information in the statements has been used to reach a conclusion which describes the relationship between two of the groups. This conclusion is also presented. The sets of statements are similar to those in examples 1 and 2 below:

Example 1: All of the teachers are cellists
Some of the hikers are teachers

Conclusion: Some of the hikers are cellists

Example 2: None of the pilots are anglers
Some of the anglers are pianists

Conclusion: Some of the pianists are not pilots

Your task is to read the statements and in each case decide whether the conclusion presented is valid or invalid. For a conclusion to be valid, it *must* be true given that the two initial statements are true. If this is not the case, then the conclusion is invalid. Your decision should be based solely on the information given and not on any prior knowledge.

Each set of statements is presented on a separate page. Because we are trying to find out how people solve these problems it is vital that we have a written record of your "working out". Therefore, we would like you to show how you have worked out the conclusion in the space provided beneath each problem. This can include diagrams, words, arrows, symbols or any other written form that helps you to use the information in the statements to reach a decision about the validity of the conclusion presented. When you have decided whether the conclusion is valid or not, indicate your decision by ticking the appropriate box alongside each problem.

On the following page the above examples are presented again. Try and work out whether the conclusion stated in each case is valid or invalid and practice representing your work on paper. Use the space beneath each example for your working out.

APPENDIX 21

Experiment 2: SPSS output for reliability analysis on reasoning behaviour questionnaire.

Analysis 1: all items included.

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	Q1	Q2	Q3	Q4	Q5
Q1	1.0000				
Q2	-.1814	1.0000			
Q3	.1464	.2541	1.0000		
Q4	.0437	.0697	-.0478	1.0000	
Q5	.1228	.1691	.0291	.5071	1.0000
Q6	.0263	.1777	.1431	.5537	.5220
Q7	-.0480	-.0595	-.0716	-.2158	-.2681
Q8	.0176	.2733	.2630	.0440	.0731
Q9	.1538	-.1267	.0360	.1618	.1714
Q10	-.1309	.7089	.2576	-.0531	.1805
Q11	-.0213	.1358	.0355	-.0534	-.0803

	Q6	Q7	Q8	Q9	Q10
Q6	1.0000				
Q7	-.3200	1.0000			
Q8	.0359	.1163	1.0000		
Q9	.2955	.0460	.0483	1.0000	
Q10	.2547	.0305	.2987	-.0947	1.0000
Q11	-.1474	.0614	.0607	-.1118	.118

	Q11
Q11	1.0000

N of Cases = 68.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
Q1	25.0000	28.7761	.0451	.1133	.4981
Q2	24.2353	24.1528	.2545	.5516	.4375

Q3	24.8235	25.9087	.2369	.1809	.4457
Q4	24.1176	25.0904	.2901	.4556	.4280
Q5	23.8676	23.5195	.3638	.3868	.3981
Q6	24.1912	23.6495	.4006	.5466	.3894
Q7	24.4412	31.4741	-.1860	.1887	.5808
Q8	23.4706	24.0439	.2822	.1689	.4269
Q9	23.2941	26.8077	.1741	.1771	.4649
Q10	26.7794	28.2640	.4245	.6017	.4468
Q11	26.5147	30.3431	-.0084	.0605	.4938

Reliability Coefficients 11 items

Alpha = .4839 Standardized item alpha = .5017

Analysis 2: Item 7 removed.

R E L I A B I L I T Y A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	Q1	Q2	Q3	Q4	Q5
Q1	1.0000				
Q2	-.1814	1.0000			
Q3	.1464	.2541	1.0000		
Q4	.0437	.0697	-.0478	1.0000	
Q5	.1228	.1691	.0291	.5071	1.0000
Q6	.0263	.1777	.1431	.5537	.5220
Q8	.0176	.2733	.2630	.0440	.0731
Q9	.1538	-.1267	.0360	.1618	.1714
Q10	-.1309	.7089	.2576	-.0531	.1805
Q11	-.0213	.1358	.0355	-.0534	-.0803

	Q6	Q8	Q9	Q10	Q11
Q6	1.0000				
Q8	.0359	1.0000			
Q9	.2955	.0483	1.0000		
Q10	.2547	.2987	-.0947	1.0000	
Q11	-.1474	.0607	-.1118	.1185	1.0000

N of Cases = 68.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
Q1	22.3676	29.5494	.0573	.1122	.6000
Q2	21.6029	24.8101	.2683	.5460	.5556
Q3	22.1912	26.5748	.2540	.1762	.5565
Q4	21.4853	25.2087	.3515	.4549	.5298
Q5	21.2353	23.3468	.4453	.3738	.4983
Q6	21.5588	23.3547	.4988	.5155	.4855
Q8	20.8382	25.4809	.2409	.1573	.5634
Q9	20.6618	27.8988	.1581	.1507	.5804
Q10	24.1471	29.2318	.4093	.5889	.5550
Q11	23.8824	31.3591	-.0241	.0605	.5950

Reliability Coefficients   10 items

Alpha =   .5808   Standardized item alpha =   .5738

Analysis 3: Item 1 removed.

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	Q2	Q3	Q4	Q5	Q6
Q2	1.0000				
Q3	.2541	1.0000			
Q4	.0697	-.0478	1.0000		
Q5	.1691	.0291	.5071	1.0000	
Q6	.1777	.1431	.5537	.5220	1.0000
Q8	.2733	.2630	.0440	.0731	.0359
Q9	-.1267	.0360	.1618	.1714	.2955
Q10	.7089	.2576	-.0531	.1805	.2547
Q11	.1358	.0355	-.0534	-.003	-.1474

	Q8	Q9	Q10	Q11
Q8	1.0000			
Q9	.0483	1.0000		
Q10	.2987	-.0947	1.0000	
Q11	.0607	-.1118	.1185	1.0000

N of Cases =   68.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
Q2	19.5294	22.2529	.3260	.5350	.5613
Q3	20.1176	25.0606	.2291	.1414	.5867
Q4	19.4118	23.4100	.3547	.4549	.5519
Q5	19.1618	21.8093	.4315	.3584	.5260
Q6	19.4853	21.5072	.5135	.5137	.5030
Q8	18.7647	23.6155	.2462	.1571	.5870
Q9	18.5882	26.3951	.1293	.1410	.6115
Q10	22.0735	27.1736	.4524	.5885	.5718
Q11	21.8088	29.4107	-.0206	.0601	.6168

Reliability Coefficients   9 items

Alpha =   .6000   Standardized item alpha =   .6011

Analysis 4: Item 11 removed.

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	Q2	Q3	Q4	Q5	Q6
Q2	1.0000				
Q3	.2541	1.0000			
Q4	.0697	-.0478	1.0000		
Q5	.1691	.0291	.5071	1.0000	
Q6	.1777	.1431	.5537	.5220	1.0000
Q8	.2733	.2630	.0440	.0731	.0359
Q9	-.1267	.0360	.1618	.1714	.2955
Q10	.7089	.2576	-.0531	.1805	.2547

	Q8	Q9	Q10
Q8	1.0000		
Q9	.0483	1.0000	
Q10	.2987	-.0947	1.0000

N of Cases =   68.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
Q2	18.9706	22.3275	.3111	.5336	.5881
Q3	19.5588	24.9666	.2260	.1410	.6085
Q4	18.8529	23.2019	.3618	.4519	.5706
Q5	18.6029	21.5564	.4427	.3574	.5431
Q6	18.9265	21.1736	.5335	.5021	.5169
Q8	18.2059	23.5689	.2402	.1571	.6105
Q9	18.0294	26.1185	.1409	.1399	.6297
Q10	21.5147	27.0893	.4417	.5859	.5927

Reliability Coefficients 8 items

Alpha = .6168 Standardized item alpha = .6463

Analysis 5: Item 9 removed.

R E L I A B I L I T Y A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	Q2	Q3	Q4	Q5	Q6
Q2	1.0000				
Q3	.2541	1.0000			
Q4	.0697	-.0478	1.0000		
Q5	.1691	.0291	.5071	1.0000	
Q6	.1777	.1431	.5537	.5220	1.0000
Q8	.2733	.2630	.0440	.0731	.0359
Q10	.7089	.2576	-.0531	.1805	.2547

	Q8	Q10
Q8	1.0000	
Q10	.2987	1.0000

N of Cases = 68.0

APPENDIX 2I cont.

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
Q2	15.1912	18.5450	.3776	.5297	.5817
Q3	15.7794	21.7864	.2324	.1406	.6256
Q4	15.0735	20.4273	.3415	.4503	.5928
Q5	14.8235	18.8639	.4246	.3557	.5630
Q6	15.1471	18.8437	.4817	.4630	.5457
Q8	14.4265	20.4572	.2447	.1475	.6299
Q10	17.7353	23.6901	.4963	.5814	.6018

Reliability Coefficients 7 items

Alpha = .6297 Standardized item alpha = .6696

Analysis 6: Item 8 removed.

R E L I A B I L I T Y A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	Q2	Q3	Q4	Q5	Q6
Q2	1.0000				
Q3	.2541	1.0000			
Q4	.0697	-.0478	1.0000		
Q5	.1691	.0291	.5071	1.0000	
Q6	.1777	.1431	.5537	.5220	1.0000
Q10	.7089	.2576	-.0531	.1805	.2547

	Q10
Q10	1.0000

N of Cases = 68.0

APPENDIX 2I cont.

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
Q2	11.5882	14.1861	.3217	.5287	.6110
Q3	12.1765	17.1326	.1657	.0994	.6572
Q4	11.4706	14.9394	.3820	.4436	.5778
Q5	11.2206	13.5178	.4714	.3554	.5378
Q6	11.5441	13.3264	.5579	.4546	.5023
Q10	14.1324	18.4449	.4569	.5689	.6057

Reliability Coefficients   6 items

Alpha = .6299   Standardized item alpha = .6648

Analysis 7: Item 3 removed.

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	Q2	Q4	Q5	Q6	Q10
Q2	1.0000				
Q4	.0697	1.0000			
Q5	.1691	.5071	1.0000		
Q6	.1777	.5537	.5220	1.0000	
Q10	.7089	-.0531	.1805	.2547	1.0000

N of Cases = 68.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
Q2	9.3382	11.8690	.2586	.5207	.6976
Q4	9.2206	11.4581	.4541	.4370	.5839
Q5	8.9706	10.2976	.5287	.3544	.5424
Q6	9.2941	10.4794	.5733	.4424	.5218
Q10	11.8824	15.4188	.4169	.5675	.6489

Reliability Coefficients   5 items

Alpha = .6572   Standardized item alpha = .6910

Analysis 8: Item 2 removed.

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	Q4	Q5	Q6	Q10
Q4	1.0000			
Q5	.5071	1.0000		
Q6	.5537	.5220	1.0000	
Q10	-.0531	.1805	.2547	1.0000

N of Cases = 68.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
Q4	6.3824	6.4785	.5609	.4185	.5798
Q5	6.1324	5.7882	.5948	.3544	.5561
Q6	6.4559	5.9532	.6463	.4349	.5143
Q10	9.0441	11.1771	.1568	.1371	.7688

Reliability Coefficients 4 items

Alpha = .6976 Standardized item alpha = .6608

APPENDIX 2J

Experiment 2: Analysis of Variance (ANOVA) tables.

Table 2J.1: Comparison of mean reasoning behaviour scale scores across strategy groups, when scale based on all 11 questionnaire items.

	Sum of squares	df	Mean square	F	Sig.
Between groups	750.32	4	187.58	8.47	.000
Within groups	1284.19	58	22.14		
Total	2034.41	62			

Table 2J.2: Scheffé Post-hoc comparisons of differences between mean reasoning behaviour scale scores, c.f. above ANOVA

Strategies		Mean Difference	Std. Error	Sig.
Verbal	Spatial	-9.86	2.03	0.00
	Mixed	-9.93	1.40	0.22
	Flow (verbal)	-3.43	1.85	0.13
	Flow (spatial)	-6.88	3.47	0.10
Spatial	Mixed	-0.07	2.04	1.00
	Flow (verbal)	6.43	2.37	0.05
	Flow (mixed)	2.98	3.77	0.81
Mixed	Flow (verbal)	6.50	3.48	0.49
	Flow (mixed)	3.06	3.68	0.95
Flow (verbal)	Flow (mixed)	-3.44	1.86	0.50

Table 2J.3: Comparison of mean reasoning behaviour scale scores across strategy groups, when scale based on 4 most reliable items.

	Sum of squares	df	Mean square	F	Sig.
Between groups	750.32	4	25.04	2.19	0.08
Within groups	1284.19	58	11.43		
Total	2034.41	62			

APPENDIX 2J cont.

Table 2J.4: Scheffé Post-hoc comparisons of differences between mean reasoning behaviour scale scores, c.f. above ANOVA

Strategies		Mean Difference	Std. Error	Sig.
Verbal	Spatial	-2.84	1.45	0.44
	Mixed	-1.70	2.49	0.98
	Flow (verbal)	-0.65	1.01	0.98
	Flow (spatial)	-3.36	1.33	0.19
Spatial	Mixed	1.14	2.71	0.10
	Flow (verbal)	2.19	1.47	0.70
	Flow (mixed)	-0.52	1.70	1.00
Mixed	Flow (verbal)	1.05	2.49	0.99
	Flow (mixed)	-1.67	2.64	0.98
Flow (verbal)	Flow (mixed)	-2.71	1.34	0.40

Table 2J.5: Comparison of strategy groups on overall performance on conclusion production and evaluation tasks: Within-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Task	21477.53	1	21477.53	225.22	0.00
Task*strategy	586.71	4	146.68	1.54	0.20
Error (task)	5531.149	58	95.37		

Table 2J.6: Comparison of strategy groups on overall performance on conclusion production and evaluation tasks: Between-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Intercept	111350.25	1	111350.25	425.57	0.00
Strategy	853.22	4	213.31	0.82	0.52
Error	15175.74	58	261.65		

Table 2J.7: Comparison of verbal and spatial reasoners on number of correct conclusions produced to SF and DF syllogisms: Within-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Form	61.16	1	61.16	59.03	0.00
Form*strategy	7.45	4	1.86	1.80	0.14
Error (form)	60.10	58	1.04		

Table 2J.8: Comparison of verbal and spatial reasoners on number of correct conclusions produced to SF and DF syllogisms: Between-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Intercept	136.44	1	136.44	97.56	0.00
Strategy	3.37	4	0.84	0.60	0.66
Error	81.12	58	1.40		

Table 2J.9: Comparison of all strategy groups in terms of number of correctly evaluated conclusions to SF and DF valid and invalid syllogisms: Within-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Form	95.86	1	95.86	44.95	0.00
Form*strategy	7.22	4	1.81	0.85	0.50
Error (form)	123.69	58	2.13		
Validity	8.76	1	8.77	11.77	0.001
Validity*strategy	2.23	4	0.56	0.75	0.56
Error (validity)	43.21	58	0.75		
Form*validity	12.17	1	12.17	8.25	0.006
Form*validity*strateg y	3.78	4	0.64	0.64	0.64
Error (form*validity)	85.58	58	1.48		

Table 2J.10: Comparison of all strategy groups in terms of number of correctly evaluated conclusions to SF and DF, valid and invalid syllogisms: Between-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Intercept	1153.17	1	1153.17	772.02	0.00
Strategy	4.42	4	1.11	0.74	0.57
Error	86.64	58	1.49		

APPENDIX 3A

Experiment 3: The 10 syllogisms presented to all participants. In each case, line 3 represents the logical conclusion.

No.	Syllogism	Figure	Form
1	Some churchgoers are not butchers All churchgoers are pilots Some pilots are not butchers	4	same
2	All smokers are cyclists Some cricketers are not cyclists Some cricketers are not smokers	3	same
3	No biologists are musicians Some vegetarians are musicians Some vegetarians are not biologists	3	different
4	All criminals are chessplayers Some criminals are not catholics Some chessplayers are not catholics	4	same
5	No gardeners are dentists All dentists are bellringers Some bellringers are not gardeners	1	different
6	All singers are athletes Some plumbers are singers Some plumbers are athletes	2	same
7	All actors are parents No chemists are actors Some parents are not chemists	2	different
8	Some drivers are cooks No politicians are cooks Some drivers are not politicians	3	different
9	Some students are farmers All farmers are painters Some students are painters	1	same
10	All lawyers are golfers All lawyers are homeowners Some golfers are homeowners	4	different

APPENDIX 3B

Experiment 3: The 16 three-term series problems presented to all participants .

Note that in each case, the two premises were presented followed by the standard question "*What is the relationship between x and y*?", where *x* and *y* are the two end terms. The third premise shown below represents one possible correct response, but inverse responses which represented the correct relationship were also accepted as correct (for example, for item 1, if the participant stated D was darker than S, that D was not as light as S, or that S was not as dark as D).

No.	Problem
1	S is lighter than P P is lighter than D S is lighter than D
2	B is weaker than W W is weaker than J J is stronger than B
3	R is smoother than X N is rougher than X N is rougher than R
4	G is happier than M G is sadder than V V is happier than M
5	T is not as thin as L L is not as thin as C T is fatter than C
6	Q is not as tall as A A is not as tall as F F is taller than Q
7	H is not as rough as Z O is not as smooth as Z O is rougher than H
8	L is not as dark as E L is not as light as I I is lighter than E

No.	Problem
9	Y is stronger than P P is not as weak as K Y is stronger than K
10	E is darker than V V is not as light as T T is lighter than E
11	X is thinner than M U is not as thin as M U is fatter than X
12	R is taller than G R is not as tall as Q Q is taller than G
13	F is not as dark as S S is lighter than W F is lighter than W
14	H is not as fat as T T is thinner than X X is fatter than H
15	N is not as rough as D J is rougher than D J is rougher than N
16	B is not as weak as P B is weaker than L L is stronger than P

APPENDIX 3C

Experiment 3: Instructions presented to participants prior to the transitive inference task.

Please read the instructions below before you begin.

Object Relationship Task

This task consists of 16 pairs of statements which describe the relationship between three objects. In each case the objects are represented by three different letters. You will be asked to describe the relationship between two of the letters. The problems are presented as shown in examples 1 and 2 below:

Example 1: J is fatter than W
P is thinner than W

What is the relationship between J and P?

Example 2: R is not as small as G
T is larger than R

What is the relationship between T and G?

Your task is to read the statements and use the information given to work out the relationship which is asked for. For instance, the relationship between J and P in example 1 above might be any one of the following:

J is fatter than P
P is thinner than J
P is fatter than J
J is thinner than P
No relationship (there is insufficient information to determine relationship)

Each pair of statements is presented on a separate page. Because we are trying to find out how people solve these problems it is vital that we have a written record of your "working out". Therefore, we would like you to show how you have worked out the conclusion in the space provided beneath each problem. This can include diagrams, words, arrows, symbols or any other written form that helps you to use the information in the statements to determine the relationship requested. When you have decided on the relationship, state that relationship clearly in writing and move on to the next page. You must only one give one answer to each problem. If you think there is insufficient information to determine a relationship, say so.

On the following page the above examples are presented again for you to practice. Try and work out the relationships that are asked for and practice representing your work on paper. You can write in the space beneath each example.

APPENDIX 3D

Experiment 3: Reasoning behaviour questionnaire presented to participants following the syllogistic reasoning task.

Note that items 1 to 8 were each presented with a five point rating scale and participants asked to give a response in the range "A lot" ...to..."Not at all".

Reasoning Style Questionnaire (Conclusion Production Task)

This short questionnaire aims to gather some further information about how you personally approached the reasoning problems you have just completed. Please answer every question. For questions 1 to 8 indicate your answer by **circling** the appropriate response. For question 9 please try to describe your overall approach as clearly as you can. If you have any difficulties or need to clarify anything, please ask the researcher at any time.

1. To what extent did you use images of shapes (e.g. circles/other shapes in spatial relationships) in deciding on a conclusion?

2. To what extent did you represent the information in linguistic form (e.g. as sentences/words) in deciding on a conclusion?

3. How often did you attempt to substitute terms/occupations from one statement to another (i.e. switch the occupations around between statements)?

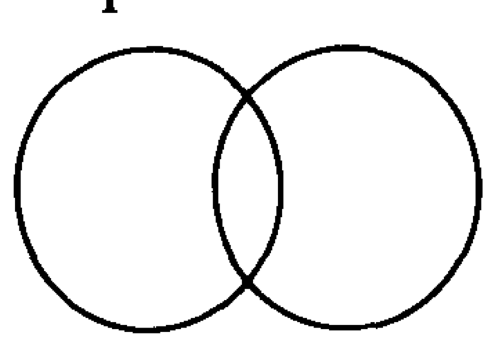
4. Did you attempt to combine the two statements into one to form a single, longer, verbal description?

5. To what extent did you attempt to reverse the position of the occupations within the statements?

6. Did you organise the occupations into groups and sub-groups/sets and sub-sets?

7. Did you work with the statements in the order they were presented (i.e. dealt with information in the first statement first, then went on to the second statement)?

8. Did you use set diagrams/Venn diagrams (similar to that shown below) either using circles or any other shape? YES/NO



9. We are trying to find out how people solve these problems. Please try to describe the way in which you approached the problems, giving as much detail as you can about what you did and why.

APPENDIX 3E

Experiment 3: Reasoning behaviour questionnaire presented to participants following the transitive inference task.

Note that items 1 to 8 were each presented with a five point rating scale and participants asked to give a response in the range "A lot" ...to..."Not at all".

Reasoning Style Questionnaire (Object Relationship Task)

This short questionnaire aims to gather some further information about how you personally approached the reasoning problems you have just completed. Please answer every question. For questions 1 to 8 indicate your answer by **circling** the appropriate response. For question 9 please try to describe your overall approach as clearly as you can. If you have any difficulties or need to clarify anything, please ask the researcher at any time.

1. To what extent did you use images depicting the physical properties of the three letters?

2. To what extent did you place the three letters in various positions on a scale or continuum according to their relationship?

3. Did the physical properties of the three letters seem to differ between problems?

4. Did you attempt to combine the two statements into one to form a single, longer, verbal description?

5. To what extent did you attempt to reverse the position of the letters within the statements?

6. Did the three letters appear to lie in a set sequence/order?

7. Did you work with the statements in the order they were presented (i.e. dealt with information in the first statement first, then went on to the second statement)?

8. Did you find that a particular letter could be cancelled out/ignored?

9. We are trying to find out how people solve these problems. Please try to describe the way in which you approached the problems, giving as much detail as you can about what you did and why.

APPENDIX 3F

Experiment 3: SPSS output from reliability analysis conducted on the 8 quantitative syllogistic reasoning behaviours questionnaire. The questionnaire items were presented in Appendix 4D.

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	SQ1	SQ2	SQ3	SQ4	SQ5
SQ1	1.0000				
SQ2	.7217	1.0000			
SQ3	.3468	.4654	1.0000		
SQ4	.3443	.4221	.4636	1.0000	
SQ5	.2785	.3463	.6293	.2900	1.0000
SQ6	.5409	.5599	.3316	.2333	.2997
SQ7	.1117	.1659	.1736	.0421	.1999
SQ8	.8567	.7607	.4295	.3936	.2516

	SQ6	SQ7	SQ8
SQ6	1.0000		
SQ7	.1025	1.0000	
SQ8	.5864	.0761	1.0000

N of Cases =   62.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
SQ1	19.1774	38.8041	.6311	.7560	.7578
SQ2	19.3871	40.4707	.7300	.6484	.7439
SQ3	18.6129	39.5854	.6052	.5373	.7627
SQ4	17.9194	45.0590	.4537	.2808	.7867
SQ5	18.7903	43.8406	.5041	.4384	.7793
SQ6	18.5806	42.0836	.5399	.3923	.7738
SQ7	17.8065	50.8472	.1803	.0646	.8209
SQ8	21.3548	50.6261	.7449	.8018	.7867

Reliability Coefficients   8 items

Alpha = .8005   Standardized item alpha = .8259

APPENDIX 3G

Experiment 3: SPSS output from reliability analysis conducted on the 8 quantitative transitive inference behaviours questionnaire. The questionnaire items were presented in Appendix 3E.

Initial Analysis (all items)

R E L I A B I L I T Y A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	TQ1	TQ2	TQ3	TQ4	TQ5
TQ1	1.0000				
TQ2	-.0027	1.0000			
TQ3	.7034	.0298	1.0000		
TQ4	.1267	.1176	.1138	1.0000	
TQ5	.2657	-.0715	.3059	.3185	1.0000
TQ6	.0397	.2056	.0613	.1976	.1458
TQ7	-.2629	.1719	-.2961	-.0905	-.4371
TQ8	-.1558	-.0623	-.0068	.1652	.2710

	TQ6	TQ7	TQ8
TQ6	1.0000		
TQ7	.1283	1.0000	
TQ8	.0149	-.2375	1.0000

N of Cases = 62.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
TQ1	19.4677	17.2039	.2923	.5358	.2636
TQ2	20.8548	23.0114	.1074	.0846	.3698
TQ3	19.5000	16.8443	.3845	.5223	.2112
TQ4	18.2903	18.9635	.3196	.1518	.2698
TQ5	18.8710	17.7864	.3130	.3421	.2579
TQ6	19.9677	19.8350	.2319	.1179	.3125
TQ7	20.3710	27.5159	-.3380	.2959	.5483
TQ8	18.0323	22.5563	.0042	.1774	.4186

Reliability Coefficients 8 items

Alpha = .3793 Standardized item alpha = .3491

APPENDIX 3G cont.

Analysis 2 (item 7 excluded)

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	TQ1	TQ2	TQ3	TQ4	TQ5
TQ1	1.0000				
TQ2	-.0027	1.0000			
TQ3	.7034	.0298	1.0000		
TQ4	.1267	.1176	.1138	1.0000	
TQ5	.2657	-.0715	.3059	.3185	1.0000
TQ6	.0397	.2056	.0613	.1976	.1458
TQ8	-.1558	-.0623	-.0068	.1652	.2710

	TQ6	TQ8
TQ6	1.0000	
TQ8	.0149	1.0000

N of Cases = 62.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
TQ1	17.6452	18.9212	.3603	.5314	.4709
TQ2	19.0323	26.3596	.0552	.0720	.5674
TQ3	17.6774	18.5172	.4596	.5173	.4240
TQ4	16.4677	21.5973	.3258	.1507	.4912
TQ5	17.0484	18.8665	.4397	.2565	.4346
TQ6	18.1452	23.2737	.1782	.0861	.5451
TQ8	16.2097	24.6274	.0685	.1510	.5840

Reliability Coefficients   7 items

Alpha = .5483   Standardized item alpha = .5168

APPENDIX 3G cont.

Analysis 3 (item 8 excluded)

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	TQ1	TQ2	TQ3	TQ4	TQ5
TQ1	1.0000				
TQ2	-.0027	1.0000			
TQ3	.7034	.0298	1.0000		
TQ4	.1267	.1176	.1138	1.0000	
TQ5	.2657	-.0715	.3059	.3185	1.0000
TQ6	.0397	.2056	.0613	.1976	.1458

	TQ6
TQ6	1.0000

N of Cases = 62.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if item Deleted
TQ1	13.4839	15.2702	.4566	.4994	.4678
TQ2	14.8710	23.3273	.0766	.0691	.6093
TQ3	13.5161	15.5981	.5031	.5121	.4460
TQ4	12.3065	19.3308	.2920	.1393	.5506
TQ5	12.8871	17.1838	.3696	.1999	.5158
TQ6	13.9839	20.4424	.1856	.0850	.5920

Reliability Coefficients 6 items

Alpha = .5840 Standardized item alpha = .5522

APPENDIX 3G cont.

Analysis 4 (item 2 excluded

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	TQ1	TQ3	TQ4	TQ5	TQ6
TQ1	1.0000				
TQ3	.7034	1.0000			
TQ4	.1267	.1138	1.0000		
TQ5	.2657	.3059	.3185	1.0000	
TQ6	.0397	.0613	.1976	.1458	1.0000

N of Cases = 62.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
TQ1	12.1452	13.9622	.4781	.4989	.4864
TQ3	12.1774	14.3778	.5175	.5105	.4666
TQ4	10.9677	18.2940	.2774	.1271	.5951
TQ5	11.5484	15.6943	.4017	.1834	.5342
TQ6	12.6452	19.6097	.1510	.0474	.6484

Reliability Coefficients 5 items

Alpha = .6093 Standardized item alpha = .5960

APPENDIX 3G cont.

Analysis 5 (item 6 excluded)

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	TQ1	TQ3	TQ4	TQ5
TQ1	1.0000			
TQ3	.7034	1.0000		
TQ4	.1267	.1138	1.0000	
TQ5	.2657	.3059	.3185	1.0000

N of Cases = 62.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
TQ1	9.9194	10.4360	.5361	.4987	.4959
TQ3	9.9516	10.9320	.5681	.5101	.4758
TQ4	8.7419	15.3094	.2339	.1037	.6920
TQ5	9.3226	12.6155	.3918	.1779	.6054

Reliability Coefficients 4 items

Alpha = .6484 Standardized item alpha = .6378

Analysis 6 (item 4 excluded)

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	TQ1	TQ3	TQ5
TQ1	1.0000		
TQ3	.7034	1.0000	
TQ5	.2657	.3059	1.0000

N of Cases = 62.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
TQ1	6.0161	6.7374	.6012	.4976	.4684
TQ3	6.0484	7.1288	.6458	.5100	.4184
TQ5	5.4194	9.6901	.3087	.0986	.8242

Reliability Coefficients 3 items

Alpha = .6920 Standardized item alpha = .6892

Analysis 7 (item 5 excluded)

R E L I A B I L I T Y A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	TQ1	TQ3
TQ1	1.0000	
TQ3	.7034	1.0000

N of Cases = 62.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
TQ1	2.6935	2.6095	.7034	.4948	.
TQ3	2.7258	3.0875	.7034	.4948	.

Reliability Coefficients 2 items

Alpha = .8242 Standardized item alpha = .8259

APPENDIX 3H

Experiment 4: Reasoning behaviour questionnaire presented participants following the transitive inference task in phase 1. Note that items 1 to 6 were each presented with a rating scale on which participants were asked to give a response ranging from "a lot" to "not at all".

Reasoning Style Questionnaire (Object Relationship Task)

This short questionnaire aims to gather some further information about how you personally approached the reasoning problems you have just completed. Please answer every question. For questions 1 to 8 indicate your answer by **circling** the appropriate response. For question 7 please try to describe your overall approach as clearly as you can. If you have any difficulties or need to clarify anything, please ask the researcher at any time.

1. To what extent did you use images depicting the physical properties of the three letters?
2. To what extent did you place the three letters in various positions on a scale or continuum according to their relationship?
3. Did the physical properties of the three letters seem to differ between problems?
4. Did you attempt to combine the two statements into one to form a single, longer, verbal description?
5. To what extent did you attempt to reverse the position of the letters within the statements?
6. Did the three letters appear to lie in a set sequence/order?
7. We are trying to find out how people solve these problems. Please try to describe the way in which you approached the problems, giving as much detail as you can about what you did and why. Continue overleaf if you need to.

APPENDIX 3I

Experiment 4: The 16 basic sentence-picture verification trial types presented to all participants. Four sets of the following items were presented, in a different random order for each participant.

True affirmative (TA)			False affirmative (FA)		
No.	**Sentence**	**Picture**	**No.**	**Sentence**	**Picture**
1	A is above B	A B	5	B is above A	A B
2	B is above A	B A	6	A is above B	B A
3	A is below B	B A	7	A is below B	A B
4	B is below A	A B	8	B is below A	B A

True negative (TN)			False negative (FN)		
No.	**Sentence**	**Picture**	**No.**	**Sentence**	**Picture**
9	A is not above B	B A	13	B is not above A	B A
10	B is not above A	A B	14	A is not above B	A B
11	A is not below B	A B	15	A is not below B	B A
12	B is not below A	B A	16	B is not below A	A B

APPENDIX 3J

Experiment 4: Instructions presented to participants prior to the computerised SPV task in phase 2.

In this experiment you will be asked to make judgements about whether a simple picture presents a true illustration of the situation described in a given sentence. In each case, you will first be presented with a short sentence describing the relationship between two objects, represented by the letters A and B, as in the example below:

Example sentence: A is above B

This will remain on the screen for as long as you need to read and understand it. When you are satisfied that you have understood the sentence, click on the button marked ready. A picture will then appear which also describes a possible relationship between A and B, for instance:

Example picture: A
B

Your task is to decide whether you think the relationship shown in the picture is true or false according to the sentence you saw previously. If you think the picture is a true representation of the sentence, click on the button labelled TRUE. If you think the picture is false, click on the button labelled FALSE (in the above example the relationship shown in the picture is true). After a short break, a warning dot will appear, after which the next sentence will be presented and the above procedure can be repeated.

Before you begin the experiment, you will be presented with a series of practice trials to get you used to the procedure. If you have any questions about what you need to do please ask the experimenter now. There will be further opportunity for questions after the practice trials. When you are ready to begin the practice, press the space bar to start. When you have completed the practice, tell the experimenter.

APPENDIX 3K

Experiment 4: SPSS output from reliability analysis on syllogistic reasoning behaviour questionnaire completed by participants.

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	SQ1	SQ2	SQ3	SQ4	SQ5
SQ1	1.0000				
SQ2	.4976	1.0000			
SQ3	.3565	.3150	1.0000		
SQ4	.2592	.4383	.3591	1.0000	
SQ5	-.1213	.1828	.4400	.2022	1.0000
SQ6	.3948	.3102	.2075	.3736	-.0961
SQ7	.1598	.1294	.1541	.0469	-.0034
SQ8	.6508	.4311	.2719	.2375	.0492

	SQ6	SQ7	SQ8
SQ6	1.0000		
SQ7	.1404	1.0000	
SQ8	.3444	.1620	1.0000

N of Cases = 70.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
SQ1	19.3714	26.3528	.4894	.5851	.6469
SQ2	19.8714	27.1571	.5736	.3915	.6265
SQ3	19.4714	28.6586	.5250	.4041	.6416
SQ4	18.7571	28.0996	.4826	.3085	.6490
SQ5	19.5143	33.7896	.1381	.3559	.7244
SQ6	19.2429	28.8242	.3987	.2706	.6704
SQ7	18.3000	33.9812	.1747	.0531	.7122
SQ8	21.9714	34.3760	.5484	.4605	.6785

Reliability Coefficients 8 items

Alpha = .7003 Standardized item alpha = .72

APPENDIX 3L

Experiment 4: SPSS output from reliability analysis on transitive inference reasoning behaviours questionnaire completed by participants.

Analysis 1 (all 6 items)

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	TQ1	TQ2	TQ3	TQ4	TQ5
TQ1	1.0000				
TQ2	.0669	1.0000			
TQ3	.3905	.0403	1.0000		
TQ4	-.0686	.0212	-.1145	1.0000	
TQ5	.0827	-.0645	-.1389	.2498	1.0000
TQ6	-.0399	.2054	-.2595	.0800	.1433

	TQ6
TQ6	1.0000

N of Cases = 70.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
TQ1	14.3286	9.8760	.1637	.1802	.1041
TQ2	15.6000	12.4754	.1102	.0631	.1724
TQ3	14.1714	13.1586	-.0061	.2351	.2512
TQ4	13.5857	11.6665	.0717	.0751	.1997
TQ5	14.0143	11.0578	.1420	.1157	.1363
TQ6	14.5143	12.1375	.0497	.1305	.2167

Reliability Coefficients 6 items

Alpha = .2130 Standardized item alpha = .1984

Analysis 2 (item 3 removed)

R E L I A B I L I T Y A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	TQ1	TQ2	TQ4	TQ5	TQ6
TQ1	1.0000				
TQ2	.0669	1.0000			
TQ4	-.0686	.0212	1.0000		
TQ5	.0827	-.0645	.2498	1.0000	
TQ6	-.0399	.2054	.0800	.1433	1.0000

N of Cases = 70.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
TQ1	11.2571	10.0778	.0121	.0255	.3324
TQ2	12.5286	11.1513	.1019	.0596	.2253
TQ4	10.5143	9.8186	.1227	.0733	.2042
TQ5	10.9429	9.1271	.2124	.0998	.1114
TQ6	11.4429	9.8155	.1563	.0724	.1725

Reliability Coefficients 5 items

Alpha = .2512 Standardized item alpha = .2662

Analysis 3 (item 1 removed)

R E L I A B I L I T Y A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	TQ2	TQ4	TQ5	TQ6
TQ2	1.0000			
TQ4	.0212	1.0000		
TQ5	-.0645	.2498	1.0000	
TQ6	.2054	.0800	.1433	1.0000

N of Cases = 70.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
TQ2	9.6143	8.3273	.0780	.0520	.3619
TQ4	7.6000	6.3884	.1987	.0652	.2401
TQ5	8.0286	6.4629	.1965	.0872	.2430
TQ6	8.5286	6.5427	.2182	.0682	.2168

Reliability Coefficients 4 items

Alpha = .3324 Standardized item alpha = .3214

Analysis 4 (item 2 removed)

R E L I A B I L I T Y A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	TQ4	TQ5	TQ6
TQ4	1.0000		
TQ5	.2498	1.0000	
TQ6	.0800	.1433	1.0000

N of Cases = 70.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
TQ4	5.9571	4.7083	.2205	.0644	.2505
TQ5	6.3857	4.5012	.2694	.0777	.1479
TQ6	6.8857	5.4360	.1410	.0226	.3998

Reliability Coefficients 3 items

Alpha = .3619 Standardized item alpha = .3597

APPENDIX 3L cont.

Analysis 5 (item 6 removed)

R E L I A B I L I T Y   A N A L Y S I S - S C A L E (A L P H A)

Correlation Matrix

	TQ4	TQ5
TQ4	1.0000	
TQ5	.2498	1.0000

N of Cases = 70.0

Item-total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
TQ4	3.2286	2.1499	.2498	.0624	.
TQ5	3.6571	2.1996	.2498	.0624	.

Reliability Coefficients 2 items

Alpha = .3998 Standardized item alpha = .3998

APPENDIX 3M

Experiments 3 and 4: Analysis of variance (ANOVA) tables. Post-hoc test tables are also included where appropriate.

Experiment 3

Analysis 1: Performance on syllogistic reasoning, by strategy and syllogism form.

Table 1.1: Within-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Form	43004.20	1	43004.20	5.35	0.03
Form* strategy	2276.93	1	2276.93	0.28	0.60
Error (form)	425744.89	53	8032.92		

Table 1.2: Between-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Intercept	453120.09	1	453120.09	77.33	0.00
Strategy	0.09	1	0.09		0.99
Error	310567.18	53	5859.76		

Experiment 4

Analysis 1: Transitive inference: comparison of strategy groups on comprehension and verification response times.

Table 1.1: Within-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
RT	710229881.80	1	710229881.8	38.01	.00
RT* strategy	8612498.41	1	8612498.41	0.46	.50
Error (RT)	1177299617	63	18687295.51		

Table 1.2: Between-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Intercept	5329840065	1	5329840065.0	1411.13	.00
Strategy	474128.78	1	474128.78	.126	.72
Error	237951923.7	63	3777014.66		

Table 1.3: Bonferroni post-hoc comparisons

Response time		Mean Difference	Std. Error	Sig.
CRT	VRT	4843.33	785.63	.00

APPENDIX 3M cont.

Experiment 4 ANOVA tables cont.

Analysis 2: Transitive inference: comparison of comprehension and verification response times for two strategy groups as a function of end-anchoring.

Table 2.1: Within-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
RT	1535910852.7	1	1535910852.7	41.16	0.00
RT* strategy	15378765.85	1	15378765.85	0.41	0.52
Error (RT)	2350781886	63	373139998.19		
End-anchoring	3320739.35	1	3320739.35	3.62	0.06
End-anch* strategy	991.46	1	991.46	0.00	0.97
Error (end-anch)	57731716.88	63	916376.46		
RT* end-anch	11494750.44	1	11494750.44	11.51	0.001
RT* end-anch* strategy	48511.11	1	48511.11	0.05	0.83
Error (RT* end-anch)	62933742.89	63	998948.30		

Table 2.2: Between-subjects effects.

	Type 111 Sum of squares	df	Mean square	F	Sig.
Intercept	10805942291	1	10805942291	14440.87	0.00
Strategy	478100.43	1	478100.43	0.06	0.80
Error	471166585.6	63	7478834.69		

Analysis 3: Transitive inference: comparison of comprehension and verification response times for two strategy groups as a function of the presence of inverse adjectives.

Table 3.1: Within-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
RT	1274282932	1	1274282932.2	35.86	0.00
RT* strategy	13714467.26	1	13714467.26	0.39	0.54
Error (RT)	2238516897	63	35532014.24		
Inverse	2580616.23	1	2580616.23	2.58	0.11
Inverse* strategy	789841.71	1	789841.71	0.79	0.38
Error (Inverse)	63098261.51	63	1001559.71		
RT* inverse	64788.54	1	64788.54	0.06	0.81
RT* inverse* strategy	1524100.67	1	1524110.67	1.40	0.24
Error (RT* inverse)	68443345.54	63	1086402.31		

APPENDIX 3M cont.

Experiment 4 ANOVA tables cont.

Table 3.2: Between-subjects effects.

	Type 111 Sum of squares	df	Mean square	F	Sig.
Intercept	10536000722	1	10536000722	1371.37	0.00
Strategy	120851.92	1	120851.92	0.02	0.90
Error	484020012.1	63	7682857.34		

Analysis 4: Sentence-picture verification: Comparison of verification response times across trial types, as a function of truth and polarity.

Table 4.1: Within-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Truth	2736910.37	1	2736910.37	74.80	0.00
Error (truth)	2524859.74	69	36592.17		
Polarity	6096809.48	1	6096809.48	32.41	
Error (polarity)	12979475.41	69	18810834		0.00
Truth* polarity	342439.41	1	342439.41	6.86	0.01
Error (truth* polarity)	3445781.51	69	49938.86		

Analysis 5: Sentence-picture verification: comparison of comprehension and verification response times by strategy group.

Table 5.1: Within-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
RT	13127068.04	1	13127068.04	36.97	0.00
RT* strategy	8797.03	1	8797.03	0.03	0.88
Error (RT)	22372492.91	63	355118.94		

Table 5.2: Between-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Intercept	347416691.3	1	347416691.26	829.18	0.00
Strategy	704655.52	1	704655.52	1.68	0.20
Error	26396345.09	63	418989.61		

APPENDIX 3M cont.

Experiment 4 ANOVA tables cont.

Analysis 6: Sentence-picture verification: Comparison of verification response times across trial types, as a function of truth, polarity and strategy group.

Table 6.1: Within-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Truth	2357492.32	1	2357492.32	0.00	0.00
Truth* strategy	23989.34	1	23989.34	0.61	0.44
Error (truth)	2463865.22	63	39108.97		
Polarity	6036476.83	1	6036476.83	30.64	0.00
Polarity* strategy	169725.09	1	169725.09	0.86	0.36
Error (polarity)	12411924.37	63	197014.67		
Truth* polarity	371480.46	1	371480.46	7.74	0.01
Truth* polarity* strategy	116941.60	1	116941.60	2.44	0.12
Error (truth* polarity)	3023775.73	63	47996.44		

Table 6.2: Between-subjects effects

	Type 111 Sum of squares	df	Mean square	F	Sig.
Intercept	454845433.1	1	454845433.07	618.46	0.12
Strategy	1187982.80	1	1187982.80	1.62	
Error	46333071.24	63	735445.58		

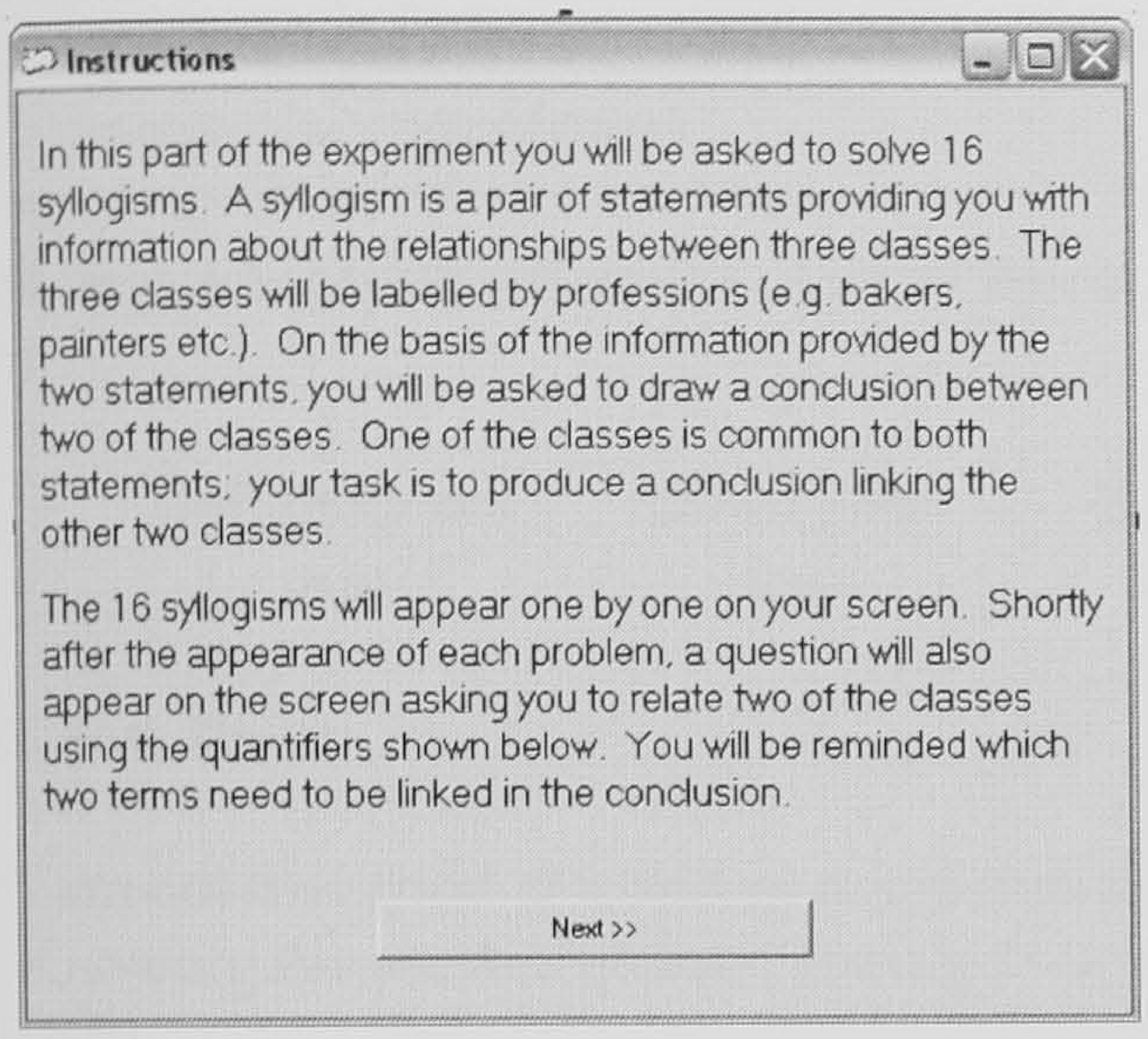
APPENDIX 4A

Experiments 5 and 6: Syllogisms presented in reasoning performance task. Problems are shown in order presented, together with figure, mental model count and conclusion form as indicators of relative difficulty.

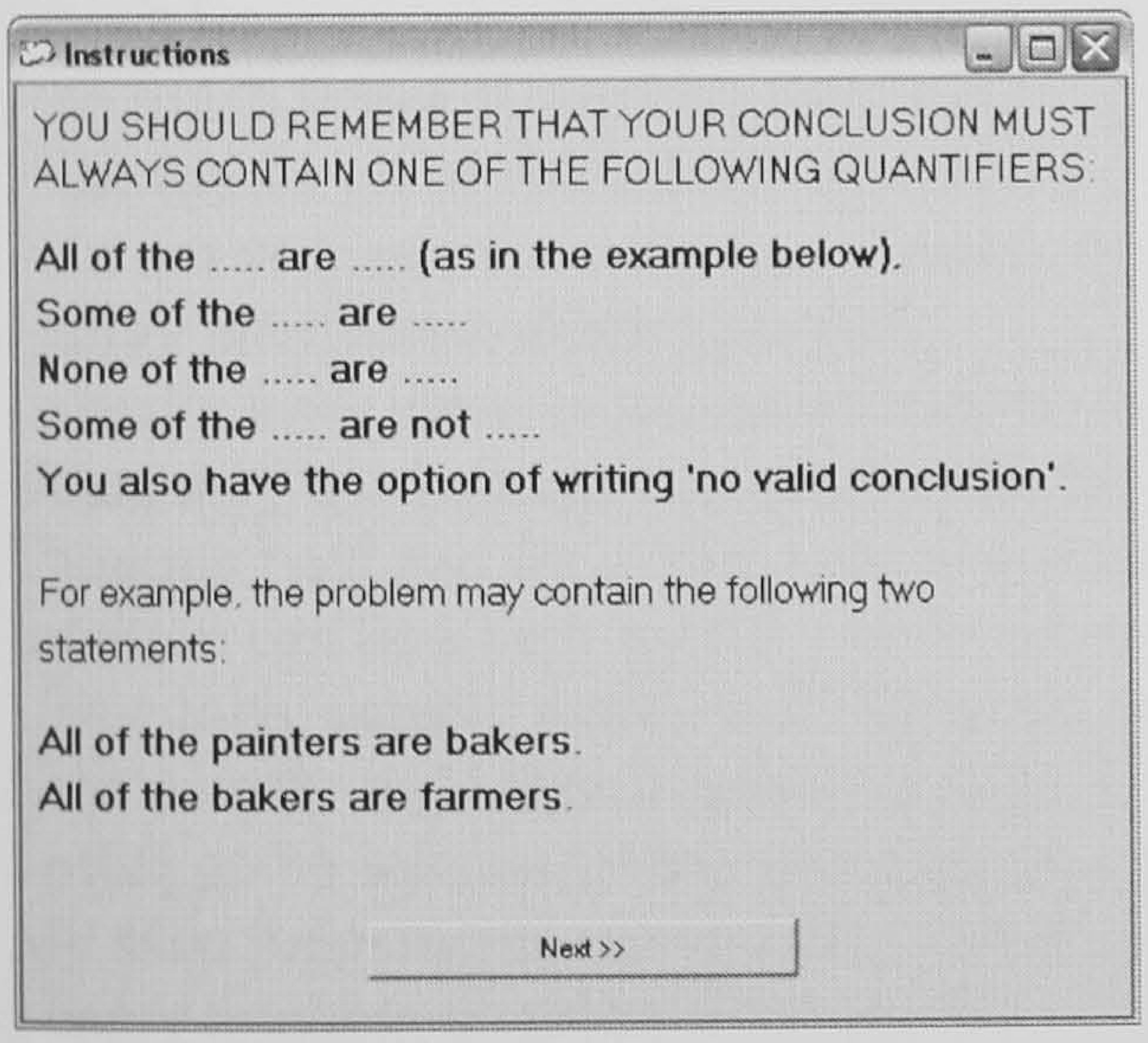
No.	Syllogism	Figure	Mental model count	Conclusion form
1	Some of the bankers are managers. All of the bankers are chefs. Some of the managers are chefs	B-A B-C	1	same
2	None of the journalists are pilots. Some of the journalists are divers. Some of the divers are not pilots	B-A B-C	3	different
3	None of the farmers are clowns. All of the butchers are clowns. None of the farmers are butchers	A-B B-C	3	same
4	None of the journalists are mechanics. All of the drivers are journalists. None of the drivers are mechanics	B-A C-B	1	same
5	All of the librarians are journalists. None of the journalists are reporters None of the librarians are reporters	A-B B-C	1	same
6	None of the jugglers are clowns. Some of the clowns are cricketers. Some of the cricketers are not jugglers	A-B B-C	3	different
7	All of the clowns are chemists. Some of the parents are clowns. Some of the parents are chemists	B-A C-B	1	same
8	Some of the athletes are singers. All of the singers are dancers. Some of the athletes are dancers	A-B B-C	1	same
9	Some of the lawyers are bankers. None of the bankers are sailors. Some of the lawyers are not sailors	A-B B-C	3	different
10	None of the bankers are teachers. Some of the sculptors are bankers. Some of the sculptors are not teachers	B-A C-B	3	different
11	None of the historians are singers. Some of the hairdressers are singers. Some of the hairdressers are not historians	A-B C-B	3	different
12	Some of the surgeons are journalists. None of the golfers are journalists. Some of the surgeons are not golfers	A-B C-B	3	different
13	All of the nurses are bankers. None of the politicians are bankers. None of the nurses are politicians	A-B C-B	1	same
14	Some of the singers are carpenters. None of the weavers are singers. Some of the carpenters are not weavers	B-A C-B	3	different
15	Some of the clowns are pianists. None of the clowns are surfers. Some of the pianists are not surfers	B-A B-C	3	different
16	All of the singers are hikers. Some of the singers are publicans. Some of the hikers are publicans	B-A B-C	1	same

APPENDIX 4B

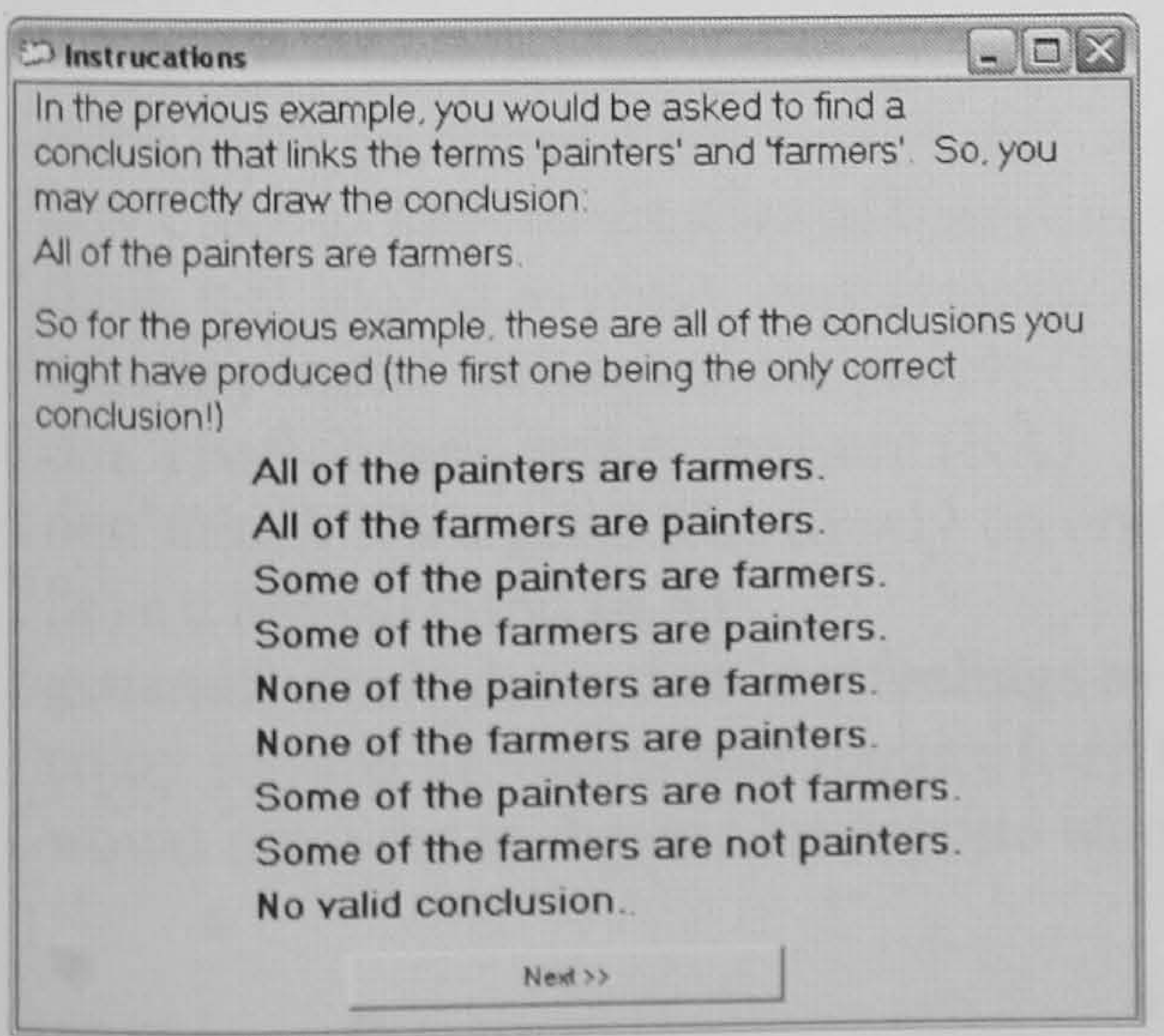
Experiments 5 and 6: Instructions and practice item presented to participants prior to computerised syllogistic reasoning performance task (as per Capon, Handley and Dennis, 2003).



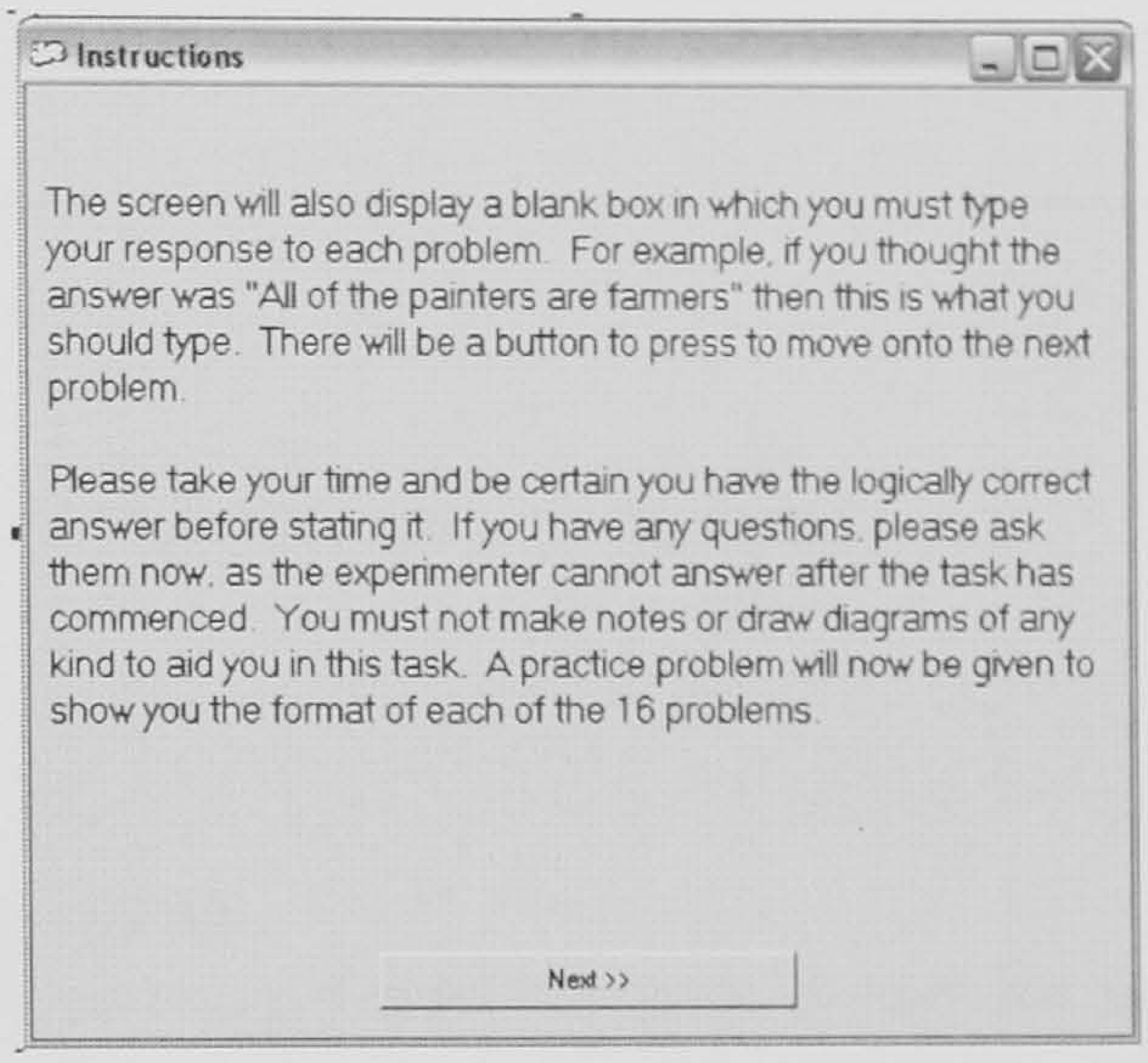
1.



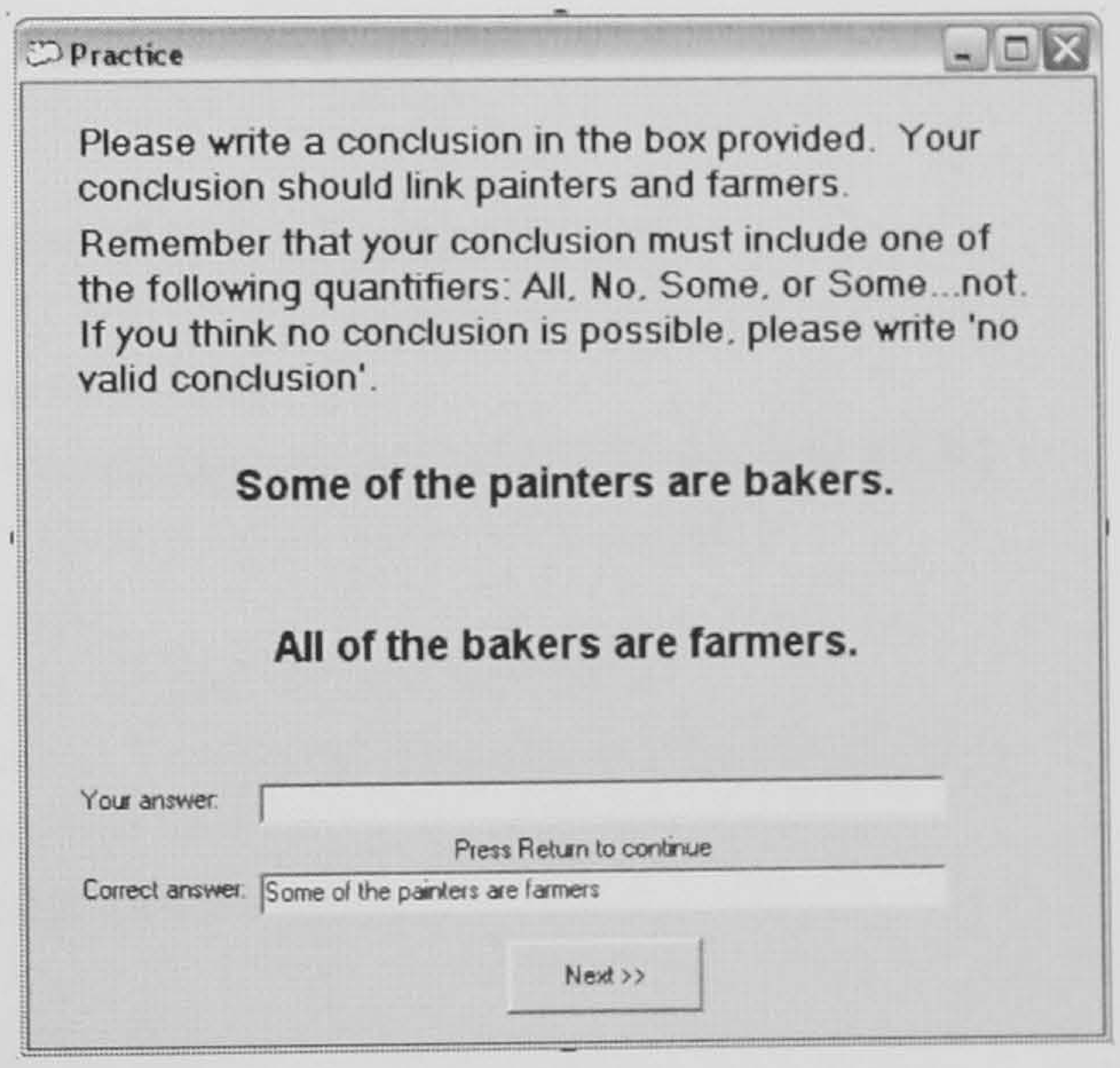
2.



3.



4.



5.

APPENDIX 4C

Experiment 5: Rational Experiential Inventory (REI) items. The construct measured by each item is presented in parentheses as per the following key.

RA: Rational ability
RE: Rational engagement
EA: Experiential ability
EE: Experiential engagement

Each item was presented with a scale as below.

1	2	3	4	5
Definitely not true	Mostly not true	Equally true and false	Mostly true	Definitely true

REI Test Items

1. I am not that good at working out complicated problems (RA)
2. Knowing the answer without having to understand the reasoning behind it is good enough for me (RE)
3. I try to avoid situations that require thinking in depth about something (RE)
4. I don't have a very good sense of intuition (EA)
5. I am much better at working things out logically that most people (RA)
6. I like to rely on my intuitive impressions (EE)
7. I am not very good at solving problems that require careful logical analysis (RA)
8. I enjoy intellectual challenges (RE)
9. I usually have clear, explainable reasons for my decisions (RA)
10. Using my "gut feelings" usually works well for me in working out problems in my life (EA)
11. Thinking hard and for a long time about something gives me little satisfaction (RE)
12. Intuition can be a very useful way to solve problems (EE)
13. Using logic usually works well for me in working out problems in my life (RA)
14. I can usually feel when a person is right or wrong, even if I can't explain how I know (EA)
15. I often go by instincts when deciding on a course of action (EE)
16. My snap judgements are probably not as good as most people's (EA)
17. I enjoy thinking in abstract terms (RE)
18. I don't like situations in which I have to rely on intuition (EE)
19. I prefer complex to simple problems (RE)
20. I suspect my hunches are inaccurate as often as they are accurate (EA)
21. I think there are times when one should rely on one's intuition (EE)
22. I don't like to have to do a lot of thinking (RE)
23. I think it is foolish to make important decisions based on feelings (EE)
24. I trust my initial feelings about people (EA)
25. I don't reason well under pressure (RA)
26. I don't think it's a good idea to rely on one's intuition for important decisions (EE)
27. I have a logical mind (RA)
28. I generally don't depend on my feelings to help me make decisions (EE)
29. I enjoy solving problems that require hard thinking (RE)
30. I would not want to depend on anyone who described himself or herself as intuitive (EE)

31. I am not a very analytical thinker (RA)
32. I tend to use my heart as a guide for my actions (EE)
33. Learning new ways to think would be very appealing to me (RE)
34. I believe in trusting my hunches (EA)
35. I have no problem in thinking things through carefully (RA)
36. When it comes to trusting people, I can usually rely on my gut feelings (EA)
37. Thinking is not my idea of an enjoyable activity (RE)
38. Reasoning thing out carefully is not one of my strong points (RA)
39. If I were to rely on my gut feelings, I would often make mistakes (EA)
40. I hardly ever go wrong when I listen to my deepest gut feelings to find an answer (EA)

Experiment 5: Matrix of correlations between all measures for entire sample (N=110)

	Verbal ability (SILS)		IQ. est. from SILS	Spatial ability (VMRT)	Cognitive style (REI)				Cognitive style (CSA)					
									Verbal-imagery			Wholist-analytic		
	Vocab.	Abstrac			Rational ability	Rational engage.	Exper ability	Exper engage.	ratio	speed	correct	ratio	speed	correct
Syllogisms	0.27**	0.24*	0.31**	0.23*	-0.03	0.04	-0.25**	-0.07	0.16	0.23	-0.05	-0.1	0.05	-0.12
SILS vocab.	-	0.32**	0.72**	0.09	-0.10	-0.03	-0.07	0.10	0.27**	0.16	-0.05	-0.02	0.22*	0.08
SILS abs.	-	-	0.86**	0.31**	0.06	0.10	-0.15	0.09	0.0.05	-0.01	0.07	0.1	-0.08	0.03
IQ	-	-	-	0.29**	-0.02	0.12	-0.13	0.12	0.16	0.07	-0.01	0.06	0.06	0.05
VMRT	-	-	-	-	-0.03	0.12	-0.13	0.11	0.06	0.08	0.08	-0.04	-0.10	0.08
RA	-	-	-	-	-	0.02	0.11	0.18	0.00	-0.02	-0.17	-0.05	-0.05	-0.11
RE	-	-	-	-	-	-	0.07	0.01	0.07	0.13	-0.10	-0.05	0.08	-0.01
EA	-	-	-	-	-	-	-	0.19*	-0.03	0.14	-0.01	-0.02	0.13	0.06
EE	-	-	-	-	-	-	-	-	0.01	-0.01	0.01	0.06	-0.01	0.00
VI ratio	-	-	-	-	-	-	-	-	-	0.41**	-0.23	-0.03	0.15	-0.21*
VI speed	-	-	-	-	-	-	-	-	-	-	-0.11	-0.15	0.41**	-0.12
VI correct	-	-	-	-	-	-	-	-	-	-	-	-0.03	-0.15	-0.35**
WA ratio	-	-	-	-	-	-	-	-	-	-	-	-	-0.30**	0.21*
WA speed	-	-	-	-	-	-	-	-	-	-	-	-	-	0.06

** sig. at 0.01 level * sig. at 0.05 level

APPENDIX 4E

Experiment 5: Correlations between cognitive style measures for verbal and spatial reasoners.

		Cognitive style (REI)				Cognitive style (CSA)					
						Verbal-imagery			Wholist-analytic		
		Rational ability	Rational engage.	Experien ability	Experien engage.	ratio	speed	correct	ratio	speed	correct
Syllogisms	Verbal	-0.37	-0.13	-0.37**	-0.18	0.17	0.04	0.01	-0.09	0.01	0.03
	Spatial	-0.18	-0.05	-0.18	-0.01	-0.16	-0.07	0.12	0.01	-0.26	0.12
RA	Verbal	-	0.01	0.07	0.14	0.07	-0.16	-0.21	-0.02	-0.09	-0.21
	Spatial	-	-0.28	0.20	0.23	-0.26	0.12	-0.12	-0.16	-0.06	0.19
RE	Verbal	-	-	0.00	0.06	-0.05	-0.05	-0.16	-0.12	0.05	0.03
	Spatial	-	-	0.17	0.17	0.06	-0.09	0.24	0.30	-0.07	0.07
EA	Verbal	-	-	-	0.14	-0.08	0.15	-0.03	-0.11	0.21	0.08
	Spatial	-	-	-	0.36*	0.06	0.22	0.15	0.18	0.04	0.21
EE	Verbal	-	-	-	-	0.04	-0.15	-0.05	0.09	-0.08	-0.09
	Spatial	-	-	-	-	-0.04	-0.01	-0.01	0.02	0.03	0.14
VI ratio	Verbal	-	-	-	-	-	0.00	-0.15	0.08	-0.05	-0.11
	Spatial	-	-	-	-	-	0.17	0.03	-0.21	0.36*	-0.01
VI speed	Verbal	-	-	-	-	-	-	0.00	-0.19	0.49**	0.28*
	Spatial	-	-	-	-	-	-	0.17	-0.04	0.54**	0.15
VI correct	Verbal	-	-	-	-	-	-	-	-0.12	-0.10	0.28*
	Spatial	-	-	-	-	-	-	-	0.08	-0.06	0.36*
WA ratio	Verbal	-	-	-	-	-	-	-	-	-0.40**	0.19
	Spatial	-	-	-	-	-	-	-	-	-0.01	0.32
WA speed	Verbal	-	-	-	-	-	-	-	-	-	0.08
	Spatial	-	-	-	-	-	-	-	-	-	0.15

** sig. at 0.01 level * sig. at 0.05 level

APPENDIX 5A

Experiment 6: Computer presented instructions for Simple Word Span task.

INSTRUCTIONS

Simple Word Span Instructions

In this part of the experiment, you must remember sets of words that will appear on the screen. The number of words to remember in each set will increase throughout the test.

An example of a set of words can be seen below:

The screen displays the word **GUITAR**

Then the screen displays the word **LORRY**

Then the screen displays

Your task is to remember the set of words, in the order they were presented. For the above example, you would write down the words:GUITAR LORRY

You are provided with a response sheet to mark down your responses to each set of words. Please have a look at the sheet now to ensure you understand how to mark your responses.

If you DO NOT understand the task, please tell the experimenter now.

If you DO understand the task, please click the button below and you will be given three practice sets of words.

Start the Practice Set

APPENDIX 5B

Experiment 6: Words presented for Simple Word Span task, shown in order and in sets as presented to all participants.

Practice Sets

Berry	Lemon	Jersey
Cattle	Singer	Apple

2 Word Sets

Wallet	Movie	Banner	Silver	Planet
Coffee	Kitten	Forest	Cannon	Table

3 Word Sets

Adult	Mother	Banker	Dinner	Flower
Carpet	Tower	Insect	Ankle	Jockey
Dollar	Hotel	Otter	Tunnel	Hammer

4 Word Sets

Bucket	Medal	Battle	Ribbon	Manor
Pepper	Organ	Sugar	Cabin	Supper
Cotton	Collar	Angel	Pocket	Dancer
Stable	Doctor	Rabbit	Bishop	Racket

5 Word Sets

Anchor	Metal	Cherry	Valley	Island
Liver	Butter	Engine	Bubble	Party
Needle	Father	Pillow	Driver	Canoe
Window	River	Statue	Circle	Basket
Mirror	Candy	Market	Locker	Letter

6 Word Sets

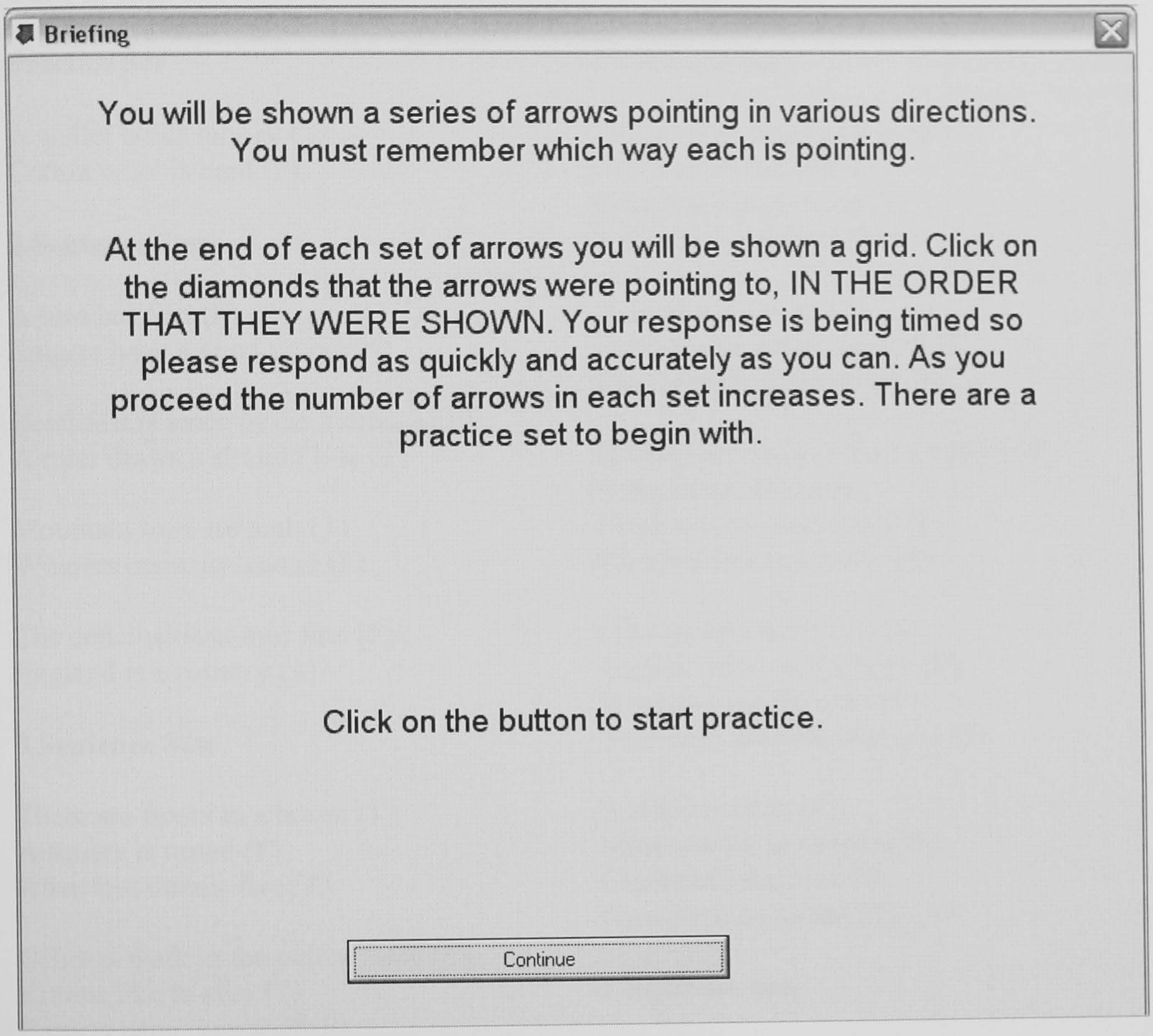
Puzzle	Cement	Bottle	Artist	Giant
Marble	Ferry	Turtle	Willow	Barrel
Tiger	Infant	Candle	Ticket	Eagle
Hunter	Python	Gravel	Summer	Rocket
Saddle	Woman	Jacket	Circus	Cellar
Basin	Lawyer	Maple	Lobby	Demon

7 Word Sets

Mutton	Cigar	Honey	Coral	Music
Sister	Rubber	Cable	Monkey	Penny
Buckle	Elbow	Prison	Shower	Olive
Office	Powder	Money	Fabric	Resort
Clover	Button	Number	Bullet	Salad
Lever	Garden	Brandy	Heaven	Alley
Paper	Linen	Water	Arrow	Walnut

APPENDIX 5C

Experiment 6: Computer presented instructions for Simple Spatial Arrow Span task.



APPENDIX 5D

Experiment 6: The Complex Verbal Sentence Span. Sentence sets in order presented to all participants with indication of correct response (T = True, F = False). Final word in each sentence to be recalled.

Practice Set

A wallet holds money (T)
Cotton wool is hard (F)

2 Sentence Sets

A bird has feet (F)
Singers have a good voice (T)

Breakfast is eaten in the morning (T)
A ruler draws a straight line (T)

Mountain tops are high (T)
Winners come in second (F)

The conclusion comes first (F)
England is a country (T)

3 Sentence Sets

There are floors in a house (T)
A square is round (F)
After four comes five (T)

Officers work in the police force (T)
Kittens like to play (T)
Carpets hang on a wall (F)

The ace of hearts is black (F)
English drive on the left (T)
Twenty months make a year (F)

Winners come in last (F)
A prince is a woman (F)
Tomorrow is in the future (T)

A penny is a coin (T)
The sun comes out at night (F)
There are toes on a hand (F)

4 Sentence Sets

The number three is even (F)
Fish live on the land (F)
Gold has a high value (T)
A lemon is a flower (F)

A dinghy is a boat (T)
Cartoon characters are real (F)
The leader is at the front (T)

Billions are smaller than a million (F)
Polar bears are white
Elephants are very small (F)
Families live in a home (T)

Lifts go up and down (T)
English drive on the right (F)
Windows can be open (T)
Memories are from the past (T)

A slope is level (F)
Miners work in school (F)
Clocks tell the time (T)
Cars drive on a road (T)

5 Sentence Sets

A brother is a girl
Keyboards are used to type (T)
Seven comes before four (F)
People live in a city (T)
Football is a sport (T)

At night time it is light (F)
Electricity is a form of power (T)
Water is found down a well (T)
Rock music is sung in church (F)
Legs are part of the face (F)

APPENDIX 5D continued.

Experiment 6: Complex Verbal Sentence Span test items.

5 Sentence Sets cont.

Judges work in court (T)
Crops grow in a wood (F)
Museums are open to the public (T)
One person makes a group (F)
Carpets cover the floor (T)

Water lights a fire (F)
Planes fly around the world (T)
Seventeen comes before three (F)
Chess is a game (T)
A frog has hair (F)

The hospital sends mail (F)
Towns are full of sand (F)
The ocean contains water (T)
An apple is a fruit (T)
Cars are made of bone (F)

6 Sentence Sets

A hand has a palm (T)
You find fur in a salad (F)
Plug rhymes with leaf (F)
Bald people have a fringe (F)
Haunted houses have a ghost (T)
Brownies go to camp (T)

Tea is kept in a flask (T)
Lions sleep in a cradle (F)
Cigarettes come in a pack (T)
The sea is made of gravy (F)
Cement is a spice (F)
You wake up with an alarm (T)

Canoeists use a paddle (T)
You sit on a stool (T)
Vodka is found in a lake (F)
Ninety seconds equals one hour (F)
Pigs take off from an airport (F)
red is the colour of lime (F)

6 Sentence Sets cont.

You keep animals in a file (F)
You eat with a sword (F)
The sea is near the coast (T)
A cinema has a spire (F)
You can mow a lawn (T)
People walk along a pier (T)

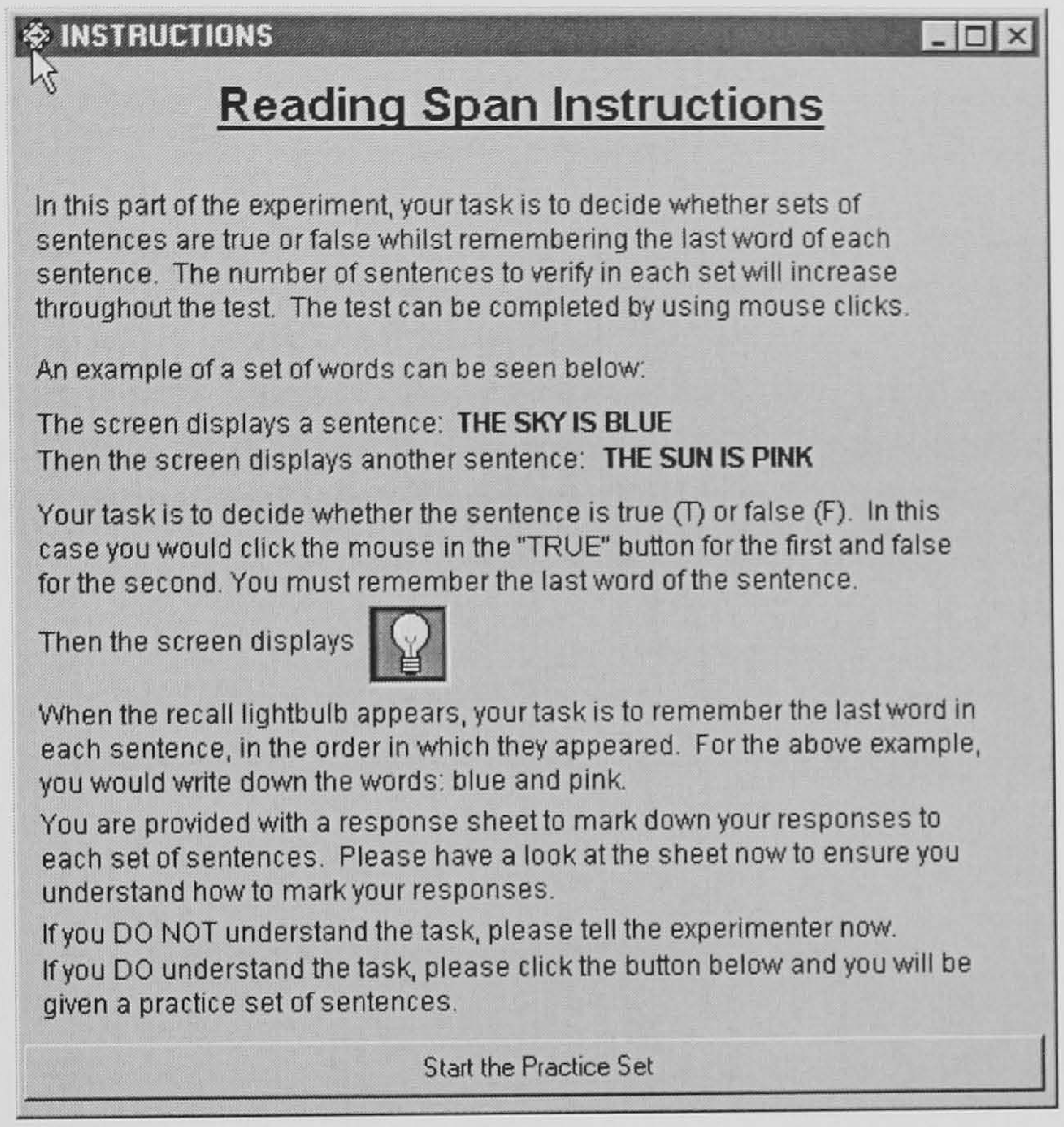
You can suck your thumb (T)
Peasants live in a palace (F)
You can drink from a goblet (T)
A daffodil grows from a bulb (T)
You stay dry in the rain (F)
Pubs sell crisps and beer (T)

APPENDIX 5E

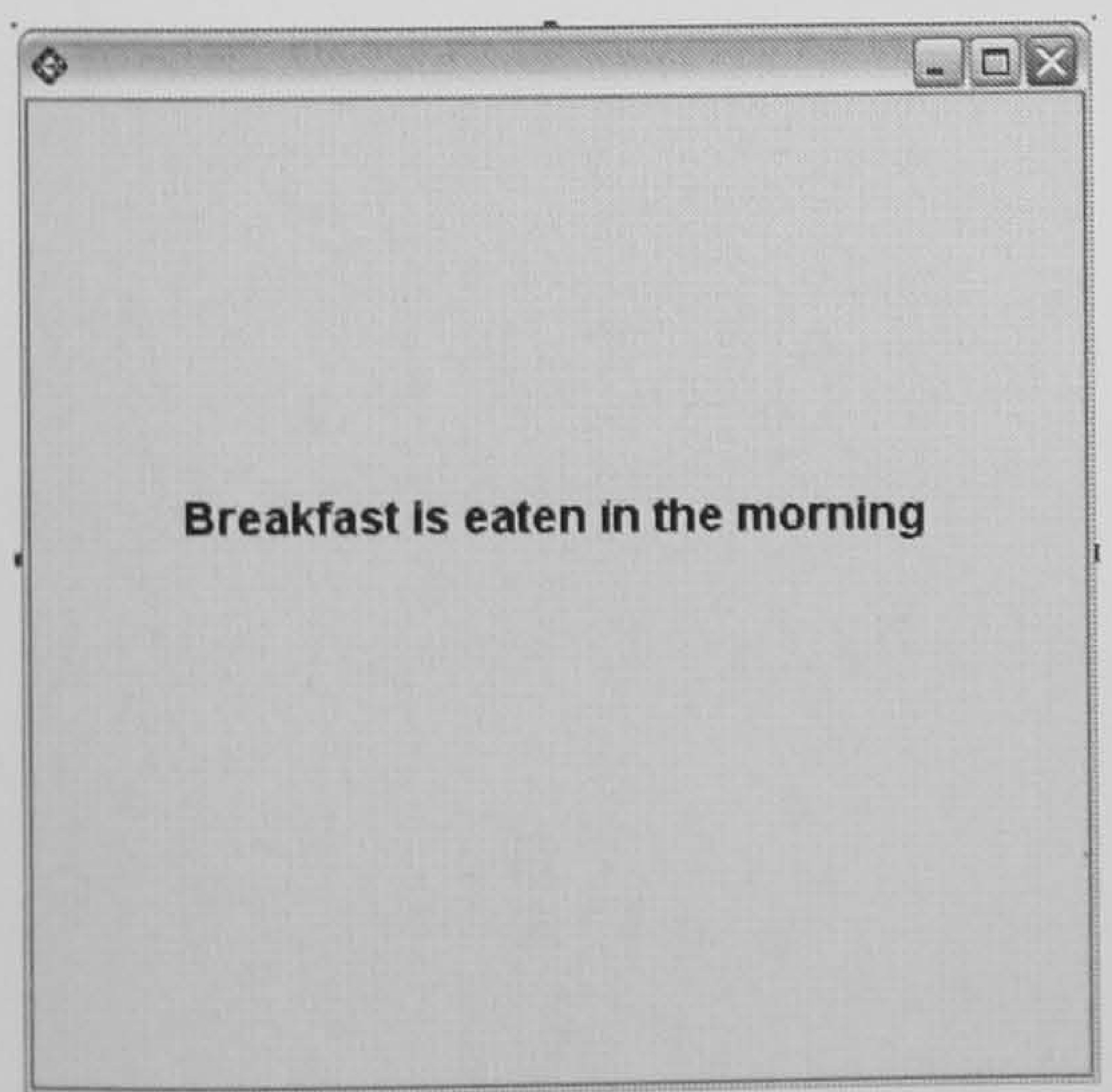
Experiment 6: The Complex Verbal Sentence Span. Written instructions and example of test screen format as presented to all participants.

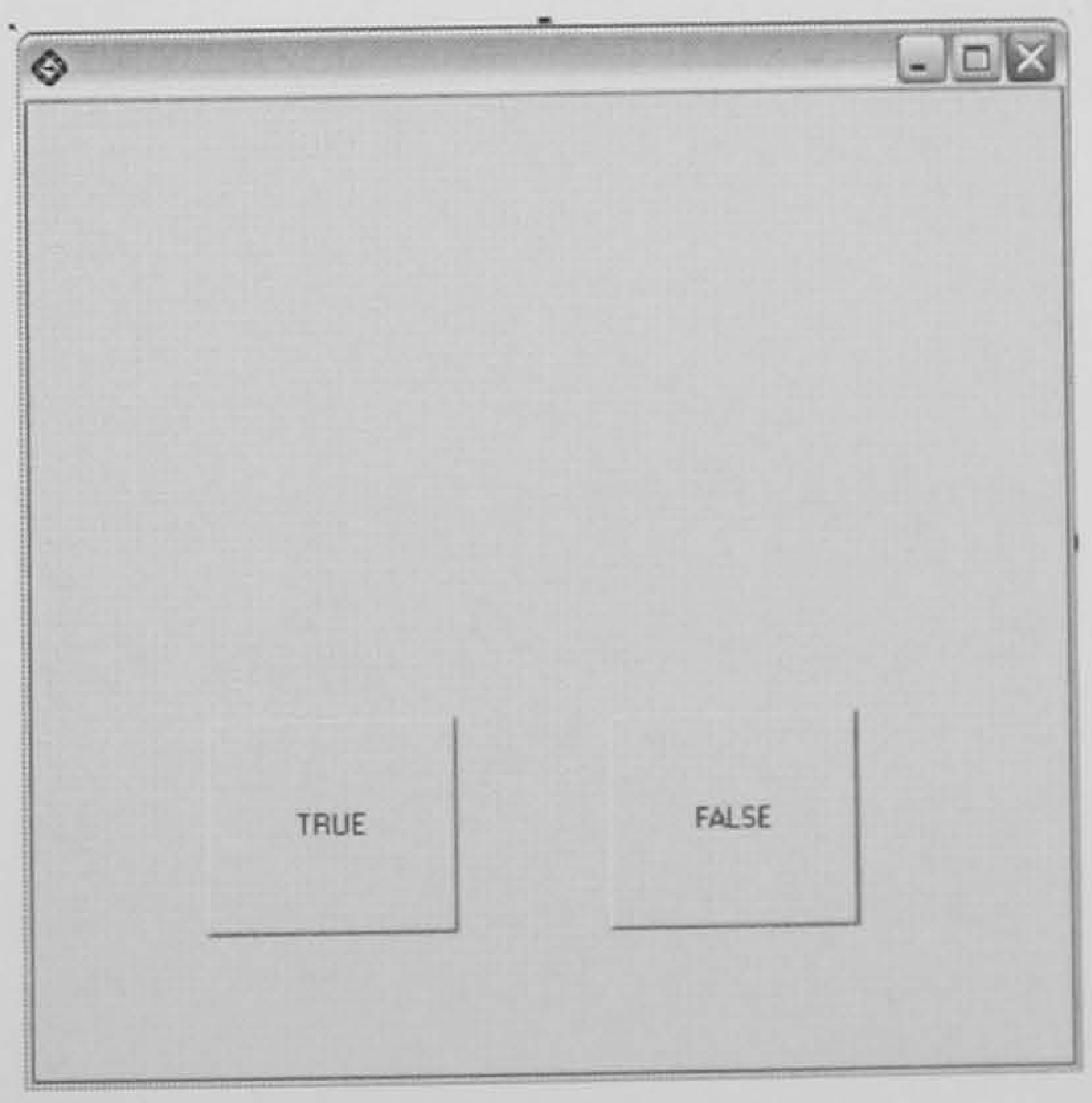
---

1. Instruction Screen



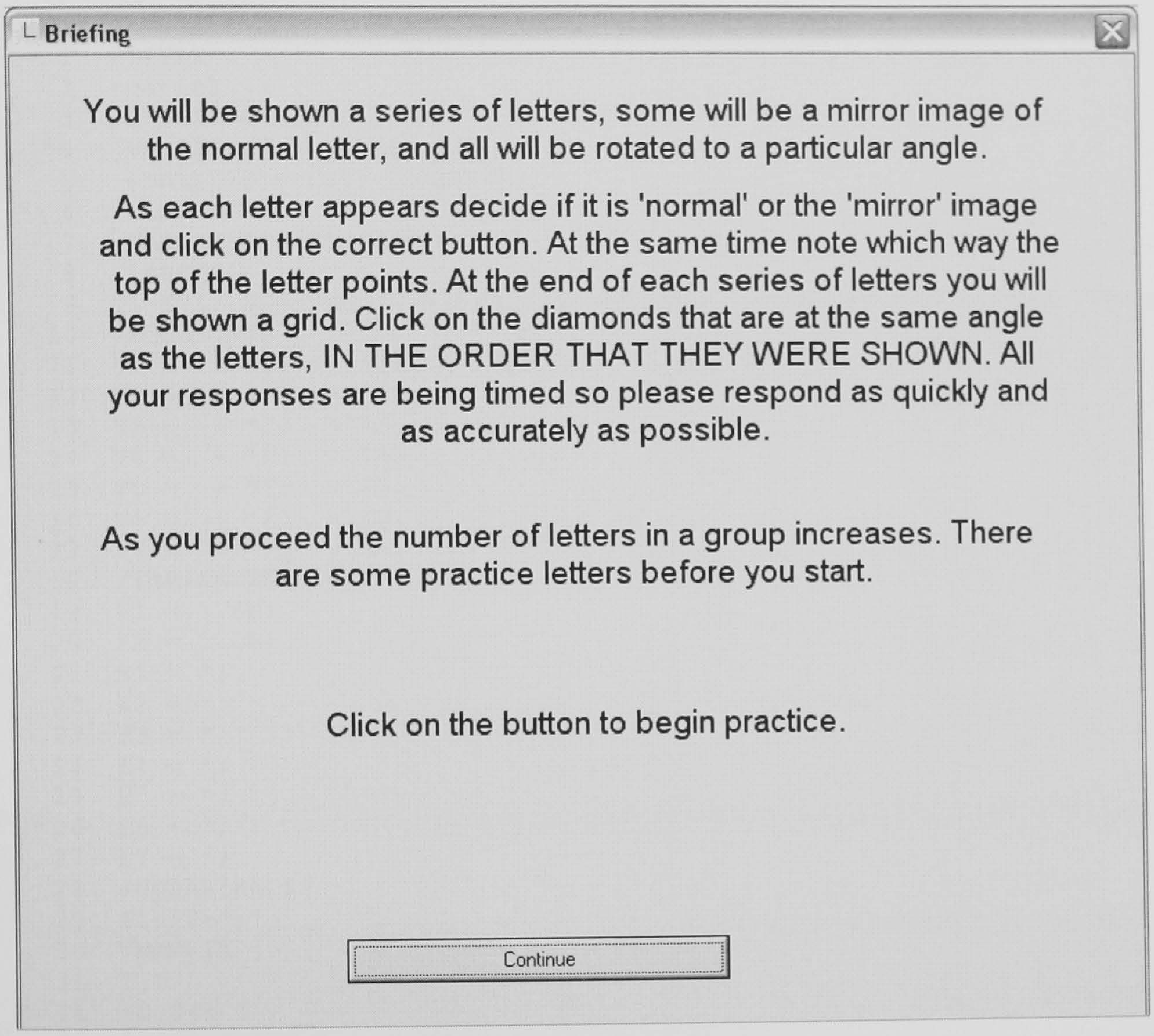
2. Typical test item presentation and response screens





APPENDIX 5F

Experiment 6: Complex Spatial Letter Span. Instruction screen presented to all participants.



APPENDIX 5G

Experiment 6: Full EQS output of Confirmatory Factor Analysis - Model 1.

```
EQS, A STRUCTURAL EQUATION PROGRAM          MULTIVARIATE SOFTWARE, INC.COPYRIGHT
BY P.M. BENTLER                     VERSION 5.7b (C) 1985 - 1998.

    PROGRAM CONTROL INFORMATION

    /TITLE
    spatial
    /SPECIFICATIONS
     VARIABLES=7; CASES=48;
     ANALYSIS = COV; GROUP=2;
     METHODS=ML;
     MATRIX=CORRELATION;
    /LABELS
    V1=SYL; V2=WORD; V3=SENT; V4=SENTERR; V5=ARROW; V6=LETTERS; V7=LETTERR;
    /EQUATIONS
    V1 =  + *F1  + *F2  + E1;
    V2 =  + *F1  + E2;
    V3 =  + *F1  + E3;
    V4 =  + *F1  + E4;
    V5 =  + *F2  + E5;
    V6 =  + *F2  + E6;
    V7 =  + *F2  + E7;
    /VARIANCES
    F1 = 1.00;
    F2 = 1.00;
    E1 = *;
    E2 = *;
    E3 = *;
    E4 = *;
    E5 = *;
    E6 = *;
    E7 = *;
    /COVARIANCES
    F1,F2=*;
    /MATRIX
    1
    -0.046 1
    0.368 0.529 1
    -0.133 -0.154 -0.223 1
    0.371 -0.022 0.256 -0.198 1
    0.192 0.137 0.297 -0.331 0.594 1
    -0.152 -0.091 -0.056 -0.099 -0.205 -0.132 1
    /STANDARD DEVIATIONS
    3.629 19.147 17.765 5.947 9.925 15.934 8.659
    /TECHNICAL
    ITR=50;
    /END

      42 CUMULATED RECORDS OF INPUT MODEL FILE WERE READ (GROUP   1)
```

```
PAGE :  2
  EQS/EM386 Licensee: Alison Bacon

    PROGRAM CONTROL INFORMATION

    /TITLE
    verbal
    /SPECIFICATIONS
     VARIABLES=7; CASES=93;
     METHODS=ML;
     MATRIX=CORRELATION;ANALYSIS = COV;
    /LABELS
    V1=SYL; V2=WORD; V3=SENT; V4=SENTERR; V5=ARROW; V6=LETTERS; V7=LETTERR;
    /EQUATIONS
    V1 =  + *F1  + *F2  + E1;
    V2 =  + *F1  + E2;
    V3 =  + *F1  + E3;
    V4 =  + *F1  + E4;
    V5 =  + *F2  + E5;
    V6 =  + *F2  + E6;
    V7 =  + *F2  + E7;
    /VARIANCES
    F1 = *;
    F2 = *;
    E1 = *;
    E2 = *;
    E3 = *;
    E4 = *;
    E5 = *;
    E6 = *;
    E7 = *;
    /COVARIANCES
    F1,F2=*;
    /MATRIX
    1
    0.357 1
    0.367 0.568 1
    -0.274 -0.265 -0.226 1
    0.301 0.260 0.219 -0.237 1
    0.204 0.362 0.373 -0.328 0.332 1
    -0.177 -0.054 -0.037 0.031 -0.282 -0.241 1
    /STANDARD DEVIATIONS
    2.701 12.721 15.678 5.630 8.779 12.857 8.692
    /CONSTRAINTS
    (1,V2,F1)= (2, V2,F1);
    (1,V3,F1)= (2, V3,F1);
    (1,V4,F1)= (2, V4,F1);
    (1,V5,F2)= (2, V5,F2);
    (1,V6,F2)= (2, V6,F2);
    (1,V7,F2)= (2, V7,F2);
    /TECHNICAL
    ITR=50;
    /PRINT
    FIT=ALL;
    /LMTEST
    /END
      93 CUMULATED RECORDS OF INPUT MODEL FILE WERE READ (GROUP   2)
```

```
  TITLE:   spatial
09/16/03        PAGE :  3
  EQS/EM386 Licensee: Alison Bacon
  COVARIANCE  MATRIX TO BE ANALYZED:   7 VARIABLES (SELECTED FROM   7 VARIABLES)
  BASED ON    48 CASES.
```

		SYL V 1	WORD V 2	SENT V 3	SENTERR V 4	ARROW V 5
SYL	V 1	13.170				
WORD	V 2	-3.196	366.608			
SENT	V 3	23.725	179.937	315.595		
SENTERR	V 4	-2.870	-17.536	-23.560	35.367	
ARROW	V 5	13.363	-4.181	45.137	-11.687	98.506
LETTERS	V 6	11.102	41.797	84.071	-31.365	93.938
LETTERR	V 7	-4.776	-15.087	-8.614	-5.098	-17.618

		LETTERS V 6	LETTERR V 7
LETTERS	V 6	253.892	
LETTERR	V 7	-18.212	74.978

```
  BENTLER-WEEKS STRUCTURAL REPRESENTATION:

        NUMBER OF DEPENDENT VARIABLES =  7
            DEPENDENT V'S :     1    2    3    4    5    6    7

        NUMBER OF INDEPENDENT VARIABLES =  9
            INDEPENDENT F'S :     1    2
            INDEPENDENT E'S :     1    2    3    4    5    6    7

        NUMBER OF FREE PARAMETERS =  16
        NUMBER OF FIXED NONZERO PARAMETERS =    9
```

```
TITLE:   verbal
09/16/03         PAGE :  4
  EQS/EM386 Licensee: Alison Bacon
  COVARIANCE  MATRIX TO BE ANALYZED:   7 VARIABLES (SELECTED FROM   7 VARIABLES)
  BASED ON    93 CASES.
```

		SYL V 1	WORD V 2	SENT V 3	SENTERR V 4	ARROW V 5
SYL	V 1	7.295				
WORD	V 2	12.266	161.824			
SENT	V 3	15.541	113.282	245.800		
SENTERR	V 4	-4.167	-18.979	-19.948	31.697	
ARROW	V 5	7.137	29.036	30.143	-11.714	77.071
LETTERS	V 6	7.084	59.207	75.186	-23.742	37.473
LETTERR	V 7	-4.155	-5.971	-5.042	1.517	-21.519

		LETTERS V 6	LETTERR V 7
LETTERS	V 6	165.302	
LETTERR	V 7	-26.932	75.551

```
  BENTLER-WEEKS STRUCTURAL REPRESENTATION:

        NUMBER OF DEPENDENT VARIABLES =  7
            DEPENDENT V'S :     1    2    3    4    5    6    7

        NUMBER OF INDEPENDENT VARIABLES =  9
            INDEPENDENT F'S :     1    2
            INDEPENDENT E'S :     1    2    3    4    5    6    7

        NUMBER OF FREE PARAMETERS =  18
        NUMBER OF FIXED NONZERO PARAMETERS =    7

  3RD STAGE OF COMPUTATION REQUIRED      6101 WORDS OF MEMORY.
  PROGRAM ALLOCATED    100000 WORDS

  DETERMINANT OF INPUT MATRIX IN GROUP  1 IS   0.23014E+14

  DETERMINANT OF INPUT MATRIX IN GROUP  2 IS   0.24603E+13

 IN ITERATION #  1, MATRIX W_CFUNCT MAY NOT BE POSITIVE DEFINITE.
  YOU HAVE BAD START VALUES TO BEGIN WITH.
  IF ABOVE MESSAGE APPEARS ON EVERY ITERATION, PLEASE PROVIDE BETTER START
VALUES AND RE-RUN THE JOB.

 IN ITERATION #  1, MATRIX W_CFUNCT MAY NOT BE POSITIVE DEFINITE.
  YOU HAVE BAD START VALUES TO BEGIN WITH.
  IF ABOVE MESSAGE APPEARS ON EVERY ITERATION, PLEASE PROVIDE BETTER START
VALUES AND RE-RUN THE JOB.
```

```
  TITLE:   spatial
09/16/03       PAGE :  5
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  PARAMETER ESTIMATES APPEAR IN ORDER,
  NO SPECIAL PROBLEMS WERE ENCOUNTERED DURING OPTIMIZATION.

  RESIDUAL COVARIANCE MATRIX  (S-SIGMA) :
```

		SYL V 1	WORD V 2	SENT V 3	SENTERR V 4	ARROW V 5
SYL	V 1	0.000				
WORD	V 2	-18.599	-22.925			
SENT	V 3	3.144	-7.034	9.550		
SENTERR	V 4	0.037	8.879	11.734	-2.571	
ARROW	V 5	2.292	-38.444	-0.643	-5.219	3.724
LETTERS	V 6	-5.581	-9.835	15.083	-21.619	2.903
LETTERR	V 7	-0.444	-1.678	9.302	-7.629	6.024

		LETTERS V 6	LETTERR V 7
LETTERS	V 6	-4.929	
LETTERR	V 7	17.414	-4.781

```
              AVERAGE ABSOLUTE  COVARIANCE  RESIDUALS    =     8.6435
     AVERAGE OFF-DIAGONAL ABSOLUTE  COVARIANCE  RESIDUALS    =     9.2160


  STANDARDIZED RESIDUAL MATRIX:
```

		SYL V 1	WORD V 2	SENT V 3	SENTERR V 4	ARROW V 5
SYL	V 1	0.000				
WORD	V 2	-0.268	-0.063			
SENT	V 3	0.049	-0.021	0.030		
SENTERR	V 4	0.002	0.078	0.111	-0.073	
ARROW	V 5	0.064	-0.202	-0.004	-0.088	0.038
LETTERS	V 6	-0.097	-0.032	0.053	-0.228	0.018
LETTERR	V 7	-0.014	-0.010	0.060	-0.148	0.070

		LETTERS V 6	LETTERR V 7
LETTERS	V 6	-0.019	
LETTERR	V 7	0.126	-0.064

```
    AVERAGE ABSOLUTE STANDARDIZED RESIDUALS              =  0.0725
    AVERAGE OFF-DIAGONAL ABSOLUTE STANDARDIZED RESIDUALS =  0.0830
```

```
  TITLE:   spatial
09/16/03        PAGE :  6
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  LARGEST STANDARDIZED RESIDUALS:

     V  2,V  1   V  6,V  4   V  5,V  2   V  7,V  4   V  7,V  6
      -0.268      -0.228      -0.202      -0.148       0.126

     V  4,V  3   V  6,V  1   V  5,V  4   V  4,V  2   V  4,V  4
       0.111      -0.097      -0.088       0.078      -0.073

     V  7,V  5   V  7,V  7   V  5,V  1   V  2,V  2   V  7,V  3
       0.070      -0.064       0.064      -0.063       0.060

     V  6,V  3   V  3,V  1   V  5,V  5   V  6,V  2   V  3,V  3
       0.053       0.049       0.038      -0.032       0.030

   DISTRIBUTION OF STANDARDIZED RESIDUALS

     ----------------------------------------
    !                          !
 20-                           -
    !                          !
    !                          !
    !                          !
    !                          !
 15-                           -
    !                          !
    !                          !
    !           *              !
    !           *              !
 10-            * *            -
    !           * *            !
    !           * *            !
    !           * *            !
    !           * *            !
  5-            * *            -
    !           * *            !
    !       *   * *            !
    !       *   * * *          !
    !       * * * * *          !
     ----------------------------------------
      1 2 3 4 5 6 7 8 9 A B C
```

	RANGE			FREQ	PERCENT
1	-0.5	-	--	0	0.00%
2	-0.4	-	-0.5	0	0.00%
3	-0.3	-	-0.4	0	0.00%
4	-0.2	-	-0.3	3	10.71%
5	-0.1	-	-0.2	1	3.57%
6	0.0	-	-0.1	12	42.86%
7	0.1	-	0.0	10	35.71%
8	0.2	-	0.1	2	7.14%
9	0.3	-	0.2	0	0.00%
A	0.4	-	0.3	0	0.00%
B	0.5	-	0.4	0	0.00%
C	++	-	0.5	0	0.00%
	TOTAL			28	100.00%

EACH "*" REPRESENTS 1 RESIDUALS

```
  TITLE:   spatial
09/16/03        PAGE :  7
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS

   SYL    =V1  =     .896*F1    + 1.091*F2    + 1.000 E1
                     .603          .628
                    1.486         1.737

   WORD   =V2  =   11.829*F1    + 1.000 E2
                    2.150
                    5.501

   SENT   =V3  =   15.806*F1    + 1.000 E3
                    2.512
                    6.292

 SENTERR =V4  =   -2.233*F1    + 1.000 E4
                     .668
                   -3.342

 ARROW    =V5  =    7.772*F2    + 1.000 E5
                    1.380
                    5.631

 LETTERS =V6  =   11.713*F2    + 1.000 E6
                    2.184
                    5.362

 LETTERR =V7  =   -3.042*F2    + 1.000 E7
                    1.117
                   -2.724
  TITLE:   spatial
09/16/03        PAGE :  8
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  VARIANCES OF INDEPENDENT VARIABLES
  ----------------------------------

                V                                F
               ---                              ---
                         I F1  -   F1                     1.000 I
                         I                                      I
                         I                                      I
                         I                                      I
                         I F2  -   F2                     1.000 I
                         I                                      I
                         I                                      I
                         I                                      I
```

```
  TITLE:   spatial
09/16/03       PAGE :  9
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  VARIANCES OF INDEPENDENT VARIABLES
  ----------------------------------

                 E                                D
                ---                              ---
 E1  -  SYL             10.450*I                                I
                         2.275 I                                I
                         4.593 I                                I
                               I                                I
 E2  -  WORD           249.599*I                                I
                        61.122 I                                I
                         4.084 I                                I
                               I                                I
 E3  -  SENT            56.223*I                                I
                        57.743 I                                I
                          .974 I                                I
                               I                                I
 E4  -SENTERR           32.952*I                                I
                         7.013 I                                I
                         4.699 I                                I
                               I                                I
 E5  -ARROW             34.371*I                                I
                        16.182 I                                I
                         2.124 I                                I
                               I                                I
 E6  -LETTERS          121.638*I                                I
                        41.711 I                                I
                         2.916 I                                I
                               I                                I
 E7  -LETTERR           70.507*I                                I
                        15.098 I                                I
                         4.670 I                                I
                               I                                I
  TITLE:   spatial
09/16/03       PAGE : 10
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  COVARIANCES AMONG INDEPENDENT VARIABLES
  ---------------------------------------

                 V                                F
                ---                              ---
                               I F2  -   F2                .373*I
                               I F1  -   F1                .168 I
                               I                          2.218 I
                               I                                I
```

```
  TITLE:   spatial
09/16/03        PAGE : 11
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)


  STANDARDIZED SOLUTION:                                                R-SQUARED



   SYL    =V1  =     .247*F1    +   .301*F2    +   .891 E1                .207
   WORD   =V2  =     .599*F1    +   .800 E2                               .359
   SENT   =V3  =     .903*F1    +   .429 E3                               .816
 SENTERR =V4  =    -.363*F1    +   .932 E4                               .131
 ARROW   =V5  =     .798*F2    +   .602 E5                               .637
 LETTERS =V6  =     .728*F2    +   .686 E6                               .530
 LETTERR =V7  =    -.341*F2    +   .940 E7                               .116

  TITLE:   spatial
09/16/03        PAGE : 12
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  CORRELATIONS AMONG INDEPENDENT VARIABLES
  ---------------------------------------

                V                                F
               ---                              ---
                           I F2   -   F2                   .373*I
                           I F1   -   F1                        I
                           I                                    I




 -------------------------------------------------------------------------------
                         E N D    O F    M E T H O D
 -------------------------------------------------------------------------------
```

TITLE: verbal
09/16/03 PAGE : 13
EQS/EM386 Licensee: Alison Bacon
MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP 2

MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

PARAMETER ESTIMATES APPEAR IN ORDER,
NO SPECIAL PROBLEMS WERE ENCOUNTERED DURING OPTIMIZATION.

ALL EQUALITY CONSTRAINTS WERE CORRECTLY IMPOSED

RESIDUAL COVARIANCE MATRIX (S-SIGMA) :

		SYL V 1	WORD V 2	SENT V 3	SENTERR V 4	ARROW V 5
SYL	V 1	0.000				
WORD	V 2	-0.059	3.494			
SENT	V 3	-0.928	1.236	-9.131		
SENTERR	V 4	-1.840	-3.150	1.202	1.101	
ARROW	V 5	1.084	-3.931	-13.907	-5.491	-2.832
LETTERS	V 6	-2.038	9.526	8.806	-14.365	-5.835
LETTERR	V 7	-1.786	6.931	12.197	-0.918	-10.271

		LETTERS V 6	LETTERR V 7
LETTERS	V 6	2.029	
LETTERR	V 7	-9.984	2.382

AVERAGE ABSOLUTE COVARIANCE RESIDUALS = 4.8734
AVERAGE OFF-DIAGONAL ABSOLUTE COVARIANCE RESIDUALS = 5.4993

STANDARDIZED RESIDUAL MATRIX:

		SYL V 1	WORD V 2	SENT V 3	SENTERR V 4	ARROW V 5
SYL	V 1	0.000				
WORD	V 2	-0.002	0.022			
SENT	V 3	-0.022	0.006	-0.037		
SENTERR	V 4	-0.121	-0.044	0.014	0.035	
ARROW	V 5	0.046	-0.035	-0.101	-0.111	-0.037
LETTERS	V 6	-0.059	0.058	0.044	-0.198	-0.052
LETTERR	V 7	-0.076	0.063	0.090	-0.019	-0.135

		LETTERS V 6	LETTERR V 7
LETTERS	V 6	0.012	
LETTERR	V 7	-0.089	0.032

AVERAGE ABSOLUTE STANDARDIZED RESIDUALS = 0.0556
AVERAGE OFF-DIAGONAL ABSOLUTE STANDARDIZED RESIDUALS = 0.0659

```
  TITLE:   verbal
09/16/03        PAGE : 14
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  2

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  LARGEST STANDARDIZED RESIDUALS:

     V  6,V  4   V  7,V  5   V  4,V  1   V  5,V  4   V  5,V  3
      -0.198      -0.135      -0.121      -0.111      -0.101

     V  7,V  3   V  7,V  6   V  7,V  1   V  7,V  2   V  6,V  1
       0.090      -0.089      -0.076       0.063      -0.059

     V  6,V  2   V  6,V  5   V  5,V  1   V  4,V  2   V  6,V  3
       0.058      -0.052       0.046      -0.044       0.044

     V  3,V  3   V  5,V  5   V  5,V  2   V  4,V  4   V  7,V  7
      -0.037      -0.037      -0.035       0.035       0.032

  DISTRIBUTION OF STANDARDIZED RESIDUALS

     ----------------------------------------
     !                                      !
   20-                                      -
     !                                      !
     !                                      !
     !                                      !
     !                                      !            RANGE       FREQ PERCENT
   15-                                      -
     !                                      !   1   -0.5  -  --         0   0.00%
     !                                      !   2   -0.4  -  -0.5       0   0.00%
     !                 *                    !   3   -0.3  -  -0.4       0   0.00%
     !                 *  *                 !   4   -0.2  -  -0.3       0   0.00%
   10-                 *  *                 -   5   -0.1  -  -0.2       5  17.86%
     !                 *  *                 !   6    0.0  -  -0.1      12  42.86%
     !                 *  *                 !   7    0.1  -   0.0      11  39.29%
     !                 *  *                 !   8    0.2  -   0.1       0   0.00%
     !                 *  *                 !   9    0.3  -   0.2       0   0.00%
    5-              *  *  *                 -   A    0.4  -   0.3       0   0.00%
     !              *  *  *                 !   B    0.5  -   0.4       0   0.00%
     !              *  *  *                 !   C     ++  -   0.5       0   0.00%
     !              *  *  *                 !   -------------------------------
     !              *  *  *                 !            TOTAL         28 100.00%
     ----------------------------------------
        1  2  3  4  5  6  7  8  9  A  B  C    EACH "*" REPRESENTS  1 RESIDUALS
```

```
  TITLE:    verbal
09/16/03        PAGE : 15
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  2

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS

   SYL    =V1  =     1.383*F1    +   .595*F2    +  1.000 E1
                      .723            .840
                     1.913            .708

   WORD   =V2  =    11.829*F1    +  1.000 E2
                     2.150
                     5.501

   SENT   =V3  =    15.806*F1    +  1.000 E3
                     2.512
                     6.292

 SENTERR =V4   =    -2.233*F1    +  1.000 E4
                      .668
                    -3.342

 ARROW    =V5  =     7.772*F2    +  1.000 E5
                     1.380
                     5.631

 LETTERS =V6   =    11.713*F2    +  1.000 E6
                     2.184
                     5.362

 LETTERR =V7   =    -3.042*F2    +  1.000 E7
                     1.117
                    -2.724

  TITLE:    verbal
09/16/03        PAGE : 16
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  2

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  VARIANCES OF INDEPENDENT VARIABLES
  ----------------------------------

                V                              F
               ---                            ---
                         I F1   -    F1                  .599*I
                         I                               .213 I
                         I                              2.818 I
                         I                                    I
                         I F2   -    F2                  .476*I
                         I                               .191 I
                         I                              2.490 I
                         I                                    I
```

```
  TITLE:    verbal
09/16/03         PAGE : 17
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  2

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  VARIANCES OF INDEPENDENT VARIABLES
  ----------------------------------

                 E                                 D
                ---                               ---
 E1  -  SYL              5.391*I                                I
                          .880 I                                I
                         6.125 I                                I
                               I                                I
 E2  -  WORD            74.472*I                                I
                        17.690 I                                I
                         4.210 I                                I
                               I                                I
 E3  -  SENT           105.220*I                                I
                        29.147 I                                I
                         3.610 I                                I
                               I                                I
 E4  -SENTERR           27.608*I                                I
                         4.197 I                                I
                         6.578 I                                I
                               I                                I
 E5  -ARROW             51.163*I                                I
                        10.615 I                                I
                         4.820 I                                I
                               I                                I
 E6  -LETTERS           98.010*I                                I
                        22.174 I                                I
                         4.420 I                                I
                               I                                I
 E7  -LETTERR           68.767*I                                I
                        10.432 I                                I
                         6.592 I                                I
                               I                                I
  TITLE:    verbal
09/16/03         PAGE : 18
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  2

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  COVARIANCES AMONG INDEPENDENT VARIABLES
  ---------------------------------------

                 V                                 F
                ---                               ---
                               I F2  -   F2                 .359*I
                               I F1  -   F1                 .121 I
                               I                           2.968 I
                               I                                 I
```

```
 TITLE:   verbal
09/16/03        PAGE : 19
 EQS/EM386 Licensee: Alison Bacon
 MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  2

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)


 STANDARDIZED SOLUTION:                                                R-SQUARED


  SYL    =V1  =    .396*F1   +   .152*F2    +   .860 E1                  .261
  WORD   =V2  =    .728*F1   +   .686 E2                                 .530
  SENT   =V3  =    .766*F1   +   .642 E3                                 .587
SENTERR  =V4  =   -.313*F1   +   .950 E4                                 .098
ARROW    =V5  =    .600*F2   +   .800 E5                                 .360
LETTERS  =V6  =    .632*F2   +   .775 E6                                 .400
LETTERR  =V7  =   -.245*F2   +   .969 E7                                 .060

 TITLE:   verbal
09/16/03        PAGE : 20
 EQS/EM386 Licensee: Alison Bacon
 MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  2

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 CORRELATIONS AMONG INDEPENDENT VARIABLES
 ---------------------------------------

               V                              F
              ---                            ---
                          I F2  -   F2                    .672*I
                          I F1  -   F1                        I
                          I                                   I



 ------------------------------------------------------------------------------
                        E N D    O F    M E T H O D
 ------------------------------------------------------------------------------
1
 STATISTICS FOR MULTIPLE POPULATION ANALYSIS

 ALL EQUALITY CONSTRAINTS WERE CORRECTLY IMPOSED

 GOODNESS OF FIT SUMMARY

 INDEPENDENCE MODEL CHI-SQUARE =       187.344 ON    42 DEGREES OF FREEDOM

 INDEPENDENCE AIC =   103.34423   INDEPENDENCE CAIC =   -62.50368
        MODEL AIC =   -21.03712          MODEL CAIC =  -131.60240
```

```
CHI-SQUARE =        34.963 BASED ON    28 DEGREES OF FREEDOM
PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS     0.17095

BENTLER-BONETT NORMED    FIT INDEX=       0.813
BENTLER-BONETT NONNORMED FIT INDEX=       0.928
COMPARATIVE FIT INDEX (CFI)       =       0.952
BOLLEN (IFI)             FIT INDEX=       0.956
McDonald (MFI)           FIT INDEX=       0.963
LISREL GFI               FIT INDEX=       0.937
LISREL AGFI              FIT INDEX=       0.874
ROOT MEAN SQUARED RESIDUAL (RMR)  =       9.753
STANDARDIZED RMR                  =       0.086
ROOT MEAN SQ. ERROR OF APP.(RMSEA)=       0.043
90% CONFIDENCE INTERVAL OF RMSEA (       0.000,       0.081)
```

ITERATIVE SUMMARY

ITERATION	PARAMETER ABS CHANGE	ALPHA	FUNCTION
1	9.767826	0.50000	1.72750
2	13.985048	0.50000	0.84900
3	11.423270	1.00000	0.48973
4	8.343430	1.00000	0.28490
5	2.922937	1.00000	0.25626
6	1.716993	1.00000	0.25268
7	0.617836	1.00000	0.25183
8	0.528792	1.00000	0.25161
9	0.207439	1.00000	0.25156
10	0.172477	1.00000	0.25154
11	0.083013	1.00000	0.25153
12	0.058729	1.00000	0.25153
13	0.031843	1.00000	0.25153
14	0.020617	1.00000	0.25153
15	0.011904	1.00000	0.25153
16	0.007374	1.00000	0.25153
17	0.004401	0.06250	0.25153
18	0.003962	0.06250	0.25153
19	0.003569	0.06250	0.25153
20	0.003214	0.06250	0.25153
21	0.002895	0.06250	0.25153
22	0.002608	0.06250	0.25153
23	0.002351	0.06250	0.25153
24	0.002118	0.06250	0.25153
25	0.001910	0.06250	0.25153
26	0.001722	0.06250	0.25153
27	0.001552	0.06250	0.25153
28	0.001400	0.06250	0.25153
29	0.001263	0.06250	0.25153
30	0.001140	0.06250	0.25153
31	0.001029	0.06250	0.25153
32	0.000929	1.00000	0.25153

```
  TITLE:
09/16/03       PAGE : 21
  EQS/EM386 Licensee: Alison Bacon


  LAGRANGE MULTIPLIER TEST (FOR RELEASING CONSTRAINTS)

  CONSTRAINTS TO BE RELEASED ARE:


          CONSTRAINTS FROM GROUP  2

          CONSTR:    1   (1,V2,F1)-(2,V2,F1)=0;
          CONSTR:    2   (1,V3,F1)-(2,V3,F1)=0;
          CONSTR:    3   (1,V4,F1)-(2,V4,F1)=0;
          CONSTR:    4   (1,V5,F2)-(2,V5,F2)=0;
          CONSTR:    5   (1,V6,F2)-(2,V6,F2)=0;
          CONSTR:    6   (1,V7,F2)-(2,V7,F2)=0;


           UNIVARIATE TEST STATISTICS:

   NO    CONSTRAINT    CHI-SQUARE   PROBABILITY
   --    -----------   ----------   -----------

    1    CONSTR:   1        0.742        0.389
    2    CONSTR:   2        2.823        0.093
    3    CONSTR:   3        1.144        0.285
    4    CONSTR:   4        1.210        0.271
    5    CONSTR:   5        0.167        0.683
    6    CONSTR:   6        1.327        0.249


         CUMULATIVE MULTIVARIATE STATISTICS              UNIVARIATE INCREMENT
         ----------------------------------              --------------------

  STEP  PARAMETER    CHI-SQUARE  D.F.   PROBABILITY       CHI-SQUARE  PROBABILITY
  ---- -----------   ----------  ----   -----------       ----------  -----------

    1  CONSTR:   2      2.823      1       0.093             2.823       0.093
    2  CONSTR:   6      4.149      2       0.126             1.327       0.249
    3  CONSTR:   4      4.797      3       0.187             0.647       0.421
    4  CONSTR:   3      5.181      4       0.269             0.384       0.535
1
  Execution begins at 16:37:00.12
  Execution ends   at 16:37:00.56
  Elapsed time =       0.44 seconds
```

APPENDIX 5H

Experiment 6: Full EQS output of Confirmatory Factor Analysis - Model 2.

```
EQS, A STRUCTURAL EQUATION PROGRAM            MULTIVARIATE SOFTWARE, INC.
  COPYRIGHT BY P.M. BENTLER                     VERSION 5.7b (C) 1985 - 1998.


    PROGRAM CONTROL INFORMATION

     /TITLE
     spatial
     /SPECIFICATIONS
      VARIABLES=7; CASES=48;
      ANALYSIS = COV; GROUP=2;
      METHODS=ML;
      MATRIX=CORRELATION;
     /LABELS
    V1=SYL; V2=WORD; V3=SENT; V4=SENTERR; V5=ARROW; V6=LETTERS; V7=LETTERR;
     /EQUATIONS
     V1 =  + *F1  + *F2  + E1;
     V2 =  + *F1  + E2;
     V3 =  + *F1  + E3;
     V4 =  + *F1  + E4;
     V5 =  + *F2  + E5;
     V6 =  + *F2  + E6;
     V7 =  + *F2  + E7;
     /VARIANCES
     F1 = 1.00;
     F2 = 1.00;
     E1 = *;
     E2 = *;
     E3 = *;
     E4 = *;
     E5 = *;
     E6 = *;
     E7 = *;
     /COVARIANCES
     F1,F2=*;
     /MATRIX
     1
     -0.046 1
     0.368 0.529 1
     -0.133 -0.154 -0.223 1
     0.371 -0.022 0.256 -0.198 1
     0.192 0.137 0.297 -0.331 0.594 1
     -0.152 -0.091 -0.056 -0.099 -0.205 -0.132 1
     /STANDARD DEVIATIONS
     3.629 19.147 17.765 5.947 9.925 15.934 8.659
     /TECHNICAL
     ITR=50;

     /END

     43 CUMULATED RECORDS OF INPUT MODEL FILE WERE READ (GROUP   1)
```

```
 TITLE:
09/16/03       PAGE :  2
 EQS/EM386 Licensee: Alison Bacon


   PROGRAM CONTROL INFORMATION

   /TITLE
   verbal
   /SPECIFICATIONS
    VARIABLES=7; CASES=93;
    METHODS=ML;
    MATRIX=CORRELATION;ANALYSIS = COV;
   /LABELS
   V1=SYL; V2=WORD; V3=SENT; V4=SENTERR; V5=ARROW; V6=LETTERS; V7=LETTERR;
   /EQUATIONS
   V1 =  + *F1  + *F2  + E1;
   V2 =  + *F1  + E2;
   V3 =  + *F1  + E3;
   V4 =  + *F1  + E4;
   V5 =  + *F2  + E5;
   V6 =  + *F2  + E6;
   V7 =  + *F2  + E7;
   /VARIANCES
   F1 = 1;
   F2 = 1;
   E1 = *;
   E2 = *;
   E3 = *;
   E4 = *;
   E5 = *;
   E6 = *;
   E7 = *;
   /COVARIANCES
   F1,F2=*;
   /MATRIX
   1
   0.357 1
   0.367 0.568 1
   -0.274 -0.265 -0.226 1
   0.301 0.260 0.219 -0.237 1
   0.204 0.362 0.373 -0.328 0.332 1
   -0.177 -0.054 -0.037 0.031 -0.282 -0.241 1
   /STANDARD DEVIATIONS
   2.701 12.721 15.678 5.630 8.779 12.857 8.692
   /TECHNICAL
   ITR=50;
   /PRINT
   FIT=ALL;
   /LMTEST
   /END


     87 CUMULATED RECORDS OF INPUT MODEL FILE WERE READ (GROUP   2)
```

```
 TITLE:                                                               spatial
09/16/03      PAGE :  3
 EQS/EM386 Licensee: Alison Bacon
 COVARIANCE  MATRIX TO BE ANALYZED:   7 VARIABLES (SELECTED FROM   7 VARIABLES)
 BASED ON    48 CASES.
```

		SYL V 1	WORD V 2	SENT V 3	SENTERR V 4	ARROW V 5
SYL	V 1	13.170				
WORD	V 2	-3.196	366.608			
SENT	V 3	23.725	179.937	315.595		
SENTERR	V 4	-2.870	-17.536	-23.560	35.367	
ARROW	V 5	13.363	-4.181	45.137	-11.687	98.506
LETTERS	V 6	11.102	41.797	84.071	-31.365	93.938
LETTERR	V 7	-4.776	-15.087	-8.614	-5.098	-17.618

		LETTERS V 6	LETTERR V 7
LETTERS	V 6	253.892	
LETTERR	V 7	-18.212	74.978

```
 BENTLER-WEEKS STRUCTURAL REPRESENTATION:

      NUMBER OF DEPENDENT VARIABLES =  7
          DEPENDENT V'S :     1    2    3    4    5    6    7

      NUMBER OF INDEPENDENT VARIABLES =  9
          INDEPENDENT F'S :     1    2
          INDEPENDENT E'S :     1    2    3    4    5    6    7

      NUMBER OF FREE PARAMETERS =  16
      NUMBER OF FIXED NONZERO PARAMETERS =   9
```

```
TITLE:                                                                          verbal
09/16/03       PAGE :  4
  EQS/EM386 Licensee: Alison Bacon
  COVARIANCE  MATRIX TO BE ANALYZED:   7 VARIABLES (SELECTED FROM   7 VARIABLES)
  BASED ON    93 CASES.
```

		SYL V 1	WORD V 2	SENT V 3	SENTERR V 4	ARROW V 5
SYL	V 1	7.295				
WORD	V 2	12.266	161.824			
SENT	V 3	15.541	113.282	245.800		
SENTERR	V 4	-4.167	-18.979	-19.948	31.697	
ARROW	V 5	7.137	29.036	30.143	-11.714	77.071
LETTERS	V 6	7.084	59.207	75.186	-23.742	37.473
LETTERR	V 7	-4.155	-5.971	-5.042	1.517	-21.519

		LETTERS V 6	LETTERR V 7
LETTERS	V 6	165.302	
LETTERR	V 7	-26.932	75.551

```
  BENTLER-WEEKS STRUCTURAL REPRESENTATION:

        NUMBER OF DEPENDENT VARIABLES =  7
            DEPENDENT V'S :     1    2    3    4    5    6    7

        NUMBER OF INDEPENDENT VARIABLES =  9
            INDEPENDENT F'S :     1    2
            INDEPENDENT E'S :     1    2    3    4    5    6    7

        NUMBER OF FREE PARAMETERS =  16
        NUMBER OF FIXED NONZERO PARAMETERS =    9


  3RD STAGE OF COMPUTATION REQUIRED      5153 WORDS OF MEMORY.
  PROGRAM ALLOCATED    100000 WORDS

  DETERMINANT OF INPUT MATRIX IN GROUP  1 IS   0.23014E+14

  DETERMINANT OF INPUT MATRIX IN GROUP  2 IS   0.24603E+13
  TITLE:                                                                       spatial
09/16/03       PAGE :  5
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  PARAMETER        CONDITION CODE
    E3,E3          CONSTRAINED AT LOWER BOUND
```

TITLE: spatial
09/16/03 PAGE : 6
EQS/EM386 Licensee: Alison Bacon
MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP 1

MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

RESIDUAL COVARIANCE MATRIX (S-SIGMA) :

		SYL V 1	WORD V 2	SENT V 3	SENTERR V 4	ARROW V 5
SYL	V 1	0.000				
WORD	V 2	-16.723	0.000			
SENT	V 3	0.000	0.000	0.000		
SENTERR	V 4	-1.099	-4.103	0.000	0.000	
ARROW	V 5	0.317	-31.799	-3.302	-8.071	0.000
LETTERS	V 6	-3.564	10.748	29.613	-27.300	0.130
LETTERR	V 7	-2.001	-9.212	1.690	-5.867	0.132

		LETTERS V 6	LETTERR V 7
LETTERS	V 6	0.000	
LETTERR	V 7	1.742	0.000

AVERAGE ABSOLUTE COVARIANCE RESIDUALS = 5.6220
AVERAGE OFF-DIAGONAL ABSOLUTE COVARIANCE RESIDUALS = 7.4960

STANDARDIZED RESIDUAL MATRIX:

		SYL V 1	WORD V 2	SENT V 3	SENTERR V 4	ARROW V 5
SYL	V 1	0.000				
WORD	V 2	-0.241	0.000			
SENT	V 3	0.000	0.000	0.000		
SENTERR	V 4	-0.051	-0.036	0.000	0.000	
ARROW	V 5	0.009	-0.167	-0.019	-0.137	0.000
LETTERS	V 6	-0.062	0.035	0.105	-0.288	0.001
LETTERR	V 7	-0.064	-0.056	0.011	-0.114	0.002

		LETTERS V 6	LETTERR V 7
LETTERS	V 6	0.000	
LETTERR	V 7	0.013	0.000

AVERAGE ABSOLUTE STANDARDIZED RESIDUALS = 0.0503
AVERAGE OFF-DIAGONAL ABSOLUTE STANDARDIZED RESIDUALS = 0.0670

```
  TITLE:                                                              spatial
09/16/03        PAGE :  7
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  LARGEST STANDARDIZED RESIDUALS:

    V  6,V  4    V  2,V  1    V  5,V  2    V  5,V  4    V  7,V  4
     -0.288       -0.241       -0.167       -0.137       -0.114

    V  6,V  3    V  7,V  1    V  6,V  1    V  7,V  2    V  4,V  1
      0.105       -0.064       -0.062       -0.056       -0.051

    V  4,V  2    V  6,V  2    V  5,V  3    V  7,V  6    V  7,V  3
     -0.036        0.035       -0.019        0.013        0.011

    V  5,V  1    V  7,V  5    V  6,V  5    V  3,V  1    V  1,V  1
      0.009        0.002        0.001        0.000        0.000

  DISTRIBUTION OF STANDARDIZED RESIDUALS

   -------------------------------------------------
   !                                                !
 20-                                                -
   !                                                !
   !                                                !
   !                                                !
   !                                                !
 15-                      *                         -
   !                      *                         !
   !                      *                         !
   !                      *                         !
   !                      *                         !
 10-                      *                         -
   !                      *                         !
   !                      *                         !
   !                      *   *                     !
   !                      *   *                     !
  5-                      *   *                     -
   !                      *   *                     !
   !                  *   *   *                     !
   !              *   *   *   *                     !
   !              *   *   *   *   *                 !
   -------------------------------------------------
      1   2   3   4   5   6   7   8   9   A   B   C
```

	RANGE			FREQ	PERCENT
1	-0.5	-	--	0	0.00%
2	-0.4	-	-0.5	0	0.00%
3	-0.3	-	-0.4	0	0.00%
4	-0.2	-	-0.3	2	7.14%
5	-0.1	-	-0.2	3	10.71%
6	0.0	-	-0.1	15	53.57%
7	0.1	-	0.0	7	25.00%
8	0.2	-	0.1	1	3.57%
9	0.3	-	0.2	0	0.00%
A	0.4	-	0.3	0	0.00%
B	0.5	-	0.4	0	0.00%
C	++	-	0.5	0	0.00%
	TOTAL			28	100.00%

EACH "*" REPRESENTS 1 RESIDUALS

```
 TITLE:                                                                    spatial
09/16/03       PAGE :  8
 EQS/EM386 Licensee: Alison Bacon
 MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)


 MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS


  SYL    =V1   =      .998*F1    +  1.130*F2    +  1.000 E1
                      .518             .572
                     1.929            1.974

  WORD   =V2   =    10.129*F1    +  1.000 E2
                     2.590
                     3.911

  SENT   =V3   =    17.765*F1    +  1.000 E3
                     1.832
                     9.695

 SENTERR =V4   =    -1.326*F1    +  1.000 E4
                      .857
                    -1.548

 ARROW   =V5   =     9.135*F2    +  1.000 E5
                     1.942
                     4.705

 LETTERS =V6   =    10.269*F2    +  1.000 E6
                     2.747
                     3.739

 LETTERR =V7   =    -1.943*F2    +  1.000 E7
                     1.361
                    -1.428
```

```
 TITLE:                                                                    spatial
09/16/03       PAGE :  9
 EQS/EM386 Licensee: Alison Bacon
 MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)
```

```
 VARIANCES OF INDEPENDENT VARIABLES
 ----------------------------------

                V                             F
               ---                           ---
                              I F1  -   F1              1.000 I
                              I                               I
                              I                               I
                              I                               I
                              I F2  -   F2              1.000 I
                              I                               I
                              I                               I
                              I                               I

 TITLE:                                                                    spatial
09/16/03       PAGE : 10
 EQS/EM386 Licensee: Alison Bacon
 MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 VARIANCES OF INDEPENDENT VARIABLES
 ----------------------------------

                E                             D
               ---                           ---
 E1  -  SYL               10.222*I                            I
                           2.178 I                            I
                           4.694 I                            I
                                 I                            I
 E2  -  WORD             264.016*I                            I
                          54.462 I                            I
                           4.848 I                            I
                                 I                            I
 E3  -  SENT                .000*I                            I
                            .000 I                            I
                            .000 I                            I
                                 I                            I
 E4  -SENTERR             33.608*I                            I
                           6.933 I                            I
                           4.848 I                            I
                                 I                            I
 E5  -ARROW               15.064*I                            I
                          29.404 I                            I
                            .512 I                            I
                                 I                            I
 E6  -LETTERS            148.431*I                            I
                          48.113 I                            I
                           3.085 I                            I
                                 I                            I
 E7  -LETTERR             71.203*I                            I
                          14.839 I                            I
                           4.798 I                            I
                                 I                            I
```

```
 TITLE:                                                                  spatial
09/16/03       PAGE : 11
 EQS/EM386 Licensee: Alison Bacon
 MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 COVARIANCES AMONG INDEPENDENT VARIABLES
 ---------------------------------------

              V                              F
             ---                            ---
                          I F2  -    F2                  .299*I
                          I F1  -    F1                  .150 I
                          I                             1.996 I
                          I                                   I

 TITLE:                                                                  spatial
09/16/03       PAGE : 12
 EQS/EM386 Licensee: Alison Bacon
 MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)


 STANDARDIZED SOLUTION:                                              R-SQUARED


  SYL    =V1  =     .275*F1    +   .311*F2    +   .881 E1               .224
  WORD   =V2  =     .529*F1    +   .849 E2                              .280
  SENT   =V3  =    1.000*F1    +   .000 E3                             1.000
 SENTERR =V4  =    -.223*F1    +   .975 E4                              .050
 ARROW   =V5  =     .920*F2    +   .391 E5                              .847
 LETTERS =V6  =     .644*F2    +   .765 E6                              .415
 LETTERR =V7  =    -.224*F2    +   .974 E7                              .050

 TITLE:                                                                  spatial
09/16/03       PAGE : 13
 EQS/EM386 Licensee: Alison Bacon
 MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  1

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)
```

```
 CORRELATIONS AMONG INDEPENDENT VARIABLES
 ---------------------------------------

                V                               F
               ---                             ---
                              I F2   -   F2                 .299*I
                              I F1   -   F1                      I
                              I                                  I

 ------------------------------------------------------------------------------
                        E N D     O F     M E T H O D
 ------------------------------------------------------------------------------
TITLE:                                                                   verbal
09/16/03       PAGE : 14
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  2

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  PARAMETER ESTIMATES APPEAR IN ORDER,
  NO SPECIAL PROBLEMS WERE ENCOUNTERED DURING OPTIMIZATION.

  RESIDUAL COVARIANCE MATRIX  (S-SIGMA) :
```

		SYL V 1	WORD V 2	SENT V 3	SENTERR V 4	ARROW V 5
SYL	V 1	0.000				
WORD	V 2	-0.706	0.000			
SENT	V 3	-0.016	4.727	0.000		
SENTERR	V 4	-1.171	1.922	5.117	0.000	
ARROW	V 5	1.966	-1.820	-6.860	-4.589	0.000
LETTERS	V 6	-2.267	3.410	8.275	-10.859	-4.168
LETTERR	V 7	-1.007	12.813	17.484	-2.820	-7.500

		LETTERS V 6	LETTERR V 7
LETTERS	V 6	0.000	
LETTERR	V 7	-1.583	0.000

```
              AVERAGE ABSOLUTE   COVARIANCE   RESIDUALS      =        3.6100
     AVERAGE OFF-DIAGONAL ABSOLUTE   COVARIANCE   RESIDUALS      =        4.8133
```

STANDARDIZED RESIDUAL MATRIX:

		SYL V 1	WORD V 2	SENT V 3	SENTERR V 4	ARROW V 5
SYL	V 1	0.000				
WORD	V 2	-0.021	0.000			
SENT	V 3	0.000	0.024	0.000		
SENTERR	V 4	-0.077	0.027	0.058	0.000	
ARROW	V 5	0.083	-0.016	-0.050	-0.093	0.000
LETTERS	V 6	-0.065	0.021	0.041	-0.150	-0.037
LETTERR	V 7	-0.043	0.116	0.128	-0.058	-0.098

		LETTERS V 6	LETTERR V 7
LETTERS	V 6	0.000	
LETTERR	V 7	-0.014	0.000

AVERAGE ABSOLUTE STANDARDIZED RESIDUALS = 0.0436
AVERAGE OFF-DIAGONAL ABSOLUTE STANDARDIZED RESIDUALS = 0.0581

TITLE: verbal
09/16/03 PAGE : 15
EQS/EM386 Licensee: Alison Bacon
MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP 2

MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

LARGEST STANDARDIZED RESIDUALS:

V 6,V 4	V 7,V 3	V 7,V 2	V 7,V 5	V 5,V 4
-0.150	0.128	0.116	-0.098	-0.093
V 5,V 1	V 4,V 1	V 6,V 1	V 4,V 3	V 7,V 4
0.083	-0.077	-0.065	0.058	-0.058
V 5,V 3	V 7,V 1	V 6,V 3	V 6,V 5	V 4,V 2
-0.050	-0.043	0.041	-0.037	0.027
V 3,V 2	V 6,V 2	V 2,V 1	V 5,V 2	V 7,V 6
0.024	0.021	-0.021	-0.016	-0.014

DISTRIBUTION OF STANDARDIZED RESIDUALS

```
  ----------------------------------------
  !                                      !
20-                                      -
  !                 *                    !
  !                 *                    !
  !                 *                    !
  !                 *                    !
15-                 *                    -
  !                 *                    !
  !                 *                    !
  !                 *                    !
  !                 *                    !
10-                 *                    -
  !                 *                    !
  !                 *                    !
  !                 *                    !
  !                 *  *                 !
 5-                 *  *                 -
  !                 *  *                 !
  !                 *  *                 !
  !                 *  *  *              !
  !              *  *  *  *              !
  ----------------------------------------
     1  2  3  4  5  6  7  8  9  A  B  C
```

	RANGE			FREQ	PERCENT
1	-0.5	-	--	0	0.00%
2	-0.4	-	-0.5	0	0.00%
3	-0.3	-	-0.4	0	0.00%
4	-0.2	-	-0.3	0	0.00%
5	-0.1	-	-0.2	1	3.57%
6	0.0	-	-0.1	19	67.86
7	0.1	-	0.0	6	21.43%
8	0.2	-	0.1	2	7.14%
9	0.3	-	0.2	0	0.00%
A	0.4	-	0.3	0	0.00%
B	0.5	-	0.4	0	0.00%
C	++	-	0.5	0	0.00%
	TOTAL			28	100.00%

EACH "*" REPRESENTS 1 RESIDUALS

TITLE: verbal
09/16/03 PAGE : 16
EQS/EM386 Licensee: Alison Bacon
MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP 2

MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

MEASUREMENT EQUATIONS WITH STANDARD ERRORS AND TEST STATISTICS

```
 SYL    =V1  =    1.169*F1    +  .287*F2    + 1.000 E1
                   .533           .569
                  2.192           .505

 WORD   =V2  =    9.514*F1    + 1.000 E2
                  1.383
                  6.877

 SENT   =V3  =   11.410*F1    + 1.000 E3
                  1.704
                  6.695

SENTERR =V4  =   -2.197*F1    + 1.000 E4
                   .647
                 -3.397
```

```
ARROW   =V5  =     4.799*F2    + 1.000 E5
                   1.099
                   4.365

LETTERS =V6  =     8.678*F2    + 1.000 E6
                   1.690
                   5.134

LETTERR =V7  =    -2.921*F2    + 1.000 E7
                   1.095
                  -2.668


 TITLE:                                                                verbal
09/16/03       PAGE : 17
 EQS/EM386 Licensee: Alison Bacon
 MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  2

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 VARIANCES OF INDEPENDENT VARIABLES
 ----------------------------------

                 V                                F
                ---                              ---
                              I F1  -   F1                1.000 I
                              I                                 I
                              I                                 I
                              I                                 I
                              I F2  -   F2                1.000 I
                              I                                 I
                              I                                 I
                              I                                 I

 TITLE:                                                                verbal
09/16/03       PAGE : 18
 EQS/EM386 Licensee: Alison Bacon
 MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  2

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 VARIANCES OF INDEPENDENT VARIABLES
 ----------------------------------

                 E                                D
                ---                              ---
E1  -  SYL                5.391*I                               I
                           .885 I                               I
                          6.092 I                               I
                                I                               I
E2  -  WORD              71.302*I                               I
                         18.568 I                               I
                          3.840 I                               I
```

```
                                  I                              I
 E3  -  SENT                115.620*I                                  I
                             27.932 I                                  I
                              4.139 I                                  I
                                    I                                  I
 E4  -SENTERR                26.871*I                                  I
                              4.198 I                                  I
                              6.400 I                                  I
                                    I                                  I
 E5  -ARROW                  54.042*I                                  I
                             10.450 I                                  I
                              5.171 I                                  I
                                    I                                  I
 E6  -LETTERS                90.002*I                                  I
                             24.870 I                                  I
                              3.619 I                                  I
                                    I                                  I
 E7  -LETTERR                67.017*I                                  I
                             10.588 I                                  I
                              6.329 I                                  I
                                    I                                  I

 TITLE:                                                                        verbal
09/16/03       PAGE : 19
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  2

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

  COVARIANCES AMONG INDEPENDENT VARIABLES
  ---------------------------------------

                 V                                F
                ---                              ---
                                    I F2  -   F2                      .676*I
                                    I F1  -   F1                      .128 I
                                    I                                5.279 I
                                    I                                      I
 TITLE:                                                                        verbal
09/16/03       PAGE : 20
  EQS/EM386 Licensee: Alison Bacon
  MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  2

  MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)


  STANDARDIZED SOLUTION:                                                 R-SQUARED


   SYL    =V1   =     .433*F1    +   .106*F2    +   .860 E1               .261
   WORD   =V2   =     .748*F1    +   .664 E2                              .559
   SENT   =V3   =     .728*F1    +   .686 E3                              .530
 SENTERR  =V4   =    -.390*F1    +   .921 E4                              .152
 ARROW    =V5   =     .547*F2    +   .837 E5                              .299
 LETTERS  =V6   =     .675*F2    +   .738 E6                              .456
 LETTERR  =V7   =    -.336*F2    +   .942 E7                              .113
```

```
 TITLE:                                                                  verbal
09/16/03       PAGE : 21
 EQS/EM386 Licensee: Alison Bacon
 MULTIPLE POPULATION ANALYSIS, INFORMATION IN GROUP  2

 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 CORRELATIONS AMONG INDEPENDENT VARIABLES
 ---------------------------------------

               V                               F
              ---                             ---
                          I F2  -   F2                  .676*I
                          I F1  -   F1                       I
                          I                                  I

 ------------------------------------------------------------------------------
                         E N D   O F   M E T H O D
 ------------------------------------------------------------------------------
1
  STATISTICS FOR MULTIPLE POPULATION ANALYSIS



  GOODNESS OF FIT SUMMARY

  INDEPENDENCE MODEL CHI-SQUARE =       187.344 ON    42 DEGREES OF FREEDOM

  INDEPENDENCE AIC =   103.34423   INDEPENDENCE CAIC =   -62.50368
         MODEL AIC =   -18.17624          MODEL CAIC =  -112.94647

  CHI-SQUARE =       29.824 BASED ON    24 DEGREES OF FREEDOM
  PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS     0.19066

  BENTLER-BONETT NORMED    FIT INDEX=       0.841
  BENTLER-BONETT NONNORMED FIT INDEX=       0.930
  COMPARATIVE FIT INDEX (CFI)      =       0.960
  BOLLEN (IFI)             FIT INDEX=       0.964
  McDonald (MFI)           FIT INDEX=       0.969
  LISREL GFI               FIT INDEX=       0.946
  LISREL AGFI              FIT INDEX=       0.874
  ROOT MEAN SQUARED RESIDUAL (RMR)  =       8.611
  STANDARDIZED RMR                  =       0.077
  ROOT MEAN SQ. ERROR OF APP.(RMSEA)=       0.043
  90% CONFIDENCE INTERVAL OF RMSEA (       0.000,        0.084)
```

ITERATIVE SUMMARY

ITERATION	PARAMETER ABS CHANGE	ALPHA	FUNCTION
1	11.302238	1.00000	3.55956
2	11.952324	1.00000	0.97489
3	8.372164	0.50000	0.84599
4	7.630892	1.00000	0.79746
5	9.102877	1.00000	0.57550
6	8.161210	1.00000	0.25979
7	6.393053	1.00000	0.22138
8	0.867268	1.00000	0.21514
9	0.531189	1.00000	0.21467
10	0.040459	1.00000	0.21458
11	0.095571	1.00000	0.21456
12	0.008350	1.00000	0.21456
13	0.017941	1.00000	0.21456
14	0.001446	1.00000	0.21456
15	0.003488	1.00000	0.21456
16	0.000473	1.00000	0.21456

```
  TITLE:
09/16/03       PAGE : 22
  EQS/EM386 Licensee: Alison Bacon
1
  Execution begins at 16:19:40.71
  Execution ends   at 16:19:40.93
  Elapsed time =       0.22 seconds
```