VISUAL AND VERBAL PROCESSING IN REASONING.

by

PHILIP GRAHAM BROOKS.

Department of Psychology, Plymouth Polytechnic

in collaboration with

University College, London.

Thesis submitted to the Council for National Academic Awards in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

September 1984.

BEST COPY

AVAILABLE

Variable print quality

ACKNOWLEDGEMENTS.

I am indebted to my supervisor, Dr. J.St.B.T. Evans for his help, encouragement and, not least, his boundless patience. His advice has proved invaluable at all stages of the research and in the production of the final draft of this thesis.

I would also like to thank Dr. P.C. Wason for his supporting role in my supervision.

Thanks are due to several of the staff in the Department of Psychology, Plymouth Polytechnic, but most notably to Dr. I. Dennis for expert criticisms offered at various stages of my work.

I would also like to acknowledge the financial support given by the Science and Engineering Research Council. Without the research grant that they awarded to my supervisors, this project would not have been possible.

Finally, I should like to thank my wife, Janet, for her tolerance and resignation, especially during the last months of my work.

DECLARATIONS.

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

2) The experiments described in this thesis form part of a more extensive research project entitled "Investigation of the Dual Process Theory of Reasoning" funded by a grant to Drs. J.St.B.T. Evans & P.C. Wason from the Science and Engineering Research Council (S.E.R.C. reference GR/A55/851) during the period from 1st September 1978 to 31st August 1981. I was employed as a Research Assistant on the project throughout this time. Although the project involved research of a collaborative nature, the experiments reported in this thesis are mainly of my own design.

3) Experiment I forms the basis of a paper entitled:

Competing with Reasoning: A Test of the Working Memory Hypothesis.

It was written jointly by J.St.B.T. Evans and myself, and was published in:

Current Psychological Research (1981)
Volume 1, part 2, pages 139 - 147.

The paper is contained in Appendix A of this thesis.

ABSTRACT.

Visual and Verbal Processing in Reasoning, by P.G. Brooks.

This programme of research, involving seven experiments, investigates Evans' (1980a; 1980b) revised version of the Dual Process theory of reasoning (Wason and Evans, 1975). A Type 2 process is characterised as verbal-rational and a Type 1 process as non-verbal and non-logical. Evans links the processes to two statistical components of observed reasoning performance. The Type 1 process reflects non-logical response biases and the Type 2 process reflects attention to the logical nature of the task.

Six experiments employ a concurrent articulation (with or without a short-term memory load) methodology devised by Baddeley and Hitch (1974) for investigating their Working Memory model. Four experiments apply this technique to conditional reasoning tasks in an attempt to disrupt the verbal Type 2 process.

Some weak evidence for the revised Dual Process theory is found. There is a tendency, marked in only one experiment, for concurrent articulation to inhibit logical performance, whilst having little effect on response biases. Unexpectedly, articulation conditions (without memory load) are characterised by _faster_ responding than silent conditions.

The results are inconsistent with Hitch and Baddeley's (1976) data and several features of their Working Memory model. Two further experiments repeat and extend their work. A number of important theoretical implications are discussed in the light of recent revisions to their theory (eg. Baddeley, 1983).

A possible connection is drawn between Type 1 and Type 2 processes and dual memory codes (Paivio, 1971; 1983) and thought systems (Paivio, 1975) of a verbal and visual nature. The hypothesis that Type 1 processes may be associated with visual mechanisms is tested by introducing a factor into three experiments to induce use of a visual code. This does not affect the Type 1 process but facilitates _logical_ performance. These results are discussed in relation to the revised Dual Process theory. An explanation in terms of a recent tricoding model for processing of pictures and words (Snodgrass, 1980; 1984) is suggested.

CONTENTS.

INTRODUCTION.

There are many different routes to the study of human psychology. The cognitive approach can be distinguished from others quite readily. Cognitive Psychology can be defined as "the scientific analysis of human mental processes and memory structures in order to understand human behaviour" (Mayer, 1981, p1). Whilst Behaviourists banished the notion of mental experience or of any unobservable events in their explanations of behaviour, the cognitive approach places emphasis on the analysis of mechanisms underlying behaviour in order to explain that behaviour. However, cognitive processes should only be postulated if they can lead to testable predictions about observable behaviour.

Cognitive psychology adopts an information processing model in its approach to understanding behaviour. In this model the human being is viewed as a processor of information somewhat analogous to a computer. Input information enters through the senses and a number of mental operations are performed on it, thus changing it, until an output is generated. The information processing model is concerned with the cognitive operations, such as coding, storing, retrieving and transforming input information, which are employed in any given situation. Testable predictions about the latency or nature of responding to particular kinds of input are made according to the specific operations which are postulated during the processing stages.

This thesis adopts a cognitive approach to the study of thinking. Whilst Behaviourists may conceive of thinking as subvocal speech (eg. Watson, 1930), it is viewed by cognitive psychologists as a form of information processing which mediates between stimulus and response. In the present case, the main focus of attention will be on that aspect of thinking known as reasoning, mainly conditional reasoning. Until recently this field has been somewhat isolated from many of the mainstream issues of concern in cognitive psychology. In the study of conditional reasoning considerable attention has been paid to the role of logic and the implications of reasoning research for rationalistic explanations of behaviour. Although these are interesting matters, the present research focus will lie elsewhere.

In the psychology of thinking dichotomies seem to abound. As Neisser (1963) observes, thought processes have been divided into two or more sorts by several distinguished theorists. Examples include the distinctions between: productive and blind, creative and constrained, autistic and realistic, primary process and secondary process, intuitive and rational, and multiple and sequential thought. Since Neisser's article, further dichotomies have arisen, for example between visual and verbal thought processes (eg. Paivio, 1975). The present research investigates the cognitive mechanisms underlying reasoning performance. It is mainly concerned with the experimental investigation of the Dual Process theory of reasoning (Wason and Evans, 1975; Evans and Wason, 1976). In a revised form, (Evans, 1980a; 1980b), it postulates reasoning processes of a verbal and non-verbal nature. These processes are linked to two orthogonal statistical components which account for performance on conditional reasoning tasks.

In order to assess the revised Dual Process theory of reasoning, the research reported here employs competing task methodology. As we shall see, such techniques are regularly used in cognitive psychology although they are not without critics. In essence the idea is that if two tasks require the use of a common mechanism then they will compete for its use, with consequent interference. In the present context, if a concurrent verbal task is performed by a subject engaged in conditional reasoning then we might expect the interference to be restricted to the verbal process leaving the non-verbal process relatively undisturbed. This kind of selective interference should be reflected in the performance data if the revised Dual Process theory is correct.

However, this project is also concerned with matters of relevance to the study of memory, imagery and more general cognitive theory. Reasoning is a complex, high-level cognitive process. It necessitates interaction between many lower-level processes involved in the comprehension, representation and manipulation of symbolic information in a working memory system. The exact nature of those processes and the representations on which they operate is of considerable interest in cognitive psychology. Also of interest is the current, unresolved debate between theorists who postulate a functional role for mental imagery in cognition and those who view imagery as an epiphenomenon resulting from a more abstract propositional representation. In this thesis an attempt is made to narrow the gap between research on reasoning and research on issues of more general interest.

The Layout of this Thesis.

This thesis is divided into three major sections. Section One contains three Chapters which review various theoretical areas of particular relevance to the project. Chapter 1 concentrates on the general issues and paradigms involved in conditional reasoning research and introduces the Dual Process theory of reasoning. Chapter 2 relates this theory to other research concerning the nature of coding processes in high-level cognition. In Chapter 3 theoretical and methodological issues are considered and a current theory of Working Memory is introduced. Section Two contains three chapters describing seven original experiments. Section Three contains two chapters which discuss the interpretation of the experiments in the light of the Dual Process theory and other literature considered in the review.

SECTION ONE.

REVIEW.

CHAPTER 1

THE PSYCHOLOGY OF CONDITIONAL REASONING.

This chapter will survey the psychological literature concerning conditional reasoning. In order to assist our understanding of this literature, a brief description of relevant aspects of the propositional calculus of logic will be given initially. Following this a number of experimental paradigms which have been used to study conditional reasoning will be outlined and several psychological studies which have utilised each of these will be reviewed. Theoretical approaches to the study of reasoning will be considered and one particular approach, arising from this research, which postulates dual thought processes in reasoning will be evaluated. Finally, recent modifications to this theory characterising discrete verbal and non-verbal thought processes will be described.

PROPOSITIONAL LOGIC AND CONDITIONAL SENTENCES

Logic is defined by Copi (1982, p3) as "the study of the methods and principles used to distinguish good (correct) from bad (incorrect) reasoning". He distinguishes this from 'the science of reasoning' which is part of the psychologists' domain. Reasoning is a particular sort of thinking in which inference takes place. Copi (1982, p5) defines 'inference' as "a process by which one proposition is arrived at and affirmed on the basis of one or more other propositions accepted as the starting point of the process". The validity (or correctness) of an inference is determined by examining the logical relationships between the propositions at the start and end points of the inference process. A proposition is either true or false. Unlike questions, commands and exclamations, only propositions can be affirmed or denied, or judged to be either true or false.

A conditional sentence of the form 'If it is red then it is a triangle' asserts a relationship between two propositions. One proposition ('it is red') is contained in the antecedent clause of the sentence, whilst the other ('it is a triangle') is contained in the consequent clause.

When considering propositional arguments, it is often convenient to strip sentences of their particular content in order to lay bare their logical form. When this is done it is conventional to substitute single letters (eg. p, q, r) for particular propositions. Consider the particular sentence: 'If it is red then it is a triangle'. The antecedent proposition can be replaced by the letter 'p' and the consequent proposition by the letter 'q'. We are then left with the conditional assertion: 'If p then q'. Any logical inference derived from this argument will be valid (ie. consistent with the laws of logic) no matter what particular content is substituted for the propositions 'p' and 'q'.

In standard logic the principle of bivalence is assumed and thus propositions are either true or false. In this system the fundamental operation of negation always reverses truth value. Thus if the proposition 'p' is true, then its negation 'not p' is false. The converse of this argument also holds such that if 'p' is false then 'not p' is true. In actual usage, however, the principle of bivalence may be considered inadequate and, as will be illustrated later, a third truth value of 'irrelevant' or 'indeterminate' is required.

In standard logic, an analysis of the sentence 'If p then q.' will reveal that four possible contingencies can be defined depending upon the combinations of truth value of the two propositions. This is illustrated below in Table 1.1 .

Proposition		Truth Table Case
p	q	Notation
True	True	TT
True	False	TF
False	True	FT
False	False	FF

Table 1.1 . The four possible combinations of truth value of two propositions used in standard logic and their notation.

In order to assess the validity of arguments arising from a conditional rule, it is essential to assign a truth value to each of the four truth table cases that can be derived from it. Unfortunately the interpretation of a conditional sentence is not entirely clear-cut and four possible truth tables have been assigned to it. Which of these is deemed appropriate for a linguistic circumstance will depend upon the particular content of the propositions and upon the context in which the conditional is used.

Logicians have distinguished Material Implication and Material Equivalence relationships. In the former relation, p implies q which means that p could never be observed without q. The relationship is false when p is true and q is false, and is true otherwise. The Material Equivalence (or bi-conditional) relationship means that p implies q and also the converse, q implies p. Therefore the relationship will be true when both p and q are true or when both p and q are false, otherwise the relationship is false. The truth tables for Material Implication and Material Equivalence are shown in Table 1.2 .

Truth Value of p	Truth Value of q	Truth Value of 'If p then q'. M.I.	M.E.
True	True	True	True
True	False	False	False
False	True	True	False
False	False	True	True

Table 1.2 . Truth Tables showing Material Implication (M.I.) and Material Equivalence (M.E.) for the rule 'If p then q'.

Although logicians often use statements such as 'If p then q' to denote implication, various other linguistic possibilities exist for this relation. Amongst these are 'q if p', 'whenever p then q', 'never p without q'. Although formally equivalent to each other, these may well entail very different psychological interpretations. Logicians suggest that the sentence 'If p and only if p then q' should be used to denote material equivalence. However, in common usage, the abbreviated form 'If p then q' is usual in both circumstances and semantic factors are used to aid precise interpretation. For instance it is obvious that the sentence 'If it is a dog then it is a mammal' does not entail its converse. A mammal may be a dog, a human or any other animal that suckles its young. However, in some circumstances, the converse form of a conditional does seem appropriate. For instance, the definitional rule 'If a person has an XY chromosone then that person is male' would seem to suggest that 'If a person is male then that person has an XY chromosone'. With conditional promises and threats (eg. 'If you mow the lawn then I'll give you five pounds') an equivalence is often assumed.

It has been argued by Kneale and Kneale (1962) that in conditional sentences where the antecedent is false, they have no application and, as a result, no truth value is appropriate. This interpretation is known as Defective Implication and goes beyond standard logic's principle of bivalence in that a third category of 'irrelevant' is required in the truth table. A truth table for Defective Equivalence can also be derived in which the FF case is considered 'irrelevant'. These defective truth tables are shown in Table 1.3 . Some experiments which apparently support defective interpretations of conditionals will be reviewed later.

Truth Value of p	Truth Value of q	Truth Value of 'If p then q'. D.I.	D.E.
True	True	True	True
True	False	False	False
False	True	Irrelevant	False
False	False	Irrelevant	Irrelevant

Table 1.3 Truth Tables showing Defective Implication (D.I.) and Defective Equivalence (D.E.) for the rule 'If p then q'.

It should be pointed out that for all four truth tables of the rule 'If p then q' the rule is true when both the antecedent and consequent are true (TT case). Also the rule is considered false in all cases when the antecedent is true and the consequent is false (TF case). In other circumstances the truth value of the rule is seen to be equivocal.

A number of inferences can be drawn from a conditional rule such as 'If p then q'. These are shown in Table 1.4.

Inferences drawn from			Validity	
"If p then q"	Given	Conclude	I.	E.
Modus Ponens (MP)	p	q	V	V
Denial of the Antecedent (DA)	Not p	Not q	F	V
Affirmation of the Consequent (AC)	q	p	F	V
Modus Tollens (MT)	Not q	Not p	V	V

Table 1.4 . Inferences which can be drawn from the rule 'If p then q' together with their validity under Implication (I.) and Equivalence (E.). (V = Valid, F = Fallacious).

Under either an Implication or an Equivalence truth table the inferences known as Modus Ponens and Modus Tollens are shown to be valid. Modus Ponens infers 'q' given 'p'. Modus Tollens infers 'not p' given 'not q'. Both of these inferences depend upon the fact that the TF truth table case is prohibited (see Table 1.2) and thus 'p' and 'not q' cannot occur together if the conditional is true.

Under the truth table for Equivalence, the FT case is also prohibited (see Table 1.2) and hence 'q' and 'not p' are also not permitted to occur together given the true conditional "If p then q". Hence the inferences known as Denial of the Antecedent (DA) and Affirmation of the Consequent (AC) are also valid under an Equivalence truth table. Thus given 'not p', 'not q' is inferred by DA. Also given 'q', 'p' is inferred by AC.

However under the truth table for implication the DA and AC inferences are shown to be fallacious. This is because a conditional denoting Implication states that 'q' must be true when 'p' is true. It does not state that 'q' cannot also be true when 'p' is false. Since the validity of the inferences is determined from the false truth table cases, they are unaffected by whether or not the truth table is defective.

So far we have only considered expressions with affirmative constituents. Obviously expressions can be derived which incorporate negative antecedents and consequents. A convenient notation will be adopted by referring to rules as AA, AN, NA or NN. These notations describe four possible combinations of affirmative and negative antecedents and consequents in the conditional sentences which are shown below in Table 1.5 .

Several experiments have manipulated the presence of negatives in conditional rules. On occasion these experiments have yielded data which have been of considerable theoretical interest. Some of these will be considered in the next section of this chapter.

Rule	Notation
If p then q.	AA
If p then not q.	AN
If not p then q.	NA
If not p then not q.	NN

Table 1.5 . The notation used to describe the four possible combinations of affirmative and negative antecedent and consequent in conditional sentences.

EXPERIMENTAL STUDIES OF CONDITIONAL REASONING

Two paradigms which have been used in experimental studies of conditional reasoning will be surveyed in this section. These are:

a) Inference tasks, and

b) Psychological truth table tasks.

a) Inference Tasks.

In this paradigm the tendency of subjects to make or withhold each of the four inferences shown in Table 1.4 is considered. Generally subjects are presented with a conditional rule together with a premise which either affirms or denies one component of the rule. For example, given the conditional rule:

If the letter is A then the number is 7,

together with the premise:

The letter is A,

the subject might be asked to state what conclusion follows, if any. With the above example, a Modus Ponens inference would lead to the conclusion that:

The number is 7.

Alternatively, the subject might be presented with a conclusion and asked whether it necessarily follows from the conditional rule and premise which have been given. Otherwise, he might be asked to assess the truth value of a given conclusion assuming that the conditional rule together with the premise which have been presented are true.

If humans reason strictly in accordance with the principles of formal logic, a position advocated by Henle (1962), then we might expect their responses to such problems to reflect their interpretation of the conditional sentence. On the one hand, given a Material Implication interpretation, subjects might be expected to endorse MP and MT as valid but to reject DA and AC as fallacious. Whereas, on the other hand, given a Material Equivalence interpretation, they would be expected to endorse all four inferences as valid. Table 1.4 illustrates these points. Group data might be expected to reflect both these interpretations and thus MP and MT should be consistently endorsed whereas DA and AC should be endorsed at some level between 0% and 100%, according to the proportion of subjects adopting a Material Equivalence interpretation.

Several experiments concentrating upon affirmative conditionals have shown that, with adult subjects, MP and MT inferences are usually endorsed and DA and AC are endorsed more frequently than not (see Wason and Johnson-Laird, 1972; Evans, 1982). Such data has sometimes led to the conclusion that most subjects interpret the conditional as an equivalence.

Let us consider the data from three experiments summarised by Evans (1982). The data shown in Table 1.6 give the percentage of adult subjects endorsing each of the four inferences in three separate studies.

Study	Inference			
	MP	DA	AC	MT
Taplin (1971)	92	52	57	63
Taplin & Staudenmayer (1973)	99	82	84	87
Evans (1977a)	100	69	75	75

Table 1.6 . The percentage of subjects endorsing each of the four inferences for an affirmative rule: 'If p then q'. (Data from Evans, 1982, table 8.1).

Taplin (1971) looked at the consistency with which subjects made each of the four inferences over a long series of thematic problems. Although only 45% of his subjects were consistent in their response, he found a tendency for all four inferences to be made. He concluded that the conditional rule was most usually interpreted as a biconditional having a truth table for Material Equivalence.

Taplin and Staudenmayer (1973, experiment I) replicated the Taplin (1971) study using abstract materials but with a higher degree of consistency. Evans (1977a) has reported data for affirmative rules which also fit the same general pattern of results. However, as will be seen, the conclusion based upon this pattern of data - that subjects generally interpret the conditional rule as an equivalence - is not necessarily justified.

A second experiment was performed by Taplin and Staudenmayer (1973, experiment 2) in which a slight procedural difference was introduced. Subjects were presented with a conditional rule relating to the combinations of letters of the alphabet which were permissible together. They were also given a premise affirming or denying one component of the rule.

In the first experiment subjects were asked to evaluate the conclusion as either 'true' or 'false' - yielding the data shown in Table 1.6 . In experiment II, however, subjects had the choice of three conclusions; 'always true', 'sometimes but not always true' and 'never true' (the word 'false' was used for half of the subjects, but this had no influence on the data). Although no specific data are recorded, Taplin and Staudenmayer report a much lower frequency of DA and AC inferences. The pattern was more consistent with an interpretation of the rule as Material Implication.

Obviously such a dramatic difference in results emanating from such a slight change in procedure emphasises the danger in extrapolating from single paradigms. However, a plausible explanation is offered by Evans (1978). He suggests that, with abstract materials such as those used by Taplin and Staudenmayer, there is no semantic basis to assist subjects in their interpretation of conditional rules as being equivalence or implication. When a less committal response category ('sometimes but not always true') is introduced, this is selected as an expression of this ambiguity. In a forced choice situation subjects opt for an equivalence interpretation on DA and AC.

There are further considerable difficulties for the interpretation of reasoning experiments which follow from the introduction of negative components into the rules. In Table 1.7 the four conditional inferences for rules involving negative components are summarised.

Rule	Inference							
	MP		DA		AC		MT	
	Given	Conclude	Given	Conclude	Given	Conclude	Given	Conclude
AA	p	q	$\bar{p}$	$\bar{q}$	q	p	$\bar{q}$	$\bar{p}$
AN	p	$\bar{q}$	$\bar{p}$	q	$\bar{q}$	p	q	$\bar{p}$
NA	$\bar{p}$	q	p	$\bar{q}$	q	$\bar{p}$	$\bar{q}$	p
NN	$\bar{p}$	$\bar{q}$	p	q	$\bar{q}$	$\bar{p}$	q	p

Table 1.7 . The four conditional inferences for rules involving negative components. ($\bar{p}$ = not p, $\bar{q}$ = not q).

Evans (1972b) required subjects to construct falsifying and verifying cases for each of the four rule types. He noted how, when AA rules are used alone, particular inferences become associated with either affirmative or negative conclusions. The acceptance or rejection rate of each inference type could well be affected by this difference. By using each rule type, the extent to which each inference produces affirmative or negative conclusions can be balanced. For instance, Table 1.7 shows that an MP inference is associated with a negative conclusion for AN and NN rules but with an affirmative conclusion for AA and NA rules.

Evans (1972c) required subjects to make inferences from rules defining which letters were allowed to be paired with which digits in imaginary letter-number pairs. Only MT and AC problems were studied in this experiment. The following is an example of one of the problems used:

Given: 1) If the letter is not G then the number is 9.

2) Not 9

Conclusion: G, Not G, Indeterminate.

The correct answer here is 'G' (a valid Modus Tollens inference), the choice 'Not G' would indicate susceptibility to the fallacious AC inference. If the minor premise had been '9' then the correct response would have been 'Indeterminate'. Table 1.8 presents the percentage of valid MT and fallacious AC inferences made with each rule type.

Rule	Inference	
	MT	AC
AA	91	32
AN	75	35
NA	38	61
NN	41	55

Table 1.8 . The percentage of valid MP and fallacious AC inferences made with each rule type. (Data from Evans, 1972c experiment 1, table 2).

Significantly more MT and less AC inferences were made on rules having affirmative antecedents. Thus subjects affirmed significantly more negative conclusions. Pollard (1979) suggests the following as three possible alternative interpretations of these data:

1) A response bias producing a preference for negative conclusions,

2) The greater confusion of NA and NN rules produces more

erroneous responses,

3) NA and NN rules tend to be interpreted as expressing Equivalence (If and only if not p then q) rather than implication. In this case the AC inference would be valid.

However, Evans (1977a) has also shown that the pattern of response to each of the four inferences is affected by the manipulation of negative components in the rules. In this study subjects were required to evaluate each of the four inferences. That is, given the major and minor premises, they had to decide whether a given conclusion followed. The systematic effects of negatives are shown in Table 1.9.

Rule	Inference			
	MP	DA	AC	MT
AA	100	<u>69</u>	75	<u>75</u>
AN	<u>100</u>	12	31	<u>56</u>
NA	100	<u>50</u>	<u>81</u>	12
NN	<u>100</u>	19	<u>81</u>	25

Table 1.9 . The percentage frequency with which each inference was made for each rule type. For clarity, decisions which entailed the acceptance of a negative conclusion have been underlined. (Data from Evans, 1977a, table II) .

Evans' data shows that the frequency of all inferences (except Modus Ponens which was smothered by a ceiling effect) varies significantly as a function of introducing negative components. As Pollard (1979) concedes, these data suggest that subjects are biased towards negative conclusions. Both of his alternative explanations would predict more DA responses on rules having negative antecedents and this was clearly not the case. However, the 100% correct acceptance of the MP inference emphasises the strong tendency to respond in accordance with the logic of the problem.

Studies of inference patterns, such as those outlined in this section have sometimes been considered as indirect measures of truth tables. Several authors have attempted to classify subjects as having a truth table for implication (when MP and MT inferences are made and DA and AC are withheld) or for equivalence (when all four inferences are made), (eg. Taplin, 1971; Taplin & Staudenmayer, 1973; Staudenmayer, 1975; Marcus and Rips, 1979). However this sort of approach can only be justified if we accept the view that people reason logically given their particular interpretation of the rules. The work of Evans (1972b; 1972c; 1977a) illustrates that the frequency with which subjects respond to conditional inference problems is distorted by a response bias producing a preference for negative conclusions.

It should be plain that the classification of subjects as possessing a particular truth table must take into account the consistency with which they conform to that particular truth table. It has already been noted that only 45% of Taplin's (1971) subjects were consistently truth-functional (ie. consistent with some kind of truth table) in their reasoning. Taplin and Staudenmayer (1973) found about 80% of subjects were consistent. In both these studies the majority were classified as 'equivalence'. However, in Taplin and Staudenmayer's second experiment when a third response choice of 'sometimes true (or false)' was added to the usual 'true' and 'false' options, only 50% were truth-functional and the majority of these were classified as 'implication'. As Evans (1982, p136) notes, quite apart from the "unimpressive" proportion of truth-functional subjects, the method of testing a subject's inferences should not influence their interpretation of the rules. He also shows how the inclusion of an indeterminate choice is logically necessary in certain circumstances. For instance, consider the problem below:

Given: 1) If the letter is H then the number is 7

2) The number is 7

Conclusion: The letter is H.

Assuming the truth of the premises, it is not possible to determine the truth of the conclusion in the above AC syllogism. Thus studies which have not allowed an indeterminate option are clearly somewhat lacking.

Staudenmayer (1975), who included an 'indeterminate' option, found 78% consistency with abstract, context-free materials but only 54.5% with concrete materials. Marcus and Rips (1979) claimed that the majority (52.5%) of their subjects' responses to conditional inferences in a variety of contexts (whilst including two and three choice response formats) were logically contradictory in the sense that no single truth-function could account for them. It is surprising that even with such low consistency levels, and it should be stressed that neither of these studies required <u>absolute</u> consistency, these authors still considered it worthwhile to classify their subjects into two types: those who interpret conditional rules as 'equivalence' (biconditional) and those who interpret such rules as 'implication' (conditional).

In any case, Evans (1982) points out that if subjects do consistently conform to a particular truth table, they are not necessarily <u>using</u> it. Some subjects may appear to be consistent by chance and others could be induced by non-logical factors to appear consistently truth-functional. These experiments can serve to illustrate that it is of paramount importance to fit the theory to the data rather than ignore data which does not fit a particular theoretical interpretation.

b) <u>Psychological Truth Table Tasks.</u>

The majority of authors referred to truth tables as either Implication or Equivalence and did not have any means of differentiating between non-defective and defective truth tables which were mentioned in the introduction to this chapter. However, Wason (1966) suggested that subjects have a Defective Implication truth table (TF??) for a conditional rule (see table 1.3). That is, when the presupposition stated by the antecedent of a conditional is unfulfilled, no association is made and the rule is regarded as neither 'True' nor 'False', but as 'Irrelevant'. As Wason (1968, p274) puts it the "assumption is that individuals are biased, through a long learning process, to expect a relation of truth, correspondence or match to hold between sentences and states of affairs" and we merely use a proposition or statement that something is false in order to make a deduction.

In a subsequent experiment, Wason (1968) used the

Truth Table Evaluation paradigm to investigate the hypothesis that subjects have a Defective Implication truth table for a conditional rule. In this paradigm subjects are presented with a conditional sentence together with examples of all four truth table cases (see Table 1.1) and are required to evaluate the rule as 'True', 'False' or 'Irrelevant'. In fact, Wason found that the pattern which most often occurred was that defined as Defective Equivalence in Table 1.3 . This result suggests that subjects were interpreting the conditional sentence as an Equivalence (or Biconditional) in this particular experiment. In support of Wason's general line of argument, when the presuppositions stated in the rule were unfulfilled (ie. the FF case) the rule was regarded as irrelevant to the situation at hand.

Another experiment, reported by Johnson-Laird and Tagart (1969), was intended to discover which truth table was psychologically appropriate for the Implication relationship. They concentrated upon four alternative linguistic forms in which implication could be expressed:

1) If p then q.

2) There isn't p, if there isn't q.

3) Either there isn't p, or there is q (or both).

4) There is never p without there being q.

It was expected that, if Wason's original hypothesis was correct, sentence 1) would be considered 'Irrelevant' when the antecedent was false. Also sentence 2), which is derived from the contrapositive 'If not q then not p', would be considered 'Irrelevant' when 'q' is true. However, sentences 3) and 4), which are not conditional sentences, would be less likely to elicit 'Irrelevant' judgements and more likely to produce the truth table for Material Implication (see Table 1.2).

Subjects were presented with examples of each of the four basic sentence types which expressed relationships between particular letters and numbers which could appear on the left- and right-hand side respectively of cards. A pack of cards with examples of all four truth table cases were given to the subject to be sorted into one of three categories: 'True', 'False' or 'Irrelevant'. The 'p' and 'q' terms were falsified in one of three different ways which, although logically equivalent as falsifications, might not be psychologically equivalent. Either an alternative letter (or number), or a geometric shape or a blank was used.

The authors did not report any differences between the three alternative ways of presenting false terms. It was found that for sentence 1) (If p then q), the most usual pattern of responding conformed to a truth table for Defective Implication, in accordance with Wason's (1966) original prediction. It was also found that, to a lesser degree, sentence 4) (There is never p without there being q) was most commonly interpreted as Defective Implication. However for the other sentences a wide range of responses was given. In summary, it was found that AA conditionals were most frequently interpreted as Defective Implication but also the linguistic form of the sentence used had a dramatic effect on the interpretation even though all sentences have the same truth table in formal logic.

It was suggested by Evans (1972b) that Johnson-Laird and Tagart, by giving the 'Irrelevant' category as a possible choice to subjects in their experiment, had introduced a strong demand characteristic for its usage. Evans decided to utilise an alternative procedure in order to check this possibility. He presented subjects with a series of abstract conditional rules concerning the relationships between various coloured shapes. An example of one such rule is:

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

Subjects were also presented with an array of coloured shapes. Their task was to construct as many verifying and falsifying cases of the given rules as possible. Since the procedure was exhaustive, Evans could infer that any logical cases which were not constructed were irrelevant. Another important innovation introduced in this experiment involved the manipulation of negative components in the rules. Although Johnson-Laird and Tagart had used negatives in some of their rules, all of the sentences expressed the same logical relationship (p implies q) and so the truth and falsity of components was confounded with affirmation and negation. By his procedure, Evans ensured that "overall the effect of instances matching (affirming) or mismatching (negating) values named in the rules should cancel out" (Evans, 1972b, p194). This is illustrated in Table 1.10 .

It can be seen from Table 1.10 that each of the four possible matching cases (pq, $p\bar{q}$, $\bar{p}q$, $\bar{p}\bar{q}$) appears just once for each of the rules but they are mapped differently onto each of the logical cases (TT, TF, FT, FF).

Rule	Logical Case			
	TT	TF	FT	FF
AA	pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$
AN	$p\bar{q}$	pq	$\bar{p}\bar{q}$	$\bar{p}q$
NA	$\bar{p}q$	$\bar{p}\bar{q}$	pq	$p\bar{q}$
NN	$\bar{p}\bar{q}$	$\bar{p}q$	$p\bar{q}$	pq

pq = double matching case $\bar{p}q$ = single mismatching case

$p\bar{q}$ = single mismatching case $\bar{p}\bar{q}$ = double mismatching case.

Table 1.10 . The combinations of affirmed and negated values constituting the four logical cases of the conditional rules.

The results of Evans' study are shown in Table 1.11 pooled over the four rules.

Logical Case	Classification		
	True	False	Irrelevant
TT	99	0	1
TF	3	80	17
FT	14	34	52
FF	33	23	44

Table 1.11 . Percentage frequency of construction of the four Logical Cases summed across the four Rules. N=24. (Data from Evans, 1972b. Table from Evans, 1982, table 8.5 - i,a).

It can be seen from Table 1.11 that the modal responses, when pooled over the four rules, support the prediction of the Defective Truth Table. This data can also be analysed according to matching case summed across the four rules. Since each matching case appears equally often for each logical case, any effect of this can be said to be non-logical. The data are arranged in this way in Table 1.12 .

Matching Case	Classification		
	True	False	Irrelevant
pq	34	52	14
p$\bar{q}$	41	33	26
$\bar{p}$q	40	27	33
$\bar{p}\bar{q}$	34	25	41

Table 1.12 . Percentage frequency of construction of the four Matching Cases summed across the four Rules. N=24. (Data from Evans, 1972b. Table from Evans, 1982, table 8.5 - i,b).

It can be seen that the percentage frequency of 'Irrelevant' items (ie. items not constructed) increases as the number of mismatches increases. Evans referred to this tendency to prefer to construct those values named in the rule as 'Matching Bias'. Its discovery emphasises the weakness of other studies of deductive reasoning which, in concentrating their attention on affirmative rules, have confounded such a factor with the truth and falsity of a rule's components. Evans (1972b) managed to measure a three-value psychological truth table without mentioning the concept of 'Irrelevance' to subjects and, as a consequence, he has avoided the criticism of a resultant demand characteristic.

In a succeeding study, Evans (1975) replicated his earlier results using a Truth Table Evaluation Task in which three possible choices were available to subjects. In this study two linguistic forms were used for the rules:

If p then q,

and p only if q.

These were referred to as IT and OI conditionals respectively. Whilst these rule forms are logically equivalent, the distributions of responses to them differed slightly, as can be seen from Table 1.13 .

Logical Case	Linguistic Form					
	IT			OI		
	True	False	Irrelevant	True	False	Irrelevant
TT	89	5	6	82	2	16
TF	9	81	9	11	58	30
FT	19	29	52	13	57	30
FF	30	11	57	44	16	40

Table 1.13 . Percentage frequency of evaluation of the four Logical Cases summed across the four Rules for Linguistic Forms of the conditional. N=48. (Data from Evans, 1975, table 1).

As can be seen, for the IT Form the data is very similar to that presented in Table 1.11 which resulted from the Construction Task. However, the modal responses to the OI Form correspond to a 'TFFT' Truth Table which is the truth table for Material Equivalence (see Table 1.2). The same data is presented, analysed according to 'Matching Bias' summed across the four Rules in Table 1.14.

Matching Case	Linguistic Form					
	IT			OI		
	True	False	Irrelevant	True	False	Irrelevant
pq	42	42	17	41	56	3
$p\bar{q}$	39	32	28	48	38	14
$\bar{p}q$	34	35	30	32	26	42
$\bar{p}\bar{q}$	32	18	50	29	14	57

Table 1.14 . Percentage frequency of evaluation of the four Matching Cases summed across the four Rules for Linguistic Forms of the conditional. N=48. (Data from Evans, 1975, table 1).

It can be seen that, for both rules, a similar tendency is present for 'Irrelevant' responding to increase as the number of mismatches increases. This tendency replicates that found by Evans (1972b) in the Construction Task. Thus it appears that the effect of 'Matching Bias' generalises to an alternative task and is not restricted simply to an 'if p then q' formulation of the conditional rule.

Another study by Evans and Newstead (1977) measured the latency of responding, as well as frequency of response, in a Truth Table Evaluation task. They were testing the psycholinguistic hypothesis that, although IT and OI forms can both be used to express Material Implication, the IT form is more natural when the antecedent event temporally precedes the consequent event and the OI form is more natural when the consequent event precedes the antecedent event. Evans and Newstead presented subjects with an IT or OI rule (relating to the order of presentation of two letters) on one field of a three-field tachistoscope. Subjects were required to push a button to display the rule on the screen and a second button push indicated that they had understood the rule and were ready to perform the reasoning task. This interval is the Comprehension time. Following the second key press, two capital letters were presented one after the other (for one second each) on the remaining two fields of the tachistoscope. The subject was required to decide whether the pair of letters 'conformed to', 'conflicted with' or was 'irrelevant to' the rule and to indicate their answer by pushing the appropriate response key. The interval between the second and third button presses was the Verification Time. Although temporal order did not significantly affect the nature of the responses made, the latency data confirmed that both types of conditional sentence were processed faster when their linguistic directionality was congruent with the temporal order of the events they described.

Their data also indicated a tendency for Comprehension and Verification latencies to increase as negatives were introduced into the rules. The effect of negatives in each component was additive. Verification latencies also increased for the more complex Truth Table Cases with the overall order being: TT < TF < FT < FF. In line with previous studies (Evans 1972b, 1975), Evans and Newstead reported that Logical and Matching tendencies were present. There was an overall Logical tendency to regard the 'TT' case as 'True' and the 'TF' case as 'False' and the effects of

'Matching Bias' were weakest on these two cases. Evans and Newstead account for this in terms of "some form of competition between the two tendencies" (Evans and Newstead, 1977, p280) - a point to which I shall return in the next section.

The Evans and Newstead (1977) study was novel in that it measured latencies on a task known to produce large variations in response frequencies. Their latency data provided additional useful information which helped in the interpretation of their results. Although the response frequency data did not show any effect of psycholinguistic temporality, significant effects were demonstrated in the Comprehension and the Verification latencies. They considered the distinction between Comprehension and Verification periods to be particularly important. The Comprehension Time measure was considered useful "for distinguishing interpretational from operational factors" - see next section - in that Comprehension latency can be "regarded as a 'pure' measure of interpretation in that it is measured prior to the commencement of any reasoning operations". Verification Time is harder to interpret "owing to the concurrent variations in response frequency" (Evans and Newstead, 1977, p281). However, in this study, Verification Latency was found to reflect Interpretational Factors, for example relating to negatives in the rules, but also revealed the effects of Truth Table Case which clearly arose in the Operational Stage.

In this section several important experiments, performed in two major reasoning paradigms, have been discussed. The effects of linguistic features were shown and various response biases were revealed. At this point it is appropriate to consider the range of theoretical interpretations which have been developed to account for data such as these.

THEORETICAL INTERPRETATIONS OF REASONING DATA.

As Evans (1972a) noted, in many reasoning experiments the arbitrary criterion of correctness as provided by the rules of formal logic has been used. This is insufficient to explain adequately the observed behaviour which constitutes experimental results because it entails the assumption that reasoning problems are perceived and solved by the subject in the same sort of manner in which they are conceived by the experimenter - as logical problems. He claims that psychological factors quite unconnected with logic have often been ignored and, as a consequence, results have been misinterpreted and faulty theories have evolved.

Evans (1972a) has distinguished three types of theories of reasoning including logical, illogical and non-logical. Each of these will be considered in relation to propositional reasoning. Perhaps the main proponent of a logical theory which I will consider is Mary Henle (1962). She proposed that reasoning essentially follows the laws of logic and that mistakes occur only when subjects misinterpret the given problem. Her claims were based on a selective analysis of the protocols of subjects who were given thematic syllogisms to solve. Specifically, she claimed that errors occurred due to premises being omitted, incorrectly interpreted, additional premises being added or to a failure to accept the logical task. Her theory led to a rationalist revival in the psychological literature emanating mainly from the USA (eg.

Staudenmayer, 1975).

An effective challenge to Henle's position has been made by Evans (1972a). He points out that there are two types of factor which are likely to influence a subject's behaviour in reasoning experiments. One of these factors relates to the subject's comprehension of the sentence forming the rule and is referred to as an interpretational factor. In addition 'task variables' should be distinguished. These refer to "the influence of certain operational requirements of the task which act independently of the subjects' interpretation of the sentences" (Evans, 1972a, p376) and are referred to as operational variables. Two striking examples of operational variables were discussed previously. One, pertaining to the preference for negative conclusions, was found in the inference task. Since it is not limited to one kind of inference nor indeed to conditionals (Roberge, 1976, finds a similar effect with exclusive disjunction) it is referred to as an operational rather than an interpretational factor. The other notable example is that of 'Matching Bias' which was discussed in relation to the truth table paradigm and which also generalises over different reasoning tasks and rule formulations (Evans, 1972b, 1975; Evans and Lynch, 1973).

In assessing the rationalist viewpoint we should consider the consistency of subject's solutions to given problems. Staudenmayer (1975, p78), for instance, writes that, according to Henle and her followers, "once an individual accepts the most plausible interpretation for him, the evaluations follow consistently and logically". In fact the studies reviewed in the previous sections of this chapter showed that such interpretations are far from consistent, particularly when abstract materials are used. In addition it has been shown that the introduction of negatives into the rules used in the inference task substantially affected the frequency with which particular inferences were endorsed. For instance, fewer AC and more MT inferences were made when the rules had affirmative antecedents - that is more negative conclusions were endorsed (Evans, 1972c). Now, whilst Henle could argue that the introduction of negatives alters the interpretation of the rule, this viewpoint is hardly tenable since both implication and equivalence interpretations of the rule require the MT inference. An alternative hypothesis, proposed by Evans (1978, p100), states that "a non-logical response bias acts against any inference in which the subject is required to infer the falsity of a component which is negative". In this case, when the consequent is negative, less DA and more MP inferences would be expected. In support of this hypothesis, Evans (1977a) found that less DA inferences were made with AN and NN rules but, unfortunately, the MP inference suffered a massive ceiling effect and was always endorsed. Another non-logical response tendency (Matching Bias) has been demonstrated in various paradigms including truth table construction (Evans, 1972b), truth table evaluation (Evans, 1975; Evans and Newstead, 1977) and in another paradigm known as the Wason Selection Task ((Evans and Lynch, 1973). The evidence observed in propositional reasoning is overwhelmingly against the extreme rationalist position advocated by Henle (1962).

The origins of certain non-logical response factors, are clearly linguistic in nature. Recently, Evans (1983a) has shown how linguistic features can even affect 'Matching Bias'. He presented subjects with a truth table task using conditional rules such as the following:

If the letter is not K then the number is 3.

The instances associated with such rules were varied for two groups of subjects. In the first group, instances employed implicit negation, as is usual with this task, to form the various logical cases . For instance the FF logical case for the above rule would be:

The letter is K and the number is 5.

The second group received instances which employed explicit negation, so that the named items in the instance always matched the items in the rule. An example is given in the following FF logical case for the above rule:

The letter is K and the number is not 3.

Evans found that the usual 'Matching Bias' effect was significantly reduced, although not completely absent, for the explicit negative group. Evans argues that the use of negatives in the instances could account for the residual 'Matching Bias' effect. After all negative statements can often be seen to cause difficulty or confusion in various tasks (see Evans 1982, chapter 3) and this could lead to greater use of the 'irrelevant' response choice as the number of negatives in the instance increased. The logical performance of the explicit negative group was also significantly improved compared with the implicit negative group and this suggests that some general facilitation occurs with explicitly negated instances. However an alternative hypothesis which could explain 'Matching Bias' is not ruled out by Evans' demonstration. The alternative explanation of 'Matching Bias' involves the possible use of visual imagery but discussion of it is deferred until the next chapter where its plausibility will be established.

In order to discuss illogical theories of conditional reasoning, it is necessary to consider research concerning another reasoning paradigm, the Wason Selection Task.

<u>The Wason Selection Task.</u>

In its original form (Wason, 1966) subjects were shown an array of four cards and were told that every card had a letter on one side and a number on the other side. Only one face of each card was revealed and these displayed a vowel (p), a consonant ($\bar{p}$), an even number (q) and an odd number ($\bar{q}$). The subject was then given the following conditional rule;

'If a card has a vowel on one side, then
it has an even number on the other side'.

He was told that this rule related only to the four cards in front of him. The subject's task was to name those cards, and only those cards, which must be turned over to discover whether the rule was true or false. The solution to this problem is p and $\bar{q}$, since only this combination can falsify the rule. However, the vast majority of subjects selected either the p card alone or the p and q cards.

Wason (1966) proposes that subjects assume a conditional rule to have three truth values: True, False and Irrelevant. Vowels with even numbers verify, vowels with odd numbers falsify and consonants with any number are irrelevant. In addition they are inclined to verify, rather than falsify, the rule since in everyday life conditionals are only used if they are true. He suggests that "in adult experience truth is encountered more frequently than falsity, and we seldom use a proposition or judgement that something is false in order to make a deduction" (Wason, 1968, p274). Although subjects could determine, prior to a selection task, which combinations of letter and number would make the rule false, their selection of the $\bar{q}$ card was not facilitated in the task itself. In fact, several attempts to simplify the task by using binary stimuli and simpler forms of the rule (Wason, 1969) have had little effect. The possible confusion of referring to 'the other side' of the card - which could be interpreted as being the side which is face downwards - was eliminated by Wason and Johnson-Laird (1970) by presenting all of the information so that it was potentially visible on the same side of the card, but to little avail. Even when therapy was introduced to induce insight into the correct solution after a selection task, by making subjects aware of the falsifying case and that selection of the $\bar{q}$ case can produce it, several subjects still declined to revise their original selections in a subsequent task (Wason, 1969; Wason and Johnson-Laird, 1970).

Johnson-Laird and Wason (1970) attempted to account for the results obtained on the selection task with an information processing model. Basically they assumed that subjects could be in one of three possible states of insight when performing the task: No insight in which subjects attempt to verify the rule, partial insight in which the necessity for falsification is combined with the desire to verify, or complete insight in which subjects only select potential falsifiers. They assume that as the subject gains insight he switches his attention from verification to falsification of the rule. However since these states are defined in terms of the combinations of cards which subjects select (p or p & q, p & q & $\bar{q}$, p & $\bar{q}$ respectively), as a consequence their definition is circular. A point that has not escaped Evans (1977b).

The previous research on the selection task had concentrated upon affirmative (AA) rules. However, the existence of 'Matching Bias' in the truth table construction paradigm (Evans, 1972b) suggested a plausible explanation of performance on the selection task without reference to verification bias (Evans and Lynch, 1973). Evans and Lynch introduced negatives into the conditional rules (see Table 1.5) used in four selection tasks and found that 'Matching Bias' exerted a powerful influence on responding. Overall there was no evidence of verification bias but a preponderance to choose logically correct values (p & $\bar{q}$) was found. This could not be explained in terms of the insight model.

However several authors have referred to their subjects' verbal protocols in defending insight models (eg.Goodwin and Wason, 1972; Bracewell, 1974; Smalley, 1974). The protocols seem to suggest that responses are due to interpretational factors or verification bias. Indeed Evans (1972a) has been criticised by Van Duyne (1973) for failing to include 'thinking aloud' protocols in his 'Matching Bias' experiments. Unfortunately such evidence as is available comes exclusively from studies which used the affirmative (AA) form of the conditional rule. In order to rectify this state of affairs, Wason and Evans (1975) performed an experiment in which the consequent of an abstract conditional rule was either negated or not. Two independent groups of 12 subjects each performed the selection task with both of these rules. One group performed the affirmative task first and the other performed the negative task first. Subjects were asked to write down the reasons for their selections or non-selections of each of the four cards. As can be seen from Table 1.15 performance was dominated by 'Matching Bias'.

Values		Rule Type					
Selected		Affirmative			Negative		
	Order	1st	2nd	Total	1st	2nd	Total
pq		6	6	12	9	6	15*
p		2	2	4	2	4	6
$p\bar{q}$		0	0	0*	0	0	0
Others		4	4	8	1	2	3
N		12	12	24	12	12	24

* = Correct response

Table 1.15 . The frequency of responses in affirmative and negative selection tasks. (Data from Wason and Evans, 1975, table 1).

However, the reasons given varied according to the logical consequence of the responses. With the affirmative rule subjects claimed that their selections were aimed at verification whereas, when the negative was introduced, they claimed that their selections were aimed at falsification. Wason and Evans' (1975) explanation of this apparent paradox is discussed in the next section.

The Dual Process Theory of Reasoning.

In explaining data from the selection task, Wason and Evans (1975) proposed a non-logical model - the Dual Process theory of reasoning - which accommodates the subjects' performance and also their own explanations of their performance. The Dual Process hypothesis postulates that performance and introspection reflect different underlying processes. Two fundamental assumptions were entailed:

1) Operational processes (Type 1) underlying reasoning performance (eg. Matching Bias) are not generally introspectible.

2) Introspective reasons (Type 2) do not reflect the underlying thought processes which caused the selections, but are rather a justification of the subject's behaviour in the context of the experimental situation and instructions.

Previously Evans (1972a) had been criticised for over-emphasising the importance of non-logical operational variables and for failing to indicate how they interact with interpretative processes in reasoning (Van Duyne, 1973). In the weaker form of their theory, Wason and Evans (1975, p150) suggest "a process of rapid continuous feedback between tendencies to respond and consciousness rather than two temporally distinct phases". This being the case, one might wonder how a response is eventually selected. However, the stronger form of their theory assumes that response determines conscious thought in which case the interpretative process is a consequence of selection behaviour. Circumstantial evidence for the hypothesis has been claimed in an inductive reasoning problem - the '2, 4, 6 problem' - investigated by Wason (1960; 1968). Here, subjects are required to discover a rule by generating triads of numbers. They are given feedback about whether the triads conform to the rule. It appears that subjects often reformulated, without awareness, the same hypothesis about the rule after the first formulation had been pronounced incorrect. Wason and Evans (1975, p152) suggest that the first hypothesis "continues to exert itself unconsciously but allows a conscious displacement to fulfil the requirements of the task".

Most of the supportive evidence comes from selection task experiments. Indeed the data of Wason (1969) has been reinterpreted in terms of dual processing (Wason and Evans, 1975). Twenty subjects were given the correct solution in a selection task and asked to give reasons why it was correct. All subjects accomplished this and it was originally inferred that subjects were prevented from imposing their own erroneous structure on the task. In the light of their theory Wason and Evans (1975, p151) predict that "'reasons' would be found to satisfy the purported correctness of any common wrong solution". This prediction was tested by Evans and Wason (1976) by giving one of several different 'solutions' to independent groups of subjects each of whom was told it was correct. Their prediction was supported and, furthermore, subjects generally expressed confidence in the correctness of their reasons.

As outlined above, the original form of the Dual

Process theory envisaged that the operational (Type 1) and interpretational (Type 2) processes ran in alternation. Unfortunately, this viewpoint could not be reconciled with another theory which explains reasoning performance in terms of a probabilistic mathematical model (Evans, 1977b).

Evans' stochastic approach was developed in order to account for the observed variability of data collected in reasoning experiments. Whereas previous models of the reasoning process had accounted for variability in terms of individual differences such as experience or intelligence, Evans (1977b) considered the alternative possibility that reasoning behaviour is intrinsically probabilistic. For instance if all subjects have a Ø.6 probability of making a certain response then about 6Ø% of a sample of subjects would make that response.

In re-analysing the data from several selection task experiments, Evans found that card selections were statistically independent. Thus previous (eg. insight) theories which attached psychological significance to particular combinations of card selections were rendered unparsimonious since, as Evans (1977b, p624) writes, " the combination of selections observed in an individual would be the result of independent stochastic processes: a 'statistical accident' of no psychological significance in itself".

The mathematical model incorporates Evans' (1972a) two factor approach and proposes that the probability of a particular response (r) reflects a combination of interpretational (I) and operational or response (R) factors. In more formal terms, it states that the probability of a particular response Pr(r) is equal to the weighted addition of

I and R factors:

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

where α is the weighting factor and:

$$0 <= \alpha <= 1$$

$$0 <= I <= 1$$

$$0 <= R <= 1.$$

Now, as Evans (1980a) points out, this sort of additive probabilistic model implies parallel rather than sequential processing. This is because the response is made either on the basis of logic or on matching depending upon the value of the weighting factor. As a consequence it is not compatible with the initial formulation of the Dual Process theory of reasoning which, as stated above, envisaged alternating processes.

The above leads to the first radical amendment of the theory. Evans (1980a) proposes that the underlying Type 1 and Type 2 processes operate in parallel. The Type 2 process is claimed to be involved <u>prior</u> to making the response and is equated with the Interpretational component of Evans' (1977b) model. Thus, it is claimed to be responsible for the logical component of reasoning behaviour. This process competes, for the control of the response, with the Type 1 process now equated with Evans' (1977b) response bias factor.

At this point, it is worth considering briefly how this approach can be used to explain the so-called 'thematic facilitation effect' which has been observed in various reasoning paradigms (see Evans, 1982; Griggs, 1983). It has been demonstrated by Evans and Lynch (1973) that, with abstract materials, selection task data reflects a combination of logical (I) and matching (R) tendencies. Several early experiments have shown that thematic content (ie. anything other than arbitrarily related symbols and forms such as those commonly used in the task - cf. Griggs, 1983), in an appropriate context leads to improved performance (eg. Wason and Shapiro, 1971; Johnson-Laird, Legrenzi and Legrenzi, 1972).

Such effects have caused problems for Piaget's classical theory of Formal Operations and have led him to make revisions to it (Piaget, 1972). However, the revised Dual Process theory can explain such effects quite adequately by proposing that, when the subject's understanding of the sentence is facilitated by thematic materials, more weighting is attached to the verbal, interpretational process than the operational one thus leading to improved performance.

It should be stated that thematic materials' effects are elusive, however, as Griggs (1983) has shown in his review. He suggests that when substantial facilitation has been observed with thematic materials, the effect could have resulted from "the cueing of familiar relevant material in long-term memory, instructions that conceivably biased subjects' strategies, and a problem context that may have changed the nature of the selection task" (Griggs, 1983, p31). However, since the effect has lead to considerable research activity and debate in the literature (eg. Manktelow, 1978; Manktelow and Evans, 1979; Pollard, 1981; Griggs, 1983; Wason, 1983), it is important that Evans' revisions to the Dual Process theory should encompass it.

In another publication (Evans, 1980b), other important modifications are suggested. For instance, Evans now refers to the Type 2 process as verbal and the Type 1 process as non-verbal rather than continuing with the somewhat vaguer conscious/unconscious distinction. Thus a verbal (Type 2) process is seen as a cause of reasoning behaviour, in parallel with a non-verbal (Type 1) process, rather than merely being adept at generating _post hoc_ rationalisations.

It has been shown how the revised theory allows the possibility of a verbal rational process acquiring control of behaviour - for instance, when realistic materials are used. However, since the verbal process is not introspectible, introspective reports are still viewed as rationalisations and as products of the Type 2 process, rather than a description of it. In certain circumstances subjects' rationalisations may appear wholly appropriate, when problems lie within their competence or experience. However, as Pollard (1979) states, this does not mean that one can use introspections to infer the process underlying behaviour in the way that Van Duyne (1973) suggests, since they are essentially _post hoc_ rationalisations (cf. Nisbett and Wilson, 1977).

Speculation as to the origin of the dual processes was made by Evans (1980b) when he suggested a link with hemispheric specialisation. Much of the evidence reviewed by Cohen (1983) suggests that the left hemisphere is specialised for verbal and the right hemisphere for non-verbal processing. Some modest evidence in support of Evans' hypothesis was apparent from a study by Golding, Reich and Wason (1974). Their subjects performed a selection task using the tactile modality, in which information was presented to the right or left hand. Subjects were given the opportunity to revise their original selections following the presentation of the solution to one of their hands. They found a tendency (just below significance) for better performance with information presented to the left hemisphere. Other stronger evidence (Golding, 1981) indicates that patients with right hemisphere lesions perform better on the selection task than those with left hemisphere lesions or than those in a control group. In a follow up study (Golding, 1980) she found that ECT administered to the non-dominant (usually right) hemisphere led to improved performance on the task relative to a control group. This evidence is suggestive that "logical performance on the selection task is normally inhibited by competing influences from the right hemisphere" (Evans, 1982, p251). This suggestion is in line with the latest revised version of the Dual Process theory.

The use of abstract materials in studying reasoning performance has been criticised by some authors who believe that _only_ thematic problems are worth studying (eg. Johnson-Laird and Steedman, 1978; Fillenbaum, 1978). Whilst the study of reasoning with thematic materials is of considerable importance, it is difficult with realistic materials to isolate the effects of the semantic context of the problem from purely logical effects. For instance, a subject's prior beliefs or prejudices are likely to influence judgement when realistic materials are used. As Evans (1982, p226) writes "the point about realistic materials is that they induce responses that are _appropriate_ to our experience, which may or may not correspond to a logical definition of validity". The use of abstract content makes it impossible for subjects to make direct use of their previous learned experience and so, as Evans (1983, p636) writes, "logical ability is, then, best assessed on abstract problems, where prior knowledge can neither help nor hinder reasoning". On this basis it is considered wise to employ abstract materials which arouse few semantic associations in experiments designed to assess the nature of the processes underlying conditional reasoning performance.

Whilst the Dual Process theory was initially developed to account for performance obtained on Wason's selection task, it is also applied to explaining performance on other reasoning paradigms (such as those considered earlier in this chapter) in which logical and non-logical processes are thought to operate. The revised formulation of the theory obviously renders it considerably broader in scope. Its increased precision also makes it more susceptible to testing and, in these senses, it is a better theory. Although several aspects of it are of considerable importance, the distinction between Verbal and Non-verbal processes is of particular interest in the present case. However, before this aspect is tested experimentally, some broader links with other ideas will be explored in order to formulate additional hypotheses about the nature of the dual processes. In the next chapter of the review, parallels will be drawn between the revised theory and other contemporary models of cognition.

CHAPTER 2

THE NATURE OF CODING PROCESSES IN HIGH-LEVEL COGNITION.

One aspect of Evans' (1980a; 1980b) revised Dual Process theory of reasoning that is particularly interesting is the distinction between competing verbal and non-verbal processes which underlie the logical and non-logical factors of conditional reasoning performance (Evans, 1977b). The notion of dual codes in the elementary stages of cognition is not uncommon (eg. Posner, 1973). However Evans' hypothesis that dual processes influence the more advanced stages is much more controversial although the idea is not without precedents.

A particular advantage of the verbal/non-verbal distinction lies in its connection with the theoretical proposals of others. Neisser (1963), for instance, proposes that a main sequence of verbal thought interacts with multiple pre-attentive processes. However in this case the main sequence is identified with consciousness whilst the multiple processes are said to be more effectively active at a pre-conscious level. A much more influential theory that distinguishes discrete verbal and non-verbal systems of thinking was derived from the Dual Coding Hypothesis originally proposed by Paivio in 1971. This theory will be considered in the following section.

PAIVIO'S DUAL CODING HYPOTHESIS.

The idea of dual cognitive systems of equal status has been most obviously considered in the study of mental imagery. In 1971, Paivio published an important book in which the influence of imaginal and verbal symbolic processes were assessed in relation to problems of meaning, perception, learning, memory and language. The empirical approach which he adopted involves the use of three types of converging operations which are all, in the words of Paivio (1971, p9), "conceptually linked by the postulated imaginal and verbal symbolic processes". These operations include:

1) Attributes of stimulus materials, with particular emphasis on their concreteness/abstractness properties,

2) Experimental manipulations, such as differential task instructions, presentation rates and task demands,

3) Individual difference variables.

The evidence gathered from this approach led to the Dual Coding Hypothesis (Paivio, 1971) which postulates the existence, in memory, of two independent but interconnected coding systems - one verbal and the other imaginal - operating in parallel.

More recently, Paivio (1975) has extended the notion of dual codes in memory to dual systems in thinking. He assumes (p147) that thinking involves a continuous interplay of non-verbal imagery with verbal symbolic processes, "which though inter-connected are functionally distinct". As Paivio (p147) writes "however else it might be characterised, thinking clearly involves taking in or encoding stimulus information, organising and storing it in memory and retrieving that information according to the requirements of a given task". The differential effects of imagery and verbal processes on each of these elements of memory led to his proposal of the Dual Coding Hypothesis.

He suggests that the imagery system is specialised for processing non-verbal information and is characterised by concrete, analogical thinking. The imagery process contributes to the richness of content, flexibility and speed of thinking. By contrast, the verbal system is characterised as an abstract, logical mode of thinking which limits memory content and flexibility but contributes logical direction to thinking. The interconnectedness of the systems "means that representations in one system can activate those in the other, so that for example, pictures can be named and images can occur to words. Independence implies, among other things, that non-verbal (imaginal) and verbal memory codes, aroused directly by pictures and words or indirectly by imagery and verbal encoding tasks, should have additive effects on recall" (Paivio and Lambert, 1981, p532 - 533). Paivio (1983, p309) identifies the imagery and verbal systems in terms of "synchronous and sequential processes, correlated with the contrast between analog and discrete representations".

There are some obvious similarities between the theories of Paivio (1975) and Evans (1980a; 1980b) in that a logical process which is verbal in nature and a non-logical, non-verbal process are believed to operate in thinking. Admittedly Evans does not claim that his non-logical process is imagery-based, but this would appear to be a reasonable possibility. The origin of 'Matching Bias' could be derived quite plausibly from the operation of a visual imagery system. The focus on values which are perceptually present could reveal the influence of a concrete, visual system of thinking which interferes with the abstract logical thought that successful conditional reasoning necessitates.

However, this kind of approach would be contested by certain theorists (eg. Pylyshyn, 1973; Anderson and Bower, 1973) who assume that the coding processes underlying memory and thought consist of abstract propositions that are neutral with respect to input modality, including the verbal/non-verbal dichotomy. Before the empirical evidence relating to Paivio's Dual Coding Hypothesis is assessed the theoretical dispute concerning the functional status of imagery will be outlined.

THE IMAGERY DEBATE

For several years there have been differing views expressed in the literature about the nature of the coding and processes underlying various cognitive acts. There are two main opposing theoretical camps in this debate and, in this section, some of the main points under consideration will be outlined.

The so-called 'Imagist' position claims a functional role for mental imagery in cognition (eg. Kosslyn and Pomerantz, 1977; Kosslyn, Pinker, Smith and Shwartz, 1979; Kosslyn, 1981). Its main opposition comes from 'Propositionalists' who suggest that all, including visual, information is internally represented by means of abstract propositions and that cognitive operations consist of their manipulation (eg. Pylyshyn, 1973; 1981) with no functional role for imagery envisaged. Both camps have cited empirical investigations, discussed the relevance of introspective accounts, proposed computer simulations and engaged in considerable philosophical discussion but still, to date, no consensus has been reached.

Many propositionalists would regard imagery as epiphenomenal - as the result of a process rather than the process itself. Although it may be tempting to call upon introspective reports to counter such a suggestion, their usefulness as explanations has, quite correctly, been questioned. For instance, Pylyshyn (1973, p2 - 3) notes that "what is available to conscious inspection may not be what plays the important causal role in psychological processes". However, other authors do not wish to ignore the fact that introspective accounts of performance on various tasks frequently include reference to mental imagery (see Paivio, 1971; Richardson, 1980a). Although such accounts by no means establish that imagery plays a functional role, Kosslyn and Pomerantz (1977) claim that, nevertheless, they constitute important corroborative evidence and, as data in their own right, should be encompassed by any comprehensive theory in this area. Other dual coding proponents stress that the phenomenal experience of imaging is by no means essential to its usefulness as an explanatory construct (Bugelski, 1977). After all, there may not be a strong correlation between conscious awareness of imagery and the efficiency of a process involving it in various tasks (see Evans, 1980b, p282).

The crux of the debate seems to rest not on the existence of the phenomenon called imagery, which very few psychologists would deny, but on what mental representations underlie it (Anderson, 1978). One opponent of the imagery position (Pylyshyn, 1973) criticises the fuzziness of the theoretical construct of 'image'. Although, as Paivio (1969) claims, imagery has been operationally defined for the purposes of empirical research in different paradigms, Pylyshyn suggests that imagists may be unjustified in assuming that the various definitions converge on an equivalent theoretical construct. He questions the inappropriate use of the 'Picture-in-the-head' metaphor. There is a persistent use, in describing imagery, of words and phrases which are more appropriate for describing pictures and the process of perception. This merely pushes all of the problems of perception onto an homunculus. Pylyshyn (1973) suggests that the metaphor is inadequate in a number of respects. For instance, images are not re-perceived as pictures are perceived since they are already interpreted to a great extent. When parts are missing from one's recollections these form meaningful units rather than being like the missing corner of a torn photograph. Furthermore he claims that the capacity needed to store just a few 'raw' picture-like images would far exceed that available in the brain. Another problem concerns the retrieval of uninterpreted images from amongst the wealth of images that people commonly claim. He dismisses the possibility of scanning prospective candidates before the 'mind's eye' since the time taken for an exhaustive search would be prohibitive and no awareness of searching is apparent but rather access appears to be direct. The alternative possibility Pylyshyn (1973, p9) suggests is that they are "tagged by some gross labels and associatively retrieved by a multiple-sort key". Pylyshyn dismisses this on the grounds that small, and even abstract, details of an event can be retrieved in fine detail without first being aware of calling up the entire scene. The point is that images behave as though they have been analysed. If this is accepted then, on grounds of economy, the 'raw' image can be dispensed with and the analysis alone can be assumed to be stored.

In countering Pylyshyn's arguments concerning the absence of a precise definition of 'image', Kosslyn and Pomerantz (1977) claim that, in the early stages of theory building, it is common and indeed advantageous to employ converging operations to investigate a construct whose defiition is not precisely formulated. Also Kosslyn and Pomerantz (1977) and Paivio (1976) refer to the 'picture-in-the-head' metaphor as a 'straw man' claiming that such a view is not held seriously as a working theory. They claim that images are more like the outputs of the perceptual system rather than their inputs. As to capacity limitations, the storage capacity of the brain remains unknown. The amount of information in an image has not been effectively defined but certainly the "relatively large, interpreted, perceptual 'chunks'" envisaged by Kosslyn and Pomerantz (1977, p59) would require less capacity than the 'raw' pictures envisaged by Pylyshyn. In any case, the issue concerning capacity limitation could be similarly addressed to alternative propositional models of representation (cf. Anderson, 1978). With regard to accessibility, the speed of searching for a particular image is simply unknown but, according to Kosslyn and Pomerantz, possibly could be facilitated by assuming that verbal or other tags are associated with images.

In reviewing the empirical evidence claimed in support of the Dual Coding approach both Pylyshyn (1973) and Anderson (1978) have been criticised for being "selective and incomplete" (Paivio, 1983, p310). Pylyshyn's (1973) critique has relied in the main on various logical arguments such as those outlined above. Even his later critique (Pylyshyn, 1981) has tended to concentrate on the experiments concerning 'mental scanning' and 'mental rotation' transformations by Kosslyn and his colleagues. In explaining the results of such studies, Pylyshyn has suggested that tacit knowledge of what should occur in analogous real-life situations is employed by subjects to draw inferences appropriate to the experimental task and imagery instructions without the utilisation of mental imagery. However Kosslyn (1981) provides strong counter-arguments and Pylyshyn's view seems inadequate as an account of the 'mental scanning' data derived in a recent study by Reed, Hock and Lockhead (1983). In any case, as will be shown, these kinds of task produce only a small amount of the overall empirical evidence claimed in support of Paivio's (1971; 1975) hypothesis.

In his assessment of the debate, Anderson (1978) has stated that both propositional and dual code models can be made to yield identical behavioural predictions and so the form of internal representation cannot be determined by appeal to behavioural data alone. The point he makes is that for any model postulating a particular internal representation, an alternative model which is behaviourally indistinguishable can be defined which uses another form of representation. This can be done by making compensatory changes in the accessing process. Although this argument is valid, Anderson's conclusion that "barring decisive physiological data, it will not be possible to establish whether an internal representation is pictorial or propositional" (Anderson, 1978, p249) appears unjustified. The reason is simply that no finite amount of data, even physiological, uniquely determines the correctness of a theory (cf. Popper, 1968). Appeals to other criteria such as parsimony and plausability, efficiency and optimality are not given sufficient weighting in Anderson's account. However, two of the major protagonists in the debate (Pylyshyn, 1981; Paivio, 1983) suggest that considerations such as generalisability and integrative value, constrainedness, predictive qualities, etc., are of considerable relevance to the dispute.

As Paivio (1983, p311) writes "dual coding and imagery based theories generally account for a wide range of findings, which cannot be handled by abstract descriptive approaches except by the addition of post hoc assumptions with each new turn in the data". Some propositionalists, such as Anderson (1978) and Kieras (1978), have even incorporated perceptual and linguistic propositions into their theories in order to accommodate exactly the same range of findings as the Dual Coding theory. However, Paivio (1983, p328) writes that "the two approaches would then become indistinguishable because the propositional model would simply be a conceptual variant or paraphrase of dual coding". Nevertheless Anderson's argument in favour of a tricode theory is worthy of consideration. He writes (p274): "it seems clear that the human must process three kinds of information: visual-spatial, verbal-sequential and abstract-propositional......the kinds of information representations optimal for these three domains are different. Therefore, it would seem that there would be a strong survival advantage pushing in the direction of three separate codes with the potential for intertranslation among them". Although seemingly unparsimonious at first, the tricode theory approach is currently gaining favour (eg. Snodgrass, 1984; Glucksberg, 1984). Notwithstanding, the imagery debate is by no means resolved.

In fairness it must be said that the propositionalists have been more reactive in their explanation of empirical data. Indeed most of the empirical data have been generated by imagery researchers and, since a major aim of any explanatory theory is prediction, this is a criticism of propositional theories. Also many of the propositional models which are put forward appear to be limited in application to a narrow range of tasks and need modification to accommodate even slight changes in procedure (eg. Clark and Chase, 1972). Nevertheless, their underlying basis is claimed to have considerable generality. The dual coding approach on the other hand seems to be supported by a wide variety of different findings. In the next few pages of this chapter some of the empirical evidence relating to this debate will be considered.

EMPIRICAL EVIDENCE AND THE IMAGERY DEBATE.

Paivio (1983) has identified about sixty reliable empirical findings that lend support to the Dual Coding theory and appear to detract from the plausibility of propositional accounts. In his own words "the classes of findings include effects of item concreteness or imagery value, pictures as compared to words as stimuli, imagery instructions in various tasks, reaction time functions in such tasks as mental comparisons and figural transformations of various kinds, modality specific interference, perceptual and memory comparisons, effects of individual differences in spatial and verbal abilities, and functional differences in the two cerebral hemispheres" (Paivio, 1983, p311). Whilst the bulk of these experiments will not be considered in detail, a selection of these and other studies will be assessed in the next few pages.

As Paivio (1983) suggests "when a perceptual task selectively disrupts performance on a concurrent mental task or vice versa, it is generally assumed that common processing systems are involved". A classic series of experiments by Brooks (1967, 1968) employ this sort of competing task methodology to demonstrate imagery suppression.

The studies of Brooks demonstrate that the secondary task can be performed during the input or response output stage. An example of input interference is given by Brooks (1967) who shows that reading and hearing sentences describing a spatial array leads to poorer immediate recall than just hearing them. The reverse is the case when nonsensical sentences are used. It appears that visual presentation interferes with the concurrent construction of an internal spatial representation. Output inteference is shown in another experiment in which written recall of a spatial message took longer than spoken recall although no difference was found for non-spatial control messages.

It is worthwhile describing one of Brooks' (1968) follow-up experiments in some detail because of the importance of the selective interference methodology to the investigations that will be reported in the present experimental chapters. In this study he required his subjects to categorise each word in a remembered, aurally presented sentence as either a noun or a non-noun. If the sentence is recalled in an articulatory manner then vocally signalling the response sequence should be more disruptive than using a different, non-articulatory, mode of responding. The subject was given one of three different ways of signalling the sequence:

1) saying 'Yes' for a noun and 'No' for a non-noun. Thus the correct sequence would be "NO", "YES", "NO", "NO", "YES", NO", "NO", "NO", "NO", "YES", for the sample sentence "a bird in the hand is not in the bush",

or 2) tapping with the left hand for each noun and the right hand for each non-noun,

or 3) pointing to a 'Y' for nouns and to an 'N' for non-nouns on an array such as that shown in Figure 2.1 . To produce the correct sequence for the sample sentence given above, the subject would point to the top "N", the second "Y", the third and fourth "N"'s etc, in the figure.

```
        Y             N
              Y  N
         Y            N
        Y              N
              Y    N
         Y             N
        Y                N
              Y       N
       Y                  N
              Y        N
                Y  N
        Y                   N
```

Figure 2.1 . A sheet such as those used for the pointing condition of Brooks' (1968) experiment I. The letters were staggered to force close visual monitoring of pointing. (From Brooks, 1968, figure 1).

In another condition of the same experiment subjects were presented with a diagram such as that illustrated in Figure 2.2 .

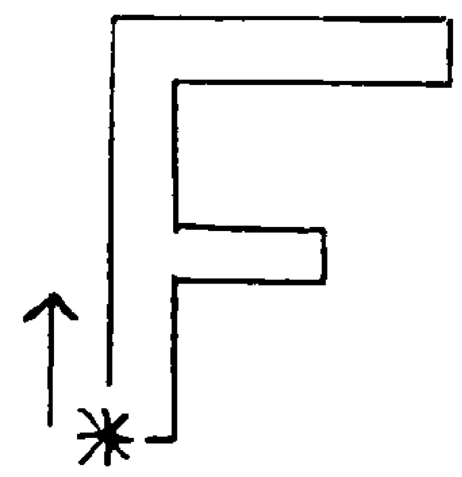


Figure 2.2 . A sample of the simple block diagrams used by Brooks (1968) experiment I. The asterisk indicates the starting point and the arrow the direction of working. (From Brooks, 1968, figure 2).

They were asked to categorise,from memory, each of the corners of block letters, such as that illustrated in the diagram, starting from the asterisk and proceeding in the direction of the arrow, as either extreme top, extreme bottom or as somewhere in between. If the letter is recalled in a visuo-spatial manner then responding visually (that is by pointing in the same manner as before) should be more disruptive than responding vocally. The results of this experiment are shown in Table 2.1 .

	Output		
Referent	Pointing	Tapping	Vocal
Sentences	9.8	7.8	13.8
Diagrams	28.2	14.1	11.3

Table 2.1 . Mean output time (seconds) for the six conditions of Brooks' Experiment I. (Data from Brooks, 1968, table I).

It can be seen from Table 2.1 that vocal output was slowest for categorising sentences whereas visually monitored, pointing output was slowest for categorising block letter diagrams. These differences were significant. These results suggest that images and percepts conflict with each other providing that they occur in the same modality and impose simultaneous demands on specific processing resources. Dual coding specifically appears to gain support from this sort of double selective interference since visuo-spatial and verbal information are apparently processed in separate modality-specific ways.

However, it should be stated that the above interpretation of Brooks' data has been questioned by Phillips and Christie (1977). They argue that, since his designs lack control conditions in which no interference is present, it is inappropriate to infer specificity in both modalities from the results. After all, performance in the control condition might not be intermediate to the visuo-spatial and verbal interference conditions. Indeed their own experiments are interpreted as showing that visualisation is interfered with when a competing task (the adding of a series of visually or aurally-presented digits) demands concurrent use of a general purpose rather than a modality-specific resource. They concede that mental addition (even of auditory digits) could involve mental imagery. Indeed individuals often claim to employ imagery in mental arithmetic tasks (Hayes, 1973). However, Phillips and Christie (1977, p648) argue that this possibility "seems unlikely, and if true would reduce the grounds for calling the visual processor 'special purpose'".

Baddeley and Lieberman (1980) report a series of experiments which clarify Brooks' results and suggest that the disruption involves a spatial processing system rather than a visual one. However this does not seem to damage Paivio's Dual Process Hypothesis "because visual imagery is assumed to include spatial information as an essential component" (Paivio, 1983 , p321). It is very difficult for the propositionalists to accommodate results such as these, since they contend that all information is processed in a unitary amodal system. Furthermore, Baddeley and Lieberman (1980, p537) suggest that the evidence in favour of a spatially-based system "does not preclude the occurrence of a parallel system or component concerned with pictorial or non-spatial visual representation". The evidence of Atwood (1971) is indicative of such a system.

Atwood's (1971) evidence is derived from a paired-associate learning task in which pairs of nouns were incidentally memorised. The nouns for recall were embedded in either highly visualisable, concrete phrases such as 'a nudist devouring a bird' or abstract phrases such as 'the intellect of Einstein was a miracle'. One group of subjects received the former kind of material and were asked to visualise the scene described. The other group received the latter kind of material and were asked to contemplate the meaning of the sentences. An interfering task was interposed in the period between presentation and recall. This task involved the presentation of a digit (either '1' or '2') to which the subject responded with the name of the digit which was not presented. This secondary task involved either the visual or the auditory modality. A control condition was also employed involving no interfering task. It was found that nouns embedded in visualisable phrases were recalled better than other nouns in the control condition. Lower recall occurred with imageable phrases under the visual interference condition relative to the auditory interference condition, whereas the reverse effect occurred with the abstract phrases. This result has been replicated and extended by Janssen (1976) who eliminated various methodological weaknesses of the original study. He found that the interference effect occurred with single nouns as well as paired associates and found that the magnitude of the effect decreased as the rated imageability of the nouns decreased.

Kosslyn (1980) suggests that visual mental images are actively generated from information stored in long-term memory. He hypothesises that they are "like displays on a cathode ray tube that are generated by a computer programme (plus data)" (Kosslyn, 1980, p5-6). However, he distinguishes these quasi-pictorial 'surface images' from their underlying 'deep representations'. The crux of the imagery debate concerns the nature and function of the 'surface images' and whether they possess emergent properties which are not manifested in the underlying representations. Kosslyn and his colleagues (Kosslyn, 1980; Kosslyn, 1981; Kosslyn, Pinker, Smith and Shwartz, 1979) provide evidence that the 'surface images' have perceptual-like functional properties. Typical experiments require subjects to construct visual images and transform or inspect them in a systematic manner.

It has been shown by Kosslyn (1976) that less time is taken to verify large rather than small properties of stated objects (eg. 'cats - head' versus 'cats - claws') when subjects are instructed to use visual imagery. However, in the absence of imagery instructions, the effect is determined by the strength of the verbal association between the object-noun and the property-noun. In the above case the smaller property ('claws'), being the stronger verbal associate, is verified more quickly.

Other experiments reviewed by Kosslyn (1980) show that image scanning time is proportional to the distance between points in a spatial image. Also, larger objects seem to 'overflow' sooner than smaller objects when subjects are requested to imagine the object at a distance and then to imagine approaching the object. These results suggest that visual images depict spatial extent which is limited by boundaries. Furthermore, the acuity of the image is claimed to decrease as one moves from the centre towards the periphery of the visual field, just as in visual percepts. The precision of Kosslyn's account is illustrated by the production of a computer simulation of the model (Kosslyn and Shwartz, 1977; Kosslyn, 1980) which has considerable heuristic value in predicting new data.

The review of Paivio (1969) has emphasised the effects of word imagery-concreteness in learning and memory. The use of concrete rather than abstract materials generally leads to superior performance. Whilst Paivio (1971) suggests that both verbal and imaginal codes are interrelated, he assumes that different kinds of stimulus material have differential access to them. This is illustrated in Figure 2.3.

Stimulus	Coding systems available	
	Imaginal	Verbal
Picture	+++	++
Concrete word	+	+++
Abstract word	-	+++

Figure 2.3 . The availability of imaginal and verbal coding systems as a function of stimulus concreteness. The number of plus signs indicates the degree of availability of the appropriate coding system. (Adapted from Paivio, 1971, figure 7-1).

Paivio suggests that highly concrete items evoke imagery more easily and facilitation in recall arises because imaginal representations serve as a supplementary memory code. However, this idea has been disputed (Anderson and Bower, 1973; Richardson, 1980b) and the alternative post hoc hypothesis they suggest is that abstract words lead to poorer performance because they are more confusible owing to their greater lexical complexity and their greater number of dictionary definitions. The imageability and concreteness of words is highly correlated, however, although differential effects on each dimension have been shown experimentally (see Richardson, 1980c). When imagery instructions are given the distinction between them breaks down and ease of learning is dominated by rated imageability.

In order to investigate Paivio's (1971) hypothesis that the superiority of recall for high over low imagery words is occasioned by their access to a supplementary imaginal code, a series of experiments was performed by Baddeley, Grant, Wight and Thomson (1975). They employed identical basic tasks and materials to those of Brooks (1968) described earlier (ie. Block 'F' and sentences such as 'bird in the hand is not in the bush') to establish that performance on a pursuit rotor tracking task involves a visuo-spatial component. It was found subsequently that memory for aurally presented visualisable material (sentences describing the location of eight digits in a four by four matrix) was impaired more than memory for otherwise equivalent, non-visualisable nonsense sentences (cf. Brooks, 1967) whilst performing on the pursuit rotor task. Thus imagery as an active control process in visual Working Memory can be disrupted by concurrent visual activity. However, recall performance in a paired-associate memory task did not demonstrate an interaction between concreteness of the materials and tracking although the usual main effect of concreteness was attained. Baddeley et al concluded that concreteness effects are not due to an imagery component.

More recently, Mathews (1983) has investigated the same hypothesis by presenting for recall lists with equal numbers of words high on both concreteness and imagibility rating scales together with some filler items. All of the words used were equivalent in familiarity ratings. During presentation of the lists, subjects were engaged in one of two concurrent visual activities. These were equivalent in their perceptual and motor demands but had been shown to differ in the degree to which they require the maintenance of information by an image-like representation. Whilst high imagery words were recalled more efficiently than low imagery words when the secondary task did _not_ involve an imagery component, there was no such advantage when the secondary task required the maintenance of visual images. This latter result is supportive of the Dual Coding account.

Further support for dual coding is given by the results of Klee and Eysenck (1973) who measured the ease of comprehension of abstract and concrete sentences which were read to subjects under visual or verbal interference conditions. The sentences they used were either meaningful or anomalous, for example:

Concrete meaningful: The veteran soldier rode the lame horse.

Concrete anomalous: The large army beat the wild pearl.

Abstract meaningful: The wrong attitude caused a major loss.

Abstract anomalous: The mere knowledge brought the true hour.

The visual interference was provided by visual presentation, between each word of the sentence, of separate five by five matrices with three of the squares blacked out. Verbal interference consisted of a separate digit being spoken in a distinctive voice between each word. After indicating via a key press whether the sentence was meaningful, the subjects had to recall the interfering stimuli. The mean comprehension latencies for concrete and abstract sentences averaged over meaningful and anomalous sentences are shown in Table 2.2 .

The interaction between interference condition and concreteness was significant and shows that comprehension latencies were longer with visual than with verbal interference for concrete sentences but _vice versa_ for abstract sentences.

These results are consistent with the idea that visual imagery is used in comprehending concrete sentences and the processing of visual matrices interferes with image formation.

	Abstract	Concrete
Verbal interference	1.35	Ø.78
Visual interference	1.18	Ø.97

Table 2.2 . Mean comprehension latencies (in seconds) for concrete and abstract sentences under visual and verbal interference conditions. (Data from Klee and Eysenck, 1973, table 1).

In spite of his earlier claim, noted above, Richardson (198Øc, p87) has changed his position somewhat and he now claims that "while concreteness is a feature of lexical organisation and not a measure of image-arousing quality of verbal material, imageability is the effective stimulus attribute determining how easily it can be remembered". Furthermore, he writes (p96) that "contemporary accounts of mental imagery, which identify stimulus imageability as a primary determiner of recall performance, are likely to be essentially correct".

This sort of interpretation remains equivocal as studies with the congenitally blind illustrate. Zimler and Keenan (1983) compared the performance of congenitally blind and sighted individuals on a paired-associate learning task. The stimulus and response referents were either both high in visual (V-V) or high in auditory (A-A) imagery, or they were mixed with one term coming from each category (A-V or V-A). Paivio (1971) has demonstrated that the imageability of the stimulus term is more critical than that of the response term in this task. If modality-specific imagery is used then blind subjects should perform worse when the stimulus term has a visual referent and the response an auditory one (V-A), rather than the other way round. Their results, which are shown in Table 2.3, were contrary to Paivio's hypothesis in that blind subjects recalled items from V-V pairs better than other items overall, although their performance in this category was worse than that of sighted subjects. There was no difference between V-A and A-V pairs and A-A pairs for blind subjects even though, according to the hypothesis, better performance in the latter condition would be expected because imagery should be invoked to both stimulus and response terms. The only other significant difference was the poorer performance of sighted subjects overall with A-A pairs. Whilst this result fails to replicate that of Paivio and Okovita (1971), it is essentially similar to other studies in the area.

	Pair type			
Subjects	A-A	A-V	V-A	V-V
Blind	.40	.40	.41	.47
Sighted	.27	.42	.41	.56

Table 2.3 . Mean proportion of pair types recalled in a paired-associate learning task for blind and sighted subjects. (Data from Zimler and Keenan, 1983, table 1).

Zimler and Keenan's second experiment, using a free recall task, showed that blind subjects performed as well as sighted ones on words grouped by colour and better on words grouped by sound. The fact that congenitally blind subjects, who cannot employ visual imagery as a supplementary memory aid, recall colour words as well as sighted subjects appears to run counter to Paivio's hypothesis and the claim of Richardson (1980c) noted above.

Consideration of Figure 2.3 illustrates the rationale for Paivio's hypothesis that pictures should be mnemonically superior to words. This is in fact the case as the extensive reviews of Paivio (1971) and Madigan (1983) illustrate. The material which they cite shows substantial improvement is occasioned in free recall by giving pictorial rather than verbal presentations of task materials. The effect is particularly marked in recognition memory when the amount of material presented is large (Standing, 1973). The superiority of pictures over words is durable and can extend up to several months. Standing also showed that with equally complex and detailed pictures, more unusual or vivid versions produced better recognition. Although increases in colour, detail and complexity seem to have no effects in recognition, they do lead to superior recall of associated verbal labels (Madigan, 1983; Madigan and Lawrence, 1980). Nisbett and Ross (1980) have suggested that the facilitative effect of 'vivid' information on judgement and inference may be the result of its greater availability in memory partially due to its imageability.

Various explanations of the above picture-word differences have been proposed. For instance, the 'levels of processing' approach of Craik and Lockhart (1972) has been invoked by Anderson (1978). However, this theory has been criticised for its circularity, amongst other things. Baddeley's (1978) critique of this theory will be discussed in the next chapter.

Another possibility derives from the study of Nelson, Reid and Walling (1976) who investigated the relative effects of visual similarity and conceptual similarity of pictorial and verbal items in a paired-associate learning task using unrelated words as responses. Items high in conceptual similarity came from the same taxonomic category (eg. 'tools' or 'animals') and items were selected so that they could be drawn to appear visually similar or dissimilar. If pictures are easier to remember than words because of the superiority of visual coding then high visual similarity should eliminate the effect. This was what occurred, visually similar pictures were no better than words at slow presentation rates, and were inferior at fast presentation rates. Nelson (1979) concludes that the picture effect arises because of the visual features of pictures, particularly their discriminability, which leads them to be represented in an inherantly superior visual code.

Further support for Paivio's account is given by the identification of orthogonal factors relating verbal abilities and imaginal and spatial abilities as measured in objective tests (Di Vesta, Ingersoll and Sunshine, 1971). There is evidence to show that objective measures can be successful in predicting performance on various tasks where imagery ability is postulated (see Ernest, 1977). This is not always the case however (eg. Richardson, 1978). Some self-report measures of imagery have even been claimed as good predictors of performance (White, Sheehan and Ashton, 1977), but many of these tests are contaminated by influences of 'social desirabilty'.

Whilst most of the evidence mentioned above and a high proportion of that cited in the literature is compatible with the Dual Coding Hypothesis, the weight of negative findings is also growing. Paivio (1983) is able to dismiss many of these as due to misinterpretations of his theory. Other findings, which do not derive significant differences according to predictions of his theory and thus seem to offer support for the propositional approach, could be dismissed possibly as being due to failure to reject the null hypothesis. He admits that modifications to the theory are required to encompass both positive and negative findings within a single conceptual framework. Unitary propositional models do not appear adequate, nor do conceptual variants which distinguish perceptual and linguistic propositions. He claims that "the real challenge to dual-coding theory is the more specific one of explaining the discrepent findings in terms that are consistent with the general assumptions of the model, including the associationistic principles on which it is essentially founded" (Paivio, 1983, p328). All things considered, the argument of Anderson (1978), stated above, for a tricode approach could be gaining ground.

In the next section of this chapter two problem solving tasks, both of which can be related to the imagery debate, will be considered.

Problem Solving Tasks and the Imagery Debate.

At this stage it is appropriate to consider Anderson's (1978) argument concerning the role of propositional representations in inference making. He considers that propositions are especially suitable because their abstract truth-bearing character means that they "only represent what is necessary to judge the validity (or plausibility) of an inference" (Anderson, 1978, p257). Pictorial representations cannot easily represent such properties as negation, except in binary circumstances. Also, as Cohen (1983) points out, the representation of categories, rather than specific examples, seems intuitively difficult for pictorial images. How can image of a particular triangle serve for thinking about the general properties of triangularity? Negation and general properties of categories can be represented easily with propositions however. Whilst these problems are particularly relevant for some kinds of logical inference, spatial and relational inferences can be extracted quite readily from a spatial represention.

The role of visual imagery in conditional reasoning has not been explored. However, there are certain classes of reasoning problem in which the role of visual imagery has been hotly contested. Two of these classes will be considered here. The first class is known variously as the linear syllogism or the three-term series problem and its solution depends upon the making of valid transitive, otherwise known as relational, inferences. The second class involves the verification of sentences against pictorial representations.

a). Transitive Inference Tasks.

Three-term series problems entail the presentation of two premises containing either the same comparative term or a comparative term and its converse, together with a question. An example of one such problem is given below:

1) Harry is taller than John

2) Harry is shorter than George

Who is shortest?

The premises in tne problem above can be represented schematically in the following manner:

1) B > C

2) B < A

where 'A' represents the most positively placed item, 'C' the most negatively placed and 'B' the middle term. Also '>' represents the comparative term when expressed positively and '<' when expressed negatively. For convenience, this notation will be used whenever appropriate.

As Johnson-Laird (1972) puts it, "the fundamental problem in making a relational inference is to set up some internal representation of the premises, be it abstract or concrete, that will allow the relation between those items, not specifically linked in a premise, to be determined". A controversy has existed for several years about whether such problems are solved in a visual way involving spatial imagery, or a purely verbal way involving an abstract propositional analysis.

Each of these alternative viewpoints will be addressed with reference to some early studies initially. Since the greatest contrast between the two approaches can be seen with relatively 'pure' imagery and propositional theories, extreme examples of both types will be considered first of all. Eventually a theory which synthesizes aspects of both approaches will be outlined.

Visual Imagery Theories.

Two theories which impute spatial strategies in the solution of three-term series problems will be discussed in this section. The seminal imagery theory belongs to DeSoto, London and Handel (1965). Their basic proposal is that a unitary representation is constructed initially from the two premises. This consists of a linearly-ordered, vertically or horizontally orientated visual image. In addition, they suggest that arrays are constructed in accordance with certain fundamental principles which they apply in predicting the relative difficulty of various problem types. These principles will now be described.

DeSoto et al have shown empirically that comparative dimensions can be assigned to vertical or horizontal arrays even when the relational term does not apparently have any ties with spatial phenomena. Certain of these assignments are fairly consistent across individuals. For instance 'good' is generally assigned to the top and 'bad' to the bottom of a vertical array. On the other hand, terms from the 'light - dark' dimension are not clearly oriented in such a manner, but show much greater individual differences. The first principle is derived from a natural directional preference to construct vertical arrays from the top downwards and horizontal arrays from left to right. Any general effect of directional preference on problem complexity will only be expected with problems on which a clear spatial preference exists. The second principle asserts that representations are easiest to construct from premises beginning with an 'end-anchor' (ie. 'A' or 'C') rather than with the middle term ('B'). How well these principles can be applied in an experimental situation to predict problem difficulty will be assessed in due course.

Whilst the theory of Huttenlocher (1968) also assumes the use of an imagined array, it differs from DeSoto et al's in certain respects. Huttenlocher appears to rely much more heavily upon adult's subjective reports, than DeSoto et al, in deriving her theory. She also draws a close analogy between production of imaginal representations in three-term series tasks and the physical construction of actual arrays in another task. In the concrete task, arrays are built by placing a third coloured block onto a ladder so that a correspondence with sentences such as "the red block is on top of the green block" is achieved. She has found in her investigations of children's performance in the physical task (Huttenlocher and Strauss, 1968) that the grammatical status (ie. Subject, Object) of the block to be moved has a considerable influence on difficulty, with grammatical subjects (in deep structure) being easiest to place. Accordingly, the principle of end-anchoring is reformulated in such terms.

Whilst the nett result of either formulation is equivalent as far as interpretation of the second premise is concerned, Huttenlocher's account does not predict a general end-anchoring effect for the first premise. It predicts that the item placed first will be the 'better' item where the relational term has an obvious spatial counterpart, otherwise it will simply be the first item mentioned in the premise.

Clark's (1969) Linguistic Theory.

Clark (1969) has criticised previous theories put forward to explain performance on three-term series problems for their lack of generality. He suggests a set of abstract linguistic processes which, he believes, can explain performance on a wide range of tasks. This linguistic theory is founded on three important psycholinguistic principles of sentence comprehension and owes much to the work of Chomsky (1965). These , he claims, can be used as a basis for predicting the relative times it takes to solve two- and three-term series problems. Since many theories based on this kind of propositional analysis have been put forward in a variety of other situations, the application of Clark's three principles to transitive inference tasks will be described below.

The first two of Clark's principles characterise the comprehension of premises, whilst the third has more to do with the nature of the question posed. The first principle is that of the primacy of functional relations. Immediately after a sentence is comprehended certain important relations specifying such things as 'Subject-of', 'Predicate-of', 'Direct-object-of' and 'Main-verb-of' (cf. Chomsky, 1965) are stored "in a more readily available form than any other kinds of information, like that of theme" (Clark, 1969, p388).

The second principle is that of lexical marking. It states that certain positive adjectives (eg. 'good', 'long', 'interesting') are stored in a more readily accessible form than their opposites. Clark cites evidence, suggesting that marked adjectives can be neutralised in certain contexts whereas unmarked ones cannot, in support of this principle. To illustrate this, consider the question 'How good is the food?' in which the an unbiased interrogator merely requests an evaluation of food, with the question 'How bad is the food?' which implies that the interrogator is already biased towards the opinion that the food is bad but is asking for the extent of its badness. Additional support for this principle is drawn from the fact that the marked member serves as the name of the full scale (the name of the 'good-bad' scale is 'goodness'), whereas the unmarked member ('bad') names only half the scale.

The third principle is that of congruence. In answering a question, the listener requires more than an understanding of the specific question as phrased. He will need to understand that other phrasings are congruent with it and, in searching memory, be guided by this in order to find the desired information and formulate an answer. Such congruency is at the level of functional relations. The listener "cannot answer the question until he finds congruent information, or until he reformulates the question so that he is able to do so" (Clark, 1969, p390).

Consider the comparative sentence below:

John is better than Dick.

According to Clark, the propositional representation of such sentences involves two base strings, eg.:

John is good. Dick is good.

containing the functional (Subject-predicate) relations, which are bound together by a comparative term (eg. 'more than'). In the example this reads:

John is more good than Dick is good.

It is claimed that the functional relations are more readily available than the comparative term. However, consideration of the way in which children encode comparative premises (cf. Donaldson, 1963) leads to a slight amendment here in that the base propositions may incorporate additional information. In the above example this would signify that 'John' is the better one of the pair, and the propositions are abbreviated to:

John is good+

Dick is good-

The principle of lexical marking suggests that, whenever possible, sentences are interpreted in their unmarked sense, since the marked sense takes longer to store and retrieve from memory. Therefore it is predicted that 'better' premises are processed faster than 'worse' premises.

The principle of congruence predicts that questions phrased in a form congruent with the representation of the premise(s), eg.:

Who is best?

will be processed more quickly than those which are not, eg.:

Who is worst?

Clark (1969) gained support for these hypotheses in an experiment in which the materials were eight two-term series problems of the general form:

A is better than B

Which is worst?

The surface structures of premise and questions were varied orthogonally with the deep structural analyses which, he proposes, they entail by incorporating negative equatives ('A isn't as bad as B') into the design. Clark suggests that his use of the 'negative equative' construction distinguishes his theory from the previous imagery theories. This is because the propositional representations of negative equative forms are radically different from comparative forms.

The application of Clark's linguistic theory to the solution of three-term series problems is obviously complex. When encoding the second premise, for instance, individuals may be required to employ a time consuming strategy to assign the appropriate ordinal relationships to the propositions stored. Clark does not give details of how individuals proceed from this stage. In performing an information processing analysis of the task, Johnson-Laird (1972) notes that one of the terms will have been encoded twice. He suggests that the fact that it must be the middle term will be recorded. Thus the extremity of the other two terms is established. Clark predicts that, when premise pairs have homogeneous relational terms, those in which the deep structural analyses are in terms of marked comparatives will be easier than those in terms of unmarked comparatives. Also congruent questions should be easier than incongruent ones. However, problems with heterogeneous relational terms involve a further complexity in that the questions will be congruent with certain base strings and incongruent with others. The theory suggests that problems are easier when the answer is embedded in a base string which is congruent with the question.

At this stage it is worth examining some experimental data from each of the above studies to see how they compare in their predictions of problem difficulty. Initially the case involving comparative problems using assymmetrical relational terms will be considered. Unfortunately DeSoto et al only reported the frequency of correct solutions. Both imagery theorists reported their data averaged over different question types. Consequently the evidence for congruency effects can only be seen in Clark's data. Also Huttenlocher's latencies were measured from the presentation of the second premise and thus minimised the contribution of the first premise to problem difficulty. The relevant data are shown in Table 2.4 .

Consideration of the predictions made by Clark's theory concerning problem difficulty will show that they are the same as made by the original imagery theory of DeSoto et al. If the unmarked adjective is that placed at the top of the imagined array, then lexical marking and 'direction of working' suggest the same order of difficulty for problems with homogeneous premises. In fact Clark's mean latencies and DeSoto et al's frequencies both show problems 1 and 1' (in Table 2.4) to be easier than problems 4 and 4'. With heterogeneous premises, predictions based on congruence and on 'end-anchoring' both lead to the prediction that problems 2 and 2' should be easier than problems 3 and 3'. This is in fact the case once again.

According to the linguistic theory, the order of premise pairs should not make any difference to difficulty so that 1 and 1', 2 and 2', etc., should be equivalent although this is not in fact the case. Clark attempts to account for this result in terms of 'compression of information' so that it is easier to handle in memory. DeSoto et al predict that the former problems would be easier in each case because the first term mentioned corresponds to the top of the imagined array and a top-down strategy is the preferred 'direction of working' with these kinds of problem.

		DeSoto et al (1965)	Huttenlocher (1968)	Clark (1969)		
Problem	Comparative:	Better-worse	Taller-shorter	Better-worse		
				Best?	Worst?	Mean
1	A>B B>C	60.5	155	542	610	575
1'	B>C A<B	52.8	135	498	552	525
2	A>B C<B	61.8	141	535	534	534
2'	C<B A>B	57.0	142	484	584	532
3	B<A B>C	41.5	157	500	602	549
3'	B>C B<A	38.3	157	612	545	577
4	B<A C>B	50.0	142	593	504	547
4'	C<B B<A	42.5	161	627	653	640
		% correct	Latencies in centiseconds			

Table 2.4 . Relative difficulty of three-term series problems from three studies. (Data from Evans, 1982, table 4.3).

Huttenlocher's theory predicts that the construction of arrays is facilitated, and hence problems are easier, when an end (eg. top or bottom) item appears as the grammatical subject of the premise. Thus whilst no differences would be expected for 2 versus 2' or 3 versus 3', 4 and 1' should be easier than 4' and 1 respectively. It should be recalled that Huttenlocher's latencies were measured from the presentation of the second premise. In the latter two comparisons, the grammatical subject of the second premise is an end item in both of the easier problem types. Although Huttenlocher's data fits well with these predictions, Clark's does not.

In his account, Clark suggests that the effects of congruency and the use of 'negative equative' premises are critical in differentiating his theory from imagery theories. However, this view has been questioned by Huttenlocher, Higgins, Milligan and Kauffman (1970) and they have provided further evidence in support of their spatial imagery interpretation. Although it was originally felt that negative equatives could be used to differentiate between the linguistic and spatial theories, it is not clear what predictions should be made from Huttenlocher's theory about problems incorporating them. Both Johnson-Laird (1972) and Evans (1982) claim that the dispute was never really resolved.

In summary, there are difficulties for both of the approaches outlined here. On the one hand the imagery theory cannot explain the effects of congruence whereas, on the other hand, the linguistic theory cannot explain the effects of premise order. Nevertheless, both kinds of theory can accommodate many important aspects of the data. Perhaps Johnson-Laird (1972) is correct in suggesting a compromise position. In solving a series of problems subjects may employ an imagery strategy at first, but "intellectual development" within the experiment might lead them to adopt a verbal one as they become more adept. It has also been suggested that "the main function of the imaginal aspect of performance may be as an aid to memory" (Wason and Johnson-Laird, 1972, p120).

Other experimenters have concluded that the type of problem material is the crucial factor. Converging operations were employed by Shaver, Pierson and Lang (1975) in their investigations. They manipulated the characteristics of the problem materials in the same way as DeSoto et al (1965) whose data suggests that the comparative dimensions above-below, better-worse and lighter-darker decrease in the degree to which they suggest an imaginal representation. Following Brooks' (1967,1968) experiments which imply that reading interferes with visuo-spatial imagery, they argue that the interference occasioned by visual versus aural presentation would be greatest for problems easiest to image and least for problems most difficult to image. Whilst they obtained main effects of Presentation Condition and Dimension in the predicted directions, the crucial interaction between these factors fell just short of significance. They also analysed the subjects' self reports concerning their strategies into Visuo-spatial, Verbal or Mixed kinds and reported a significant correlation of this with the subjects' assessments of which Presentation Condition seemed more difficult for each Dimension. Unfortunately, as Evans (1982) stresses, this correlation should include reasoning performance rather than two introspective measures if it is to add useful information to the debate.

A remarkable example of over-reliance on introspective reporting in this context is given by Quinton and Fellows (1975). Their subjects' self-reports concerning a whole series of problems could be classified into two main types - 'thinking' and 'perceptual'. When subjects were asked to implement their most effective strategy in a subsequent series, those claiming 'perceptual' strategies were most efficient. Furthermore two new subjects trained in the use of 'perceptual' strategies were faster than any of the previous subjects who had, supposedly, discovered this technique for themselves. Quinton and Fellows have summised that all psychologists need to do in order to resolve various debates concerning the processes underlying various cognitive tasks is to ask the subjects themselves. This naive approach is severely criticised by Evans (1976) not only with reference to the unreliability of introspective reports concerning processes, but also because Quinton and Fellows assumed that a causal connection could be inferred from purely correlational data. As Wason and Evans (1975; Evans and Wason, 1976) have noted, verbal reports may simply be rationalisations after the event.

A recently reported study by Newstead, Manktelow and Evans (1982) also used an interference technique based on Brooks (1967) in investigating the role of imagery in linear orderings. In essence, this study had its antecedents in the work of Potts (1972, 1974). Potts (1974) presented subjects with a passage in which details were embedded concerning the relative ordering of adjacent pairs of terms from a four-term linear series (A>B, B>C, C>D). They were subsequently required to assess the truth value of all possible pairwise comparisons of terms that could be derived validly or otherwise from the passage. Subjects were found to be more accurate and faster in responding to remote inferences (eg. A>D) rather than to those which had been actually presented in the passage. This is the 'symbolic distance' effect and it has been suggested (eg Lawson, 1977) that it is best accounted for in terms of an imagined integrated array rather than a propositional representation of the relationships presented in the passage. Newstead et al presented passages for the linear ordering task either visually or aurally to different groups of subjects. A second comparable task concerning set-inclusion materials (eg. All A's are B's, All B's are C's, All C's are D's) was also presented. Imagery is not thought to play a part in the latter task. If imagery was involved then a Presentation Modality by Task interaction should be found (cf Brooks, 1967). Very slight evidence in favour of an imagery position was derived in that an interaction between the above two factors and a third factor concerning the rated imageability (high vs low) of the materials was significant. The authors argue that their test was possibly biased against an imagery explanation since a subsidiary mental arithmetic task which could disrupt certain kinds of imaginal representation was interposed between the presentation of the passages and presentation of inferences for evaluation. Unfortunately, their second experiment changed the modality-specific interference technique to the more complex one of visual or auditory shadowing whilst also removing the interpolated mental arithmetic task for some groups of subjects. On this occasion no evidence supportive of the use of mental imagery was derived. Although they use their results to argue against the imagery position, the authors also concede that their interpretation is based on the null hypothesis which, of course, cannot be proven.

An alternative model suggested by Sternberg (1980) incorporates a mixture of linguistic and spatial components and is worthy of consideration. Sternberg suggests that the premises are first encoded propositionally as in the linguistic model. Next, the premises are encoded into separate spatial arrays. These arrays finally become integrated into a single representation when the pivot term (ie. that mentioned in both premises) is identified. The construction of the integrated array begins with the terms of the first premise and ends with those of the second premise. Next the subject reads the question. If it contains a marked adjective then an increase in response time occurs, owing to linguistic complexity and also to the difficulty of seeking a response at the less-favoured end (usually the bottom) of the array. The response may be available quickly if the subject's active 'mental location' (as determined by the position in the array of terms in the second premise) coincides with the answer. Otherwise the subject will mentally search the array for a response. Although it is not intended to discuss this model in any greater depth, it can be seen to owe much to the authors of the previous theories. Its recent origin serves to indicate that the debate between imaginal and propositional representation on three-term series tasks is still very much alive. Also it is of interest to point out that it does not rest on the assumption of a unitary mental representation.

Now, evidence concerning the existence of individual differences in task strategy will be considered. Shaver, Pierson and Lang (1975) found a significant correlation between scores on three tests of visuo-spatial ability and reasoning performance. Evans (1982, p62) questions the relevance of this result in the following manner: "since both measures can be assumed to load on general IQ there is no reason to assume that this arises from an imagery component". A more recent study of Sternberg and Weil (1980) is of relevance here.

Sternberg and Weil believe that the efficiency of the particular strategy employed in linear reasoning tasks depends upon the pattern of a subject's verbal and spatial abilities. Subjects were split into three experimental groups and two of these were trained in a particular method and given practice in its application. Visualisation training involved instruction about how to construct and use spatial arrays. Algorithmic training involved instruction in the elimination strategy identified by Quinton and Fellows (1975). In this case subjects were told to read the final question first. Then, "if the answer to the first statement is not contained in the second statement, the answer to the first statement is the correct response to the entire problem....If the answer to the first statement is contained in the second statement, then the answer choice in the second statement is the correct response to the entire problem".

Groups were asked to use the particular strategy in which they had been trained to solve a set of three-term series problems. The Control Group was asked to devise their own strategy. Whereas Algorithmic training led to better performance overall, Visualisation training had no effect on performance relative to no training at all. An intercorrelational analysis of the mean response latencies for each problem type between Groups supports the view that Visualisation and Control Groups were using a similar strategy in solving the problems and that untrained subjects "rely on visualisation at some point during the solution process" (Sternberg and Weil, 1980). These results were claimed in support of Sternberg's (1980) mixed Linguistic-Spatial model of linear syllogistic reasoning.

In this experiment Sternberg and Weil agree that it is possible that, in spite of instructions, subjects within each experimental group were not using homogeneous strategies. Consequently their data were assigned to one of four new groups on the basis of which provided the best mathematical fit to their individual data. The groups were Mixed (82 subjects), Linguistic (15 subjects), Spatial (15 subjects) and Algorithmic (32 subjects). Solution latencies within these groups were correlated with two orthogonal factor scores derived from the subjects' performance on two verbal and two visuo-spatial tests of ability and an interaction was demonstrated. Whereas the Linguistic Group correlated highly with verbal ability but not with spatial ability and the opposite was the case for the Spatial Group, the Algorithmic Group correlated weakly (although significantly) with verbal ability but only marginally with visuo-spatial ability. The Mixed-model Group correlated significantly with both verbal and visuo-spatial factors. Sternberg and Weil (1980, p234) claim that, for the Visuo-spatial and Linguistic models, their data "provide the first external validating evidence that subjects actually use the representations that proponents of the models claim subjects use when solving series problems by those models".

b). <u>Sentence-Picture Verification Tasks.</u>

The sentence-picture verification task typically involves the presentation of a simple sentence describing the location of two specified items, for example:

Star is above plus.

A picture is also presented in which the specified items are spatially depicted in a manner which either conforms (ie. $\overset{*}{+}$) or conflicts (ie. $\overset{+}{*}$) with the verbal description. Of course, the sentence and picture can be presented in either order and can be presented sequentially or simultaneously. The subjects' task is to decide whether or not the sentence and picture match. The reaction time data obtained with this paradigm have been explained in terms of a pure propositional model (eg. Clark and Chase, 1972; Carpenter and Just, 1975) and also in terms of a model involving an interaction between propositional and imaginal representations (eg. Beech, 1980). These opposing viewpoints will be assessed by considering two specific models in this section.

Clark and Chase (1972) devised four types of sentence-picture combination by varying the truth value of the sentence and the polarity of its verb. Examples of these combinations are shown below:

True affirmative (TA)	Star is above plus.	$\overset{*}{+}$
False affirmative (FA)	Plus is above star.	$\overset{*}{+}$

False negative (FN) Star is not above plus. $\overset{\ast}{+}$

True negative (TN) Plus is not above star. $\overset{\ast}{+}$

They constructed an equivalent set of sentences using the relationship 'below' and varied the spatial location of the items in the pictorial design (ie. $\overset{\ast}{+}$, $\overset{+}{\ast}$) to make sixteen distinct problem displays. In their experiments, Clark and Chase varied the relative locations of sentences and pictures by presenting one to the left and the other to the right of a tachistoscope screen, or vice versa. They directed their subjects, via instructions, to attend to either the sentence or the picture initially and measured verification latencies from initial presentation of materials to the subjects' responses which were made by pushing appropriate 'True' or 'False' buttons.

Two models were derived which were used to predict the relative difficulty of the four basic problem problem types as measured by response latencies. The models differed in their applicability to sentence-first (model A) or picture-first (model B) conditions. In devising these models Clark and Chase made the important _a priori_ assumption that "for a sentence and a picture to be compared they must be represented, ultimately, in the same mental format" (Clark and Chase, 1972, p473) which, it is claimed is propositional in nature. Both of the models incorporate four sequentially ordered information processing stages, the order of which depends on the particular model concerned. The total response latency is derived from the simple addition of times spent at each stage. The four stages are:

1) Formation of the mental representation of the sentence,

2) Formation of the mental representation of the picture,

3) Comparison of the two representations,

4) Response production.

However, the third stage is split into additional substages in both models according to discrete comparison processes which will be defined.

The sentence-first model has the ordering of stages as outlined above with the coding of the picture contingent upon the coding of the sentence. Clark and Chase argue that it is easier to encode the term 'above' than 'below' and the use of the latter term will increase processing time in stage 1 by a constant amount ('a' milliseconds). They also assert that the encoding of negatives will increase processing in this stage by a constant amount ('b' milliseconds). These claims are based on linguistic considerations and the data of Clark (1969) which was considered earlier (see also Chase and Clark, 1971). The encoding of pictures is assumed to be in a propositional form based on the relational term used to encode the sentence. However, since no linguistic features are present in the pictures it is concluded that all pictures in this task take the same time to encode.

In comparing the representations of the sentence and picture it is suggested that a truth index is used. During the comparison stage this index can flip between 'true' and 'false' settings. The identity of i) inner propositions (embedded strings) and ii) outer propositions (embedding strings) of the sentence and picture are assessed. Consider the most complex sentence-picture combination which is given below. Alongside the surface form of the sentence and picture are shown their respective propositional representations:

	Surface form	Propositional representation
Sentence:	Star is not below plus	(False(Star below plus))sen
Picture:	* +	(Plus below star)pic

It is supposed that the initial setting of the truth index is 'true', under the supposition that the sentence is true unless evidence is found to the contrary. Since the subject terms ('star' and 'plus') embedded in the two inner propositions mismatch, the truth index is flipped from its initial setting to a value of 'false'. This change is assumed to take time ('c' milliseconds). Next the outer propositions are compared. The representation for the picture contains an implicit 'true' in the outer proposition although it is not shown. Once again a mismatch is apparent and so the truth index is flipped back again to 'true' taking time ('d' milliseconds).

For the final stage the outcome of the truth index is translated into a response - in the above case a push of the 'true' button. The total time taken to respond is the basic time to perform the operations common to all sentence-picture combinations ('tØ' milliseconds) plus the extra time occasioned with sentences employing negatives and the 'below' term, together with any extra time taken in in the comparison stage when the truth index requires changing. In the above example the total reaction time is 'tØ + a + b + c + d' milliseconds, whereas with the simplest sentence-picture combination ('star is above plus', * over +) this would be just 'tØ' milliseconds.

Clark and Chase (1972) predicted that verification latencies on the sentence-first task would be accounted for by four parameters. These are 1) Below time ('a'), 2) Negation time ('b' + 'd'), 3) Falsification time ('c') and 4) Base time ('tØ'). The observed and predicted data from Clark and Chase's (1972) first experiment are shown in Table 2.5 . The sub-components of the latency value for each problem type are illustrated.

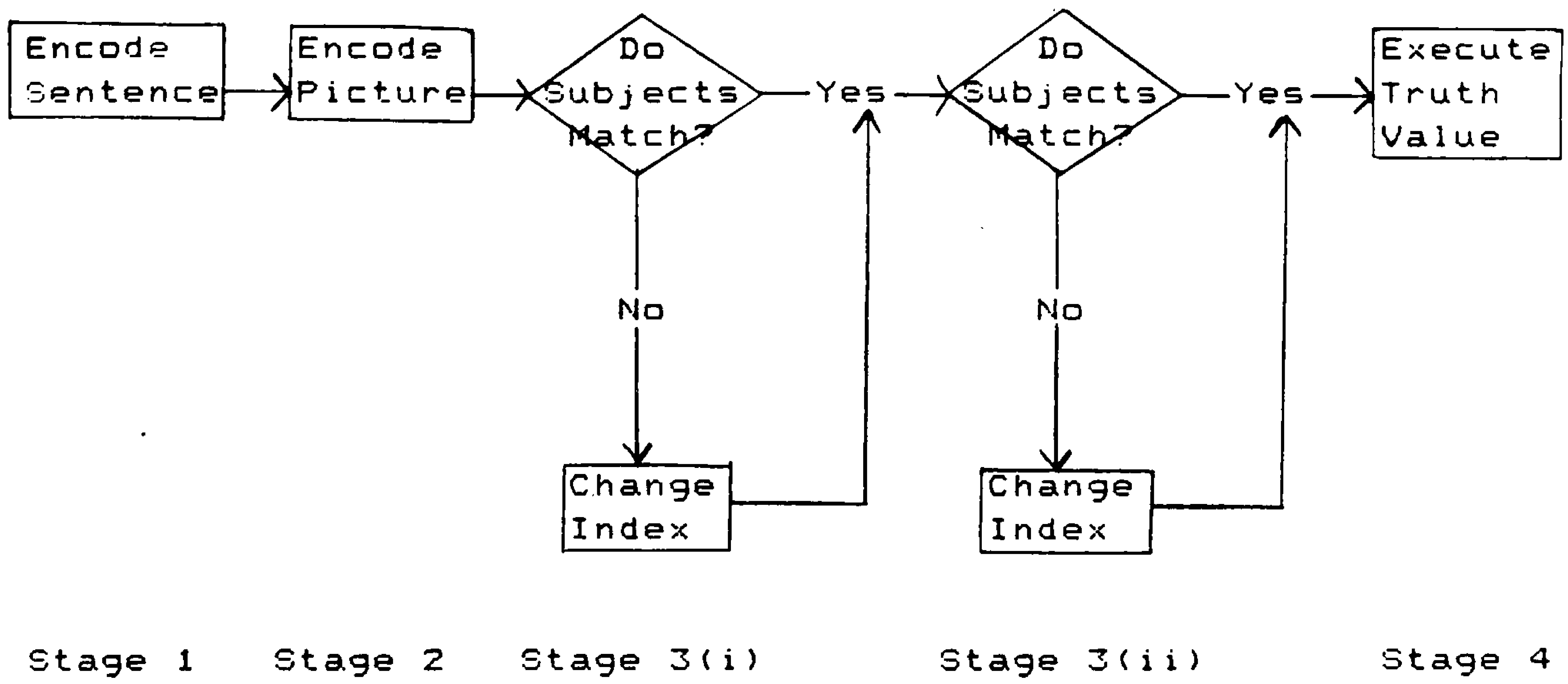
Problem	Sentence	Picture	Components	O.	P.
TA above	Star is above plus	* +	tØ	1744	1763
TA below	Plus is below star	* +	tØ+a	1875	1856
FA above	Plus is above star	* +	tØ +c	1959	195Ø
FA below	Star is below plus	* +	tØ+a +c	2Ø35	2Ø43
TN above	Plus is not above star	* +	tØ +b+c+d	2624	2635
TN below	Star is not helow plus	* +	tØ+a+b+c+d	2739	2728
FN above	Star is not above plus	* +	tØ +b +d	247Ø	2448
FN below	Plus is not below star	* +	tØ+a+b +d	252Ø	2541

Table 2.5 . Breakdown of latencies (in milliseconds) observed (O.) and predicted (P.) by the sentence-first model (A) for eight types of problem. (Data from Clark and Chase, 1972, table 2).

Clark and Chase argue that, in general, pure visual imagery models run into logical difficulties because it is not possible to construct a single image to correspond with negative sentences. However, they admit that in the case of their own experiments a single visuo-spatial representation could be predicted because of the binary nature of the particular materials employed. They claim that a pure imagery model would still not be able to explain why positive and negative sentences behave differently with respect to truth and falsity. For instance, in a pure imagery model 'star is above plus' and 'plus is not above star' could both be represented as $\overset{*}{+}$. However, the truth of the sentence can not be established until the image constructed from the sentence is compared with that of the picture, "therefore, the original form of the sentence - whether it was positive or negative - cannot enter into the comparison process: if true is faster than false for positive sentences, then true must also be faster than false for negative sentences" (Clark and Chase, 1972, p499). However, there was a highly significant Truth Value by Sentence Polarity interaction which disconfirmed the pure imagery model. Beech (1980) has also argued that a pure imagery model is not compatible with the data of Clark and Chase (1972, experiment 1). However he is not willing to dismiss <u>all</u> imagery models as they do. He proposes a model combining visual imagery and propositional representations which predicts exactly the same data as that of Clark and Chase. The sentence-first model of Clark and Chase (1972) and Beech's (1980) model combining imaginal and propositional coding are shown in Figure 2.4 .

The application of Beech's model will be considered using, once again, the most complex problem ('Star is not below plus', $\overset{*}{+}$). The sentence is encoded as an image ($\overset{+}{*}$) which is associated with a negative proposition. The time taken to encode this includes an additional amount of time ('r' milliseconds) because of the use of the marked term ('below'). Extra encoding time ('s' milliseconds) is also necessitated because the sentence includes a negative. In stage 3, when the

a) Clark and Chase's (1972) pure propositional model A.



b) Beech's (1980) mixed propositional/imagery model.

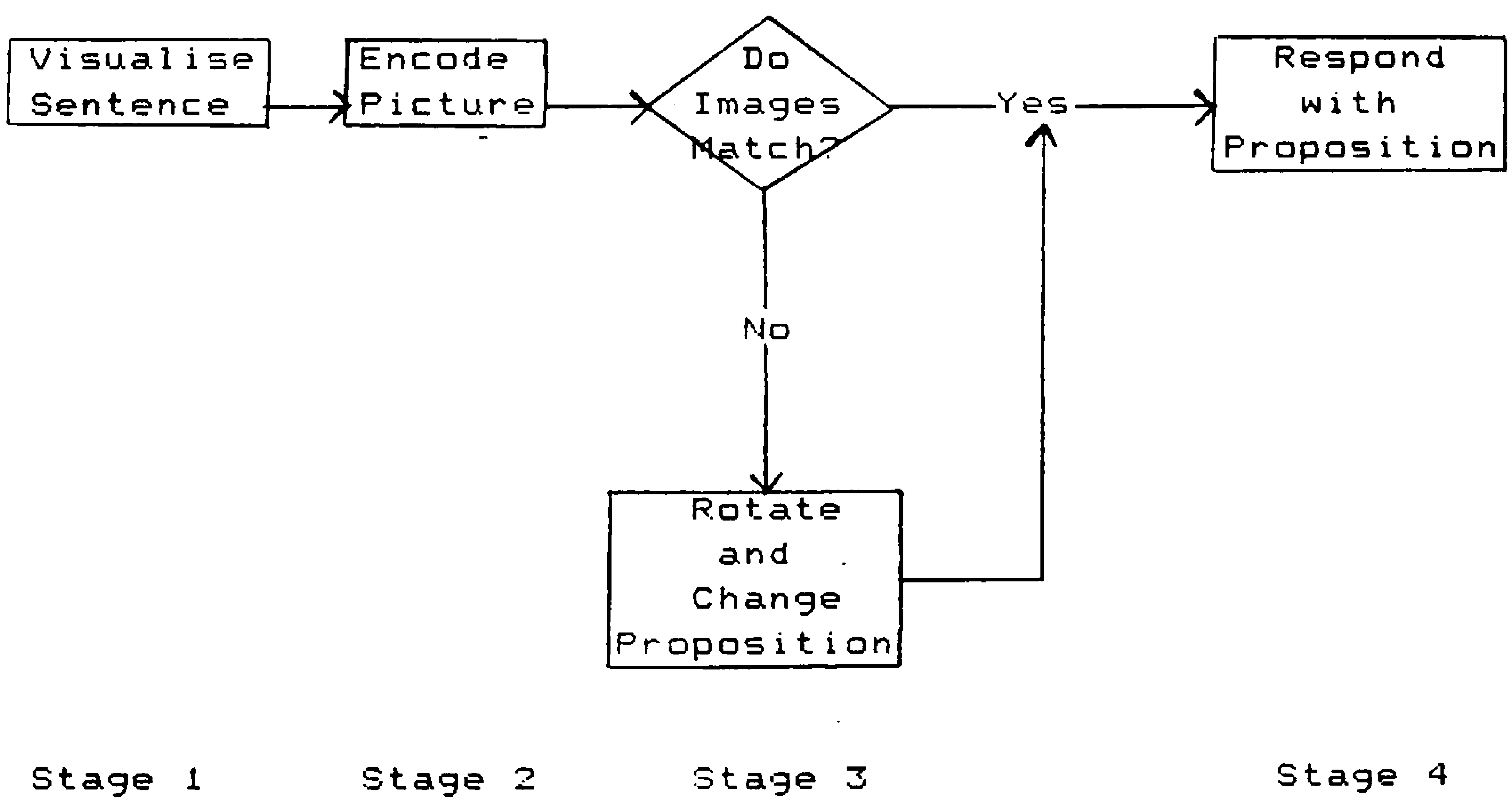


Figure 2.4 . The discrete processing stages involved in the sentence-picture verification task (sentence-first condition) according to a) the propositional model of Clark and Chase (1972), and b) the mixed propositional/imagery model of Beech (1980). (Figures from Clark and Chase, 1972, table 1, and Beech, 1980, figure 2).

images are compared, they do not match and so a rotation occurs to the image generated by the sentence. In view of this rotation the associated proposition is changed from negative to positive. These operations take additional time ('t' milliseconds). The rotation that Beech proposes need not involve imagined items circling one another which might be expected to take considerable time even with simple stimuli (eg. Cooper and Shepard, 1973). Instead it could involve one item sliding vertically over the other which might occur more rapidly. Finally the images match and, since the proposition associated with the sentence is positive, the subject responds 'true'. In a similar manner to Clark and Chase, Beech assumes a certain base time corresponding to 'tØ'. In addition it can be seen that 'r' and 'a', 't' and 'c', and 's' and 'b+d' are equivalent for the two models and thus both predict the same data. However, Beech rejects the model of Clark and Chase on the grounds of parsimony since it requires an extra processing stage.

However, an adequate explanation of sentence-picture verification needs to be able to account for circumstances in which the picture is presented or attended to before the sentence. In fact Clark and Chase modify their model substantially to accommodate such a task. Apart from the obvious reversal of the order of the first two stages, they can no longer assume that the encoding of the picture is contingent upon the relationship depicted in the sentence. Instead they suggest that pictures are always encoded propositionally in terms of the simpler ('above') relationship. However, their comparison stage also needs modification because, unlike the previous model, the relationship depicted in the embedded string (ie. 'above' or 'below') might not match that of the picture which is always in terms of the simpler relationship. In this model the comparison stage is broken down into three substages. It is suggested that first of all the subjects of the embedded strings are checked for identity and, if they do not match the sentence is transformed from '(star above plus)sen' to '(plus below star)sen' or vice versa. This requires extra time ('e' milliseconds). In the second comparison substage the propositions in the embedded string are compared and, if they do not match, the truth index is flipped occasioning extra time ('f' milliseconds). In the final comparison substage the embedding strings are compared and, if they do not match, the truth index is flipped requiring extra time ('d' milliseconds) as previously. Thus two further parameters ('e' and 'f') are required to account for a relatively small change in the nature of the task.

Clark and Chase were able to account for differences in the relative difficulties of problems within the sentence-first and picture-first conditions of their second experiment using models 'A' and 'B' respectively. However they did not explain the mean difference between the base times of both conditions. The base time for the picture first condition (t1=1793 milliseconds) was significantly larger than that of the sentence-first condition (t0=1603 milliseconds) by an average of 190 milliseconds.

Whilst Beech (1980) could account for the relative difficulty of problems within tasks as effectively as Clark and Chase (1972), he also has difficulty in accounting for the mean difference between conditions. This is because no difference should occur whether pictures or sentences are encoded first as both are transformed into images prior to comparison. However, Beech put forward a plausible post hoc suggestion that visual interference occurs when the picture is presented first and its associated image disrupts the more difficult task of translating the sentence into a visual image thus entailing longer processing time.

The above interpretations serve to indicate how, in the sentence-picture verification paradigm, explanatory models can be constructed which are founded upon abstract propositional or imaginal codes. Indeed MacLeod, Hunt and Mathews (1978, p506) argue that "untrained subjects will attack a task with a variety of strategies"...and their data collected on the same task indicates "that strategy choice is a predictable function of subject abilities as measured by psychometric tests". They found that for one group of subjects - apparently those using a linguistic strategy - reaction time scores on the task correlated highly with verbal abilities but not with spatial ability whereas the reverse was the case for those using an imagery strategy. In a follow-up study Mathews, Hunt and MacLeod (1980) found that the differences were most obvious when subjects were instructed to use an imagery strategy, but not when a linguistic strategy was required.

Consideration of the literature reviewed in this section shows that the nature of representations and processes underlying various tasks is by no means resolved. It has been argued that the data from two problem solving tasks involving a) transitive inference and b) sentence-picture verification can be explained linguistically. However, interpretations of these data in terms of models employing mixed propositional and imaginal codes are, if anything, more effective in each case. The studies of Sternberg and Weil (1980) and of MacLeod and his colleagues (1978; 1980) indicate the same general finding that both linguistic and spatial processes are involved. In both cases a clear relationship was found between the strategy employed and ability scores on psychometric tests. As Sternberg and Weil (1980, p234) claim "the effectiveness of a given strategy....depends on one's pattern of abilities" and this may be predicted from scores on psychometric tests (MacLeod, Hunt and Mathews, 1978; Mathews, Hunt and MacLeod, 1980). It appears that mental imagery can play a functional role in various problem solving tasks although it is not necessarily employed.

THE POSSIBLE NATURE OF PROCESSES IN REASONING.

Whilst the dual coding approach of Paivio and his colleagues (see Yuille, 1983) remains the subject of theoretical debate, it has generated considerable data in its support. However, the dispute between imagists and propositionalists concerning the nature of internal representations and the processes operating upon them is by no means resolved. Recently arguments have been put forward in support of three separate codes which have functional roles to play in various situations (eg. Anderson, 1978; Glucksberg, 1984). The elements of a tricode theory with multidirectional mapping relationships have been outlined by Cohen (1983) and these are illustrated in Figure 2.5 .

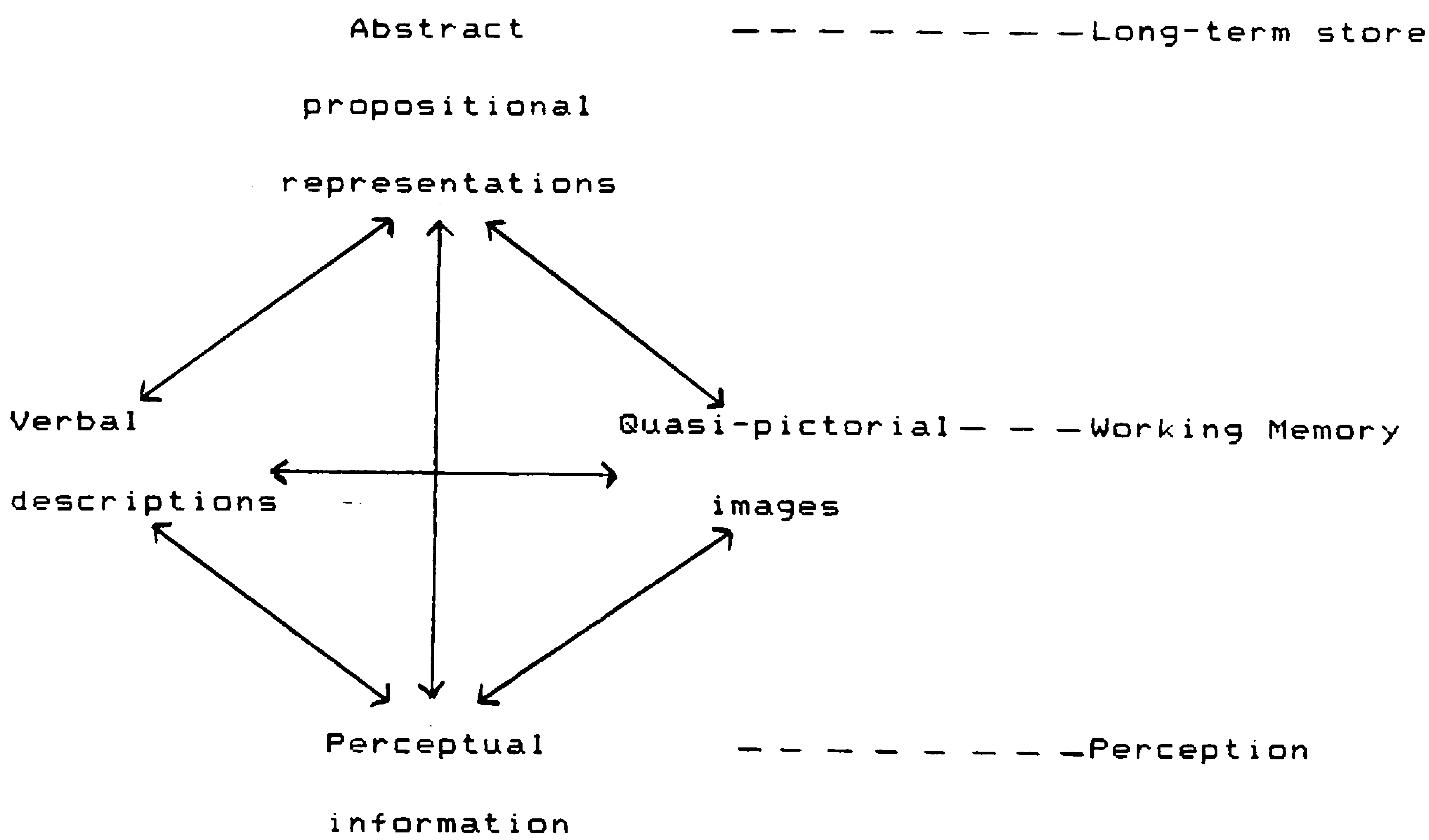


Figure 2.5 . The elements of a tricode theory showing multidirectional mapping relationships. From Cohen (1983) figure 13.

Snodgrass (1980; 1984) includes aspects of both dual coding and propositional coding theories in her tricoding approach. She suggests that verbal and visually-based imagery systems are capable of accessing an underlying propositional system. Unlike Paivio's model "the meaning of images is not contained in the image system, but in the propositional system, and interconnections between the two systems are typically made via the propositional system rather than directly" (Snodgrass, 1984, p17). Although her model will not be given detailed consideration at this stage, it will feature prominently in the concluding part of this thesis. At this stage we are more concerned to formulate hypotheses, additional to those of Evans (1980a; 1980b) about the nature of the processing underlying conditional reasoning performance.

This review has shown that the major alternative to verbal or linguistically-based coding is a system based on mental imagery. In the case of Paivio, his theory of dual codes in memory (Paivio, 1971) has been extended leading to the proposal of dual systems in thinking (Paivio, 1975). In Paivio's theory the verbal system is seen as logical and sequential, whilst the imagery system is seen as capable of synchronous analogical thinking. Evans' (1980a; 1980b) modified theory also postulates dual processes which account for conditional reasoning performance. A verbal process is supposed to account for logical performance on conditional reasoning tasks and this could be allied with Paivio's verbal system. Could the non-verbal process of Evans, which is supposed to be responsible for non-logical aspects of performance, be allied with Paivio's imagery system?

The plausibility of distinct verbal and non-verbal processes operating in high level cognition can be related to work on hemispheric specialisation. In a review of work based on a variety of methodologies Cohen (1983) points to increasing evidence that the left hemisphere is specialised for verbal cognitive functions, whilst the right hemisphere carries out non-verbal including visuo-spatial functions. This evidence provides an obvious base for specific verbal and imagery-based systems. It will be remembered that in the concluding part of chapter 1 empirical evidence was mentioned which supports the link between verbal and non-verbal processes in reasoning and hemispheric specialisation (Golding, 1980; 1981; Golding, Reich and Wason, 1974).

It appears plausible that the non-verbal process hypothesised by Evans (1980b) corresponds to an imagery-based system. If this is the case then it is implied, perhaps, that 'Matching Bias' has an imagery component. Increasing the ease with which an imaginal representation can be constructed, by manipulating the concreteness or pictorial qualities of the material employed in conditional reasoning, might be expected to lead to a corresponding increase in 'Matching Bias'. This possibility will be explored experimentally in Chapter 4.

CHAPTER 3

THEORETICAL AND METHODOLOGICAL IMPLICATIONS OF ARTICULATORY SUPPRESSION.

As outlined in chapter 1, Evans (1980a; 1980b) has suggested modifications to the Dual Process Theory of Reasoning (Wason and Evans, 1975) to explain the pattern of responses that are given in reasoning tasks. He suggested that two different thinking processes operate in parallel on such tasks. Evans (1980b) allies logical performance to the operation of a verbal reasoning process which competes with a non-verbal process responsible for various response biases. In the previous chapter, parallels were drawn between this hypothesis and another theory (Paivio, 1975) which attributes reasoning to verbal systems of thinking and more analogical, less sequential thought to non-verbal (possibly imagery related) systems.

Several studies, referred to in chapter 2, have employed selective interference techniques to identify the nature of processes underlying cognitive performance. It is thought that such a methodology would be appropriate to distinguish the verbal and non-verbal processes that Evans (1980b) postulates. Initially it was decided to attempt to interfere selectively with the supposed verbal process whilst leaving the non-verbal process undisturbed. A particular technique, known as articulatory suppression, might be expected selectively to disrupt the verbal process and it could, therefore, be a useful tool in the present experimental series. In this chapter some relevant theoretical and methodological aspects of it will be discussed.

Articulatory suppression can be defined as the articulation of something irrelevant by subjects as they perform on a criterion task. This is thought to prevent the articulatory system from playing its usual role in the criterion task. Levy (1971, p124) has claimed that "if the speech apparatus is physically engaged in verbal processing, the S cannot simultaneously be 'internally articulating'".

Although it is conceded that the verbal thought process hypothesised by Evans (1980b) might not involve articulatory mechanisms, the possibility of such involvement does appear worthy of investigation. After all, the importance of speech-based processes in cognition has been the subject of extensive investigation. Several early theories have assigned verbal processes a central role in thought.

For instance, the behaviourist approach of Watson (1930) assumed that implicit speech contributed the essential ingredient in thinking. A Russian tradition stemming from Pavlov refers to language as a 'secondary signalling system' which confers on man the unique ability to regulate his own behaviour. A development of this tradition can be found in the writings of Luria (1959). In similar vein, Vygotsky (1962) has suggested that the egocentric speech typical of young children becomes internalised by adulthood and forms the basis of logical thinking. Sokolov (1972) has reviewed much of the recent Russian literature which attempts to demonstrate the importance of implicit speech processes in thought. Furthermore, evidence from electromyographic techniques shows that electrical activity in muscles required for speaking accompanies silent reading (McGuigan, 1970), problem solving and other cognitive tasks (McGuigan, 1966; Sokolov, 1972).

Articulatory processes have been implicated in a variety of situations including, for instance, short-term memory, word recognition and reading. Some of the relevant studies will be considered below. After considering this literature, the effectiveness of articulatory suppression as a technique will be assessed.

ARTICULATORY SUPPRESSION AND SHORT-TERM RECALL.

Many authors, whose experiments will be discussed in due course, have concluded that short-term memory is speech-based and, consequently, visually presented words require translation into a speech-like code before their integration into meaningful sentences (ie. reading) can occur. Some authors have deduced that visually-displayed verbal items, presented for short-term recall, are preferentially encoded in a speech-based form.

This conclusion has been drawn because, for instance, phonological similarity amongst the items presented visually for short-term recall inhibits the subject's ability to retain the order in which the items were presented (Conrad, 1962; Conrad, 1964; Conrad & Hull, 1964). A speech-based code is presumed to be established by the subject's vocalisations at the time of presentation and, if subjects are required to produce irrelevant vocalisations at this time, the use of such a code should be prevented and the effect of phonological similarity would be attenuated. If the items to be recalled are presented visually, in which case speech-recoding is thought to be entirely prevented by enforced irrelevant articulation, then the phonological similarity effect should be eliminated altogether.

Murray (1967, 1968) studied the effect of minimising articulation in order to compare immediate serial recall of visually and aurally presented lists of letters. He either allowed the subject to articulate the item presented for recall or suppressed articulation by requiring the subject to vocalise the competing word 'the' twice every second during presentation. It was found that, when suppression was employed, recall was substantially reduced and the effect of phonological similarity was removed for visually presented lists but there was little effect with auditory lists.

Peterson and Johnson (1971) made a more strenuous effort than Murray to minimise articulation in their experiment. Their subjects were required to count from one to nine repeatedly. This sequence, which involves a variety of sounds, was uttered at a faster rate (nine items every two seconds) than Murray had required. They found that the superiority of serial recall for acoustically dissimilar material (eg. K,M,R,S,W,Y) over acoustically similar material (eg. B,C,D,P,T,V) was not apparent under suppression when the material was presented visually. Under auditory presentation such a superiority was established.

The specific nature of the speech-based code was examined more fully by Conrad (1964). He noticed in his experiments on the memory span for consonant sequences that errors often sounded similar to the target item although visual presentation was used. When the same letter vocabulary was identified against a background of white noise, the confusions which arose were highly correlated with those letters confused in recall. This suggested to Conrad that items were stored in short-term memory in an <u>acoustic</u> manner.

This conclusion was investigated further by Hintzman (1967) who indicated that at least two hypotheses about the nature of short-term memory for visually presented materials were consistent with Conrad's data. First, the visual stimuli could have caused a strongly associated auditory image to be formed. Alternatively the subject could have subvocally pronounced the items to be remembered. In the latter case articulatory features of the items to be stored are particularly important. Hintzman attempted to determine whether the basis of the effect was auditory or articulatory. He classified six consonant sounds by voicing and place of articulation. A voiced consonant requires vibration of the vocal chords (eg. B,D,G) whereas an unvoiced one does not (eg. P,T,K). Place of articulation refers to the part of the mouth which is constricted during production of the sound. Thus consonants can be articulated in the front (eg P,B), middle (eg. T,D) or back (eg. K,G) of the mouth. If the basis of short-term memory is auditory then acoustic confusions should be influenced by the same variables that cause auditory perceptual confusions. Miller and Nicely (1961) demonstrated that these generally occur within the voiced and unvoiced categories and are not typically influenced by place of articulation. However if the basis of short-term memory for visually presented material is mediated by an articulatory code then acoustic confusions should be influenced by place of articulation in addition to voicing.

Hintzman constructed nonsense syllables each beginning with one of the letters P, T, K, B, D and G and ending with AV or AF. Subjects were visually presented with thirty random sequences of six syllables each. After each sequence the subject was required to write down the items presented. It was found that memory confusions were contributed to both by voicing and place of articulation. The conclusion drawn was that short-term memory for visually presented verbal information is mediated by an articulatory rather than an auditory image. Conrad's (1970) subsequent investigations of recall in the congenitally deaf support this view.

Another researcher, Levy (1971), has assessed the changes in performance which arise from variations in overt acoustic and articulatory activity on the recall of aurally and visually presented material from short-term memory. Her subjects were given an immediate probed recall test of items which were presented visually with or without a simultaneous auditory presentation. The availability of articulatory information was varied by requiring subjects to mouth the stimulus items or an irrelevant word ('hiya') during presentation. Articulatory suppression led to poorer performance except when acoustic information was available. Therefore, it was suggested, both acoustic and articulatory features might be encoded in short-term memory.

It has been pointed out by Richardson, Greaves and Smith (1980) that the evidence concerning the effect of articulatory suppression on phonological material is equivocal. For instance, although Levy (1971) varied the phonological similarity of her stimulus materials, as was done in the studies of Murray (1967; 1968) and Peterson and Johnson (1971), she reports only a main effect of this factor and not the expected interaction with auditory/visual presentation modes. Richardson, Greaves and Smith claim, after a close inspection of Peterson and Johnson's results, that the use of visual presentation did not merely remove the auditory source of phonological information and thus reduce overall performance. In fact, although the effect of phonological similarity was attenuated with visual presentation, whether or not subjects were counting aloud, overall performance was actually improved in comparison to that with auditory presentation, although the presentation times were shorter (2.2 versus 5 seconds) in the visual condition. It was suggested that visual presentation induced subjects to employ a more efficient method of encoding the stimulus material which was distinct from a phonological representation. The experiments of Hiles (1973) suggest that a visual code might have been utilised.

Hiles (1973) reported three experiments on recognition memory in which memory and test letters were presented in either upper or lower case. When vocalisation was suppressed and the memory and test items were in different cases, subjects were slower and less accurate in their recognition. Subjects who did vocalise showed no such effect. A fourth experiment showed that visual similarity of target and test letters increased recognition time under suppression whereas acoustic similarity increased recognition time when relevant vocalisation was required. Thus, Hiles claimed, if vocalisation is suppressed, visual material can be encoded as a visual representation in short-term memory.

Richardson et al (1980) report a study which attempts to clarify the picture concerning the effects of articulatory suppression and phonological similarity on serial recall of visually presented letters and words. They found that the effect of phonological similarity was reduced to insignificance when subjects were concurrently articulating irrelevant information. They claim that, taken together, the empirical observations support the theoretical conclusion that the employment of phonological coding in short-term memory "requires the availability of either acoustic or articulatory information" (p419 - 420). Moreover they suggested that articulatory suppression abolishes the use of phonological coding with visually presented material.

Another secondary task which has been utilised to prevent speech recoding is shadowing. In an attempt to demonstrate the existence of modality specific short-term stores, Kroll, Parks, Parkinson, Bieber and Johnson (1970) required subjects to shadow a stream of letters to occupy the period between presentation and testing of a single letter which was to be remembered. The target letter was presented amongst the shadowed letters although it was differentiated by being spoken in a different voice or presented visually. Even thirty seconds following presentation, visual target letters were better retained than auditory ones. Incidentally, this result is in accordance with the subjects' claimed strategies of remembering visually presented letters as images but rehearsing aurally presented ones subvocally. The above result has often been taken to mean that shadowing prevents phonological recoding which would normally take precedence over a visual code.

In the previous pages several studies, mainly performed in the 1960's and early 1970's, were considered which employed an interference methodology, known as articulatory suppression, in an attempt to determine the nature of short-term memory. Given a whole series of disparate sensory buffer stores such as echoic and iconic memories, the question addressed was what is the preferred modality of the short-term system into which they feed? Most of the authors cited have concluded that the short-term store is phonologically based although there has been some dispute about whether it involves specifically acoustic or articulatory mechanisms.

Such investigations of memory were typical of that period. One of their major concerns was to distinguish the properties of short-term and long-term storage. The memory system was viewed as a relatively passive entity used for storage of information which might or might not be used in other cognitive tasks. However, since then, memory research has been dominated by a more functional approach in which information storage is viewed as a central part of the dynamic system necessary for information processing. This sort of approach will be considered next.

THE THEORY OF WORKING MEMORY.

At this point it is considered pertinent to consider the development of an influential theory which claims that memory occupies a central role in human information processing. The authors of this theory, Baddeley and Hitch (1974), claim to have identified a number of interacting subsystems of memory whose functions in various laboratory and everyday tasks have been explored. As a result of these explorations the notion of a general purpose Working Memory system has been developed. In essence the Working Memory system that has been proposed consists of an executive component as a central processor, which interacts with a variety of cognitive subsystems such as an articulatory rehearsal loop, long-term storage and the sensory stores (Baddeley and Hitch, 1977). The theory of Working Memory was intended to provide a broad unity between several previously distinct areas of investigation in memory and cognition. The development of this theory will be considered below.

The term 'working memory' is taken from the well known model of memory which was presented by Atkinson and Shiffrin (1968, 1971). In their view, raw environmental information first enters a modality specific register. It is further processed by being read into a phonologically-based, limited capacity, short-term store. The short-term system was conceived as a temporary working memory in which the raw information is joined by relevant associations from a semantically-based, long-term store. For example a visually presented word will be associated with its verbal name and meaning. The short-term system was responsible for directing the flow of information into and out of long-term storage, for decision making and functioned as a controlling executive system.

However, as Baddeley (1976) states, many problems for this model occured when modality effects were found in short-term storage (Murdock and Walker, 1969). Also neuropsychological evidence was inconsistent with the idea that the short-term system is essential for long-term learning and retrieval. Shallice and Warrington (1970) described a patient whose short-term span was grossly impaired but whose long-term learning ability appeared quite normal.

As an alternative to the modal approach, the levels of processing theory of Craik and Lockhart (1972) was developed. Here the emphasis was away from separate stores with different coding processes and towards a theory of memory which is viewed as "a continuum from the transient products of sensory analysis to the highly durable products of semantic-associative operations" (p676). In other words a deeper analysis would lead to a more persistent memory trace although material could be retained by "recirculating information at one level of processing" (p676).

However Baddeley (1978) suggests that this approach is theoretically barren for a number of reasons. First, experimenters within this framework have usually relied upon an intuitively plausible assumption of the hierarchy of levels of various types of processing. The failure to find an independent measure of 'depth' (eg. Craik and Tulving, 1975) has limited the theory's usefulness. There is no experimental evidence which supports the idea of a continuum of processing rather than a series of discrete domains. The basic assumption of the theory that deeper processing leads to a more durable trace is not necessarilly true. Indeed, apparently superficial characteristics of stimuli, for example the location of a written word on a page of text (Rothkopf, 1971) can lead to a very persistent memory. The assumption of a simple hierarchy of levels, for instance with words, from physical features to phonological features to a semantic level, is not born out experimentally by the work of Marcel (1983). He has shown that subjects who are unaware of the orthographic properties of printed words can nevetheless be influenced by their semantic properties.

The alternative Working Memory approach, favoured by Baddeley and Hitch (1974; Hitch and Baddeley, 1976), analysed human memory into specific subcomponents and investigated the contribution of these to a number of tasks. This approach has been developed in many subsequent articles by the authors which will be considered in due course. In the next few pages the series of experiments which led to its inauguration will be described. First, a little more context will be considered.

Although the short-term store has often been assumed to play a vital role in many cognitive tasks, little experimental evidence had been provided for this view. Baddeley and Hitch (1974) endeavoured to determine experimentally whether the same Working Memory system is involved in a number of of tasks including reasoning, language comprehension and learning.

There was much disagreement at that time about the nature of the short-term store which, Baddeley and Hitch (1974) believed, stemmed from the different paradigms which were used in its investigation. Two main approaches were apparent. As shown in the previous section, studies of the memory span suggested a limited capacity store which was concerned with order information and was closely associated with speech. On the other hand, evidence relating to the recency effect in immediate free recall (the tendency to recall the last few items in a list first and best, eg. Postman and Phillips, 1965) showed a resistance to semantic or speech-based variables (Glanzer, 1972) whilst also suggesting a limited capacity store.

Baddeley and Hitch (1974) decided to concentrate upon the point of agreement between paradigms by attempting to demonstrate the limited capacity of the Working Memory system which they postulated. The effect of a concurrent memory load upon performance in reasoning, comprehension and free recall should be to absorb some of the limited storage capacity of the system thus occasioning substantial disruption.

The verbal reasoning task which was used (Baddeley and Hitch, 1974; Hitch and Baddeley, 1976) was derived by Baddeley (1968) and has been claimed to be "typical of a wide range of studies on reasoning" (Baddeley, 1976, p170). In this task subjects are presented with a sentence which describes the order of occurence of two letters together with an instance consisting of the two letters mentioned. They are required to decide whether the sentence is true or false. For example, given the sentence 'A is followed by B' and the two letters 'B A', the subject should respond 'True'. A series of such problems were given whilst the subject was concurrently holding up to six items in memory. If interference occurred then it could be assumed that both tasks were making demands upon a single, limited capacity system. The complexity of the reasoning problems was varied by manipulating various features of the sentence (voice, truth value and polarity). It was hypothesised that, since more complex problems should require increased space in Working Memory, they would be more susceptible to interference. That is, if the store was really of limited capacity.

Their data indicated that, whilst subjects could hold up to two items in memory without impairment in reasoning latency or accuracy, interference was found when six items were retained. However, the increase in reasoning latency was roughly constant across problem types. Since a memory preload technique had been used in these studies, it was possible that the increase in latency was due to subjects consolidating memory items by rehearsal prior to reasoning rather than due to strictly concurrent interference.

Consequently they adopted a new procedure in an attempt to eliminate this possibility. Subjects were required to rehearse the memory items aloud repeatedly whilst reasoning. Other conditions were included in which overlearnt sequences ('the the the' or 'one two three four five six') were recited continuously in order to assess the effects of purely articulatory interference. Using this new procedure, they found a non-significant tendency for articulation per se to slow down performance, although a previous investigation had achieved significance with an alternative overlearnt sequence (Hammerton,1969). The solution latencies from Baddeley and Hitch's (1974) study (experiment III) are shown in Figure 3.1 .

The results illustrated in Figure 3.1 suggested to Baddeley and Hitch that verbal reasoning, like the memory span task, might have an articulatory component. The fact that a substantially greater interference effect due to memory load occurred, which interacted with problem type such that more complex problems suffered more disruption, was taken as support for the existence of a Working Memory system consisting of a limited capacity work-space which could be flexibly allocated either to storage or processing.

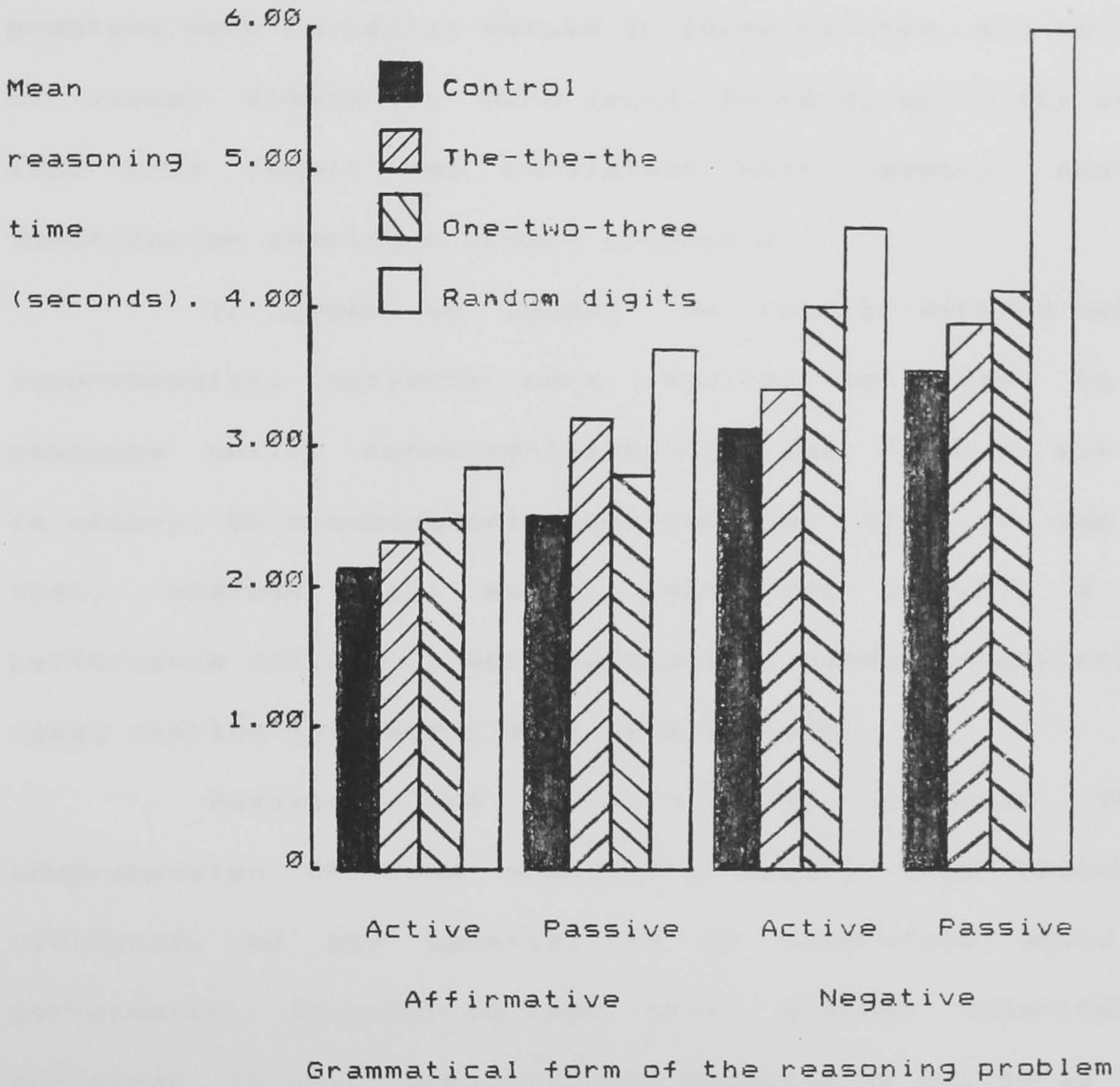


Figure 3.1 . The effect of types of concurrent articulation on mean reasoning solution latencies. (From Baddeley and Hitch, 1974, figure 2).

Baddeley and Hitch's next experiment investigated the effects of phonological and visual similarity of the letter pairs used in the verification task. Phonological similarity is known to have effects on the memory span (see previous section) but little effect of visual similarity has been found (Baddeley, 1966). In fact, phonological similarity had comparable effects in the reasoning task. When the letters in the sentence to be verified were similar phonologically, fewer problems were correctly solved in three minutes, and no effects of visual similarity were found. Baddeley and Hitch proposed that this result was consistent with memory span and verification sharing a common component.

In order to assess the role of Working Memory in comprehension, subjects were required to listen to prose passages whilst concurrently holding one, three or six digits in memory. On a subsequent comprehension test it was found that, whereas small memory loads led to only a slight performance deficit, a more substantial decrement occured with loads nearing the memory span of six items.

Baddeley and Hitch (1974) reasoned that if comprehension utilised short-term memory then phonological similarity of the material to be understood would impair performance. In order to test this, another experiment was performed in which subjects were required to decide as quickly as possible whether sentences, half of which contained a high proportion of phonologically similar words, were possible (ie. grammatical and meaningful) or not. In this case they found that, although a post-test showed phonologically similar sentences took no longer to read, subjects did take longer to make acceptability judgements.

Thus verbal reasoning and comprehension studies had both yielded data in support of the Working Memory hypothesis. These have shown that the system suffers from the demands of a near-span additional memory load and is disrupted by the presence of phonological similarity.

Subsequently, the effects of memory preload, concurrent memory load and articulatory suppression in free recall learning were investigated. In this way it was hoped to implicate the Working Memory system in a broader range of tasks than sentence verification and comprehension. It was also of interest to assess its role in long-term learning.

Free recall of aurally presented lists of sixteen unrelated words was studied under a memory preload of zero, three or six digits presented visually. Since the memory span and the recency effect were both thought to make demands on a common short-term store, a dramatic interference was expected in this task. However, a memory preload did not have any effect on the characteristic recency effect which was obtained, as expected, with immediate free recall and abolished, as usual, after a thirty second delay. The effect of memory load was restricted to the long-term component of recall, with increased load leading to poorer retention. Since subjects might have suceeded in transfering the preload items into long-term storage by the time that the word list had been fully presented, their lack of interference upon recency was not considered surprising.

In an attempt to eliminate this artefactual explanation another experiment was performed which was similar to the previous one except that a concurrent load was imposed. Once again, the recency effect obtained with immediate recall was unaffected by memory load although the long-term storage component showed the usual deficit as memory load increased. Thus Baddeley and Hitch (1974) were drawn to the conclusion that the recency mechanism is independent of the mechanism involved in the memory span task.

Finally, Baddeley and Hitch (1974) refer to an unpublished study by Richardson and Baddeley which examined the effects of articulatory suppression upon free recall of visually or aurally presented word sequences. Suppression impaired retention as much for early serial positions as for recently presented items and appeared to be most marked for visually presented items. This result is cited as being consistent with the idea of a Working Memory operating on phonologically coded information and transferring it to long-term storage.

In summary, Baddeley and Hitch (1974) showed that phonological similarity, articulatory suppression and a concurrent memory load each, individually, impaired performance in verbal reasoning, comprehension and free recall. They claimed that these results constituted "prima facie evidence for the existence of a working memory system which plays a central role in human information processing" (p237). However, they made more specific theoretical proposals about the nature of the system. This was because their results showed that, despite disruption, subjects were still capable of reasoning, comprehending and learning even when a near span memory load of six items was retained and were hardly affected at all by a two or three item load. In order to encompass these findings it was suggested that part of the Working Memory system consists of a relatively passive phonological response buffer of limited capacity. This could be used to supplement a fairly flexible central executive processor which is required for information processing, decision making and transfer to long-term storage. Subsequently, Baddeley, Thomson and Buchanan (1975) explored the capacity and the characteristics of the phonological response buffer proposed by Baddeley and Hitch (1974). They tested the subjects' memory span for words of varying length as measured by number of syllables. Subjects were required to recall lists of aurally presented words in their serial order and it was found that the span was less for longer words. However the polysyllables differed from the monosyllables in terms of their abstractness, imageability and in other ways. These linguistic factors were controlled for in an experiment of similar design in which the words to be remembered were names of countrys and thus came from the same linguistic category. Similar results to the previous study were obtained.

However, the two experiments described above did not determine whether the word length effects were due to temporal duration or because of their articulatory complexity. In an attempt to resolve this issue, Baddeley et al (1975) tested retention for two groups of words which were matched for number of syllables but which differed in the time they usually take to utter. An example of a long duration word is 'cyclone' and a short duration word is 'bishop'. Subjects were able to recall more items from the lists of short duration words. Even when the words were presented visually and subjects were requested to rehearse the items aloud, fewer long words (as measured by articulation and reading rate) were recalled. In fact it was shown that subjects could recall as many words as they could read out in about 1.6 seconds.

The above results suggested that the phonological buffer is time based and is analogous to a tape loop of about 1.5 seconds duration (eg. Hitch and Baddeley, 1977, p1Ø3). The relationship of memory span and articulation rate together with the existence of a word-length effect even with visually presented material implicates an <u>articulatory</u> rather than an auditory memory system. The latter system has also been ruled out by the finding that reducing the temporal duration of the auditory stimuli, via a computer prgramme which removes sections of the spoken item without reducing its intelligibility, does not lead to improved memory performance for compressed compared with normal digits (Hitch and Baddeley, 1977).

Further evidence suggesting an articulatory system in the word length effect was provided when it was shown that, with visual presentation, the word length effect disappears under articulatory suppression. The effect of suppression, as noted earlier, depends upon presentation modality (Levy, 1971; Peterson and Johnson, 1971) and, as we shall see later, on whether articulatory suppression continues during output. However, in their final experiment, Baddeley et al (1975) found that with auditory presentation the word length effect remains even under suppression.

Although the modality specific result is not consistent with the view that suppression blocks the articulatory system upon which the word length effect is dependent, Baddeley et al (1975) have explained it. They argue that the transformation of a visual stimulus into a phonological representation is stopped by articulatory suppression and thus access to the articulatory loop is prevented. Auditory material is, however, presumed to be already established into a phonological code.

Eventually the articulatory rehearsal loop was related to the buffer system underlying speech production (Morton, 1970) and its relation to the central executive was stated more clearly. It has been claimed that the verbal loop "is able to store a limited amount of speech-like material in the appropriate serial order" (Baddeley, 1976, p176) and its capacity has been shown to be "limited by time rather than by the amount of information or number of events" (Baddeley, 1976, p178). It has been suggested that the articulatory loop functions as a 'slave system' of limited capacity which supplements the central processor on occasion.

The articulatory loop which is subject to temporal decay is described by Hitch (1980, p168) as an "actively controlled system for retaining sub-vocal output" which distinguishes it from "a relatively passive 'input register' underlying recency" which is item-based rather than time-based (cf. Waugh and Norman, 1965). However, Baddeley (1976, p178) suggests that "since STM in general appears to be limited in number of items held rather than time, it seems unlikely that the central executive component of the system is also time-based". Interference with the central processor would only be observed when the loop's capacity is exceeded and the executive has to devote some of its information processing capacity to storage by alternative means.

Although the bulk of tasks considered in the previous pages have concerned verbal material, some consideration should be given to peripheral components of the Working Memory system which are specialised for storing and manipulating visual material. The sensory storage of visual information in iconic memory is subject to masking and has a large, though instantaneous capacity (see Baddeley, 1976; Coltheart, 1983; Neisser, 1967). However, there is a growing body of evidence which points to the existence of another visual store of longer duration (see Phillips, 1983). This, unlike iconic memory, has a limited capacity and is not subject to masking.

Unpublished evidence cited by Baddeley and Hitch (1974), which showed that visual recognition of pictures could compete with a concurrent mental arithmetic task, led them to propose the existence of a peripheral memorial component based on the visual modality. However, the case for a single central processor, forming the core of the Working Memory system and implicated in both visual and verbal memory tasks, is maintained. The results of Baddeley, Grant, Wight and Thomson (1974), Baddeley and Lieberman (1980) and Oakhill and Johnson-Laird (1984) all seem to indicate a spatial system specifically. This is referred to as the 'visuo-spatial scratch pad'. Baddeley (1980) even suggests that the dichotomy between propositional and analogical views of imagery (discussed in chapter 2) may be a false one. He writes (p20) "it is for example quite likely that the scratch pad is a device which takes propositional codes from long-term memory and manipulates and displays them <u>via</u> an analogical peripheral system".

Some recent modifications to the model of Working

Memory will be considered next. The notion of the articulatory loop, as described above, has been able to account for a wide range of experimental findings relating to phonological similarity, word length and various effects of articulatory suppression on immediate memory, word length and similarity providing the tasks employ visual presentation. However, as discussed previously, with auditory presentation anomalous results of articulatory suppression on phonological similarity and word length are found. If the word length and phonological similarity effects stem from the articulatory process, then articulatory suppression would be expected to prevent the effects occurring regardless of presentation modality. Although several authors have argued for the existence of an auditory or acoustic component in short-term memory, which could explain these discrepant results, none was envisaged within the original Working Memory model except for the relatively peripheral Precategorical Acoustic Store of Crowder and Morton (1969). However, recent evidence presented by Baddeley and his colleagues, which will be discussed below, now argues in favour of such a store.

The first study of relevance is that of Vallar and Baddeley (1982) who investigated short-term retention of visually presented meaningless trigrams under various interfering conditions. Recall was not seriously impaired after a five second interval under articulatory suppression, commenced immediately following presentation, although there was a decrement after fifteen seconds relative to a silent control condition. However, the Peterson and Peterson (1959) technique of counting backwards by threes had a drastic effect at both intervals. Vallar and Baddeley argued the possibility that, with visually presented letters, the most crucial function of the articulatory loop is in converting the visual stimulus into a phonological code. Although they reasoned that this should be disrupted if articulatory suppression commenced at input, a second experiment employing such a procedure showed no disruption after five or fifteen seconds retention, relative to a silent condition. They argue that the disruption engendered by the counting technique in the first experiment might have occurred due to its increased demands on the limited capacity central executive component. However, they speculated that the maintenance of the visually presented information could have occurred in an acoustic store which "is sufficiently powerful to take over the role of subvocal rehearsal, at least in the memory span situation" (p59). The manner in which visually presented information is translated into a form suitable to enter such a store when articulation is prevented is not stated.

A series of experiments by Salamé and Baddeley (1982) are also relevant at this point. They investigated the effects of auditory interference on immediate memory for visually presented digits. They found that unattended speech impaired subjects' performance compared with a silent control condition, even though instructions were given to ignore the speech. The effect is not at a semantic level however, since similar effects accrue whether the unattended speech consisted of meaningful or nonsense words. Since the effect occurred whether or not the unattended speech was synchronised with the visual presentation of digits it is unlikely to be due to attentional distraction.

Another experiment (Salamé and Baddeley, 1982, experiment 3) was designed to test the possibility that unattended material was gaining access to the articulatory loop and thus disrupting subvocal rehearsal of the visual material. Articulatory suppression - saying 'the' repeatedly - was employed in order to prevent availability of the loop for rehearsal. Under these conditions the effect of unattended speech disappeared. Salamé and Baddeley argue that there appears to be some conflict, manifested in the articulatory loop, between the inner speech occasioned by subvocal rehearsal and irrelevant speech which the subject is attempting to ignore. Their next experiment (4) examined the possibility that irrelevant speech is somehow gaining access to the loop and therefore disrupting performance. If this were the case then irrelevant speech containing words of longer duration should cause more disruption. This is because the capacity of the loop is time-limited and and sensitive to the effect of word length (Baddeley et al, 1975). In fact an equivalent amount of disruption occurred with long and short duration words.

The alternative possiblity considered was that unattended speech causes disruption by feeding into another memory store which can also be fed by the articulatory loop. Since white noise had previously been shown to produce less drastic impairment than speech, the effect could depend on phonological similarity between unattended speech and the visual digits. Salamé and Baddeley's (1982) final experiment established that this is indeed the case. More disruption was occasioned when phonologically similar materials were used in both the unattended speech and the materials presented visually for remembering. Salamé and Baddeley's results support the assumption that two separate memory systems exist - "one based on phonological coding and accessible either through audition or articulation, whilst the second is used to store visually presented material" (p161). This is inconsistent with the original Working Memory model (Baddeley and Hitch, 1974) since that assumes the existence of an articulatory store alone. The latest model of Working Memory with which the author is familiar (Baddeley, 1983) revises the concept of the articulatory loop. The nature of this is discussed below.

The revised concept of the articulatory loop "assumes a phonological input store supported by an articulatory control process" (Baddeley, Lewis and Vallar, 1984, p249). This concept of the articulatory loop was tested in a series of five experiments by Baddeley et al (1984). Initially they performed three experiments which replicated Murray's (1968) finding by showing that articulatory suppression did not eliminate the phonological similarity effect with auditory presentation and written recall. This was the case whether suppression occurred just at input or both at input and output so that the possibility that the effect emerged during a silent recall period was eliminated. It was concluded that the phonological similarity effect reflects the nature of the coding in a phonological input store. Auditory material is automatically registered in this store regardless of articulation conditions, but with visually presented material, suppression - which prevents articulation during the presentation period, removes the phonological similarity effect.

A further two experiments were performed to examine the word length effect which is assumed to depend on articulatory rehearsal. Since long duration words can be rehearsed less quickly, their memory trace is less frequently refreshed and hence recall is diminished relative to short words under silent conditions. Under articulatory suppression no advantage for short duration words would be expected regardless of presentation modality. An apparent contradictory result reported by Baddeley et al (1975) showed that with visual presentation the word length effect was abolished but this was not the case with auditory presentation. However, it should be pointed out that Baddeley et al (1975) did not extend suppression into the spoken recall period. The word length effect could have been emerging at that time if subjects were rapidly shifting information from an auditory to an articulatory mode before recalling the material. In contrast, Baddeley et al (1984) found the word length effect to be virtually abolished under suppression maintained during input and written recall with both visual <u>and</u> auditory presentations. The above results were argued to support the revised concept of an articulatory loop which, Baddeley et al (1984) suggest, explains a wide range of findings in a parsimonious manner.

The Working Memory framework is proving extremely fruitful but, as would be expected with any influential theory, it is not without critics (eg. Monsell, 1984; Klapp, Marshburn and Lester, 1983; Richardson, 1981; Richardson, 1984). Some criticisms will be considered in Chapter 7, after the present series of experiments has been described. In the years following its inception, the Working Memory model has been considerably refined. This is not really surprising in view of the range of experimental situations that have subjected it to testing and analysis. Indeed Hitch (1980, p157) argues that it is necessary "to consider evidence from as wide a variety of sources as possible in order to provide 'converging operations' for testing models and hypotheses". Hitch's (1978) application of the model in studying the processes underlying mental arithmetic serves to demonstrate its generality in a relatively everyday task. Investigations into the role of Working Memory in another everyday task, reading, will be discussed in the next section.

The present section has shown how the concept of Working Memory (Baddeley and Hitch, 1974; 1977; Hitch and Baddeley, 1976) involves a modification of the modal account of memory favoured by earlier investigators. Baddeley and Hitch are not content with the narrow approach which views memory simply as a system necessary for the recall of previous events. Whilst this is undoubtedly an important function of memory, they argue, along with other authors, that many cognitive tasks (including problem solving, understanding speech and reading for instance) require some interplay with the memory system. Their intention is made clear by Hitch (1980). It is to avoid the detailed, but very limited, study of a narrow range of recall and recognition paradigms and to pursue a more ecologically valid course. Of course, useful information can still be gained from traditional memory paradigms, but the overall aim, engendered in the Working Memory approach, is to determine the function of memory in general cognition.

The hope was that theoretical links between previously separate topics in cognitive psychology would become established through the Working Memory approach and that, within this broader context, questions about the nature of basic mechanisms would be formulated more easily. Since research and theory development in conditional reasoning has suffered harshly through its isolation from more popular cognitive areas, the present author can only endorse this laudable aim. The relevance of the present series of experiments will be considered in the context of Baddeley and Hitch's theory.

ARTICULATORY SUPPRESSION AND READING.

The studies reviewed earlier in this chapter have suggested that irrelevant articulation hampers the phonological coding of visually presented materials. This has encouraged its use as a technique in the investigation of the processes involved in reading. In a very early investigation of the role of 'inner speech' in reading, Pintner (1913) showed that a secondary counting task did not impair the subjects' comprehension of text which they were reading simultaneously. Unfortunately the rate of articulation was not specified in this case although this would appear to be a crucial consideration. For instance, Peterson and Peterson (1959) showed that when a similar distracting task is used in a short-term memory paradigm, in which subjects count in the interval between presentation and recall of test material, the rate of counting determines the amount recalled.

Many subsequent authors have studied the role of phonological recoding in visual word recognition and reading by requiring subjects to engage in an irrelevant articulatory task and measuring the extent of the interference caused. In one study Levy (1975) required subjects to learn sets of three short (seven word) unrelated sentences (eg. 'the attractive girl kissed the surprised fellow'). They were then presented with a test sentence and were asked to indicate whether it was identical to any of the previous three. On half of the occasions no difference was apparent. However, on the other trials, a difference was found. One of three types of change was possible:

1)Lexical - in which a synonym was substituted for one of the nouns,

2)Semantic - in which subject and object nouns were interchanged,

3)Filler - in which a verb or an adjective was changed.

Subjects were required to perform the task under one of four experimental conditions;

a)Visual Silent - in which subjects read the sentence silently,

b)Visual Vocalised - in which subjects read the sentence aloud,

c)Visual Suppressed - in which subjects read the sentence whilst concurrently articulating the numbers from one to ten repeatedly,

d)Auditory - in which the subject listened to the sentence whilst remaining silent.

Levy found that articulatory suppression reduced the ability to detect lexical and semantic changes whilst no comparable effect was found in the auditory condition. She concluded that concurrent articulation prevented the formation of a phonological representation necessary for comprehension.

Most of the evidence suggests that individual words can be understood without phonological recoding, although some authors claim that recoding often occurs prior to lexical access (Levy, 1978; Underwood and Holt, 1979; Kleiman, 1975; Baddeley, 1979). However, phonological recoding appears to be useful in order to retain words, phrases or sentences in Working Memory until comprehension has occured (Kleiman, 1975).

Kleiman (1975, experiment III) had subjects shadowing aurally presented digits whilst making decisions about short visually presented sentences. One of four types of judgement was required. For the first three types, an initial target word was written above the sentence. The decisions were;

1) <u>Phonological</u> - Is there a word in the sentence which sounds like the target?

2) <u>Graphemic</u> - Is there a word which looks like the target in its spelling?

3) <u>Category</u> - Is there a word which names a member of the category named by the target?

4) <u>Acceptability</u> - Do the five words, in the order written, form a semantically acceptable sentence?

He found that shadowing disrupted Phonological decisions more than either Graphemic or Category decisions and concluded that accessing the meanings of word pairs does not require phonological recoding. In his previous experiment (Kleiman, 1975, experiment II), Graphemic decisions about phonologically similar (eg. BLAME FLAME) and dissimilar (eg. COUCH TOUCH) materials were made with or without shadowing. Since shadowing hindered both types of decision equally, he concluded that speech recoding is not essential for Graphemic decisions. However, when judgements about the acceptability of the five word strings were needed, a marked disruptive effect of shadowing was found. It was suggested that this sort of Acceptability decision required the simultaneous evaluation and integration of several word units and entailed a similar storage load to that entailed in normal sentence comprehension. Kleiman (1975) argues that his data is consistent with a working memory model containing both a visual and an articulatory store. Phonological recoding, he suggested, is required to keep word units available in working memory for more extensive processing such as that involved in reading. However comprehension occurs via a direct visual to meaning route when the information load is small, for instance when words can be monitored individually.

More recently Baddeley (1979) has assessed the possible role of Working Memory in reading. He has pointed out that Kleiman's (1975) implication that the articulatory loop would be required for normal reading, is equivocal. Shadowing, Baddeley suggests, is a demanding task which is likely to occupy the loop as well as requiring some contribution from the central executive. This claim is supported by Kleiman's data which indicated that performance is impaired with Graphemic and Category decisions although this was less than that arising with Phonological and Acceptability judgements. Therefore the extra disruption could be due to Acceptability decisions making greater demands upon the central executive rather than the loop. Also, with Kleiman's Acceptability judgements, close monitoring of word order was often necessary if a correct decision was to be made. Baddeley suggests that this sort of task might require use of the articulatory loop to a greater extent than normal reading would.

Indeed, Baddeley, Eldridge and Lewis (1981) have also criticised Kleiman (1975) methodologically suggesting that, with his first three decison types, an implicit memory task could have been involved in that the subject holds a target word in memory whilst processing the sentence. Furthermore they argue that "a subject about to make a phonological judgement would be most likely to maintain the word phonologically, whereas the graphemic and semantic judgements might be more likely to be supported by visual and semantic representations" (Baddeley et al, 1981, p441). As a consequence, his effects might have arisen due to the interference between the shadowing and phonologically maintaining the target word, rather than shadowing and reading. Even Kleiman's 'Acceptability' condition, which was not open to the criticisms above, was not beyond reproach. After all, shadowing is a demanding task which is likely to act not only as an articulatory suppressor but also to impart a large cognitive load on Working Memory. The decrement in this condition, it was suggested, arises because of information overload rather than the suppression of articulatory coding.

Baddeley (1979), in assessing the circumstances in which the loop would be utilised, cites a study of simple sentence verification by Hitch and Baddeley in which a 'Yes/No' decision to a sequence such as 'Wasps have oars' is requested. It was found that, whereas a memory load disrupted performance, concurrent articulation per se had no effect relative to silent controls. This result suggests that the loop is not needed for reading or comprehension of simple sentences. Baddeley refers to another unpublished study with Simmonds investigating the effect of suppression on reading of a wide range of passage types. Although an occasional increase in reading speed was noted, suppression had little effect on performance.

The sort of results noted above have induced Baddeley in his later writing to de-emphasise the contribution of the loop for normal reading and comprehension. However he has specified occasions when the loop would be beneficial. Phonological comparisons benefit through usage of the loop although they can be made without it (Baddeley, 1979). Certain difficult material, particularly where order information is required, would probably require the loop (Baddeley, 1979; Baddeley, and Lewis, 1981). Therefore on tasks requiring attention to these kinds of information, articulatory suppression would be expected to drastically disrupt performance at least with visually presented materials.

AN APPRAISAL OF THE ARTICULATORY SUPPRESSION TECHNIQUE.

At this stage it is necessary to appraise techniques which have been utilised to assess the contribution of various processes in cognition. As has been shown in this and the previous chapter, the methodology known as selective interference appears to be commonly employed for this purpose.

In essence this methodology relies upon the fact that it is sometimes difficult to execute two tasks concurrently although each alone is undemanding. This difficulty is sometimes explained in structural terms by claiming, for instance, that both tasks are competing for use of a specific perceptual or motor mechanism (eg. Brooks, 1968). On other occasions the difficulty of dual task performance has been explained in terms of competition for a general, limited capacity resource (eg. Broadbent, 1958). Kahneman (1973) incorporates both structural and capacity considerations within his explanations of interference tasks in relation to attention and effort.

In addition to the classic imagery suppression studies of Brooks (1967; 1968), described in some detail in the previous chapter, another example of structural interference between imagery and a percept in the same modality is given by Segal and Fusella (1969). They found that the detection of a weak signal was more difficult if subjects were imagining a visual scene. Detection of a weak tone was more difficult if subjects were imagining a sound. The interference was presumed to occur because of the competition of the internal image and the external signal for a common process.

On occasions when concurrent task interference is due to capacity limitations then it should arise whether or not the tasks require the use of common mechanisms. Kahneman (1973) has reviewed the literature with the above points in mind and has claimed to find considerable support for them. An example which he quotes is that of Posner and Rossman (1965). They required subjects to retain three letters for a brief period whilst they were engaged in concurrent mental tasks of varying difficulty. They found that there was a decrease in the efficiency of retention as task difficulty increased. Kahneman interprets this result by assuming that rehearsal of the memory load demands considerable attention or effort. Whenever the interpolated task pre-empts attention, the rehearsal will be disrupted and retention will suffer.

Recently Allport (1980) has disputed the usefulness of concurrent interference tasks in investigations of capacity limitation. He states that this is because most if not all such theories which impute a limited capacity resource are unfalsifiable since they permit an escape clause. "If two tasks can in fact be performed concurrently, each independently of the manipulated difficulty of the other, then perforce one or both of the tasks must be 'automatic'. That task is therefore, by definition, irrelevant to the evaluation of the theory" (Allport, 1980, p.143).

However, Allport has not denied the utility of concurrent interference tasks in assessing the nature of cognitive structures or processes. Indeed he suggests that all dual task interference arises through structural/process limitations. In assessing the nature of the dual processes which are hypothesised to underlie conditional reasoning performance (Evans, 1980b), competing task methodology seems to be appropriate. The issue now to be addressed is whether articulatory suppression is an appropriate competing task for the present purpose.

Most of the studies discussed in this chapter have utilised articulatory suppression as an interference technique. Besner, Davies and Daniels (1981) have reviewed the evidence concerning its effects and are not satisfied that the assumptions upon which it is based have received sufficient critical attention. For instance they claim that Kleiman's conclusions might be ill-founded since they ignore "the distinction that can be drawn between phonology used for accessing the lexicon (prelexical phonology) and phonology that can be retrieved from the lexicon (postlexical phonology) after access for the purpose of naming, or other phonological operations" (Besner et al, 1981, p418). They point out the logical possibility that rhyme judgements could be made using either prelexical or postlexical phonology. If rhyme judgements require the use of postlexical phonology, Kleiman's finding that they are affected by suppression is entirely irrelevant to the question of whether prelexical phonology is utilised in making decisions of synonymity. It is quite possible that prelexical and postlexical phonology are not affected by the same variables.

An experiment which Besner et al (1981) consider to be of importance in assessing the locus of suppression effects is that of Martin (1978). She studied the effects of concurrent articulation on performance of a Stroop colour word interference task (Stroop, 1935). The Stroop task involves presenting a subject with a row of letters printed in coloured ink. The subject is asked to name the colour of the ink. When the letters form a word, latency of response is increased. If the letters spell a colour (eg. Red) then the response is delayed even further (Dyer, 1973).

Martin (1978) assumed that the Stroop interference should be reduced if the colour words were made more difficult to read by engaging in irrelevant concurrent articulation which, she assumed, inhibits the speech recoding of visually presented material. If (prelexical) phonological recoding is necessary for lexical access of visually presented words then its prevention should lead to improved performance on the Stroop test. In support of her hypothesis she found that Stroop interference was attenuated under concurrent articulation.

If this result is genuine, then Besner et al (1981) claim that concurrent articulation could be used with some confidence as a technique for studying the role of prelexical phonology in reading. Martin's (1978) result is particularly surprising since it has been argued by Coltheart (1978) that a visual access route to the lexicon is faster than a phonological access route. It is claimed that this would lead to an expectation of Stroop interference from the colour word condition (due to interfering lexical access of the colour word via a visual route) even under the suppression condition. Besner et al therefore decided to attempt a partial replication of Martin's experiment but did not find any reduction of the Stroop effect under concurrent articulation.

Besner et al (1981) admit that their failure to replicate Martin (1978) is not conclusive. Indeed, a recent study by Chmiel (1984) has indicated that a possible methodological difference, based on the labelling of bins into which cards are sorted, could account for the differences between their results. Using patch labels, as in Martin's study, he replicated her results showing a decrease in the Stroop effect under concurrent articulation relative to a silent control condition. However, when labels consist of colour names printed in black ink, Besner et al's result is replicated in that no reduction in the Stroop effect occurs under concurrent articulation, in fact there is a slight increase in efficiency. Unfortunately it is not known which labelling technique Besner et al used. Chmiel's data, collected on Stroop and reverse Stroop tasks (in which the cards have to be sorted according to the name of the colour-word rather than the ink colour) suggest that different codes are used based on the labelling of bins into which cards have to be sorted. He claims (p218) that "concurrent articulation must, at least, have an effect on postlexical phonology" but it is unlikely to affect phonology prelexically.

Besner et al's (1981) experiments tend to support this view. They required a task in which either a prelexical or postlexical phonological code must be used in order to test whether it is affected by suppression. In their third experiment they required subjects to work speedily through lists of ten irregular word pairs or ten non-word pairs and tick only those pairs which rhymed. Half of the items within each list did so. In half of the trials, concurrent articulation of irrelevant information was required whilst on other occasions silence was maintained.

If Kleiman's (1975) claim that prelexical phonology is affected by suppression is correct, then rhyme judgements of non-words should be more affected by suppression than those of non-regular words. After all, Besner et al stress, rhyme judgements with non-words must rely exclusively upon prelexical phonological recoding. On the other hand the rhyming characteristics of irregular words can only be determined by accessing the lexicon visually and retrieving information postlexically. This is the case since the grapheme-phoneme correspondence rules do not apply to irregular words. The results of their experiment show that the latency effects of concurrent articulation were confined to the word condition. The lack of any latency effect of suppression in the non-word condition suggests that prelexical phonology was not affected by it. However, the effect of suppression on error rates was common to words and non-words.

The interpretation given by Besner et al (1981) of what articulatory suppression does is at odds with that of other authors (Kleiman, 1975; Martin, 1978). They suggest that phonology used to gain access to the lexicon is not affected by suppression. However Besner et al decided to assess its effects on another task before coming to a conclusion. Subjects had to indicate whether the items in a list sounded like real words or not under silent and suppression conditions. This task requires access of the internal lexicon on the basis of purely phonological information. Although no latency effect was observed, there was an effect on errors such that performance was worse under suppression.

In order to eliminate the possibility that the rate of articulatory suppression was the important variable, a final experiment was performed using a similar task in which this rate was slowed (from the fastest possible rate to one of about 170 utterances per minute). The latency data were not supportive of the interpretation that subjects were alternating between the suppression and phonological lexical decision tasks. Since there was no effect of suppression on reaction time or errors, they concluded that suppression does not specifically prevent or impair the derivation of a prelexical phonological code.

If articulatory suppression is to be used as an experimental tool then a more precise understanding of what it does is required. Besner et al (1981) attempt this in their closing paragraphs. They compared the results of their final experiment with those of Baddeley, Thomson and Buchanan (1975) in which the suppression rate and materials were similar. It will be remembered that Baddeley et al observed that the superiority of recall of short over long words was eliminated with visually presented material under suppression. Their conclusion, that suppression prevents the translation of visual material into a phonological code is not consistent with the data from Besner et al's final experiment which required the use of a phonological code.

Besner et al (1981) give an alternative interpretation that suppression prevents the translation of print into an articulatory rather than a phonological code. Furthermore they suggest that phonological recoding which is not affected by suppression may be used sometimes for lexical access. However, they write "if the text is difficult and requires elaborate syntactic processing, or information about specific wording, then comprehension may involve the maintenance of phonological information in a short-term buffer for several seconds.....Such maintenance may be mediated by an articulatory code that is sensitive to suppression" (Besner et al, 1981, p432).

It should be clear from reading chapter 1 that conditional reasoning tasks generally lead to high error rates and could therefore be said to involve difficult material. The positioning of negatives and the relative ordering of terms in antecedent and consequent are vital factors for consideration if a correct solution is to be achieved. The fact that errors and latencies increase systematically with the introduction of negatives in the rule (eg. Evans, 1977a; Evans & Newstead, 1977) points to their increasing syntactic difficulty. Conditional reasoning problems appear to satisfy the various criteria of Besner et al (1981) which were given above. Consequently their solution may involve the maintenance of phonological information in a short-term buffer. If Besner et al are correct and this information is mediated by an articulatory code then it should follow that performance on such problems, when presented visually, would be drastically affected by articulatory suppression performed at an appropriate rate.

The decision about what is an appropriate articulation rate in suppression studies is not an easy one. On the one hand, slow articulation rates may allow subjects to intersperse an occasional subvocal rehearsal whereas, on the other hand, overly fast rates "may have the drawback that more general impairment occurs, possibly due to the involvement of the central executive" (Baddeley et al, 1984, p247). Besner et al (1981) demonstrate that 170 utterances per minute is too slow a rate to cause interference, but maximum rate is likely to be too fast in view of the argument stated above. The experiments reported in the next three chapters employ concurrent articulation at a constant rate similar to that used by Baddeley and Hitch (1974; Hitch and Baddeley, 1976), ie. between 240 and 300 utterances per minute. This should be an appropriate rate, after all the development of the Working Memory hypothesis relied on several experiments which successfully employed articultory suppression at such a rate.

Another important methodological point arises from the studies of Baddeley et al (1984). They show that, in their memory tasks using auditory materials, the effects of articulatory suppression depend upon its occurrence at presentation and recall. Other studies have shown that, with visual materials, suppression at the presentation stage alone is sufficient to engender interference (eg. Baddeley, Thomson and Buchanan, 1975). Several of the experiments described in the next three chapters will employ articulatory suppression at a constant rate of between 240 and 300 items per minute. Although visual presentation of materials is employed, whenever suppression is used, it is maintained throughout each experimental trial.

Finally, let us restate the position of Evans (1980a; 1980b). In developing the Dual Process theory of reasoning (Wason and Evans, 1975), he has allied logical perfomance on reasoning tasks with a verbal process which, he suggests, competes with a non-verbal process responsible for various non-logical response biases. If the verbal process is articulatory in nature then we might expect the use of articulatory suppression to diminish its influence on the resulting data in a conditional reasoning task. Some of the experiments contained in the next few chapters will investigate this possibility.

SECTION TWO.

EXPERIMENTS.

CHAPTER 4

EXPERIMENTS I AND II.

EXPLORING THE NATURE OF PROCESSES IN REASONING.

The experiments reported in this chapter were designed to test the Dual Process theory of reasoning (Wason and Evans, 1975) in its revised version (Evans, 1980a; 1980b). The original theory, together with Evans' revisions, has been discussed extensively in previous chapters and possible links with Paivio's (1975) theory of dual systems of thinking have been explored. Essentially, the research reported here investigates the possibility that the data observed in reasoning experiments reflect two distinct types of thought.

In brief, Evans developed the original Dual Process theory of reasoning by supposing that the logical component in reasoning data is due to a verbal (Type 2) process and the non-logical component is due to a non-verbal (Type 1) process. It is suggested that both the Type 1 and the Type 2 processes are influential prior to a subject making his response in a reasoning task. In order to test this theory it was decided to adopt a selective interference methodology. At this early stage the decision was made to concentrate upon the verbal/non-verbal distinction outlined above. Accordingly, an interference task was sought which would selectively disrupt one of these types of process.

It is recognised that the attribution of logical reasoning to a verbal process associates it with language although not necessarily with speech functions. However the existence of an important tradition implicating implicit speech as an essential ingredient in thinking (eg. Watson, 1930; Luria, 1959; Vygotsky, 1962; Sokolov, 1972) assisted in the decision to concentrate on disrupting articulatory processes. In this context the accepted interference task is that of concurrently speaking aloud irrelevant verbal material. This technique is known as articulatory suppression and much of the literature pertaining to its use in investigating the role of implicit speech in problem solving and the role of articulatory processes in short-term memory, word recognition and reading has been considered in chapter 3.

The specific procedure adopted was similar to that of Baddeley and Hitch (1974; Hitch and Baddeley, 1976) in their investigations of the Working Memory hypthesis. This is discussed in chapter 3. In investigating their hypothesis with respect to reasoning, Hitch and Baddeley (1976, experiment III) required subjects to solve a verbal reasoning problem under one of three conditions:

1) Control - no competing task,

2) Articulation - subjects said either 'the - the - the' or '1 - 2 - 3 - 4 - 5 - 6' repeatedly whilst performing the task,

3) Memory - subjects were presented with a different six digit number at the start of each trial and they were required to recite it aloud continually. This adds a short-term memory load to the concurrent articulation task.

Although the present series of experiments was motivated differently to theirs, a similar procedure to that descibed above was adopted. The particular problems used in conditional reasoning paradigms, such as we are concerned with, are much more complex than those of Hitch and Baddeley and they are known to produce systematic errors under normal conditions (see chapter 1).

Whereas Hitch and Baddeley's prime dependent measure was the latency of responding, the present study was more concerned with the nature of responses made. It was expected that the verbal interference conditions would selectively disrupt the logical component of performance relative to the non-logical component. However since Hitch and Baddeley observed interference in latency scores with their problems, a similar measure was employed in the present study. In view of the exploratory nature of the research, it was decided that no directional prediction would be made for statistical purposes.

EXPERIMENT I

The Effect of Articulatory Suppression on Conditional Inference.

The first experiment employed a set of reasoning problems in which subjects are invited to make each of four inferences, for each of four problems, thus producing sixteen distinct logical problems (as shown in Table 1.7). As we noted in chapter 1, only two of these inferences (MP and MT) are considered valid in formal logic. We should also bear in mind, however, that in natural usage the conditional sentence is sometimes used to express equivalence rather than implication. In this case all four inferences are valid.

It can be seen from Table 1.7 that each of the four inferences results in an affirmative conclusion on two rules and a negative conclusion on two rules. Previous research, which was discussed in chapter 1, has shown that, all else being equal, subjects prefer to accept negative conclusions (Evans, 1972c, 1977a; Roberge, 1978; Pollard and Evans, 1980). This has been interpreted as a non-logical response bias, and according to Evans (1980b) should be under the control of non-verbal processes. Hence, this negative 'Conclusion Bias' should not be vulnerable to the verbal interference tasks.

In the first experiment, performance of subjects on the sixteen problems was assessed under similar conditions to those of Hitch and Baddeley (1976) except that one of the Articulation Groups ('the-the-the') was dropped.

METHOD.

Design

Three experimental groups, each consisting of six male and six female subjects, were tested on an inference task using conditional rules.

In the Control group subjects were required to remain silent during the task. In the Articulation group subjects were instructed to repeat the heavily overlearnt counting sequence 'one-two-three-four-five-six' repeatedly at a rate of between four and five words per second. In the Memory group the subjects heard a spoken sequence of six random digits at the start of each trial. They were required to speak the sequence repeatedly at a rate of between four and five words per second. In this condition alone the sequence to be articulated was changed on each trial.

Within groups, subjects were required to evaluate each of the 16 types of inference shown in Table 1.7. To assess the effects of practice, 3 blocks of the 16 problems using different lexical content were constructed. A different randomised order of presentation of problems within each block was used during each session. Also the order of presentation of the 3 blocks was varied systematically within groups. Each subject thus received a total of 48 problems.

Subjects

Thirty six students at Plymouth Polytechnic, having no previous experience with this type of task, served as subjects on a paid volunteer basis. They were tested individually.

Task and Materials

Subjects were presented with the two premises of each argument, together with the appropriate valid or fallacious conclusion (cf. Table 1.7). They were required to decide whether or not the conclusion necessarily followed logically from the premises.

All of the arguments concerned shape-colour relationships. One of four shapes (Triangle, Circle, Square or Diamond) together with one of four colours (Red, Blue, Yellow or Green) were named in systematically randomised combinations for each problem.

The materials may be illustrated with the following sample problem which uses an AC inference with an AN rule:

Given:

1. IF IT IS NOT A TRIANGLE THEN IT IS NOT RED

2. IT IS NOT RED

Conclusion:

IT IS NOT A TRIANGLE

Problems were presented on a two-field tachistoscope whose onset and offset was synchronised with an automatic timer.

The subjects' task was to decide whether or not the conclusion necessarily followed logically and to signal his response by pressing a toggle switch to indicate 'YES' or 'NO'.

Procedure

Subjects in all three experimental groups were given the following written instructions initially, which were read aloud to them by the experimenter:

"Instructions

In this experiment I am interested in the ability of people to make logical inferences without the benefit of formal training. I am not, however, concerned with assessing the intelligence of individuals. Your data will be treated as confidential and reported as a component of general statistics averaged over a number of people.

You will be given a series of reasoning problems concerning imaginary coloured shapes. On each problem you will be given a rule which defines a relationship between the shape and the colour of possible figures. The rules may or may not contain negatives. For example:

IF IT IS AN OVAL THEN IT IS PINK

or IF IT IS A RECTANGLE THEN IT IS NOT ORANGE

or IF IT IS NOT A CRESCENT THEN IT IS BROWN

or IF IT IS NOT A HEXAGON THEN IT IS NOT PURPLE

For each problem you will be shown one such rule on a card, followed by a second statement relating either to the shape or to the colour of a figure which conforms to the rule.

Beneath this will be a conclusion which may or may not necessarily follow logically from the rule and the statement.

If you think the conclusion necessarily follows please press the 'YES' key, if not, press the 'NO' key.
Although you will be timed, it is more important to be accurate than fast, so please do not rush on the problems.
Have you any questions?"

Following this, instructions concerning the practice session were read aloud to the subjects. These were:
"In order to check that you understand the procedure and to familiarise you with the equipment, I will give you some practice problems first of all. These will be identical in format to the test problems except that the phrasing of the rules will be different".

Then eight practice trials using disjunctive rules, and different shape-colour combinations, but an otherwise identical format were presented to them, in randomised orders, for evaluation.

Immediately preceding the experimental trials, the following instructions, dependant upon which group subjects had been randomly assigned to, were read aloud to them:

Control Group

"Do you understand what you have to do?
We will now start on the main problems. I will give you the signal 'READY, START' as in the practice session....."

Articulation Group

"Do you understand what you have to do?
We will now start on the main problems.
I would like you to carry out an additional task whilst solving these problems. Please say aloud the sequence of numbers '1 2 3

4 5 6'. Speak them repeatedly at an even pace, like this (A demonstration was given), when I give the signal 'READY, START'. You may stop counting as soon as you have indicated your answer to the problem by pushing the key.

I cannot tell you the purpose of this procedure at this stage, but I will be happy to discuss it with you after the experiment. Do you understand what you have to do?

I will give you the signal 'READY, START' as in the practice session....."

Memory Group

"Do you understand what you have to do?

We will now start on the main problems.

I would like you to carry out an additional task whilst solving these problems. Please say aloud the sequence of numbers which I will give you after the 'READY' signal. Speak them repeatedly at an even pace, like this (A demonstration was given), when I give the signal 'READY, NUMBER'. You may stop counting as soon as you have indicated your answer to the problem by pushing the key.

I cannot tell you the purpose of this procedure at this stage, but I will be happy to discuss it with you after the experiment. Do you understand what you have to do?

I will give you the signal 'READY, NUMBER' as in the practice session....."

When the experiment was completed the subjects were debriefed. In the Articulation and Memory Groups the problems were presented immediately after subjects had commenced articulation. The experimenter was vigilant for any perceptible drop in articulation rates throughout the experimental trials, and prompted subjects to maintain the required rate as and when necessary.

In all the experiment took between thirty and forty minutes per subject.

RESULTS.

Separate analyses of variance were performed on the response frequency and latency data arising from this experiment. These will each be discussed below.

Response Frequencies

Since the known response bias on these problems relates to the polarity (affirmative/negative) of the conclusion presented, the analysis of variance was organised in terms of this factor, pooling data from different rules. Four other factors included in the analysis were Groups (3 levels), Sex, Inferences (4 levels) and Blocks (3 levels). Groups and Sex were between subjects factors and the remaining factors were within subjects. Conclusion Type and Inference are the non-logical and logical factors respectively. Since there were no significant effects including either Sex or Blocks, the percentage frequencies shown in Table 4.1 are pooled over these factors.

It was expected that logical performance might be interfered with by either Articulation or Memory load conditions. However, the main effect of Groups was not significant ($F_{2,30}=0.68$). In fact, although there was a highly significant main effect of Inference ($F_{1,30}*=37.65$, $p<0.001$), this factor did not interact significantly with Groups ($F_{2,30}=1.58$). The general pattern of acceptance rate over the four inferences was similar to previous studies of conditional inference (eg. Evans, 1977a).

Group	Conclusion	Inference				
		MP	DA	AC	MT	Mean
CONTROL	Affirmative	97	46	76	43	66
(Mean=73)	Negative	97	63	86	76	81
	Mean	97	54	81	60	
ARTICULATION	Affirmative	90	46	69	35	60
(Mean=67)	Negative	93	68	69	67	74
	Mean	92	57	69	51	
MEMORY	Affirmative	90	67	88	58	76
(Mean=76)	Negative	85	58	82	76	75
	Mean	88	63	85	67	
Mean		92	58	78	59	

Table 4.1 . The percentage of 'YES' responses (arguments accepted) in each condition of Experiment I, broken down by polarity of conclusion. Each point is based on 12 subjects. Total N=36.

The predicted non-logical bias is that negative conclusions should be accepted more than affirmative conclusions. A significant main effect of Conclusion Type confirmed this prediction ($F_{1,30}=20.76$, $p<0.001$) but this factor interacted significantly with two others.

A Conclusion Type X Inference interaction ($F_{1,30}$=10.25, $p<0.01$) is not surprising, since previous research has failed to find a 'Conclusion Bias' on the MP inference where performance is normally near to 100% correct (ie. Evans, 1977a). In this case the effect of Conclusion Type seems restricted to the DA (one-tailed t_{45}=2.47, $p<0.01$) and MT (one-tailed t_{45}=6.77, $p<0.001$) inferences, however. Much more surprising was a significant Conclusion Type X Groups interaction ($F_{2,30}$=5.58, $p<0.01$). It appears that the non-logical factor is the one affected by verbal interference. Breakdown analysis revealed that the 'Conclusion Bias' effect was significant for Control (one-tailed t_{30}=4.08, $p<0.001$) and Articulation (one-tailed t_{30}=3.90, $p<0.001$) Groups only.

* Degrees of freedom for conservative F tests (Edwards, 1967) have been used throughout the experiments reported in this thesis.

Response Latencies

Response latencies were analysed using a 4 X 4 X 3 X 2 X 3 split plot analysis of variance based on a logarithmic transformation. There were three within subjects factors: Rules, Inferences, Blocks, and two between subjects factors: Sex and Groups.

The data is summarised in Table 4.2 . Since there were no sex differences, data from male and female subjects have been combined. Also the scores are not broken down by Blocks in Table 4.2 since, although there was a significant reduction in response time with practice ($F_{1,30}$=18.51,

$p<0.001$), this did not interact with other factors.

Group	Rule	Inference				
		MP	DA	AC	MT	Mean
CONTROL	AA	6.03	8.30	5.94	7.33	6.90
(Mean=8.59)	AN	6.67	10.00	9.85	8.71	8.81
	NA	6.96	9.87	8.28	10.08	8.80
	NN	7.64	10.01	9.85	11.85	9.84
	Mean	6.83	9.55	8.48	9.49	
ARTICULATION	AA	4.23	5.70	4.29	6.08	5.07
(Mean=6.13)	AN	5.36	6.95	6.73	5.71	6.19
	NA	4.83	6.96	5.98	7.94	6.43
	NN	5.61	7.22	7.41	7.11	6.84
	Mean	5.01	6.71	6.10	6.71	
MEMORY	AA	9.39	11.70	9.37	10.26	10.18
(Mean=11.63)	AN	9.29	11.92	11.36	13.22	11.45
	NA	9.34	13.83	12.83	14.64	12.66
	NN	9.53	12.67	13.53	13.13	12.22
	Mean	9.39	12.53	11.77	12.81	
Mean		7.07	9.59	8.77	9.67	

Table 4.2 . Mean solution latencies (seconds) for each condition in Experiment I. Each mean is based on 12 subjects. Total N=36.

Both the type of rule used and the type of inference required affect the psychological complexity of the task and might be expected to affect response latency. A significant effect of Rules was observed ($F_{1,30}=14.97$, $p<0.001$), due to a tendency for response times to increase with the introduction of negative components as found previously (Evans, 1977a; Evans and Newstead, 1977). The Inference factor was also significant ($F_{1,30}=24.89$, $p<0.001$) with the marginal means showing a similar pattern to that found by Evans (1977a).

Neither of these factors, however, interacted with Groups, the respective F ratios were both less than 1. The Groups factor was itself significant ($F_{2,30}=9.66$, $p<0.001$), but in an unexpected manner. The means shown at the left of Table 4.2 suggest that while the Memory Group produced the slowest latencies as expected, the <u>fastest</u> were observed in the Articulation Group.

The nature of the Groups effect was assessed further by computing the mean logarithmic time for each subject, and comparing each pair of groups with two-tailed t tests. The Memory Group was significantly slower than Control ($t_{22}=2.27$, $p<0.05$), and the Articulation Group was significantly faster than Control ($t_{22}=2.14$, $p<0.05$). Not surprisingly the difference between Articulation and Memory was highly significant ($t_{22}=4.42$, $p<0.001$).

(The Analysis of Variance tables for Experiment I are shown in Appendix B).

DISCUSSION.

The results of Experiment I are surprising for a number of reasons. From the outset it was expected that the conditions imposing articulatory suppression would selectively interfere with the logical (Inference) factor rather than the non-logical (Conclusion Bias) factor. Thus an interaction of Groups and Inference was expected. However the pattern and extent to which inferences were accepted was not affected by the articulatory condition to which subjects were assigned but was consistent with previous studies throughout.

Owing to the ambiguity of the conditional sentence, we can only clearly classify acceptance of MP and MT inferences as correct whether an implication or equivalence interpretation of the rule is adopted. The percentage correct on these inferences combined are:

Control 79%, Articulation 71%, Memory 78%.

Since there is no significant difference between these rates of acceptance, there is little support for the hypothesis that articulatory suppression (with or without memory load) selectively affects the logical factor.

As a whole the results show a typically high incidence of logical errors. The normal systematic non-logical bias to prefer negative conclusions was observed and the expected interaction of conclusion with inference type is in line with the previous literature. However, the non-logical factor was not expected to interact with Groups although, surprisingly, such an interaction was observed with 'Conclusion Bias' completely absent in the Memory Group. Although this finding appears to contradict Evans' (1980b) suggestion that response biases are mediated by non-verbal processes, we can see that it is the presence of a memory load rather than articulation per se that appears responsible. Since the usual effects of inference type were observed in this condition this result cannot be explained by the suggestion that subjects were guessing in this Group. The finding that an increased memory load does not seem to affect a fairly complex task of distinguishing inferences but can affect a fairly low level response bias is difficult to interpret. Perhaps speculation is unwise at this point and we should wait to see if a comparable result will occur with another response bias in an alternative reasoning paradigm in the next experiment.

Let us now consider the latency data. The finding of slower responses in the Memory Group is not surprising and this result is consistent with the data of Hitch and Baddeley (1976) and their theory of Working Memory. However the finding of significantly faster responding in the Articulation Group appears to be inconsistent with their data. Hitch and Baddeley found that interference conditions significantly increased latencies and this interference increased with more complex problems which made more demands upon working memory.

The model of working memory consists essentially of a central executive plus an articulatory loop which is used for subvocal rehearsal. The effect of articulatory suppression is supposedly to prevent use of the loop. Although Baddeley (1976, 1979) has suggested that some activities (eg. skilled reading) do not involve use of the loop and are consequently not slowed down by articulatory suppression, he has not suggested conditions under which a 'speeding-up' effect should occur.

Other effects observed in the present experiment appear inconsistent with the model of Hitch and Baddeley (1976). For instance none of the manipulations affecting problem complexity (eg. presence of negatives and inference type) interacted with the Groups factor as their model predicts, since more complex problems demand more space in working memory.

One problem of interpreting latency data on these tasks arises because subjects are making variable responses. Although there was no significant effect of Groups on the frequency analysis a check was made of possible confounding. Over all problems the mean time for a 'Yes' response was 8.27 seconds, and for a 'No' response 10.95 seconds. On a significant majority of problems the 'No' response was slower ($p<0.02$, two-tailed binomial test). However, this could not explain the difference in reasoning speed observed between the Groups. The percentage of slow, 'No' responses was greatest in the Articulation Group and least in the Memory Group. Any effect of this bias, then, was exactly opposite to the Groups effect observed.

The results of Experiment I are interesting and surprising then for a number of reasons. However no detailed speculation will occur until these effects have been replicated on another conditional reasoning paradigm.

EXPERIMENT II

The Effect of Articulatory Suppression on Truth Table Evaluation with Conditionals, Employing Verbal or Pictorial Instances.

In view of the surprising nature of the results discussed above, it was decided to design a similar experiment using an alternative reasoning paradigm, of similar complexity, known as Truth Table Evaluation. A detailed review of the psychological literature pertaining to this task is given in chapter 1. This change of paradigm should permit some generalisation of the effects observed in Experiment I.

Truth table evaluation bears some resemblance to sentence-picture verification tasks which have been discussed in chapter 2. Subjects are presented with a conditional rule, together with an instance, and they are asked to evaluate the truth value of the rule. The extra complexity engendered by the conditional rule demands the use of three response categories: True, False and Irrelevant. This new paradigm also permits response times to be split into Comprehension and Verification periods as was achieved in the study of Evans and Newstead (1977).

On truth table evaluation tasks the operation of another non-logical response bias, called 'Matching Bias', can be observed. This bias relates to the inclination of subjects to regard an instance which matches the values named in the rule as relevant to the rule. That is they will perceive it as True or False. Other instances which do not match the values named in the rule tend to be regarded as Irrelevant to it.

Whilst Experiment I investigated the hypothesis of Evans (1980b) that logical responses may be mediated by a verbal thought process, in the present case it was decided to investigate an additional hypothesis also. This is that non-logical responses may be mediated by a non-verbal thought process (Evans, 1980b). It will have been noticed from chapter 2 that the main alternative to verbal processes that have been considered in cognitive psychology is a system based on a visual code (Paivio, 1971; 1975). The possible influence of visual imagery in reasoning processes was discussed previously.

A recent interpretation of 'Matching Bias', which was discussed in chapter 1, suggests that it may have a linguistic basis (Evans, 1983a). However, an interpretation of it in terms of visual imagery processes is still quite attractive particularly in view of the parallels drawn in chapter 2 between the modified Dual Process theory of Evans (1980a; 1980b) and the dual coding approach to cognition (eg. Paivio, 1971; 1975). The focus on values which are perceptually present could reveal the influence of a concrete visual system of thought, which interferes with the verbal/abstract thought required for a logical solution.

In the following experiment a strong attempt will be made to encourage the use of visual processes. According to Paivio (1971), visual processes are liable to influence processing of verbal material when it is concrete and imageable. It is considered plausible that the 'Matching' tendency is imagery related and this could be facilitated when pictorial presentation of the instance is used. Given a rule such as:

IF IT IS A TRIANGLE THEN IT IS RED,

subjects could tend to form a visual image of a red triangle, and be biased to consider only instances including one or preferably both of these features. This would also lead them to disregard the logical significance of negative components, which is a feature of the 'Matching Bias' phenomenon. Thus it was decided to manipulate the nature of the instance in an attempt to encourage the use of an imagery related process.

In a verbal condition a TT case for the above rule would be produced by printing the words 'red triangle' and in a pictorial condition by a picture of a red triangle. As usual, false values are generated by alternatives such that a TF case could be a verbal description or a picture of a yellow triangle for instance. If the presentation of instances in the pictorial mode encourages the use of a visual code, then the same increased tendency to focus on matching values might be expected.

The second experiment continues the investigation of the articulation effect and it is expected, following Experiment I, that the overt rehearsal of an overlearnt sequence will lead to faster responding.

<u>METHOD.</u>

<u>Design</u>

Sixteen subjects in each of three experiemntal Groups, were tested on a truth table evaluation task using conditional rules. Control, Articulation and Memory groups were required to perform the reasoning task concurrent with a subsidiary task as specified in Experiment I.

The Groups were further subdivided according to whether pictorial or verbal instances were to be evaluated. Each subject attempted 16 distinct logical problems (four rules X four logical cases, cf Table 1.10). The problems were presented in three randomised Blocks, each containing the sixteen logical forms, so that each subject received forty-eight problems in all. Subjects were permitted to give one of three types of response. They could decide that the instance conformed to the rule (equivalent to a truth table value of 'True'), conflicted with the rule (equivalent to a truth table value of 'False'), or else was irrelevant to the rule. In addition to the type of response made, latencies of comprehension (CT) and verification (VT) were recorded.

Subjects

Forty-eight students of Plymouth Polytechnic, having no previous experience with this type of task, served as subjects on a paid volunteer basis. They were tested individually.

Task and Materials

Subjects were presented with a conditional rule on one field of a three-field tachistoscope. When they had read and understood it, they pushed a switch causing a particular instance of a coloured shape (either pictorial or written) to be superimposed under the rule. The subjects' task was to evaluate whether the instance conformed to, conflicted with, or was irrelevant to the rule, and to signal their response by pushing the appropriate 'Conforms', 'Conflicts' or 'Irrelevant' button. The tachistoscope was coupled to automatic timers so that CT and VT could be measured.

All of the problems concerned shape-colour relationships. The materials may be illustrated with the following sample problem which uses an FT logical case (written form) with an NA rule:

Rule: IF IT IS NOT A TRIANGLE THEN IT IS GREEN

Instance: GREEN TRIANGLE

Procedure

Subjects in all Groups were given the following preliminary written instructions, which were read aloud to them by the experimenter:

"Instructions

This is an experiment to see how well people who are untrained in formal logic are able to understand sentences which define logical relationships. If you have had any training in formal logic please inform the experimenter.

The experiment is not intended as a test of your intelligence. Your data will be treated as confidential and will be reported only as a component of general statistics averaged over a number of people".

The particular logical task was then described in the following manner:

"You will be presented, on the screen, with a series of rules defining the relationship between the colour and the shape of possible figures.

The rules may or may not contain negatives. For example:

IF IT IS AN OVAL THEN IT IS PINK

or IF IT IS A RECTANGLE THEN IT IS NOT ORANGE

or IF IT IS NOT A CRESCENT THEN IT IS BROWN

or IF IT IS NOT A HEXAGON THEN IT IS NOT PURPLE

When you are satisfied that you understand each rule you should

move the toggle switch (marked 'X') to the left or to the right. This will cause a particular instance of a coloured shape to appear on the screen underneath the rule.
You have to decide whether that coloured shape conforms to, conflicts with, or is irrelevant to that rule and to indicate your decision by pushing the appropriate button firmly.
Although you will be timed it is more appropriate to be accurate than fast, so please do not rush on the problems.
Have you any questions?"

The experimenter then read instructions, concerning the practice session, to the subjects. These were identical to those used in Experiment I. Following these, eight practice problems were given. These differed from the experimental ones in linguistic form (disjunctive rather than conditional) and in the colours and shapes used.

Immediately prior to the experimental trials subjects received appropriate verbal instructions, again identical to those used in Experiment I, dependant upon which Group (Control, Articulation or Memory) they had been randomly assigned to.

In the Articulation and Memory Groups the problems were presented immediately after subjects commenced articulation. Once again, the experimenter was vigilant for any perceptible drop in articulation rate during the trials and prompted subjects to maintain the required rate as and when necessary.

When the experiment was completed subjects were debriefed. In all the experiment took about 35 - 40 minutes per subject.

RESULTS.

As with Experiment I, separate analyses of variance were performed on the response frequency and latency data arising from Experiment II. Since the typical effects of 'Matching Bias' are shown in the frequency of responding and linguistic effects (eg. presence of negaives in the rule) are shown in the latency of responding (Evans and Newstead, 1977), appropriate factors are introduced into the respective analyses of variance. These analyses are discussed below.

Response Frequencies

The modal responses to each logical case were in line with the 'defective' truth table. That is the TT cases were most often said to conform to the rule (89%), TF cases to conflict with the rule (82%), whereas FT and FF cases were most frequently described as 'Irrelevant' (56% and 60% respectively). On the basis of previous literature these responses represent the subjectively correct answers, and any interference in logical performance expected due to a competing task of a verbal nature might be expected to lower their frequency in favour of a more random use of the three available response categories.

In order to assess this, separate analyses of variance were performed on the frequency of modal responses to each logical case. The factors in each analysis were Groups, Instance, Matching Case and Blocks (3 X 2 X 4 X 3), the last two factors being within subjects. The main effect of Blocks was significant for each of the analyses (TT case - $F_{1,42}=5.66$, $p<0.025$; TF case - $F_{1,42}=6.56$, $p<0.025$; FT case $F_{1,42}=7.34$, $p<0.01$; FF case - $F_{1,42}=14.74$, $p<0.001$), showing a steady increase in subjectively correct responding with practice. Since Blocks did not interact significantly with any other factor, the percentage of correct responses for each analysis is collapsed over it in Table 4.3 .

Table 4.3 . Percentage frequency of responses to each of the four logical cases in Experiment II conforming to a defective truth table. Each point is based on 8 subjects. Total N=48.

(i) <u>TT as 'Conforms'</u>

Group	Instance	Matching Case				
		pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$	Mean
CONTROL	Verbal	100	100	88	83	93
(Mean=95)	Pictorial	100	100	96	92	97
	Mean	100	100	92	88	
ARTICULATION	Verbal	96	83	88	50	79
(Mean=85)	Pictorial	100	96	96	75	92
	Mean	98	90	92	63	
MEMORY	Verbal	100	83	88	63	83
(Mean=86)	Pictorial	96	88	96	79	90
	Mean	98	85	92	71	
Mean		99	92	92	74	
(Rule)		(AA)	(AN)	(NA)	(NN)	

(iv) TF as 'Conflicts'

Group	Instance	Matching Case				
		pq	p$\bar{q}$	$\bar{p}$q	$\bar{p}\bar{q}$	Mean
CONTROL	Verbal	96	92	83	79	91
(Mean=89)	Pictorial	100	96	88	79	91
	Mean	98	94	85	79	
ARTICULATION	Verbal	92	88	50	42	68
(Mean=79)	Pictorial	96	100	92	71	90
	Mean	94	94	71	56	
MEMORY	Verbal	79	92	63	50	71
(Mean=77)	Pictorial	83	96	75	75	82
	Mean	81	94	69	63	
Mean		91	94	75	66	
(Rule)		(AN)	(AA)	(NN)	(NA)	

(iv) FT as 'Irrelevant'

Group	Instance	Matching Case				
		pq	p$\bar{q}$	$\bar{p}$q	$\bar{p}\bar{q}$	Mean
CONTROL	Verbal	54	67	75	92	72
(Mean=67)	Pictorial	46	63	58	83	63
	Mean	50	65	67	88	
ARTICULATION	Verbal	13	21	25	58	29
(Mean=40)	Pictorial	29	38	54	83	51
	Mean	21	29	40	71	
MEMORY	Verbal	42	58	50	67	54
(Mean=59)	Pictorial	63	71	54	71	65
	Mean	52	65	52	69	
Mean		41	53	53	76	
(Rule)		(NA)	(NN)	(AA)	(AN)	

(iv) FF as 'Irrelevant'

Group	Instance	pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$	Mean
		Matching Case				
CONTROL	Verbal	54	75	75	83	72
(Mean=70)	Pictorial	54	67	67	83	68
	Mean	54	71	71	83	
ARTICULATION	Verbal	4	33	25	71	33
(Mean=44)	Pictorial	33	50	58	79	55
	Mean	19	42	42	75	
MEMORY	Verbal	46	54	67	75	60
(Mean=65)	Pictorial	63	63	83	71	70
	Mean	54	58	75	73	
Mean		42	57	63	77	
(Rule)		(NN)	(NA)	(AN)	(AA)	

Matching Case produced a significant effect in all four analyses, in each case in line with 'Matching Bias' (TT case $F_{1,42}=15.80$, $p<0.001$; TF case $F_{1,42}=16.77$, $p<0.001$; FT case $F_{1,42}=14.38$, $p<0.001$; FF case $F_{1,42}=16.81$, $p<0.001$). The effect was that determinate responses (to TT and TF) decreased with more mismatches, whilst irrelevant responses to FT and FF increased with more mismatches. In none of the analyses did the expected interaction between Instance and Matching Case occur. The extent of 'Matching Bias' was no greater for pictorial than for verbal instances.

There was evidence of interference with responding due to the competing tasks. In all four analyses Groups exerted a significant effect. In the case of TT and TF, the correct responses were less frequent in both the Articulation and the Memory Groups relative to Control (TT case $F_{2,42}=3.74$, $p<0.05$; TF case $F_{2,42}=4.82$, $p<0.025$). In the FT and FF analyses, however, the decrease in subjectively correct 'Irrelevant' responses, appears to be largely restricted to the Articulation Group (FT case $F_{2,42}=3.82$, $p<0.05$; FF case $F_{2,42}=3.53$, $p<0.05$).

Unexpectedly, the Instance factor proved to be significant on the TT ($F_{1,42}=6.21$, $p<0.025$) and TF ($F_{1,42}=11.89$, $p<0.01$) cases. More correct responding was made when the instance was presented in a pictorial rather than a verbal manner. The means were:

TT case   Pictorial 93%, Verbal 85%

TF case   Pictorial 88%, Verbal 75%.

Response Latencies

Both CT's and VT's were submitted to analyses of variance based on logarithmic transformation. In the case of CT's there was one between subject factor (Groups) and two within subject factors (Rules and Blocks). In the VT analysis there was an additional between subject factor of Instance, and an additional within subject factor of Logical Case (these last concern only what is presented after the CT period is completed).

The mean CT's are shown in Table 4.4 . Although the Blocks factor was significant ($F_{1,45}=6.88$, $p<0.025$), indicating that latency reduces with practice, it did not interact with any other factor and is not shown in the table.

Group	Rule				
	AA	AN	NA	NN	Mean
CONTROL	3.30	4.16	3.88	5.32	4.17
ARTICULATION	3.17	3.79	3.83	4.40	3.80
MEMORY	6.61	7.86	7.89	8.16	7.63
Mean	4.36	5.27	5.20	5.96	

Table 4.4 . The mean comprehension latencies (seconds) in Experiment II. Each mean is based on 16 subjects. Total N=48.

The effect of Rules was highly significant ($F_{1,45}=42.85$, $p<0.001$) showing the expected increase with the addition of negative components. This is in line with the results of Evans and Newstead (1977).

As with Experiment I, the Groups factor was significant ($F_{2,45}=3.95$, $p<0.05$) with the order of means:

Articulation < Control < Memory.

When each pair of Groups were compared with t-tests (one tailed) it was found that, whereas the Memory Group was significantly slower than Control ($t_{30}=1.79$, $p<0.05$) and Articulation ($t_{30}=2.32$, $p<0.025$), the difference between Articulation and Control fell short of significance ($t_{30}=1.05$).

There was also a significant interaction between Rules and Groups in the CT analysis ($F_{2,42}=3.86$, $p<0.05$). Breakdown analysis indicated that while Rules had a significant effect on each Group individually, only the Control Group showed the expected additive effect of negatives. In the Articulation and Memory Groups the double affirmative rule was significantly more quickly processed than the other rules, but a second negative added no significant extra difficulty to that caused by one in either component.

The mean VT's are shown in Table 4.5 . Again a significant but non-interacting Blocks factor ($F_{1,42}=58.16$, $p<0.001$) is not included. The Instance factor had no significant effects in this analysis and so, for simplicity, it also is omitted from the table.

There were significant main effects of Groups ($F_{2,42}=4.50$, $p<0.025$), Rules ($F_{1,42}=68.25$, $p<0.001$) and Logical Case ($F_{1,42}=39.46$, $p<0.001$). The last two effects again replicated the findings of Evans and Newstead (1977) showing the additive, increasing effect of negative components, and the significantly slower responding overall to the FT and FF cases. However Logical Case interacted significantly with Groups ($F_{2,42}=4.66$, $p<0.025$) and inspection of Table 4.5 reveals that the effect of Logical Case is less marked in the interference groups. Also, whereas TT and TF cases tend to be slower in the Memory group compared with Control, the trend is reversed for the FT and FF cases.

The Groups factor produced the same order of means as in the Comprehension latencies, with the Articulation Group faster than the Control Group. One-tailed t-tests indicated that the Articulation Group was significantly faster than both Control ($t_{30}=1.89$, $p<0.05$ and Memory ($t_{30}=3.33$, $p<0.005$).

Group	Rule	Logical Case				
		TT	TF	FT	FF	Mean
CONTROL	AA	1.97	2.64	4.43	2.79	2.96
(Mean=4.22)	AN	2.82	2.89	5.11	4.38	3.80
	NA	2.93	3.43	5.68	6.24	4.57
	NN	4.34	4.13	7.15	6.57	5.55
	Mean	3.02	3.27	5.59	5.00	
ARTICULATION	AA	1.73	2.53	2.97	3.04	2.57
(Mean=3.16)	AN	2.50	2.76	2.77	4.26	3.07
	NA	2.58	3.37	3.11	3.93	3.25
	NN	3.53	3.48	4.28	3.75	3.76
	Mean	2.59	3.04	3.28	3.74	
MEMORY	AA	2.43	3.17	4.07	3.50	3.29
(Mean=4.18)	AN	3.76	4.41	3.97	4.80	4.24
	NA	4.12	3.49	5.18	4.10	4.22
	NN	4.01	4.35	5.08	6.39	4.96
	Mean	3.58	3.86	4.58	4.70	
Mean		3.06	3.39	4.48	4.48	

Table 4.5 . The mean verification latencies (seconds) in Experiment II. Each mean is based on 16 subjects. Total N=48.

Finally, as in Experiment I, a check was made for a possible relationship between the type of response made and its associated (verification) latency, which was calculated for each condition. In the present case, there was no evidence of any significant relationship between them.

(The Analysis of Variance tables for Experiment II are shown in Appendix C).

DISCUSSION.

The present results are compatible with the Dual Process Theory of Reasoning (Wason and Evans, 1975) as modified by Evans (1980a; 1980b). Evans suggested that the logical component of reasoning performance was mediated by a verbal process. Disruption of this verbal component by a verbal interference task might be expected to selectively impair logical performance. If Evans' other suggestion, that non-logical 'Matching' responses are mediated by a non-verbal process, is correct then no disruption of 'Matching Bias' would be expected due to verbal interference. Both Articulation and Memory Groups did appear to disrupt logical performance on the TT and TF cases and also on the modal 'Irrelevant' responses to the FT and FF cases. However 'Matching Bias', attributed to a non-verbal process, was not affected by the presence of verbal interference tasks.

It was expected that the Instance factor (verbal vs pictorial) would influence the use of visual imagery and hence interact with the 'Matching' tendency. However, it appears that 'Matching Bias' is no more marked for pictorial instances than for verbal ones. Type of instance did have a rather surprising effect on logical performance. Subjects made significantly more correct evaluations of TT and TF cases when the instance was pictorial. It is possible, however, that the poorer performance in the verbal condition occurred for a special reason. For, whilst the shape was described in the antecedent and the colour in the consequent of the rule (eg. 'IF IT IS A TRIANGLE THEN IT IS RED'), the verbal instance, following grammatical convention, described these features in reverse order with the colour before the shape (eg. 'RED TRIANGLE'). Consequently the verbal instance was incongruent with the rule. Thus, the facilitation of conditional reasoning with pictorial rather than verbal instances could be an artefact.

The latency analyses yielded comparable results to those of Experiment I with respect to differences between Groups. In both Comprehension and Verification periods, the articulation of an overlearnt sequence led to faster responding than in the silent Control Group. The addition of a six-digit memory load caused response latencies to increase relative to Control. However, t-test comparisons revealed that the slowing effect of the Memory Load was significant only for CT, whilst the speeding up effect of articulation per se was significant only for VT.

The overall effect of Memory Group on CT are generally in line with the Working Memory hypothesis of Baddeley and Hitch (1974). The fact that the Memory Group were slower than Control in CT could, quite simply, be due to subjects rehearsing the 'novel' digits several times to commit them to memory prior to engaging in the reasoning task. The diminished effect in VT could have occurred because repetition enabled some of the information to be memorised <u>before</u> they actually attempted the reasoning task.

The facilitatory effect under articulation is of considerable interest. This result apparently conflicts with the findings of Hitch and Baddeley (1976). There were no interactions of Articulation Group with problem complexity factors which are comparable to those observed by Hitch and Baddeley. The significant interaction on VT between Groups and Logical Case was opposite to that expected according to their Working Memory hypothesis with the effect of memory load being less rather than more marked on the more complex logical cases.

The facilitatory effect of articulation was characteristic mainly of the verification period in the present experiment. It thus appears that the effect is characteristic of the reasoning process rather than of the time taken to read and understand the sentence.

However it also appears that some kind of speed/error trade-off may be occuring since the subjectively correct, defective truth table responses were significantly reduced in the Articulation Group for all four logical cases. It is somewhat surprising that a speed/error trade-off should occur since the instructions explicitly emphasised the importance of accuracy rather than speed of responding. Indeed another finding argues against this sort of interpretation. In the Memory Group interference on response frequencies was least marked for FT and FF cases although these particular cases were, if anything, associated with faster VT's than the Control Group. However, in the conditional inference study (Experiment I), the Articulation Group made most errors, with Control and Memory at a similar lower level, although these differences were not significant. Possibly the Articulation subjects are induced to speed up for some reason, with consequent deterioration in accuracy.

In the present Experiment, as in the previous one, instructional emphasis was placed upon accuracy of responding rather than speed. The latency effects could therefore be slightly difficult to interpret. The Control subjects might be inclined to spend rather longer than necessary on the task due to the instruction; "It is more important to be accurate than fast, so please do not rush on the problems". If the concurrent articulation task were found to be aversive (as Baddeley suggests in a personal communication) then the subjects might speed up to avoid it. The lesser evidence of interference in the Memory Group, who are also articulating, could arise because the memory task forces them to go slower, giving more time to consider the problem. This explanation implies that articulatory suppression per se may not inhibit logical reasoning, unless it induces subjects to spend less time on the task. The observation that subjects can speed up under certain circumstances during concurrent articulation is still damaging to the original Working Memory hypothesis.

GENERAL DISCUSSION.

In summary then the two experiments presented here have provided some support for Wason and Evans' Dual Process theory of reasoning as modified by Evans' subsequent writings (Evans, 1980a; 1980b). For instance, in Experiment II the logically correct classifications of TT and TF were significantly reduced under concurrent articulation (with and without a memory load). Also the modal 'Irrelevant' response to

FT and FF was reduced in the Articulation Group. In line with prediction the 'Matching Bias' effect, attributed to a non-verbal process, was not affected by the presence of verbal interference tasks.

The results of Experiment I appear to be less encouraging. For example Groups did not interact with Inference rate. However, if we look closer at the two inferences which are logically necessary, regardless of interpretation,a slight trend in the expected direction is observed. MP dropped from 97% Control to 92% Articulation and MT dropped from 60% to 51%. The awkward group to explain is Memory, who inexplicably showed no 'Conclusion Bias'. However, it is concurrent articulation which is essentially verbal, and the presence of this in itself (without memory load) does not alter 'Conclusion Bias', while it is associated with slightly more logical errors.

In Experiment II, subjects were given either a verbal description of the instance to be evaluated, or a pictorial display of it. Although it was expected that the pictorial instance might have encouraged the use of visual imagery and leadto an increase in the 'Matching' tendency, no such effect was observed. Surprisingly, significantly more responding in accordance with logic was found with pictorial presentations. However, there was a possible artefactual explanation - the incongruency of feature order in the rule and the instance - which could have led to this result. Therefore further discussion of the effect is considered inappropriate until this deficiency has been eliminated. In chapter 6, Experiments VI and VII will follow up this curious result.

The present results also appear to present some problems for the theory of Working Memory proposed by Baddeley and Hitch (1974; Hitch and Baddeley, 1976). The faster latencies of responding under concurrent articulation (without a memory load) relative to silent controls is anomalous. Furthermore the lack of interaction between the presence or absence of a memory load and degree of problem complexity is incompatible with their model. The next chapter contains three experiments (III - V) which follow up this particular issue.

CHAPTER 5

EXPERIMENTS III TO V.

FOLLOWING UP THE CURIOUS EFFECTS OF SUPPRESSION.

The Working Memory framework has been useful in explaining data collected from many experimental tasks and, subsequently, a number of testable hypotheses have emerged. The involvement of a Working Memory system in verbal reasoning has been discussed in some detail by Hitch and Baddeley (1976) and much of chapter 3 has focussed on their work.

Using a relatively simple sentence verification task, Hitch and Baddeley (1976) found small but reliable effects of phonological similarity. When concurrent articulation of an overlearnt sequence was required, verification accuracy was not affected although the tendency for latencies to increase was marginally significant. Nevertheless, Hammerton (1969) has shown a significant decrement due to articulation of another overlearnt sequence ('Mary had a little lamb') repeatedly whilst solving similar problems. Articulation with a memory load significantly increased latencies whilst not affecting accuracy. The effects which they observed interacted with sentence complexity. This was mainly due to the imposition of a concurrent memory load.

Their data were argued to be compatible with the idea of a Working Memory system with a central executive of limited capacity interacting with a peripheral articulatory loop which plays a relatively minor role in verbal reasoning. The interactive effects of concurrent memory load with sentence complexity factors were explained by the suggestion that "the STM and verification tasks appear to compete for a pool of limited capacity since it is reasonable to assume that the complicated grammatical transformations take up extra capacity while the STM task occupies a roughly constant amount" (Hitch and Baddeley, 1976, p616).

In his review of Working Memory and reading, Baddeley (1979, p367) writes that whilst "the articulatory loop is not essential for normal reading and comprehension", there are certain conditions under which it is employed. Probably included amongst these are tasks involving phonological comparisons, which he notes, citing Kleiman (1975), are slowed under suppression. Also, Baddeley (1979, p.355) suggests that the articulatory loop would be utilised for tasks "where strict word order is crucial to comprehension". In a subsequent paper, Baddeley and Lewis (1981, p.127) write that, in the comprehension of prose, subvocalisation does allow the subject "to process complex material more accurately".

These conclusions are strengthened by Baddeley, Eldridge and Lewis's (1981) evidence of interference in the accuracy (rather than latency) of judgements of anomalous word order in sentences under articulatory suppression. Both Besner, Davies and Daniels (1981) and Baddeley et al suggest that what suppression suppresses is an <u>articulatory</u> rather than a phonological code, which primarily affects accuracy, rather than latency of responses (see Chapter 3).

In view of these ideas, it is unfortunate that previous attempts to investigate the role of suppression in verbal reasoning have employed sentence verification tasks (eg.

Hitch and Baddeley, 1976) in which errors are few, and latency the prime measure of interest. However, conditional reasoning tasks are sensitive to the effect of experimental manipulations on both accuracy and latency measures. Considering the above, the results obtained with the conditional inference task used in Experiment I are surprising. Although the task involved in that experiment was of a basically similar design to the sentence verification task used by Baddeley and Hitch (1974) and Hitch and Baddeley (1976), little support for their ideas was found. In Experiment II, whilst a significant increase in errors was found under suppression with and without a memory load, it is argued that the data were incompatible with those of Hitch and Baddeley in other respects which will be detailed below.

Generally, conditional reasoning tasks involve more difficult material than Baddeley and Hitch's task; as is indicated by their high error rates (see Chapter 1). In conditional reasoning tasks word order is crucial, the positions of antecedent and consequent and the placement of negatives are essential features for consideration if a logically correct solution is to be achieved. All in all, on conditional reasoning tasks we should expect to find substantial interference due to concurrent articulation of irrelevant material if Baddeley's contentions are correct.

In fact no significant impairment of conditional reasoning performance was found under articulation in Experiment I. Furthermore no interaction between the presence or absence of a memory load and the degree of problem complexity emerged from that experiment. Curiously, subjects performed significantly <u>faster</u> under concurrent articulation (without a memory load) than in a silent control group. This 'speeding-up' effect was also observed in Experiment II, but a significant difference between Articulation and Control Groups was only achieved with verification latencies on that occasion. Although there was some disruption of reasoning performance due to articulation in Experiment II, one might have expected the deficit to be of greater magnitude in view of Baddeley's (1979; Baddeley and Lewis, 1981) claims.

As mentioned above, the Working Memory hypothesis predicts an interaction between articulation group and problem complexity factors, since both tasks are assumed to impose a competing load on a limited capacity central executive. However, the only interactions between these two factors, on latency or accuracy, in either of these studies occurred in Experiment II. A significant interaction between Groups and Rules was apparent from the CT analysis. If due to complexity, then according to Hitch and Baddeley (1976), the additive effect of negatives (see Chapter 1 and Evans and Newstead, 1977) should be <u>more</u> marked under interference conditions, particularly when a memory load is imposed. In fact the interaction arose because the addition of a second negative into the rules produced no extra difficulty for the Articulation and Memory Groups, although it had led to increased comprehension latency for the Control Group. In the VT analysis Logical Case and Groups interacted significantly but, on this occasion, the effect of problem complexity was reduced for both articulatory groups relative to Control. Furthermore the imposition of a concurrent memory load lead, if anything, to faster responding than Control for the most complex FT and FF cases. These are <u>not</u> the sort of interactions that the Working Memory model predicts.

In view of the surprising results summarised above, it was decided to investigate this effect more closely. In Experiment III a specific manipulation was introduced to test a possible interpretation of the finding of accelerated solution times under concurrent articulation. As Baddeley and Lewis (1981, p109) have noted "the early stages of reading were and are associated with speaking, with reading aloud preceding silent reading". It is suggested (eg. Mattingly, 1972, p135) that comparatively few "high speed readers are somehow able to go directly to a deep level of language, omitting the intermediate stages of processing to which other readers and all listeners must presumably have recourse". It seems plausible that an habitual dependence upon an intermediate phonological process would recur with complex verbal reasoning tasks. If concurrent articulation interferes with the use of this phonological code in the manner proposed by Kleiman (1975), then subjects would have to operate in an alternative code such as a visual or semantic one. Such a code may permit <u>faster</u> processing of the material than the repressed phonological one. On the other hand, if the accuracy of rhyme judgements is impaired by suppression as the Besner et al (1981) study suggests, then this might be expected to increase the frequency of logical errors observed in this condition.

In order to test this hypothesis, the truth table evaluation task used in Experiment II was modified to require attention to different characteristics of the stimuli, which were words. In one condition the sound of the words was relevant, in another the meaning, and in a third condition the colour of the ink in which they were printed.

EXPERIMENT III

The Effect of Articulatory Suppression and Coding Requirements on Truth Table Evaluation with Conditionals.

METHOD.

Design

Three concurrent load Groups with twelve subjects in each were tested on a truth table evaluation task using conditional rules. As in the previous experiments Control, Articulation and Memory Groups were required to perform the conditional reasoning task with a concurrent task imposed as detailed for Experiment I.

As in Experiment II, within Groups subjects attempted 16 distinct logical problems (see Table 1.10). In the present case, however, these were presented in each of three blocks corresponding to three experimental Conditions. The nature of the problem materials within each block was such as to influence the type of code used to store and compare the rules and instances. Thus each Condition required subjects to concentrate upon a particular Visual, Rhyming or Semantic aspect of the problem materials in order to make their decisions. The six different possible orders of Conditions were used twice within each experimental Group. A different randomised order of presentation of problems within each Condition was used for each presentation.

Subjects

Thirty-six undergraduate students of Plymouth Polytechnic, having no previous experience of this type of task, served as subjects on a paid volunteer basis. They were tested individually.

Task and Materials

Subjects were presented with a conditional rule on one field of a three-field tachistoscope. The rule described the relationship between a pair of words. They referred either to the colours of the inks in which the words were written (Visual Condition), the sounds of the words (Rhyming Condition) or the semantic categories to which the words belonged (Semantic Condition). Examples of each Task Modality Condition are given below:

Visual Condition

IF THE LEFT WORD IS COLOURED YELLOW
THEN THE RIGHT WORD IS COLOURED GREEN

Rhyme Condition

IF THE LEFT WORD RHYMES WITH GLOW
THEN THE RIGHT WORD RHYMES WITH VOTE

Semantic Condition

IF THE LEFT WORD IS A PART OF THE BODY
THEN THE RIGHT WORD IS A TYPE OF ANIMAL

When subjects had read and understood the rule, they pushed a switch causing a particular instance of a pair of words (printed in coloured inks) to be superimposed beneath it. One word appeared to the left and the other word to the right of the screen.

The subjects' task was, as in Experiment II, to evaluate whether the instance conformed to, conflicted with, or was irrelevant to the rule and to signal their response by pushing an appropriate button.

The ink colour, rhyming characteristic and semantic categories of the words were manipulated to generate the various logical cases.

Sixteen different words were used in the instances in various paired combinations. Each word was selected from one of four semantic categories and it also fitted into one of four (orthographically dissimilar) rhyming categories. Each of these words was printed consistently in one of four coloured inks throughout the Experiment. The stimulus words are shown in Appendix D.

As in Experiment II, the tachistoscope was connected to automatic timers so that latencies of response, split into Comprehension Time (CT) and Verification Time (VT), could be measured. Of course, the particular responses made to each problem were recorded.

Procedure

Subjects in all conditions were read the standard set of preliminary instructions as used in Experiment II. They were told that the experiment was concerned with people's ability to understand logical relationships but was not intended as a test of intelligence.

The particular logical task was then described in the following manner:

"You will be presented, on the screen, with a series of rules defining the relationship which exists between various pairs of words. The rules may or may not contain negatives. For example:

IF THE LEFT WORD IS 'BALL' THEN THE RIGHT WORD IS 'LEAF', or

IF THE LEFT WORD IS 'TIN' THEN THE RIGHT WORD IS NOT 'BAG', or

IF THE LEFT WORD IS NOT 'HOUSE' THEN RIGHT WORD IS 'POLE', or

IF THE LEFT WORD IS NOT 'SOUP' THEN THE RIGHT WORD IS NOT 'ANT'

When you are satisfied that you understand each rule you should move the toggle switch (marked 'X' in front of you) to the left or to the right. This will cause a particular instance of a pair of words to appear on the screen underneath the rule. One word will appear on the left and one word on the right of the screen. You have to decide whether that instance conforms to, conflicts with, or is irrelevant to the given rule and to indicate your answer by pushing the appropriate button firmly.

On different occasions I will require you to attend to either:

1) the colours of the words in the instance,

or 2) the rhyming characteristics of the words in the instance,

or 3) the meanings of the words in the instance.

Although you will be timed, it is more important to be accurate than fast, so please do not rush on the problems.

Have you any questions?"

Following this the subjects were given instructions relating to the practice session. This session comprised of a set of 12 practice trials, in three blocks of 4 problems each. Each block concentrated upon either the Visual, Rhyme or Semantic modality. The orders of presentation of the three practice blocks was consistent with the order of the experimental Conditions for all subjects. The practice problems used disjunctive rules and different materials from those used in the experimental trials.

Immediately prior to the experimental trials subjects were given instructions (as in Experiments I and II) depending upon which concurrent load Group (Control, Articulation or Memory) they had been randomly assigned to. As previously, the experimenter ensured that the articulation rate was held constant during the trials.

When the Experiment was completed subjects were debriefed. In all, the Experiment took between thirty and forty minutes per subject.

RESULTS.

Response frequencies and latencies are discussed separately in the sections below. As with Experiment II, appropriate factors are introduced into the frequency and latency analyses of variance to investigate the effects of 'Matching Bias' and linguistic features respectively.

Response Frequencies

The modal responses for the four logical cases are shown in Table 5.1 . These subjectively 'correct' answers are in line with the truth table for defective implication (see Chapter 1) as were those observed in Experiment II.

It was expected that the concurrent articulation task performed whilst S's were evaluating the problems would lower the frequency of the subjectively 'correct' answers. This disruption should be particularly obvious in the Rhyme modality where the articulation of an irrelevant sequence of digits might be expected to disrupt the formation or use of a phonological code involved in making rhyme judgements.

Separate analyses of variance were performed on the modal responses to each logical case. The factors were Groups

(Control, Articulation and Memory), Conditions (Visual, Rhyme and Semantic) and Matching Case. The latter two factors were within subjects.

Table 5.1 . Percentage frequency of responses to each of the four logical cases in Experiment III conforming to a defective truth table. Each point is based on 12 subjects. Total N=36.

(i) _TT as 'Conforms'_

Group	Condition	Matching Case				
		pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$	Mean
CONTROL	Visual	92	92	75	67	81
(Mean=86)	Rhyme	100	92	83	83	90
	Semantic	92	92	83	83	88
	Mean	95	92	80	78	
ARTICULATION	Visual	100	100	100	100	100
(Mean=90)	Rhyme	83	83	100	75	85
	Semantic	100	92	75	67	83
	Mean	94	92	92	81	
MEMORY	Visual	83	83	75	67	77
(Mean=81)	Rhyme	83	100	83	67	83
	Semantic	100	100	83	50	83
	Mean	89	94	80	61	
Mean		93	93	84	73	
(Rule)		(AA)	(AN)	(NA)	(NN)	

(ii) _TF as 'Conflicts'_

Group	Condition	pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$	Mean
		Matching Case				
CONTROL	Visual	83	83	83	75	81
(Mean=80)	Rhyme	83	92	67	75	79
	Semantic	83	75	75	83	79
	Mean	83	83	75	78	
ARTICULATION	Visual	100	100	83	58	85
(Mean=81)	Rhyme	67	92	67	50	69
	Semantic	83	100	75	100	90
	Mean	83	97	75	69	
MEMORY	Visual	83	92	75	67	79
(Mean=78)	Rhyme	75	100	67	83	81
	Semantic	83	67	75	67	73
	Mean	80	86	72	72	
Mean		82	89	74	73	
(Rule)		(AN)	(AA)	(NN)	(NA)	

(iii) FT as 'Irrelevant'

Group	Condition	Matching Case				
		pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$	Mean
CONTROL	Visual	17	47	25	58	35
(Mean=35)	Rhyme	25	25	17	50	29
	Semantic	25	33	50	50	40
	Mean	22	35	31	53	
ARTICULATION	Visual	17	25	17	50	27
(Mean=29)	Rhyme	25	8	25	50	27
	Semantic	8	42	25	58	33
	Mean	17	25	22	53	
MEMORY	Visual	42	75	67	58	60
(Mean=52)	Rhyme	33	50	42	50	44
	Semantic	33	58	58	58	52
	Mean	36	61	56	55	
Mean		25	40	36	54	
(Rule)		(NA)	(NN)	(AA)	(AN)	

(iv) FF as 'Irrelevant'

Group	Condition	Matching Case				
		pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$	Mean
CONTROL	Visual	33	42	50	58	46
(Mean=38)	Rhyme	42	42	33	33	38
	Semantic	25	25	25	50	31
	Mean	33	36	36	47	
ARTICULATION	Visual	33	42	42	67	46
(Mean=37)	Rhyme	25	25	33	42	31
	Semantic	8	42	33	50	33
	Mean	22	36	36	53	
MEMORY	Visual	50	50	50	83	58
(Mean=57)	Rhyme	50	33	67	75	56
	Semantic	42	67	58	67	58
	Mean	47	50	58	75	
Mean		34	41	44	58	
(Rule)		(NN)	(NA)	(AN)	(AA)	

Contrary to expectations neither Groups nor Conditions produced a significant main effect and neither did those factors interact in any of the four analyses. Thus concurrent articulation did not interfere with subjects' conditional reasoning performance even when they were required to attend to phonological (rhyming) characteristics of the words. The only significant effects in the frequency analyses were manifestations of the 'Matching Bias' effect. Determinate responses - on TT and TF - decreased as mismatches in the instance increased, while irrelevant responses - on FT and FF - showed an increasing trend. The effect was significant for TT ($F1,33=8.59$, $p<0.01$), FT ($F1,33=8.78$, $p<0.01$) and FF ($F1,33=5.90$, $p<0.025$), but fell short of significance in the TF analysis ($F1,33=2.93$). In none of the analyses, however, did the effect of 'Matching Bias' interact with Conditions or any other factor.

In essence, then, reasoning responses were in accordance with previous research (see, for example, Evans and Newstead, 1977) <u>regardless</u> of the presence of concurrent articulation or the manipulation of type of encoding required to solve the problem.

<u>Response Latencies</u>

The mean Comprehension Times are shown in Table 5.2. A logarithmic transformation was applied and the data submitted to a 3 X 3 X 4 analysis of variance, with the factors Groups, Conditions and Rules.

There was a significant effect of Rules ($F1,33=4.95$, $p<0.05$) and a significant interaction of Rules with Groups ($F2,32=3.57$, $p<0.05$). However the nature of this interaction was opposite to that observed by Hitch and Baddeley (1976). Whilst, for the Control Group an increase in latency occurred with the addition of negative components in the rules (in line with Evans and Newstead, 1977), breakdown analysis indicated that the introduction of a second negative added no extra difficulty in the Articulation Group. In the Memory Group, however, the effect of negatives disappeared completely.

Group	Rule	Condition			
		Visual	Rhyme	Semantic	Mean
CONTROL	AA	4.55	5.33	5.30	5.06
(Mean=5.52)	AN	4.65	5.65	5.29	5.20
	NA	4.90	6.08	6.37	5.78
	NN	5.45	7.13	5.56	6.05
	Mean	4.89	6.05	5.63	
ARTICULATION	AA	3.71	4.54	5.12	4.46
(Mean=5.01)	AN	4.27	5.74	5.23	5.08
	NA	4.55	6.36	5.63	5.51
	NN	4.23	5.42	5.29	4.98
	Mean	4.19	5.52	5.32	
MEMORY	AA	5.95	6.97	6.68	6.53
(Mean=6.52)	AN	5.92	7.21	6.04	6.39
	NA	5.89	7.20	6.33	6.47
	NN	5.64	7.62	6.84	6.70
	Mean	5.85	7.25	6.47	
Mean		4.98	6.27	5.81	

Table 5.2 . Mean Comprehension latencies (seconds) for each condition in Experiment III. Each mean is based on 12 subjects. Total N=36.

There was also a significant main effect of Conditions ($F_{1,33}=11.59$, $p<0.01$) which seems mainly due to faster times in the Visual Condition.

The effect of Groups was not significant ($F_{2,33}=1.96$), but the order of means was the same as in Experiments I and II: Articulation faster than Control faster than Memory. Groups did not interact with Conditions ($F_{2,33}=0.25$). Indeed the order of means across the three Concurrent Load Groups was the same for Visual, Rhyme and Semantic Conditions (see Table 5.2).

Verification Times were also submitted to analysis of variance on logarithmic transformation. The factors were the same as in CT's with the addition of Logical Case as a within subjects factor on four levels. This factor refers to what is presented in the instance when the verification period is begun. Since there were no significant interactions the presentation of means is simplified in Table 5.3 .

The means for the expected interaction between Groups and Conditions is shown in Table 5.3 (i). The interaction was not significant ($F_{2,33}=2.01$), and the general pattern of results was similar to that found with CT's. Conditions was significant ($F_{1,33}=13.24$, $p<0.01$) with latencies again fastest in the Visual Condition. Although Groups was not significant ($F_{2,33}=0.88$), Articulation was, once again, associated with faster latencies than Control or Memory for all Conditions. The Memory Group was also associated with faster latencies than Control for the Visual and Semantic though not for the Rhyme Condition.

Table 5.3 . The mean Verification latencies (seconds) in Experiment III. Total N=36.

i) Groups X Conditions (Each mean is based on 12 subjects).

Group	Condition			
	Visual	Rhyme	Semantic	Mean
Control	5.18	5.73	5.86	5.59
Articulation	3.90	4.74	4.29	4.31
Memory	4.81	6.80	4.93	5.51
Mean	4.63	5.75	5.03	

ii) Logical Case (Each mean is based on 36 subjects).

TT	TF	FT	FF	Mean
3.97	4.74	5.72	6.12	5.14

iii) Rules (Each mean is based on 36 subjects).

AA	AN	NA	NN	Mean
4.34	4.79	5.60	5.83	5.14

There was a significant main effect of Logical Case ($F_{1,33}=19.13$, $p<0.001$) and of Rules ($F_{1,33}=22.72$, $p<0.001$). Inspection of the means in Table 5.3 (ii) and (iii) reveals that these effects were as expected and are consistent with previous research (eg. Evans and Newstead, 1977).

Finally, the mean VT's associated with making each type of response (Conforms, Conflicts, Irrelevant) were calculated for each Condition. There was no evidence of any significant relation between response type and latency.

(The Analysis of Variance tables for Experiment III are shown in Appendix E).

DISCUSSION.

It was very surprising that Groups did not interact significantly with Conditions in any of the analyses since, following Kleiman's (1975) findings, the concurrent articulation of an irrelevant sequence was expected to disrupt the Rhyme Condition particularly when compared with the Visual Condition in the present Experiment.

Although there was not a main effect of Groups on comprehension or verification latencies, their order followed a similar pattern to that observed in the previous two experiments in the CT analysis. In the VT analysis a similar pattern was observed at least as far as the Articulation and Control Groups were concerned. Irrelevant concurrent articulation without a memory load was associated with faster responding than the silent Control Group. As with Experiment II, on VT the Memory Group also tended to be faster overall than Controls.

The interaction of Rules with Groups on the CT analysis yielded a result somewhat similar to that observed in Experiment II. In the present case, increasing the complexity of the rule by the addition of negatives had the expected additive effect in the Control Group (see Evans and Newstead, 1977). In the Articulation Group, the addition of a second negative component did not lead to increased latencies. However, in the Memory Group, no effect of negatives was observed on latencies. This is inconsistent with the Baddeley and Hitch (1974) model of Working Memory.

Baddeley and Hitch's original attempts to demonstrate effects of Working Memory on sentence verification used a memory preload technique. However, in this case, it was argued that subjects were able to carry out the two tasks in alternation and so only very weak evidence in favour of the Working Memory hypothesis was gained.

They subsequently adopted the interference methodology (which has been utilised in most of the present Experiments) because it appears to make concurrent demands upon the memory system. Therefore the technique provides a more adequate test of the Working Memory hypothesis. In assessing their model, Hitch and Baddeley (1976, p613) claim that the crucial test is "whether memory load produces interference over and above that which is attributable to articulatory activity alone". Clearly considering the error data from the present Experiment, it does not.

However, besides the obvious differences between the natures of the conditional reasoning task employed in the present case and the sentence verification task (Hitch and Baddeley, 1976, experiment III), there were other differences in procedures adopted. Although the subjects' rates of articulation were kept constant at about 4 - 5 items per second in both experiments, the instructional emphasis in the present case was on accuracy rather than speed of response. Hitch and Baddeley (1976) gave more emphasis on the importance of speed. The present procedure also differed from that of Hitch and Baddeley in that separate subjects experienced the Control, Articulation and Memory Groups.

It was therefore decided to attempt a replication of Hitch and Baddeley's (1976) third experiment whilst maintaining the present instructional emphasis on accuracy and a between group design for the articulation factor. In this way it should be possible to tell whether the discrepancies between the results of Experiments I - III and Hitch and Baddeley (1976) are due to differences in the nature of the reasoning task or to differences in the experimental instructions and procedure. In addition, their problem materials were to be extended by including either words or letters in the instances.

EXPERIMENT IV

A Replication and Extension of Hitch and Baddeley (1976, experiment III).

It is possible that the discrepancies between the results of the previous experiments and those of Hitch and Baddeley (1976, experiment III) were due to differences in the nature of the tasks, or to other factors such as population differences. Newstead (1979) has illustrated how established experimental results are sometimes difficult to replicate with other subject populations.

In Experiment IV, Hitch and Baddeley's sentence verification task will be repeated whilst keeping as close as possible to the procedure employed in the previous experiments.

Consequently, instructions will be kept as similar as possible, and a between rather than a within Group design will be maintained. It is also of interest to determine whether Hitch and Baddeley's results will extend to words as well as letters in the instances and to tachistoscopic presentation.

METHOD.

Design

The verification task used in the present experiment was similar to that used by Hitch and Baddeley (1976, experiment III).

A sentence and an accompanying instance of a letter-pair (or word-pair) were presented simultaneously. The sentence described a particular order of the letters (or words) named in the instance. The subject's task was to indicate whether the letters (or words) named in the instance were in the order described by the rule. For example, when presented with:

A follows B

BA

the correct response would be 'True'. If the instance had been 'AB', the correct response would have been 'False'.

Sentence difficulty was manipulated by varying the polarity and voice of the verb. Also both true and false instances were used with each sentence type. An equivalent set of problems were constructed using the verb 'precedes' instead of 'follows'. Thus a block of sixteen problems concerning letter-pairs and a similar block of sixteen problems concerning word-pairs were devised. The order of presentation of blocks was counterbalanced. A different randomised order of problems within each block was used for each presentation.

Eight subjects were run in each of three experimental Groups (Control, Articulation and Memory). The response made was recorded together with the response latency for each problem.

Subjects

Twenty-four students of Plymouth Polytechnic, having no previous experience with this particular task, served as subjects on a paid volunteer basis. They were tested individually.

Task and Materials

Subjects were presented with the verification task on a two-field tachistoscope whose onset and offset was synchronised with an automatic timer. They were required to decide whether the sentence described the instance correctly or not, and to respond 'True' or 'False' as appropriate. Subjects were presented with two separate blocks of sixteen problems each.

The particular letter-pairs (or word-pairs) used (shown in Appendix F) were different for each problem. None of the combinations of letters (or words) used within any problem rhymed, and they were selected so as to be visually dissimilar. The combinations selected did not appear to have any obvious semantic connotations.

Procedure

Subjects in all Groups were given preliminary written instructions similar to those used in previous experiments. The particular task which they were to perform was described in the following manner:

"You will be presented, on the screen, with a series of rules defining the relationship which exists between various pairs of letters (or words).

The rules may or may not contain negatives.

For example:

P follows W,

or W does not follow P,

or W is followed by P,

or P is not followed by W.

For each problem you will be shown one such rule on a card. Underneath this will be an instance containing the particular letters (or words) mentioned in the rule.

For example:

WP or PW

Your task is to read each sentence and to decide whether it is a true or false description of the instance given.

If you think that the rule describes the instance correctly press the switch towards 'TRUE'. If you think that the rule does not describe the instance correctly press the switch towards 'FALSE'.

Although you will be timed, it is more important to be accurate than fast so please do not rush on the problems.

Have you any questions?"

Subjects were then given eight practice problems of a fairly similar nature (using the relationships 'above' and 'below') in order to familiarise them with the problem format and response keys.

Immediately prior to the experimental trials the three Groups received their separate instructions as in the previous experiments. Of course, the experimenter ensured that the articulation rates remained constant at about 4 - 5 items per second throughout the trials.

When the experiment was completed, subjects were debriefed. In all the experiment took about 25 minutes per subject.

RESULTS.

Response latencies were logarithmically transformed and then submitted to a 3 X 2 X 2 X 2 X 2 analysis of variance with repeated measures on the last four factors. The factors were Groups, Letters/Words, Truth Value, Sentence Voice and Polarity. The error data were submitted to a similar analysis.

Response Latencies

The mean solution latencies are shown in Table 5.4 . Since there was no significant main effect or interaction involving Letters/Words, the table is collapsed over this factor.

There was a significant main effect of Groups ($F_{2,21}=4.73$, $p<0.025$). The order of means was the same as that observed by Hitch and Baddeley (ie. Control faster than Articulation faster than Memory). Two-tailed t tests revealed a significant difference between Memory and Control ($t_{14}=2.78$, $p<0.025$) Groups only, although the difference between Articulation and Memory approached significance ($t_{14}=1.80$, $p<0.10$).

The variables of problem structure produced the effects on problem difficulty expected from previous research.

Thus negatives were processed more slowly than affirmatives ($F_{1,21}=158.07$, $p<0.001$) and passives more slowly than actives ($F_{1,21}=16.82$, $p<0.001$). There was an interaction between Truth Value and Polarity ($F_{1,21}=12.35$, $p<0.01$) indicating that problems containing negatives but requiring a 'False' answer were easier than those equiring a 'True' answer with the reverse holding for affirmatives. This finding is similar to that of Hitch and Baddeley and is typical of sentence-referent matching tasks (eg. Trabasso, 1972). There was also a significant interaction between Sentence Voice and Polarity ($F_{1,21}=10.75$, $p<0.01$).

Group	True				False			
	Affirmative		Negative		Affirmative		Negative	
	Act.	Pass.	Act.	Pass.	Act.	Pass.	Act.	Pass.
Control (Mean=3.34)	2.60	2.99	3.92	3.84	3.00	3.15	3.60	3.59
Articulation (Mean=3.93)	2.56	3.07	5.01	4.87	2.89	3.63	4.48	4.92
Memory (Mean=5.12)	3.52	4.17	5.80	6.11	3.64	5.32	6.34	6.06
Mean	2.89	3.41	4.91	4.94	3.18	4.03	4.81	4.86

Table 5.4 . Mean solution latencies (seconds) for Experiment IV broken down by Group, Truth Value, Polarity and Sentence Voice. Each mean is based on 8 subjects. Total N=24.

According to the Working Memory hypothesis, variables affecting problem difficulty should tend to interact with Working Memory load. Although no interactions in support of this hypothesis have occurred in the conditional reasoning problems of Experiments I to III, there was one such significant interaction in the present experiment. This was between Groups and Polarity ($F2,21=6.18$, $p<0.01$). The presence of a memory load increased the relative difficulty of negative statements.

A four-way interaction of Groups with Truth Value, Sentence Voice and Polarity ($F2,21=3.77$, $p<0.05$) can probably be dismissed as spurious in view of the large number of F ratios in the analysis.

Response Frequencies

The error data were submitted to analysis of variance with tha same factors as previously. The percentage means are shown in Table 5.5 with the cells, once again, collapsed over Letters/Words since this did not produce a significant main effect or interaction in the analysis.

As with Hitch and Baddeley's analysis, there was no significant main effect of Groups ($F2,21=2.27$) on error frequencies, although the magnitude of the F ratio was comparable with theirs. There was a main effect of Sentence Voice ($F1,21=9.33$, $p<0.01$) and this factor interacted with Truth Value ($F1,21=4.65$, $p<.05$).

There was also a main effect of Polarity ($F1,21=18.84$, $p<0.001$) together with a significant interaction of Polarity with Truth Value ($F1,21=7.66$, $p<0.025$). This latter result echoed that found in the latency analysis.

Group	True				False			
	Affirmative		Negative		Affirmative		Negative	
	Act.	Pass.	Act.	Pass.	Act.	Pass.	Act.	Pass.
Control (Mean=17)	3	13	22	44	9	13	9	19
Articulation (Mean=13)	3	6	25	25	3	13	19	13
Memory (Mean=25)	6	19	19	41	28	34	25	25
Mean	4	13	22	37	13	20	18	19

Table 5.5 . Error percentages for Experiment IV broken down by Group, Truth Value, Polarity and Sentence Voice. Each point is based on 8 subjects. Total N=24.

As with Hitch and Baddeley's data, the correlation between errors and response latencies was positive and significant (rho=0.83, t_{10}=4.65, p<0.001) when computed across the twelve conditions (Groups X Polarity X Voice). This suggests that speed/error trade-off was not responsible for changes in response latency.

(The Analysis of Variance tables for Experiment IV are shown in Appendix G).

DISCUSSION.

As can be seen, these results replicate those of Hitch and Baddeley (1976, experiment III) in essential respects.

In addition the present results establish that similar effects accrue from using either word-pairs or letter-pairs as stimuli even when presented tachistoscopically. More importantly, it is evident that when using this rather simple sentence verification task with instructions and conditions of presentation similar to Experiments I to III, concurrent articulation without a memory load does not speed up solution latencies. The general order of latency means observed in Experiment IV was Control faster than Articulation faster than Memory.

Thus it seems that the 'speeding-up' effects of concurrent articulation without a memory load, observed in the conditional reasoning experiments, cannot be due to the instructional emphasis on accuracy. Nor can they be attributed to differences between the populations sampled by the present author and by Hitch and Baddeley (1976), or to the use of a between rather than within subject design for the articulatory (Groups) factor. It appears that the difference lies in the reasoning task itself.

It should therefore follow that, if a sentence verification task is performed in which a large load is placed on the articulatory system, a marked disruptive effect of concurrent articulation of overlearnt material would be observed. Surprisingly, such an effect was not observed with the conditional reasoning task requiring rhyme judgements, and presumably phonological recoding, described in the previous experiment.

In Experiment V subjects were run under three conditions using the 'words' version of the Hitch and Baddeley reasoning task with modifications to require attention to visual, semantic or phonological characteristics (rhyme judgements). A fourth condition, with no particular code required, was also run as a control. In the last condition the results should replicate those of Hitch and Baddeley (1976, experiment III) and of Experiment IV. The 'coding' conditions permit a further test of the hypothesis investigated in EXperiment III, namely that the presence and absence of concurrent articulation should interact with the encoding required by the reasoning task. The failure to observe this interaction in Experiment III might be due to the complexity of the task and perhaps due to the coding requirements affecting only a small component of the total latencies. The Hitch and Baddeley task is simpler and quicker, and arguably more dependent upon sentence comprehension than reasoning per_se. Presumably it is the heavier comprehension component which accounts in some way for the discrepancy between results of their experiments and our Experiment IV on the one hand, and the results of Experiments I to III on the other hand.

Experiment V provides a more powerful test of the Baddeley (1979)/Kleiman (1975) hypothesis that articulatory recoding is necessary in order to make rhyming judgements about visually presented words. The slight increase in latency observed on their tasks under concurrent articulation should be increased in the rhyming condition, since this will increase the load on the articulatory sub-system. Also, if attention to the colour of words induces subjects to utilise a visual code, then we might expect the interference caused by concurrent articulation to diminish.

EXPERIMENT V

The Effect of Articulatory Suppression and Coding Requirements on Sentence Verification.

METHOD.

Design

In this experiment subjects were required to perform a sentence verification task similar to that presented in Experiment IV. However, on this occasion the nature of the problem materials was manipulated in four separate Conditions. The latter three Conditions (2,3 &4) were such as to influence the type of code used to store and compare the sentences and instances in a similar manner to Experiment III. In one of the Conditions (1) no particular code was emphasised specifically.

All of the sentences referred to the order of a pair of words. They described either:

1) the particular words in the instance,

or 2) the colours of the inks in which they were written,

or 3) the rhyming characteristics of the words,

or 4) the semantic categories to which they belonged.

The mean sentence length of each of the latter three sentence types was carefully controlled and was 10.75 words in each case, so that any observed effect due to Conditions could not be explained simply in terms of sentence length. The mean length of sentences in the first category was 4.75 words. An example of each sentence type is given below:

1) TRAM FOLLOWS GOAT,

or 2) A WORD COLOURED RED FOLLOWS A WORD COLOURED YELLOW,

or 3) A RHYME WITH GRAMME FOLLOWS A RHYME WITH VOTE,

or 4) A MEANS OF TRANSPORT FOLLOWS A TYPE OF ANIMAL.

Given the instance:

GOAT TRAM

the correct response would be 'True' whereas given the instance:

TRAM GOAT

the correct response would be 'False' to each of the above four sentences.

As in the previous experiment, sentence difficulty was manipulated by varying the polarity and the voice of the verb. Both true and false instances were used with each sentence type and an equivalent set of problems using the verb 'precedes' instead of 'follows' were constructed.

Thus the four Conditions corresponded to four blocks of sixteen problems, each block concentrated on one of the above types of problem material. The stimulus words are shown in Appendix D and are identical to those used in Experiment III. The particular word-pairings used were varied for each problem. Twelve different orders of presentation of Conditions were used within each experimental Group. These were randomly selected, without replacement, from the twenty-four possible orderings. A different randomised order of problems within each Condition was used for each presentation.

Twelve subjects were run in each of three experimental Groups - Control, Articulation and Memory. The response made was recorded together with the latency for each problem.

Subjects

The subjects were thirty-six students of Plymouth Polytechnic. They had no previous experience with this sort of task and were tested individually. They were paid for their participation.

Task and Materials

Subjects were presented with the sentence verification task on a two-field tachistoscope set up as in Experiment IV. They were asked to decide whether sentences described associated instances correctly or not and to make the appropriate 'True' or 'False' responses. Conditions was a within subject factor and their order of presentation was such that all Conditions appeared equally often in each blocked position. The order of problems within each Condition was randomised.

Procedure

All subjects were given similar preliminary instructions to those used in the previous experiments. The particular task was then described by the following written instructions which were read aloud to the subjects by the experimenter:

"You will be presented, on the screen, with a series of rules defining the relationship which exists between various pairs of words. The rule may or may not contain negatives. For example:

DOOR FOLLOWS TREE,

or TREE DOES NOT FOLLOW DOOR,

or TREE IS FOLLOWED BY DOOR,

or DOOR IS NOT FOLLOWED BY TREE.

For each problem you will be shown one such rule on a card. Underneath this will be an instance of a pair of words. For example:

TREE DOOR   or   DOOR TREE

Your task is to read each sentence and to decide whether it is a true or a false description of the instance given. If you think that the rule describes the instance correctly press the switch towards 'TRUE'. If you think that the rule does not describe the instance correctly press the switch towards 'FALSE'. On different occasions I will require you to attend to either:

1) the actual words named in the instance,

or 2) the colours of the words in the instance,

or 3) the rhyming characteristics of the words in the instance,

or 4) the meanings of the words in the instance.

Although you will be timed, it is more important to be accurate than fast so please do not rush on the problems.

Have you any questions?"

Subjects were then given four blocks of four practice problems. Each block corresponded to one of the four types of material to be used in the experimental Conditions. However, the practice problems used the relationship 'above' or 'below'. Thus subjects were familiarised with the problem format and response keys.

Immediately prior to the experimental problems, subjects were given their separate instructions, as in Experiments I to IV, depending upon the Group (Control, Articulation or Memory) to which they had been randomly assigned.

Subjects were debriefed when they had completed the experimental problems. In all the task took about thirty-five minutes per subject.

RESULTS.

Response latencies were logarithmically transformed and submitted to a 3 X 4 X 2 X 2 X 2 analysis of variance. The factors were Groups, Conditions, Truth Value, Sentence Voice and Polarity. The last four were within subject factors. Error data were submitted to a similar analysis.

Response Latencies

Table 5.6 shows the mean latencies broken down by Groups, Conditions, Polarity and Sentence Voice.

There was a significant main effect of Conditions ($F_{1,33}=88.67$, $p<0.001$). Breakdown analysis showed that whilst the Actual Words and Visual Conditions could not be differentiated, these gave rise to significantly faster responses than Semantic judgements which in turn were significantly faster than Rhyme judgements. It is interesting to note that the conditional reasoning tasks of Experiment III led to a similar effect of Conditions on response latency.

Variations in problem structure produced similar effects to those observed in Experiment IV. There were main effects of Sentence Voice ($F_{1,33}=21.62$, $p<0.001$) and of

Group	Condition	Affirmative		Negative		
		Act.	Pass.	Act.	Pass.	Mean
	Actual Word	2.88	3.44	4.12	4.46	3.73
Control	Visual	2.84	3.4Ø	4.Ø8	4.31	3.66
(Mean=4.48)	Rhyme	4.75	4.83	6.84	7.13	5.89
	Semantic	3.89	4.18	5.13	5.45	4.66
	Mean	3.59	3.96	5.Ø4	5.34	
	Actual Word	2.77	3.51	4.37	5.ØØ	3.91
Articulation	Visual	3.56	3.89	4.87	4.49	4.2Ø
(Mean=5.Ø4)	Rhyme	5.96	5.31	7.5Ø	6.8Ø	6.39
	Semantic	4.91	4.58	6.1Ø	7.Ø7	5.67
	Mean	4.3Ø	4.32	5.71	5.84	
	Actual Word	3.39	4.Ø1	4.33	5.13	4.22
Memory	Visual	3.52	4.57	4.13	4.79	4.25
(Mean=5.28)	Rhyme	6.98	5.89	7.Ø3	7.79	6.92
	Semantic	5.5Ø	4.67	6.Ø9	6.62	5.72
	Mean	4.85	4.79	5.4Ø	6.Ø8	
Mean		4.25	4.36	5.38	5.75	

Table 5.6 . Mean solution latencies (seconds) in Experiment V broken down by Groups, Conditions, Polarity and Sentence Voice. Each mean is based on 12 subjects. Total N=36.

Polarity ($F_{1,33}=151.63$, $p<0.001$), and a Polarity by Truth Value interaction ($F_{1,33}=12.84$, $p<0.01$). An inspection of the means for these factors showed that they were in accordance with Trabasso (1972) in that True sentences were easier than False ones for Affirmative cases only. There was also a Sentence Voice by Conditions interaction ($F_{1,33}=7.01$, $p<0.025$). Further analysis indicated that the difficulty of Passives over Actives was present only for the Actual Word and Visual Conditions.

The expected main effect of Groups, found by Hitch and Baddeley (1976) and in Experiment IV, was not significant ($F_{2,33}=0.60$), although the direction of the overall means was again Control faster than Articulation faster than Memory. The interaction between Groups and Conditions was again absent ($F_{2,33}=0.55$), confirming the finding on the Conditional reasoning tasks used in Experiment III. The effect of Conditions on response latency was thus highly similar in Experiments III and V.

The failure to repeat the Hitch and Baddeley effect of the Groups factor is underlined by another finding. Although there was again a significant Groups by Polarity interaction ($F_{2,33}=4.84$, $p<0.025$) it was of the wrong sort, ie. the negation effect was <u>least</u> marked in the Memory Group. However, the latency findings may be confused by the appearence of error differences which are discussed below.

<u>Response Frequencies</u>

The error percentages, broken down by Groups, Conditions, Polarity and Sentence Voice, are shown in Table 5.7.

Group	Condition	Affirmative		Negative		
		Act.	Pass.	Act.	Pass.	Mean
	Actual Word	6	15	8	17	12
Control	Visual	6	6	8	13	8
(Mean=11)	Rhyme	8	10	10	21	13
	Semantic	8	10	6	19	11
	Mean	7	10	8	18	
	Actual Word	8	19	15	25	17
Articulation	Visual	15	15	15	13	14
(Mean=18)	Rhyme	8	21	29	35	23
	Semantic	8	13	17	38	19
	Mean	10	17	19	28	
	Actual Word	15	23	21	23	20
Memory	Visual	13	19	21	35	22
(Mean=24)	Rhyme	25	23	27	27	26
	Semantic	21	21	31	33	28
	Mean	19	22	25	30	
Mean		12	16	17	25	

Table 5.7 . Error percentages for Experiment V broken down by Groups, Conditions, Polarity and Sentence Voice. Each point is based on 12 subjects. Total N=36.

In the analysis of variance of error data there was a significant effect of Groups ($F_{2,33}=4.98$, $p<0.025$). The nature of this was inspected using two-tailed t tests. It was found that the Memory Group were significantly less accurate than Controls ($t_{22}=3.06$, $p<0.01$) with the Articulation Group intermediate. However, the difference between Articulation and Memory failed to reach significance ($t_{22}=1.36$) and the difference between Articulation and Control ($t_{22}=1.82$, $p<0.10$) was just short of significance.

Perhaps then, the Working Memory load is having its effect on errors rather than latency in this experiment. If so we would expect Groups to interact with factors affecting problem difficulty. Several such factors produced significant main effects in the error analysis: Polarity ($F_{1,33}=13.86$, $p<0.001$), Sentence Voice ($F_{1,33}=11.51$, $p<0.01$) and Truth Value ($F_{1,33}=5.63$, $p<0.025$) in addition to a Polarity by Truth Value interaction ($F_{1,33}=9.80$, $p<0.01$). As with the latency data, this result was roughly in accordance with Trabasso (1972). However <u>none</u> of these factors interacted significantly with Groups.

There was no main effect of Conditions on response frequency and, as in the latency analysis, the interaction between Groups and Conditions was almost non-existant ($F_{2,33}=0.35$). This surprising result could have been due to the use of Conditions as a within subject factor. It was, therefore, decided to analyse the first Condition employed with each subject but no disparate evidence was apparent although, of course, this analysis reduced the number of subjects within each Condition by a factor of four. Thus it seems that, even when rhyme judgements are required, no extra disruption due to rapid concurrent articulation of irrelevant material is entailed.

Finally, correlations were performed between errors and response latencies across the twelve cells Groups X Polarity X Sentence Voice independently for each of the Conditions. In all four cases these were significant and positive (Actual Word rho=0.70, t10=3.06, p<0.01; Colour rho=0.66, t10=2.76, p<0.01; Rhyme rho=0.63, t10=2.57, p<0.025; Semantic rho=0.72, t10=3.31, p<0.005). These results suggest that speed/error trade-off was not responsible for changes in response latency within any of the four experimental Conditions.

(The Analysis of Variance tables for Experiment V are shown in Appendix H).

DISCUSSION.

Experiment V failed to replicate Hitch and Baddeley's (1976, experiment III) results even though a similar sentence verification paradigm was employed. Although a main effect of Groups was observed on error frequency, this factor did not interact significantly with any of the factors affecting sentence complexity. Although the main effect of Groups was not significant in the latency analysis, there was an interaction of this factor with sentence complexity. However, the nature of this interaction was not compatible with Hitch and Baddeley's model. Indeed, the effect of polarity was less marked under memory load.

The most surprising aspect of Experiment V was the lack of any significant interaction between Groups and

Conditions in either latency or error analyses. These results suggest that Hitch and Baddeley's contention that verbal reasoning requires the interplay of the articulatory system and the central executive is not well-founded. The subsequent claims of Baddeley (1979) that the articulatory loop is likely to be utilised when rhyme judgements are involved or when word order is crucial, were not supported in Experiment V. This aspect of the present results will be taken up in the next section where other, more recent experiments of Baddeley and his colleagues will be considered.

GENERAL DISCUSSION.

In general the findings of Experiments III and V reported in the present chapter, together with the data of Experiments I and II, are difficult to reconcile with those of Hitch and Baddeley (1976). They have interpreted the effects of concurrent articulation on the verification of sentences of the form 'A precedes B' by supposing that this task requires the use of the articulatory loop. They further propose that access to the articulatory loop is disrupted by concurrent articulation. Baddeley has also suggested that the articulatory loop assists in tasks requiring retention of word order in complex sentences (Baddeley, 1979; Baddeley and Lewis, 1981; Baddeley, Eldridge and Lewis, 1981). This would lead to interference in either errors or latencies of conditional reasoning under articulatory suppression. There was also the suggestion that suppression might affect either the latency (Kleiman, 1975) or the accuracy (Besner et al, 1981) of rhyme judgements. All in all, several aspects of the present results are difficult to reconcile with these proposals.

Concurrent articulation of irrelevant material did not lead to slower reasoning latencies relative to a silent control condition in the conditional reasoning task described in Experiment III. This is in spite of the fact that word order is an essential component of such a task. In fact such trends as were present were for <u>faster</u> responding in the Articulation Group when compared with the silent Control condition. These trends lend support to the results of Experiments I and II where a similar effect achieved significance. Experiment IV indicates that the differences between these results and those of Hitch and Baddeley are unlikely to be due to differences in subject populations or to the comparison between articulation and control conditions being made on a between, rather than a within, subject basis. This is because Experiment IV replicated Hitch and Baddeley's results in essential respects with Articulation Group manipulated as a between subject factor using subjects selected from the same parent population as Experiments I to III.

Furthermore Hitch and Baddeley indicate that a necessary prediction of their model of Working Memory is that the effect of a concurrent memory load should interact with factors affecting problem complexity. These include Polarity which manipulates the presence of negatives in the sentence to be verified and Sentence Voice which manipulates the use of the active and passive voices. According to their hypothesis, a short-term memory load, such as that imposed in the present series of experiments, would occupy space in Working Memory thus reducing the space available to deal with more complex problems. Although the results of their experiments and

Experiment IV were in accord with this prediction, Experiments I and II, reported in the previous chapter and Experiment III which all used conditional reasoning paradigms generally failed to show the predicted interaction. Even Experiment V which used a sentence verification task of a similar nature to their own did not lend support to their account.

Finally, if the articulatory loop is used in making rhyme judgements, it would be expected that conditions which involve such judgements would be more affected by concurrent articulation or memory load than conditions which do not. Therefore, a marked interaction between Articulation Group and Coding Condition should have been observed on either errors or latencies in Experiments III and V. However no such interaction was observed in either study. Indeed later findings (eg. Baddeley and Lewis, 1981) have already undermined the view that articulatory coding may assist rhyme judgements and the present results lend support to these.

It should be appreciated then that the results so far have presented several difficulties for the Working Memory hypothesis proposed by Baddeley and Hitch (1974) and developed in their subsequent writing. These issues will be discussed in greater depth in chapter 7.

Another aspect of the results of Experiments III and V merits further consideration. In both cases significantly faster latencies were associated with materials emphasising the use of a visual as against a phonological or semantic code. In Experiment V measures were taken to rule out the possibility that the effect was simply one of sentence length. It is of interest that faster latencies were observed in both comprehension and verification periods of Experiment III and this suggests that comparisons are in some respect easier if they are made in a visual code. This result is of a similar nature to that observed in Experiment II where the presentation of a pictorial instance, compared with a verbal description, facilitated performance in conditional reasoning. However, on that occasion it affected the nature rather than the speed of responding. Since a confounding factor was present in Experiment II which could explain that result, a more detailed consideration of these effects will be delayed until further evidence is gathered.

CHAPTER 6

EXPERIMENTS VI AND VII.

EXAMINING THE ADVANTAGES OF PICTURES OVER WORDS.

The experiments in this chapter were designed to follow up a particularly curious aspect of one of the studies described in chapter 4. In that chapter, Experiment II was designed to test the revised Dual Process theory of reasoning (Wason and Evans, 1975; Evans, 1980a; Evans, 1980b). This theory suggests that conditional reasoning performance is the result of dual cognitive processes operating in parallel. In the developed theory, the Type 1 process, which accounts for the non-logical aspect of reasoning data, was postulated as non-verbal in nature whereas the Type 2 process, accounting for logical performance, was supposed to be of a verbal nature (Evans, 1980a; Evans, 1980b).

An articulatory suppression technique, similar to that used by Hitch and Baddeley (1976), was employed in an attempt to selectively disrupt the verbal process which could be articulatory in nature. It was also anticipated that the presentation of instances in a pictorial as against a verbal manner would lead to a selective increase in the non-logical response tendency ('Matching Bias') which it was thought could be imagery related. However, both of these manipulations led to surprising results.

Articulatory suppression was associated with faster responding than a silent Control Group although some disruption of logical performance was also achieved. The implications of this result as it pertains in particular to the Working Memory model of Baddeley and Hitch (1974; Hitch and Baddeley, 1976) were followed up in three experiments presented in chapter 5. In Experiment VI, one possible interpretation of the 'speeding-up' effect under articulatory suppression will be examined.

The two experiments in the present chapter, which employ a Truth Table Evaluation paradigm similar to that used in Experiment II, were designed to follow up the effects of pictorial versus verbal presentation of the instance in conditional reasoning. In Experiment II, it was surprising that, in the pictorial condition, a facilitaion of logical reasoning performance and no increase in 'Matching Bias' was achieved. However, at that stage, a possible artefactual explanation was considered plausible.

In summary, the problems presented in Experiment II referred to the combinations of shape and colour which were permissible with particular conditional rules of the form:

IF IT IS A TRIANGLE THEN IT IS RED.

Following presentation of such a rule, subjects were given either a verbal description of an instance such as:

RED TRIANGLE

or an otherwise equivalent pictorial display. They were asked to decide whether it conformed to, conflicted with or was irrelevant to the rule.

On both TT and TF truth table cases the pictorial condition was associated with superior logical performance. However, the artefactual explanation of poorer performance in the verbal condition points to the incongruency between the orders of colour and shape terms in the rule and the instance. Whilst the verbal instance follows grammatical convention by referring to the colour before the shape (eg. 'RED TRIANGLE'), the rule describes features in the reverse order with the shape in the antecedent, and the colour in the consequent clause. Of course it is possible to construct an eqivalent congruent, but grammatically unconventional, verbal instance such as:

TRIANGLE RED

It is also possible to describe the colour in the antecedent and the shape in the consequent of the rule:

IF IT IS RED THEN IT IS A TRIANGLE

When this is done, the grammatical verbal instance (eg. RED TRIANGLE) is congruent, whilst the ungrammatical instance (eg. TRIANGLE RED) is incongruent with the rule. Presumably, with pictorial instances, neither grammaticality nor congruency are relevant since features could be extracted in an optional order.

Experiment VI repeats Experiment II with three modifications. The first of these allowed the comparison of congruent as well as incongruent verbal instances to pictorial ones. Secondly, in the interests of simplicity, the Memory Group was dropped from this investigation. Thirdly, the instructions are altered to give more emphasis on the need for fast responses. This is to take account of Baddeley's suggestion (personal communication) that subjects in the Articulation Group are speeding up in order to avoid an aversive articulatory task. If Baddeley is correct, this should speed up Controls to a similar level to the Articulation Group.

If error differences arise from a speed/error trade-off this should also eliminate error differences between the Groups (particularly noticeable in Experiment II), unless there is a genuine interference due to articulatory suppression.

EXPERIMENT VI

The Effect of Articulatory Suppression on Truth Table Evaluation with Conditionals, Employing Congruent-Verbal, Incongruent-Verbal and Pictorial Instances.

METHOD.

Design

Eighteen subjects were run, in each of two Groups - Control and Articulation - on a Truth Table Evaluation task using conditional rules. Each subject attempted thirty-two problems in two randomised blocks, each of which contained the sixteen logical forms shown in Table 1.10 .

In one block the colour was described in the antecedent and the shape in the consequent of the rule whereas, in the other block, the reverse was the case. The order of presentation of the blocks was balanced. Each Group was divided into three subgroups according to the type of instance given - Congruent verbal, Incongruent verbal or Pictorial. Thus for the Congruent instance subgroup the colour-shape rule lead to a grammatical colour-shape instance, whereas the shape-colour rule lead to a non-grammatical shape-colour instance. The reverse was the case for the Incongruent instance subgroup and, of course, an identical pictorial instance was associated with both types of rule in the Pictorial instance subgroup.

As in Experiments II and III, comprehension and verification latencies, together with type of response -

'Conforms', 'Conflicts' or 'Irrelevant' - were recorded for each problem.

Subjects

Thirty-six students of Plymouth Polytechnic, having no previous experience with this type of task, served as paid volunteer subjects. They were tested individually.

Task and Materials

The logical task was the same as in Experiments II and III. Although the materials were of a similar nature to those used in Experiment II, they were modified in some respects.

The conditional rules used shape-colour relationships and were expressed with reference to the colour before the shape on half the problems and in the reverse order on the other half. In the Pictorial Condition the sentence was followed by a picture of a coloured shape. In the other two Conditions the instance was verbal and either Congruent or not. For example, given the rule:

IF IT IS GREEN THEN IT IS A CIRCLE,

a Congruent FT instance would be:

RED CIRCLE,

whereas an Incongruent FT instance would be:

CIRCLE RED.

Procedure

The procedure and instructions were essentially similar to those of the Control and Articulation Groups of Experiment II, with minor modifications to accomodate the changes in materials.

However, the sentence referring to speed and accuracy used in the previous experiments was altered. Instead of "Although you will be timed......" (cf. Procedure, Experiments I to V), the subjects were told, "Since you will be timed please answer the problems as quickly as possible consistent with high accuracy".

In all the experiment took about 30 minutes per subject.

RESULTS.

As with the previous studies, analyses of response latencies and response frequencies will be described in separate sections.

Response Frequencies

As in Experiments II and III, the modal responses conformed to a defective truth table overall. Once again, separate analyses of variance were performed to each logical case on that basis. The factors were Groups (on 2 levels), Instance (3 levels), Feature Order in Rule (2 levels) and Matching Case (4 levels). The percentage of responses conforming to a defective truth table are shown in Table 6.1 .

Feature Order in Rule was far from significance in the TT, TF and FF analyses. However, in the FT analysis, it achieved significance ($F_{1,30}=5.32$, $p<0.05$) showing more 'Irrelevant' responding overall when the shape was mentioned before the colour in the rules. However this factor did not interact significantly with Instance, or any other factor, in the frequency analyses. Matching Case produced significant main effects, of the type consistent with 'Matching Bias', in the analyses of all four logical cases (TT case - $F_{1,30}=11.79$,

Table 6.1 . Percentage frequency of modal responses to each of the four logical cases in Experiment VI conforming to a defective truth table. Each point is based on 6 subjects. Total N=36.

(i) <u>TT as 'Conforms'</u>

Group	Instance	Feature Order in Rule	Matching Case: pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$	Mean
	Congruent	colour-shape	100	100	83	50	83
	(Mean=81)	shape-colour	100	83	83	50	79
Control	Incongruent	colour-shape	100	83	100	67	88
(Mean=85)	(Mean=83)	shape-colour	100	83	67	67	79
	Pictorial	colour-shape	100	100	83	67	88
	(Mean=92)	shape-colour	100	100	83	100	96
	Congruent	colour-shape	100	100	67	50	79
	(Mean=79)	shape-colour	100	83	67	67	79
Articulation	Incongruent	colour-shape	100	67	100	83	88
(Mean=83)	(Mean=85)	shape-colour	100	83	83	67	83
	Pictorial	colour-shape	100	100	83	67	88
	(Mean=83)	shape-colour	100	67	67	83	79
Mean			100	87	81	68	
(Rule)			(AA)	(AN)	(NA)	(NN)	

(ii) TF as 'Conflicts'

Group	Instance	Feature Order in Rule	pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$	Mean
	Congruent	colour-shape	100	100	83	17	75
	(Mean=73)	shape-colour	83	100	83	17	71
Control	Incongruent	colour-shape	83	83	67	67	75
(Mean=81)	(Mean=77)	shape-colour	100	100	83	33	79
	Pictorial	colour-shape	100	100	100	83	96
	(Mean=92)	shape-colour	100	100	83	67	88
	Congruent	colour-shape	100	83	50	33	67
	(Mean=60)	shape-colour	67	83	17	50	54
Articulation	Incongruent	colour-shape	67	100	83	67	79
(Mean=74)	(Mean=73)	shape-colour	83	100	17	67	67
	Pictorial	colour-shape	83	100	100	50	83
	(Mean=88)	shape-colour	100	100	100	67	92
Mean			89	96	72	51	
(Rule)			(AN)	(AA)	(NN)	(NA)	

(iii) FT as 'Irrelevant'

Group	Instance	Feature Order in Rule	Matching Case pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$	Mean
	Congruent	colour-shape	0	17	0	83	25
	(Mean=27)	shape-colour	0	17	0	100	29
Control	Incongruent	colour-shape	0	0	33	50	21
(Mean=31)	(Mean=25)	shape-colour	0	17	33	67	29
	Pictorial	colour-shape	50	0	17	67	34
	(Mean=42)	shape-colour	17	67	33	83	50
	Congruent	colour-shape	0	33	17	83	33
	(Mean=33)	shape-colour	0	17	33	83	33
Articulation	Incongruent	colour-shape	0	0	17	67	21
(Mean=27)	(Mean=29)	shape-colour	17	33	17	83	37
	Pictorial	colour-shape	0	17	17	50	21
	(Mean=19)	shape-colour	0	0	17	50	17
Mean			7	18	19	72	
(Rule)			(NA)	(NN)	(AA)	(AN)	

(iv) FF as 'Irrelevant'

Group	Instance	Feature Order in Rule	Matching Case pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$	Mean
	Congruent	colour-shape	0	0	17	83	25
	(Mean=27)	shape-colour	17	0	17	83	29
Control	Incongruent	colour-shape	0	50	0	67	29
(Mean=40)	(Mean=29)	shape-colour	0	33	17	67	29
	Pictorial	colour-shape	0	67	67	83	54
	(Mean=63)	shape-colour	67	67	83	67	71
	Congruent	colour-shape	0	17	50	67	33
	(Mean=35)	shape-colour	33	17	33	67	37
Articulation	Incongruent	colour-shape	0	67	17	83	42
(Mean=32)	(Mean=44)	shape-colour	17	17	50	100	46
	Pictorial	colour-shape	0	0	17	50	17
	(Mean=17)	shape-colour	0	0	17	50	17
Mean			11	28	32	72	
(Rule)			(NN)	(NA)	(AN)	(AA)	

$p<0.01$; TF case - $F_{1,30}=16.55$, $p<0.001$; FT case - $F_{1,30}=41.43$, $p<0.001$; FF case - $F_{1,30}=42.04$, $p<0.001$). In the FF analysis Matching Case interacted significantly with Instance ($F_{2,30}=3.53$, $p<0.05$), although the nature of the interaction is not that predicted in the introduction to this experiment. If anything, 'Matching Bias' is rather less marked in the Pictorial Instance condition.

The Instance factor achieved significance in the TF analysis ($F_{2,30}=10.11$, $p<0.001$). Multiple comparisons using t-tests revealed that the Congruent and Incongruent Instance conditions could not be distinguished significantly ($t_{30}=1.61$), although both of these conditions led to significantly poorer performance than the Pictorial Instance Condition (Congruent vs Pictorial $t_{30}=4.44$, $p<0.001$; Incongruent vs Pictorial $t_{30}=2.83$, $p<0.005$). The mean percentage of correct responses in each Instance condition were:

Congruent 67% Incongruent 75% Pictorial 90% .

Although this factor was far from significant in the TT analysis ($F_{2,30}=0.74$), the direction of means suggested a similar Pictorial advantage. The means for TT were:

Congruent 80% Incongruent 84% Pictorial 88% .

In all four analyses the Articulation Group showed a slight, but non-significant, disadvantage compared to Control.

One other highly significant effect was a Groups by Instance interaction in the FF frequency analysis ($F_{2,30}=5.79$, $p<0.01$). If Congruent and Incongruent are collapsed this emerges quite clearly as a cross-over interaction between Groups and Verbal/Pictorial Instance. 'Irrelevant' responses are considerably higher for Pictorial in Control and markedly lower in Articulation. Although the interaction is not significant in the FT analysis, a similar direction of effect is observed.

Response Latencies

The Comprehension and Verification latencies were submitted to a logarithmic transformation. CT's were analysed using a 2 X 2 X 4 split plot analysis of variance with the factors Groups, Feature Order in Rule and Rules. The last two factors were within subjects. Since there was no significant main effect or interaction involving the Feature Order in Rule factor, this is omitted from the presentation of Comprehension latencies in Table 6.2 .

Group	Rule				
	AA	AN	NA	NN	Mean
Control	2.85	3.72	3.61	4.11	3.57
Articulation	3.01	3.65	3.51	4.16	3.58
Mean	2.93	3.69	3.56	4.14	3.58

Table 6.2 . Mean Comprehension latencies (seconds) in Experiment VI. Each mean is based on 18 subjects. Total N=36.

With CT's, the only significant effect was Rules ($F_{1,34}=29.71$, $p<0.001$) and this showed that, as negative components were added, latencies increased. The Groups factor was not significant ($F_{1,30}=0.04$). On this occasion the mean latency for Control was 3.57 seconds and for Articulation 3.58 seconds.

The analysis for VT's was similar to that for CT's except for the addition of the Logical Case factor (within subjects) on 4 levels, together with Instance Condition (between subjects) on 3 levels. These additions represent manipulations incurred with the presentation of the instance.

The mean VT's are illustrated in Table 6.3 . Since there was no main effect of Instance (F2,30=1.16) or of Feature Order in Rule (F1,30=0.69) and these factors did not interact significantly with any other factor, the table is collapsed over these factors in the interests of simplicity of presentation.

Although there was not a significant main effect of Groups (F1,30=1.72), the means were Control 3.54 seconds and Articulation 2.98 seconds which is in line with the speeding-up effect of Experiments I and II, and in line with the non-significant trend of Experiment III. Inspection of the individual cells in Table 6.3, shows that 14 of the 16 cases presented fall in the predicted direction.

With VT's, there was a significant effect of Rules (F1,30=20.32, p<0.001) indicating that the addition of one extra negative led to increased latencies. However, there was not an additional increase when a second negative was introduced into the rule. There was also a significant main effect of Logical Case (F1,30=24.73, p<0.001) indicating that latencies for TT cases were faster than all other cases and TF latencies were faster than FF. There was a significant interaction between these two factors (F1,30=6.91, p<0.025). It can be seen that the usual order of difficulty of Rule form is slightly out of line for the FF case.

Group	Rule	Logical Case				
		TT	TF	FT	FF	Mean
Control	AA	1.99	2.73	3.93	3.38	3.01
(Mean=3.54)	AN	2.69	3.00	3.80	4.97	3.61
	NA	2.86	3.37	4.53	4.47	3.81
	NN	3.11	3.61	4.39	3.86	3.74
	Mean	2.66	3.18	4.16	4.17	
Articulation	AA	1.67	2.60	2.83	2.54	2.41
(Mean=2.98)	AN	2.47	2.40	3.29	3.96	3.03
	NA	2.29	3.09	2.86	3.57	2.95
	NN	2.79	3.94	3.50	3.90	3.53
	Mean	2.31	3.01	3.12	3.49	
Mean		2.48	3.09	3.64	3.83	

Table 6.3 . Mean Verification latencies (seconds) in Experiment VI. Each mean is based on 18 subjects. Total N=36.

Once again, the mean VT's associated with making each type of response (Conforms, Conflicts and Irrelevant) were calculated for each cell condition. Again there was no evidence of any significant relationship between response type and latency.

(The Analysis of Variance tables for Experiment VI are shown in Appendix I).

DISCUSSION.

In Experiment II the effect of concurrent articulation was to significantly reduce the frequency of responses in line with the defective truth table for all four logical cases. The direction of this effect is the same in the present experiment although the effect falls short of significance in all four analyses. According to the speed/error trade-off argument this reduced interference is to be expected, if the altered emphasis towards speed in the instructions have induced more similar response latencies in the two Groups.

Inspection of Tables 6.2 and 6.3 reveals this to be the case for both comprehension and verification latencies. Although the mean for Articulation is faster than Control in most of the corresponding conditions for verification latencies, the overall difference is very small and far from significant in the analysis of variance. The only significant effect in the comprehension latency analysis was Rules and this 'complexity' factor failed to interact with Groups. With verification latencies, Rules and Logical Case were both significant as was the interaction between these factors, but, once again, no interaction with Groups was apparent.

In order to assess this interpretation more carefully, it is necessary to make a comparison between the solution latencies and percentage of subjectively correct responses, in accordance with a defective truth table, for directly comparable conditions in Experiments II and VI. The relevant details are given in Table 6.4. Each mean is the average of Pictorial and Incongruent Verbal instances associated with conditional (shape-colour) rules (ie. those rules mentioning the shape in the antecedent and the colour in the consequent clauses).

a).Response latencies.

Experiment	latency	Control	Articulation
II	CT	4.17	3.80
	VT	4.22	3.16
	Total	8.39	6.96
VI	CT	3.45	3.64
	VT	3.90	3.21
	Total	7.35	6.85

b).Percentage subjectively correct responses.

Experiment	Control	Articulation
II	80	62
VI	65	55

Table 6.4 . a). The solution latencies (seconds) and b). the percentage of subjectively correct responses, in accordance with a defective truth table, for comparable Control and Articulation conditions in Experiments II and VI. Each mean is the average of Pictorial and Incongruent Verbal instances associated with conditional (shape-colour) rules (ie. those mentioning the shape in the antecedent and the colour in the consequent clauses).

The results of Experiment VI give marginal support to the view that the faster performance of subjects under concurrent articulation in Experiment II was at the expense of increased errors. Inspection of Table 6.4 shows that instructional emphasis on speed in the present experiment reduced latency differences between Control and Articulation Groups. The error differences were also reduced, but were far from eliminated. These results suggest that articulatory suppression does not intrinsically disrupt reasoning unless it induces faster responding. This point will be discussed more fully in Chapter 7.

In Experiment II a significant facilitation of TT and TF evaluations in the Pictorial Condition was found but it was possible that this occurred as an artefact of the incongruence of the verbal instances with the rules. The present experiment has indicated that some genuine facilitation is apparent with pictorial instances. It has confirmed that neither the congruency of the verbal instance with the rule, nor the grammaticality of the instance is crucial to the effect. If the grammaticality of the instance had been important then a marked interaction between the Feature Order in Rule factor and the Instance factor would have arisen. This interaction was not significant in any of the analyses.

On the face of it, it appears to be difficult to reconcile this aspect of the results with the suggestion of Evans (1980b) that the logical component of reasoning performance is mediated by a verbal process. Why should a pictorial instance, rather than a verbal one, facilitate a logical verbal process? The highly significant interaction between Groups and Instance observed in the FF analysis and the similar trend in the FT analysis may provide a clue. If a drop in 'Irrelevant' responding in these cases is regarded as interference, then it appears that pictorial instances are more subject to interference under concurrent articulation than verbal ones. This curiosity is at least of a similar kind to the general facilitation of TT and TF responding with pictorial instances. Perhaps the pictorial instances did not induce a pictorial code after all. This issue will be deferred until the final disussion.

The final experiment was designed to determine whether the presentation of features pictorially rather than verbally necessarily leads to facilitation. It is possible that the facilitation in the previous experiments was due to another confounding characteristic associated with the instances. Whereas the pictorial instances always comprised of two relevant features, colour and shape, conjointed into a 'gestalt', the verbal instances consisted of two discrete words. It was decided to examine the effect when this 'gestalt' property of pictorial presentation was not present.

For Experiment VII, conditional reasoning problems similar to those used in the previous studies of truth table evaluation were constructed. On this occasion one of three types of instance (Pictorial, Verbal or Split-Pictorial) was presented for evaluation. The Pictorial and Verbal instances were similar to those used previously and were expected to yield equivalent findings. The Split-Pictorial instances consisted of pictorial representations of the features presented discretely side by side. That is a blob of colour was presented alongside an outline shape. If the facilitation is due to pictorial qualities per se then the Split-Pictorial instances should be evaluated more easily than Verbal ones although they should not be noticeably harder than conjoint Pictorial instances. If the effect is due to the conjointness of the Pictorial features then the Split-Pictorial instances should be harder than the Pictorial instances although they should not be noticeably easier than Verbal ones.

EXPERIMENT VII

The Effect of Congruent-Verbal, Pictorial and 'Split-Pictorial' Instances on Truth Table Evaluation with Conditionals.

METHOD.

Design

Sixteen subjects were run, in each of three experimental Groups, on a truth table evaluation task using conditional rules concerning the relationship between particular colours and shapes. The Groups were differentiated according to the type of instance which was to be evaluated. One Group was given Verbal instances. A second Group was given Split-Pictorial instances in which the colour-feature and the shape-feature were presented separately to the left and to the right of a tachistoscope screen. The conditional rule and the instances were presented simultaneously. For the Verbal and Split-Pictorial Groups the features of the instance were presented in a grammatical form, congruent with the rule (see Experiment VI). The third Group was given 'gestalt' Pictorial instances with identical rules.

All subjects attempted sixteen distinct logical problems (4 Rules X 4 Logical Cases. cf Table 1.10) presented in two randomised Blocks, so that each subject attempted thirty-two problems in all.

As in the previous experiments of this kind, subjects were required to decide whether the instance conformed to, conflicted with or was irrelevant to the rule.

In addition to the particular response made, the latency of responding was measured.

Subjects

Forty-eight students of Plymouth Polytechnic served individually as paid volunteers. They had no previous experience of this type of task.

Task and Materials

Subjects were presented with a conditional rule on one field of a two-field tachistoscope. Beneath this rule was a particular instance of a Verbal, Split-Pictorial or conjoint Pictorial nature.

As in Experiments II, III and VI, the subject's task was to evaluate the instance against the rule and to signal his response by pushing the appropriate 'Conforms', 'Conflicts' or 'Irrelevant' button.

The tachistoscope was coupled to an automatic timer so that the response latency could be measured. On this occasion overall solution latency, rather than separate CT's and VT's (cf Experiments II, III and VI), were taken.

All of the problems concerned colour-shape relationships. They may be illustrated with the following example which uses an AA rule beneath which are illustrated the three alternative kinds of instance conforming to a TT Logical

Case:

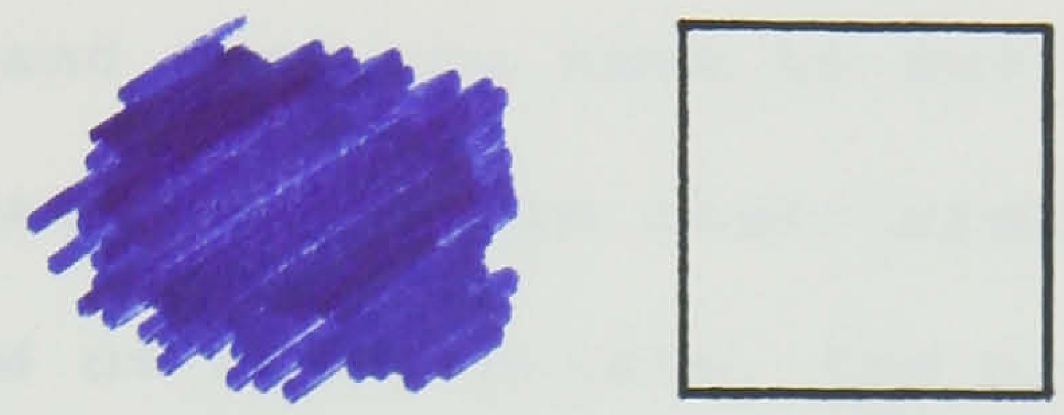
Given the rule:

IF THE COLOUR IS BLUE THEN THE SHAPE IS A SQUARE

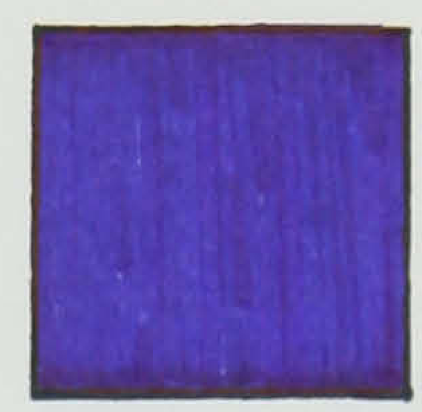
A Verbal instance would be:

BLUE SQUARE

whereas a Split-Pictorial instance would be:



and a conjoint Pictorial instance would be:



Procedure

After preliminary written instructions similar to those given in the previous experiments, the subjects were introduced to the task with the following instructions which were read aloud to him by the experimenter:

"Instructions

You will be presented, on the screen, with a rule defining the relationship between a colour and a shape. The rule may or may not contain negatives. For example:

IF THE COLOUR IS PINK THEN THE SHAPE IS AN OVAL,

or IF THE COLOUR IS ORANGE THEN THE SHAPE IS NOT A RECTANGLE,

or IF THE COLOUR IS NOT BROWN THEN THE SHAPE IS A CRESCENT,

or IF THE COLOUR IS NOT PURPLE THEN THE SHAPE IS NOT A HEXAGON.

When the rule appears on the screen you should read it very carefully. Underneath the rule on the screen will be an instance of a particular colour and a particular shape.

Your task is to decide whether that instance conforms to, conflicts with, or is irrelevant to that rule and to indicate your decision by pushing the appropriate button firmly. Please think carefully before making your decision.
Although you will be timed, it is more important to be accurate than fast, so please do not rush on the problems.
Do you understand what you have to do?"

Subjects were then instructed for a practice session which consisted of eight trials. The problems differed from the experimental ones in linguistic form (disjunctives rather than conditionals) and in the colours and shapes used.

When they had finished, subjects were debriefed. In all, this experiment took about twenty-five minutes per subject.

RESULTS.

Response frequencies and latencies were treated to separate analyses as described below.

Response Frequencies

As previously, in Experiments II, III and VI, the modal responses conformed to a defective truth table overall. Consequently separate analyses were performed to each Logical Case on that basis. There was a between subject factor of Instance (on 3 levels), and two within subject factors of Blocks (2 levels) and Matching Case (4 levels). The percentage of responses conforming to a defective truth table are shown in Table 6.5 , pooled over Blocks which was not significant in main effect or interaction.

Table 6.5 . Percentage frequency of modal responses to each of the four Logical Cases in Experiment VII conforming to a defective truth table. Each point is based on 16 subjects. Total N=48.

(i) <u>TT as 'Conforms'</u>

Instance	Matching Case				
	pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$	Mean
Verbal	100	100	91	59	88
Split-Pictorial	100	75	75	53	76
Pictorial	100	100	97	75	93
Mean	100	92	88	63	
(Rule)	(AA)	(AN)	(NA)	(NN)	

(ii) <u>TF as 'Conflicts'</u>

Instance	Matching Case				
	pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$	Mean
Verbal	91	94	78	78	85
Split-Pictorial	94	100	66	63	81
Pictorial	100	97	69	84	88
Mean	95	97	71	75	
(Rule)	(AN)	(AA)	(NN)	(NA)	

(iii) FT as 'Irrelevant'

Instance	Matching Case				
	pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$	Mean
Verbal	13	28	19	84	36
Split-Pictorial	19	47	22	75	41
Pictorial	13	47	13	66	35
Mean	15	41	18	75	
(Rule)	(NA)	(NN)	(AA)	(AN)	

(iv) FF as 'Irrelevant'

Instance	Matching Case				
	pq	$p\bar{q}$	$\bar{p}q$	$\bar{p}\bar{q}$	Mean
Verbal	13	56	38	84	48
Split-Pictorial	38	41	44	66	47
Pictorial	16	38	31	66	38
Mean	22	45	38	72	
(Rule)	(NN)	(NA)	(AN)	(AA)	

Matching Case achieved significance in all four analyses (TT case - $F_{1,45}=19.27$, $p<0.001$; TF case - $F_{1,45}=13.32$, $p<0.001$; FT case - $F_{1,45}=45.59$, $p<0.001$; FF case - $F_{1,45}=21.93$, $p<0.001$). The trends were generally in line with previous research (eg Evans and Newstead, 1977) and other experiments reported here. Determinate responses decreased as the number of matching values in the instance increased, whilst 'Irrelevant' responses showed the opposite tendency. However this trend was slightly out of line with the FT case.

The effect of Instance only achieved significance in the TT analysis ($F_{2,45}=5.01$, $p<0.025$). The means are shown in Table 6.6(i). This effect was more closely examined using t-tests. The Pictorial and Verbal Groups could not be significantly differentiated ($t_{45}=0.99$). However, the slight trend was the same as in Experiments II and VI showing a slight advantage for the Pictorial over the Verbal Group. Using two-tailed t-tests it was found that both of these Groups were associated with significantly more accurate performance than the Split-Pictorial Group (Pictorial vs Split-Pictorial $t_{45}=3.10$, $p<0.01$; Verbal vs Split-Pictorial $t_{45}=2.11$, $p<0.05$). In the TF analysis there was a very slight trend in the same direction although this was far from significant ($F_{2,45}=0.75$). The means are shown in Table 6.5(ii).

Responses Latencies

The response latencies were submitted to a logarithmic transformation and were submitted to a 3 X 2 X 4 X 4 analysis of variance with the factors Instance, Blocks, Logical Case and Rules. The last three factors were within subjects. Table 6.6 illustrates the mean solution latencies.

Instance	Rule	Logical Case				
		TT	TF	FT	FF	Mean
Verbal	AA	4.43	4.79	6.91	5.57	5.43
(Mean=7.Ø7)	AN	5.73	5.65	7.43	9.63	7.11
	NA	5.45	7.17	7.64	9.Ø7	7.33
	NN	7.48	8.Ø7	9.44	8.64	8.41
	Mean	5.77	6.42	7.86	8.23	
Split-Pictorial	AA	3.63	4.15	6.21	5.94	4.98
(Mean=6.34)	AN	5.6Ø	4.85	7.37	7.33	6.29
	NA	6.31	6.6Ø	6.33	7.56	6.7Ø
	NN	6.74	7.74	7.59	7.46	7.38
	Mean	5.57	5.84	6.88	7.Ø7	
Pictorial	AA	4.65	4.65	7.87	8.33	6.37
(Mean=8.97)	AN	6.42	6.64	11.9Ø	1Ø.42	8.84
	NA	6.14	8.48	9.54	13.36	9.38
	NN	8.84	1Ø.25	12.77	13.21	11.27
	Mean	6.51	7.51	1Ø.52	11.33	
Mean		5.95	6.59	8.42	8.87	

Table 6.6 . The mean solution latencies (seconds) in Experiment VII, broken down by Instance, Logical Case and Rules. Each mean is based on 16 subjects. Total N=48.

In the interest of simplicity, Table 6.6 has been collapsed over the Blocks factor. Whilst Blocks reached significance ($F_{1,45}=36.44$, $p<0.001$) by demonstrating the beneficial effect of practice, it did not interact significantly with any other factor.

There was a significant main effect of Rules ($F_{1,45}=79.71$, $p<0.001$), the introduction of negatives causing the expected additive increases in latency. A main effect of Logical Case ($F_{1,45}=36.83$, $p<0.001$) indicated, in line with previous experiments, that responses were slower for the more complex logical cases. There was a significant interaction between these two factors ($F_{1,45}=6.77$, $p<0.025$). These effects are in line with Evans and Newstead (1977) and with the other truth table evaluation experiments reported here.

The main effect of Instance did not quite achieve significance ($F_{2,47}=2.71$, $p<0.10$). However a close inspection of Table 6.6 will reveal that there was a tendency, in most of the cells, for Pictorial instances (Mean solution time 8.97 seconds) to be evaluated more slowly than Verbal instances (7.07 seconds) which were slower than Split-Pictorial ones (6.34 seconds).

Finally, the mean response latencies associated with each type of response ('Conforms', 'Conflicts' and 'Irrelevant') were calculated for each cell. As previously, no evidence of any significant relationship between response type and latency was found.

(The Analysis of Variance tables for Experiment VII are shown in Appendix J).

DISCUSSION.

Experiment VII replicated the finding that Pictorial instances are evaluated more accurately than Verbal instances, although at a non-significant level. However this was only the case when the Pictorial instance displayed the features for evaluation in a conjoint or 'gestalt' manner. When the pictorial features are displayed discretely, as in the Split-Pictorial condition, they appeared to be significantly more difficult than both alternative modes, at least for the TT Logical Case. The overall tendency was similar, though not significant, for the TF case.

Consideration of the introduction to this experiment leads one to the conclusion that the facilitation engendered by conjoint pictorial instances is not due to their pictorial qualities per se. It appears to be more closely linked in some way to the conjoint or 'gestalt' properties with which they are associated.

Unfortunately, there was a non-significant indication in this study that the response latencies were fastest in the Split-Pictorial condition and slowest in the Pictorial one, with those in the Verbal condition lying somewhere in-between. As a consequence it is worth exploring the possibility that the results are simply due to the effect of a speed/error trade-off.

A close examination of the response latencies in Table 6.6 shows that the largest overall differences between the three Instance Conditions occurred with the FT and FF Logical Cases. Whilst the span of this tendency, between the fastest (Split-Pictorial) and slowest (Verbal) Instance Conditions, in the TT and TF cases was only Ø.94 and 1.67 seconds respectively, in the FT case this was increased to 3.64 seconds and in the FF case it was a massive 6.14 seconds. Now, a glance at Table 6.5(iii) and (iv) will reveal that, if anything, less 'Irrelevant' responses were made in the Pictorial Condition with both of these Logical Cases. If it is accepted that 'Irrelevant' is the subjectively correct response with the FT and FF Logical Cases, as the literature reviewed in chapter 1 suggests, then a simple speed/error trade-off account would lead one to expect more such responses in conditions where the time advantage appears greatest. Since this was not the case, it is suggested that the effect of Instance on frequencies was not simply the result of a trade-off between speed and accuracy of response in Experiment VII.

GENERAL DISCUSSION.

The main aim of the experiments presented in this chapter was to shed further light on an interesting effect that was first noticed in Experiment II. In that experiment it was shown that Pictorial instances led to more subjectively correct responses than verbal instances in the truth table evaluation task. On that occasion it was possible that the result was artefactual and was related to the inconsistent ordering of colour and shape features presented in the conditional rule and the verbal instance. In Experiment VI, various manipulations were introduced into a similar task in order to rule out this criticism.

Once again, Pictorial instances led to superior performance over Verbal ones. The relationship between ordering of features in the rule and instance was not important to the effect since the two Verbal Instance conditions in which this was varied were not significantly differentiated from each other. There appears, then, to be some genuine facilitation in this task when pictorial rather than verbal instances are used.

A possible reason for the effect was investigated in Experiment VII. The verbal instances presented in the previous studies were colour-words paired with shape-words. Whereas these represent discrete features with verbal presentation, the pictorial conditions combined these features to make diagrams of coloured shapes. It was possible that it was this conjointedness, rather than pictorial qualities _per se_, that in some way led to superior reasoning performance. The results of Experiment VII lend support to this explanation for when the features are represented pictorially, but in a discrete manner, performance was _inferior_ to that found both in the Verbal and also the conventional Pictorial conditions.

It seems that discrete features in conjoint pictorial instances are more salient or discriminable than in alternative verbal instances. However, although consistent with the present data, this explanation appears to be at variance with Seymour's (1979) discussion of the relative characteristics of pictorial gestalts versus verbal symbols. He writes that "conceptually, the distinction is between a representation which has the characteristics of a literal, image-like gestalt, and one which is in the form of a logical structure specifying relationships among labelled dimensions" (Seymour, 1979, p239). Whilst conveying referentially synonymous material, he maintains that a pictorial representation emphasises 'global' properties whereas a verbal description emphasises the analysis of the object into component parts. If this were the case it would have been expected that the written instances would lead to superior logical performance to the conjoint pictures with Experiments II, VI and VII. This is because truth table evaluation with conditional rules <u>demands</u> an analysis of the component dimensions of the instances. It is difficult to reconcile the present findings with Seymour's assumptions about the unanalysed nature of the pictorial code.

The present results are also surprising in view of the revised version of Wason and Evans' (1975) Dual Process theory of reasoning. In its revised form, Evans (1980b) aligned logical performance with verbal processes and non-logical performance with non-verbal processes. In chapter 2 some parallels were drawn between this theory and that of Paivio (1975) who distinguishes between imagery and verbal systems of thinking. Paivio suggests that whereas the imagery system is specialised for processing non-verbal, concrete information and is analogical in nature, the verbal system is specialised for abstract logical thought. In view of these theories, how could it be that a verbal logical reasoning process is facilitated by a pictorial rather than verbal presentation?

There are other aspects of pictorial examples, engendered by their gestalt qualities, that are worthy of some consideration. For instance, the facilitating effect of pictures might be related to realism effects in reasoning. There is a considerable body of evidence to show that certain concrete content in an appropriate context can lead to improved performance in reasoning (see Wason and Johnson-Laird, 1972; Evans, 1982). Some of the ideas contained in a recent book by Nisbett and Ross (1980) might be relevant at this juncture.

Nisbett and Ross argue that people give inferential weight to information in proportion to its vividness. They define vividness as "the emotional interest of information, the concreteness and imageability of information, and the sensory, spatial and temporal proximity of information" (Nisbett and Ross, 1980, p62). They suggest that more vivid information is "likely to be disproportionately available for influencing inferences at any time after the information is initially encountered. The inferential impact of more vivid information usually is apparent immediately upon receiving the information, however, as well as after a delay" (Nisbett and Ross, 1980, p62). They contend that "the vividness of information exerts a disproportionate impact on inferences via processes quite separate from memory" (Nisbett and Ross, 1980, p45). In the present case the literal, gestalt nature of a Pictorial instance might render it relatively concrete when compared to discrete descriptions whether they are of a Verbal or pictorial nature.

As this discussion illustrates there are several important issues which are raised by the present experimental findings. The attempt to resolve these and to relate these matters to the revised Dual Process theory of reasoning will be deferred until chapter 8.

SECTION THREE.

FINAL DISCUSSION.

CHAPTER 7

INTERPRETATION OF THE EXPERIMENTS IN RELATION TO THE THEORY OF WORKING MEMORY.

In the present chapter some implications of the experiments reported in the previous three chapters will be discussed in so far as they are relevant to the model of Working Memory proposed by Baddeley and Hitch (1974). Account will also be taken of the substantial refinements and modifications which have been detailed in the subsequent work of the original authors and their colleagues. According to Baddeley (1981) the Working Memory system that was initially proposed was not intended as a 'predictive' model. He writes that "we (Baddeley and Hitch) were sure that WM would prove far more complex than our original conception, and that given the state of our knowledge, any attempt to make a rigidly specified predictive model was bound to fail. What we proposed was much more in the spirit of a tentative map of new terrain, giving broad guidelines and suggesting areas for more detailed explanation. The evaluation of this type of theory rests on its fruitfulness in generating new knowledge and fresh insights" (Baddeley, 1981, p18 - 19).

The experiments reported in the previous three chapters were performed during the period from 1978 to 1981. The original intention of the research was to assess the nature of the processes underlying conditional reasoning performance, in view of the theoretical statements of Evans (1980a; 1980b). It was considered that the interference methodology employed by

Baddeley and Hitch (1974; Hitch and Baddeley, 1976), which involved articulatory suppression, would be a useful technique to adopt in order to test Evans' claims. However, the results of the initial experiment were found to have important implications for the theory of Working Memory, quite apart from their relevance to Evans' proposals. As a consequence, aspects of the succeeding studies incorporated elements designed to test features of the contemporary model of Working Memory.

THE DEVELOPED THEORY OF WORKING MEMORY.

Inevitably since the experiments were completed, major refinements to the Working Memory model have proved necessary. The latest version of the model with which the author is familiar is that of Baddeley (1983). In this version Baddeley describes two of the major subcomponents of the system, together with the central executive - "the core of the system that is responsible for coordinating information from the subsidiary slave systems. The central executive is assumed to function as a limited capacity attentional system capable of selecting and operating control processes and strategies" (Baddeley, 1983, p315). The concept of the articulatory loop is modified and is now "assumed to consist of two components, a phonological input store and an articulatory rehearsal process involving subvocal speech" (Baddeley, 1983, p316). This development was considered necessary in order to accommodate various results (eg. Salamé and Baddeley, 1982) referred to in chapter 3. The visuo-spatial scratch-pad has also been more clearly described since the investigations of Baddeley and Lieberman (1980), amongst others.

Baddeley (1976, 1979, 1983) is clearly committed to the view that concurrent articulation of irrelevant material prevents access to the articulatory loop and thus prevents subvocal rehearsal of other relevant material that is presented visually. As Baddeley (1983, p317) writes "it is assumed that the phonological store can be accessed either by subvocal speech, an optional strategy, or directly through auditory speech input, an obligatory process". Indeed registration in the store is obligatory with auditory presentation whether or not the subject is engaged in subvocal rehearsal, but, with visual presentation, "such registration occurs only if the subject is able to subvocalise the items as they are presented" (Baddeley, 1983, p317). He also suggests that "under normal conditions, visually presented items will be recoded phonologically so as to take advantage of this supplementary storage" (Baddeley, 1983, p317).

The sentence verification experiments of Baddeley and Hitch (1974; Hitch and Baddeley, 1976) and of Hammerton (1969) have employed articulatory suppression as a secondary task. In the former studies a marginally significant disruption was produced by it relative to a silent control condition. In the latter study the effect was significant at a conventionally accepted level. These data suggest that verbal reasoning may have an articulatory component. Indeed, in a subsequent study, Baddeley and Hitch (1974, experiment IV) endeavoured to disrupt any short-term storage component of their reasoning task by manipulating orthogonally the phonological and visual similarity of the letters concerned in the problems. Examples of the letter-pairs used are: M - C low phonological and low visual similarity, T - D high phonological and low visual similarity, X - Y low phonological and high visual similarity and B - P high phonological and high visual similarity. It has been shown that the short-term memory span for verbal materials is dependent upon phonological coding (Conrad, 1964; Wicklgren, 1965; Levy, 1971 - see review Chapter 3). Visual similarity is usually found to have little or no effect on the memory span for letters. Baddeley and Hitch's results showed the expected significant effect of phonological similarity with little effect of visual similarity. They found that significantly fewer correct solutions were achieved in three minutes when the problems contained phonologically similar materials. On the basis of these results Baddeley and Hitch (1974, p210 - 211) argue that "the verbal reasoning task does require the use of phonemically coded information, and although the effect (of phonological similarity) is small, it is highly consistent across S's". The fact that verification latencies on the basic task increase considerably when a concurrent memory load of six digits is held, the increase being greater for more complex sentences, led Baddeley and Hitch to a further conclusion. The trade-off between verification latency and additional storage load "suggests that the interference occurs within a limited capacity 'work-space', which can be flexibly allocated either to storage or to processing" (Baddeley and Hitch, 1974, p209).

Baddeley (eg. 1979) has suggested that the articulatory loop would be employed in tasks where word order is crucial to comprehension. He has also suggested that, in comprehending prose, subvocalisation aids the accurate processing of complex information (Baddeley and Lewis, 1981).

Furthermore the work of Besner, Davies and Daniels (1981) and Baddeley, Eldridge and Lewis (1981) suggest that articulatory suppression prevents the formation of an articulatory code rather than a phonological code. The effects of suppression are most evident on accuracy rather than latency measures in various tasks. Nevertheless, Baddeley and Hitch (1974; Hitch and Baddeley, 1976) did not find any disruption of response frequencies under various articulatory conditions in their sentence verification task. In fact the error rates remained at a consistently low level (about 5 - 10%) throughout all the conditions. The effects of articulatory conditions were confined to the latency measure. Of course the failure to cause disruption to solution rates could be attributed to a ceiling effect with this relatively simple task.

THE WORKING MEMORY MODEL AND CONDITIONAL REASONING.

Several of the experiments presented in the last section of this thesis have employed conditional reasoning paradigms. Although they are of a basically similar design to the simpler sentence verification task, there are several important differences. Unlike the simpler task in which errors are few and where latency is the more sensitive measure, conditional reasoning tasks have been shown to be sensitive to various experimental manipulations on both response latency and frequency measures (see Chapter 1 and Evans, 1982). Furthermore, with these tasks, the order of terms and the placement of negatives are crucially important matters for consideration. In fact, the complexity of these problem solving tasks is evidenced by the relatively high error rates which are typical of them even under normal circumstances. All of these considerations should lead one to expect, in view of the claims of Baddeley noted above, very substantial reliance on the articulatory loop with visually presented conditional reasoning problems. The effects of articulatory suppression on conditional tasks should be drastic either on response latencies or, more especially on response frequencies. Finally, the imposition of a concurrent memory load should produce very clear evidence of an interaction with problem complexity factors since these should compete for use of the limited capacity central executive component of the Working Memory system.

A summary illustrating the main effects of articulatory Groups on response latencies for the conditional reasoning experiments (I, II, III and VI) and the simpler sentence verification experiments (IV and V) is given in Table 7.1(a). The main effects of articulatory Groups on response frequencies is given in Table 7.1(b).

In Experiment I, a conditional inference task was performed, using visual presentation, under various articulatory suppression conditions. The results as a whole show a high incidence of logical errors as is characteristic of conditional reasoning tasks. However there was no evidence to show that loads imposed on Working Memory increases the frequency of reasoning errors. This is surprising in view of the complexity of the reasoning task used. Whilst Baddeley and Hitch (1974; Hitch and Baddeley, 1976) found little disruption of performance in terms of errors on sentence verification, they reconciled their data with the Working Memory model with

Table 7.1 . The main effects of articulatory Groups (Control, Articulation, Memory) on Conditional Inference (CI; Experiment I), Truth Table Evaluation (TTE; II, III and VI) and Sentence Verification (SV; IV and V).

(a). Mean solution latencies (seconds).

Experiment		Instructional Emphasis		Group CON.	ART.	MEM.	Variance Ratio	P
I	(CI)	Accuracy		8.59	6.13	11.63	F2,30=9.66	<0.001
II	(TTE)	Accuracy	ct	4.17	3.80	7.63	F2,45=3.95	<0.05
			vt	4.22	3.16	4.18	F2,42=4.50	<0.025
III	(TTE)	Accuracy	ct	5.52	5.01	6.52	F2,33=1.96	NS
			vt	5.59	4.31	5.51	F2,33=0.88	NS
VI	(TTE)	Speed	ct	3.57	3.58	--	F1,34=0.04	NS
			vt	3.54	2.98	--	F1,30=1.72	NS
IV	(SV)	Accuracy		3.34	3.93	5.12	F2,21=4.73	<0.025
V	(SV)	Accuracy		4.48	5.04	5.28	F2,33=0.60	NS

ct = comprehension time, vt = verification time.

b). Percentage correct responding.

Experiment		Instructional Emphasis	Group CON.	ART.	MEM.	Variance Ratio	P
I	(CI)	Accuracy	79	71	78	Since separate analyses were performed for each logical case, these F ratios are not available.	
II	(TTE)	Accuracy	80	62	72		
III	(TTE)	Accuracy	60	59	67		
VI	(TTE)	Speed	65	55	--		
IV	(SV)	Accuracy	83	87	75	F2,21=2.27	NS
V	(SV)	Accuracy	89	82	76	F2,33=4.98	<0.025

The percentage correct figures for Experiment I combine only the unambiguous (MP and MT) logical cases. The figures for Experiments II, III and VI combine 'subjectively correct' responses to the four logical cases, conforming to a truth table for defective implication.

reference to their latency data.

In the present case, latencies were slowed in the concurrent memory load group relative to silent controls in accordance with Baddeley and Hitch. However, latencies were significantly faster than controls under concurrent articulation *per se*. This contrasts with the data of Baddeley and Hitch and that of Hammerton (1969). In both of these cases an increase in reasoning latency occurred under irrelevant articulatory conditions when no memory load was imposed. Whilst certain activities, such as skilled reading, are not slowed down by suppression because they do not require the loop (Baddeley, 1979), the results of Experiment I cast doubt on suggestions that complex, novel verbal reasoning problems require the use of the articulatory loop.

The latency data also deviate from those of Baddeley and Hitch for other reasons. Increased disruption should occur for more complex problems when a concurrent memory load is imposed because of increased competition for the limited-capacity central executive store. The presence of negatives is a manipulation of problem complexity, and as expected the more negated problems were more slowly processed. However this complexity factor did not interact with articulatory Groups in the manner predicted by the Working Memory hypothesis. Similarly the significant effects of Inference on latency (consistent with previous literature) might have been expected to interact with Groups but did not.

The surprising effects of Groups on reasoning latency obtained in Experiment I were replicated in Experiment II which employed a different conditional reasoning paradigm known as Truth Table Evaluation. On this occasion latencies were able to be split into comprehension times (CT) and verification times (VT) as had been done previously by Evans and Newstead (1977). In fact Groups produced significant effects on both time periods. In both cases the mean times were fastest under concurrent articulation without a memory load, as in Experiment I, but this effect only achieved significance for VTs.

The overall effect of the Memory group is in line with the Working Memory hypothesis. However, the expected increase in latencies under memory load conditions was only observed on CTs. It was argued in the discussion of Experiment II (Chapter 4) that the diminished effect of the memory load in VTs could be explained as a result of the subjects rehearsing the 'novel' digits several times during the preceding CT period. Some of the information could thus have been committed to memory before subjects actually attempted to solve the reasoning task. Once the list is memorised, the memory load condition reduces to reasoning with concurrent articulation. This sort of interpretation is not inconsistent with the Working Memory hypothesis of Baddeley and Hitch.

However, another aspect of the latency data in Experiment II appears to be in conflict with the Working Memory hypothesis. There were two significant interactions between Groups and problem complexity factors in this experiment. In the CT analysis Rules and Groups interacted significantly. Although the usual additive effect of negatives in the rules was apparent in the Control group, the effect was less rather than more marked under concurrent articulation both with and without a memory load. In the VT analysis there was a significant interaction between Logical Case and Groups. Once again the effect of problem complexity was reduced for the articulatory groups.

In the present context, the apparent facilitation under Articulation is perhaps the most interesting aspect of the first two experiments. The fact that it is characteristic of VT, as indicated by the significant difference between Articulation and Control Groups found in Experiment II, suggests that the effect is on the process of reasoning rather than on the time taken to read and understand the conditional rule. On this basis it could be argued that the sentence verification task used by Baddeley and Hitch is mainly one of comprehension whereas conditional problems require more operations in the verification stage.

A possible explanation of the 'speeding-up' effect of Articulation relative to Control will now be considered. It will be remembered that Groups exerted a significant effect on logical responding in Experiment II. In the analyses of all four logical cases performance was better, in that more subjectively correct defective truth table responses were given, under Control conditions and the worst performance was observed under Articulation. Similarly in Experiment I, the direction of errors, though not significant, was for the Articulation group to make most with Control and Memory at a similar level. One possible explanation for these results is that Articulation subjects are induced to speed up for some reason, with consequent cost in reasoning accuracy. Baddeley (personal communication) has suggested that subjects might find the articulation process aversive and speed up to avoid it. He thinks that the instruction "please do not rush on the problems", used in all but one of the seven experiments reported here, may induce subjects to take longer over the problems than they might feel to be necessary. The lesser evidence of interference in the Memory group, who are also articulating, presumably arises because the memory task forces them to go slower, and gives more time to think about the problems. If this explanation is correct, it implies that articulatory suppression per se may not inhibit logical reasoning, unless it induces the subject to spend less time on the task.

In Experiment VI a conditional reasoning task, essentially similar to that used in Experiment II, was employed. However, amongst other changes that need not concern us here, the instructions required subjects to "answer the problems as quickly as possible consistent with high accuracy". If Baddeley is correct, latency differences between the Control and Articulation groups should be considerably reduced. Then, if error differences are the result of a speed/error trade-off, these should also be eliminated between the Groups, unless the interference due to articulatory suppression is genuine. On the face of it, the results of Experiment VI give marginal support to this view, since instructions emphasising speed reduced both latency and error differences between Articulation and Control groups. This suggests that articulatory suppression per se does not disrupt reasoning, unless it induces faster responding.

It is not, however, usual to observe speed/error trade-off relationships on complex reasoning tasks. On the contrary, increased errors are normally associated with <u>longer</u> latencies. We must therefore ask why concurrent articulation should accelerate responses in the first place? Baddeley's suggestion that subjects are avoiding the aversive effects of the competing task is not very plausible. In Experiment II, it would have to be supposed that subjects were willing to sacrifice accuracy, contrary to instructions. Another consideration is that, unless suppressed, we would expect some form of subvocalisation to be present on these tasks. Electromyographic studies generally find evidence of micromovements and electrical potentials in the speech organs during problem solving, especially when the problems are novel or complex (cf. McGuigan, 1966; Sokolov, 1972). Let us suppose that such subvocalisation is helpful, but tends to slow down the thought process. The effect of articulatory suppression would, then, be to speed up solution times <u>and</u> cause some loss of accuracy. Certainly the results of Experiments I and II are compatible with this suggestion, so far as Control and Articulation groups are concerned. As was previously suggested, the longer latencies in the Memory group probably reflect an initial period of rehearsal and registration.

Why, then, did the differences in latencies and errors reduce to insignificance in Experiment VI? If the instructions are effective in making subjects go faster, then it may be that Control subjects are forced to dispense with the luxury of subvocalisation. According to this view, the change of instructions should hinder the performance of the Control subjects more than that of the Articulation subjects, who are already denied access to an articulatory strategy. Comparisons between the two experiments (see Table 6.4) are in line with this hypothesis. Mean latencies of Control subjects were 1.04 seconds faster in Experiment VI than Experiment II, whereas the difference for Articulation subjects was only 0.11 seconds. The mean percentage of responses conforming to the defective truth table dropped from 80% to 65% between the Control groups, and from 62% to 55% between the Articulation groups.

On balance, then, the present results suggest that the use of implicit speech slows down problem solving to some benefit. However, it appears that this conclusion may not be generalisable across tasks, since Sokolov (1972) reports that articulatory suppression increases the latency of mental arithmetic, when inexperienced subjects are used. On the other hand, he does report cases of accelerated solution times under suppression for highly skilled and practiced subjects who may have less need of articulatory processes. Anagram solving does not appear to suffer under concurrent articulation (Peterson, 1969), and neither is it speeded up, in the light of an unpublished experiment by Evans and Brooks. It may be that an articulatory strategy is simply irrelevant to this particular task.

Where does this leave the articulatory loop and the Working Memory model? The general tone of the above discussion appears compatible with Baddeley's (1979) amended concept of the loop as an optional control strategy. It is necessary, however, to examine the discrepencies between the present results and those of Baddeley and Hitch (1974; Hitch and Baddeley, 1976). First, the subjects in the conditional reasoning tasks speeded up under articulatory suppression, whereas theirs did not. Second, their latency data showed interactions between interference tasks and linguistic complexity factors, which the present data does not. Both discrepencies could arise from the fact that their relatively simple task was primarily a comprehension task, whereas conditional tasks have a greater reasoning requirement. The Control groups in the present experiments had a much higher base rate time, which is well beyond that required to read the sentences. It could well be that the conditional tasks induced a relatively lengthy verbal reasoning process - with accompanying subvocalisation - with much more scope for speeding up. In support of this interpretation, the latency difference between Control and Articulation was significant for VT but not for CT in Experiment II. Whilst this difference was not significant for either analysis in Experiments III and VI, the tendency in both studies was more marked on VT. Also, the interactions with linguistic complexity on Baddeley and Hitch's task could arise from comprehension processes that would account for a relatively small component of conditional reasoning latencies.

Nevertheless, if the interference under suppression - most noticeable in Experiment II - is due to impairment of Working Memory function, then it should be related to the difficulty of the problems. Analyses of variance were run separately for the four truth table cases in Experiment II (cf. Table 4.3), but it does appear that the magnitude of the interference was greater for FT and FF than for TT and TF. The drop in defective truth table responses averages 27% in the former cases and only 10% in the latter. It is certainly reasonable to regard the TT and TF cases as easier, in the sense that most psychologists studying conditional reasoning believe that subjects are predisposed to consider cases where the antecedent condition is fulfilled (see Chapter 1 and Evans, 1982, chapter 8).

EFFECTS OF SUPPRESSION ON SPECIFIC CODING REQUIREMENTS.

It was argued above that subjects may habitually subvocalise on complex conditional reasoning tasks, but that articulatory suppression causes subjects to dispense with this strategy. This could account for subjects speeding up under suppression conditions. A specific manipulation was introduced into Experiment III to test this interpretation of the data.

It had been suggested by Baddeley (1979), in view of Kleiman's (1975) finding that rhyme judgements are slowed under suppression, that the articulatory loop would be required when making judgements of phonological similarity. This claim has been withdrawn in view of subsequent data (eg. Baddeley and Lewis, 1981; Besner, Davies and Daniels, 1981). However Experiment III was designed to test the possibility that habitual dependence on a slow phonological process would recur with complex conditional reasoning tasks. Different conditions of this experiment required judgements of a visual, phonological or semantic nature. In view of Kleiman's data, it is possible that the condition requiring rhyme judgements would not be associated with faster responding under suppression. On the other hand if the accuracy of rhyme judgements is impaired by suppression, as the Besner et al study suggests, then an increase in logical errors might have been expected in this condition.

In fact the interaction, in Experiment III, between articulatory Group and task Conditions failed to reach significance in any of the analyses of response frequency. Similarly, the analyses of response latency failed to achieve significance. However, the latencies on both CT and VT produced the same overall pattern of results with Articulation faster than Control for all three task conditions. The Memory group also tended to be faster than Control on VT for Visual and Semantic conditions, but this was not the case with Rhyme judgements. The only interaction between Groups and a linguistic complexity factor occurred, as with Experiment II, in the CT analysis. As previously, there was less evidence of the additive effect of negatives in the Articulation groups, with the effect of negatives disappearing completely for the Memory group.

At this stage, it was considered expedient to attempt a replication of Baddeley and Hitch's original study (Baddeley and Hitch, 1974, experiment III; Hitch and Baddeley, 1976, experiment III) in order to check whether procedural differences or differences between the subject populations studied were responsible for the differences between their results and those of Experiments I to III. Experiment IV was successful in replicating their results with tachistoscopic presentation, using instructions which emphasised accuracy rather than speed of responding, and with Groups as a between rather than a within subject factor. Similar results were obtained whether word-pairs or letter-pairs were used as stimulus materials. The mean solution latencies showed the

Control group to be fastest, the Memory group slowest, with the Articulation group somewhere in between. There was a significant interaction between Groups and Sentence Polarity which showed that the relative difficulty of negative statements was increased under memory load.

As mentioned earlier, the sentence verification task used by Baddeley and Hitch has been claimed to require utilisation of phonologically coded information, although the disruptive effect of articulatory suppression does not always achieve significance. However, if increased load were to be placed on the articulatory system then the effects of suppression should be more severe. It is surprising that no such effect was observed in Experiment III with conditional reasoning when rhyme judgements, and presumably phonological coding, were required. However, since it has already been shown that conditional tasks respond differently to interference methods, it was appropriate to repeat the manipulation of coding conditions on the Baddeley and Hitch task.

In Experiment V subjects were required to attend to visual, rhyming or semantic characteristics of stimulus words on the simpler sentence verification task. A fourth condition, with no particular code required, was also included. The suppression Groups were identical to those previously employed. Experiment V was thought to provide a more powerful test of the hypothesis that articulatory recoding enhances the ability to make rhyme judgements about visually presented words. Once again, an interaction between coding Condition and articulatory Group was expected. The effects of suppression were expected to be considerably more marked in the Rhyme condition. In the

Visual condition, where subjects might have been induced to use a visual code, a diminished effect of suppression was anticipated. The Condition in which no particular code was specified was simply expected to produce results equivalent to those of Baddeley and Hitch, and the replication study (Experiment IV).

Whilst the main effect of Groups on response latency was not significant in Experiment V, the general trend was the same as in the Baddeley and Hitch task. However the failure to repeat the Hitch and Baddeley effect on latencies is underlined by another finding. Although there was a significant interaction between Groups and Polarity, it was of the wrong sort with the negation effect being least marked for the Memory group. Could these results be reconciled with those of Baddeley and Hitch on the basis of frequency data? The effect of Groups was significant in the analysis of response frequencies in the direction that would be expected by Baddeley and Hitch. However, there were no interactions of the sort that the Working Memory model predicts in the analysis of response frequencies. The expected interaction between Groups and Conditions did not materialise in the analyses of either response latencies or frequencies of Experiment V. Experiment V provides no evidence in support of the claim that these versions of the sentence verification task require the use of phonologically coded information, or that they demand the interplay of the articulatory loop and the central executive component of Working Memory.

It appears that the results of Experiments III and V are not compatible with Baddeley's (1979) suggestion that rhyme judgements require the use of the loop. In view of more recent evidence, this difficulty is quite simply dealt with. Baddeley's claim rested primarily on Kleiman's finding that rhyme judgements were slowed by concurrent shadowing of a series of digits. However, this might be due to the perceptual rather than the articulatory component of the shadowing task. Baddeley and Lewis (1981) have reported the results of a series of experiments which demonstrate that there is no more than a minimal effect of concurrent articulation on rhyme judgements. Besner, Davies and Daniels (1981) did obtain effects of concurrent articulation on both latency and accuracy of rhyme judgements but judgements of homophony (AIL - ALE) and pseudohomophony (KRAYDEL - TRAYDEL) showed effects only on accuracy. These effects were obtained in experiments in which subjects were instructed to articulate 'as quickly as you can'. When the effect on pseudohomophony judgements of articulation at 170 w.p.m. was studied, the accuracy effect which had been obtained previously was absent. These data, together with those of Experiments III and V, suggest that the substantial and robust effects that would be expected if rhyme judgements depended on the articulatory loop do not occur. Clearly rhyme judgements require the use of some sort of phonological code but seemingly not the one on which the articulatory loop depends.

It is appropriate at this point to reconsider the relevance of this series of experiments to the hypothesis that concurrent articulation interferes with some slow habitual phonological (or articulatory?) process which is not essential to carrying out the reasoning task accurately. It was previously suggested that if this hypothesis is valid then the reduced reasoning latencies should be absent from a condition in which the use of a phonological code is enforced by a requirement to make rhyme judgements. However, since rhyme judgements are not disrupted by concurrent articulation, the form of phonological coding they enforce must be different from that referred to by the hypothesis above since the latter is *ex hypothesi* suppressed by concurrent articulation. Hence Experiments III and V do not constitute a crtical test of the hypothesis.

It has been shown how one of the difficulties which the present results create for the proposals of Baddeley and Hitch (1974) and Baddeley (1979) can be resolved by abandoning the assumption that rhyme judgements rely on the articulatory loop. Baddeley has already abandoned this assumption in his more recent papers (Baddeley and Lewis, 1981; Baddeley, 1983). For example he writes "although articulatory suppression appears to prevent phonological coding, as indicated by the phonological similarity effect, it does not hamper a subject's ability to judge whether two written words rhyme or not" (Baddeley, 1983, p318). He speculates that, since most people can still 'hear' an inner voice despite suppression, some form of auditory imagery might be involved.

The effects of articulatory suppression in the present experiments are hard to reconcile with the traditional idea (see Chapter 3) that 'inner speech' is *essential* to problem solving thought. Whilst this view has its origins in the work of both Pavlov and Watson, it is particularly influential in the appoaches of psychologists such as Luria

(1959) and Vygotsky (1962). According to this viewpoint it should surely have been expected that drastic interference with logical reasoning performance would occur under articulatory suppression.

At this point let us reconsider the Russian work, reviewed by Sokolov (1972), which has been offered in support of the traditional 'inner speech' approach. In his book several techniques which have been employed to suppress articulation in problem solving experiments are described. These include physically clamping the lips and tongue as well as a concurrent articulation technique similar to that used in the present experiments. Although disruption of performance on many cognitive tasks is found, there were also a few cases reported of accelerated solution of mental arithmetic problems under concurrent articulation. However, these were generally restricted to well-practiced subjects. Similar acceleration was noted on some rare occasions when subjects were translating foreign text but, once again, this usually occured with experienced individuals.

In general, 'speeding-up' under concurrent articulation only occured rarely with practiced subjects on relatively simple, stereotyped tasks. However, those results do not parallel the present findings in that, in the present case, naive subjects were attempting complex problems, and the 'speeding-up' effect did not interact with the practice (Blocks) factor in any of the experiments. It is reasonable to assume that an articulatory strategy is commonly employed in problem solving tasks because it is well-learned and habitual. However, its habitual nature might also lead to its utilisation on some occasions when a more efficient alternative strategy would have been appropriate.

A REASSESSMENT OF THE THEORY OF WORKING MEMORY.

One general conclusion that can be drawn from the present research is that the articulatory loop is not essential to logical thought. For example, the correct classification of TT and TF in Experiments II, III and VI is well above chance level in the Groups given articulatory suppression. Also the effects of articulatory suppression per se lead, if anything, to faster reasoning latencies with only a marginal drop in reasoning accuracy. The possibility that a speed/error trade-off was responsible occasioned by the instructional emphasis on accuracy rather than speed, particularly in Experiment II, received only marginal support from Experiment VI.

The other remaining difficulty for the Working Memory hypothesis is that the effect of a concurrent memory load on conditional reasoning does not generally produce interactions with problem complexity of the kind that their model predicts. There was only a hint of this in Experiment II where the more complex logical cases appeared to be more severely affected by suppression. Indeed even the results of one of the sentence verification studies (Experiment V) do not appear to be compatible with the Working Memory hypothesis on this count.

Therefore in so far as some of Baddeley and Hitch's (1974) results lack generality, they provide a doubtful basis for a general model of short-term memory and in particular for the calculus of information processing capacity which is incorporated in their model. The present results lead to the suggestion that the articulatory processes described by Baddeley may not play the central role in verbal reasoning that was originally thought. Although conditional reasoning appears to satisfy many of the criteria that should implicate an interaction between the articulatory loop and the central executive components of Working Memory (eg. complex verbal information requiring attention to the ordering of words), few of the results that have arisen in conditional reasoning paradigms appear to be compatible with the Working Memory model.

Baddeley admits that the central executive "represents the most complex aspect of working memory and the most difficult to analyse and conceptualise" (Baddeley, 1983, p77). This is why most of his research has concentrated on the more peripheral 'slave' components. In this way he has gradually reduced the number of functions that need to be assigned to the central executive. He has recently claimed that "it may ultimately prove unnecessary to assume a central processor.....if control is exercised by the interaction of the various subsystems without recourse to a central controller" (Baddeley, 1983, p77 - 78). Some modifications to the model appear to be needed in order to explain why dual tasks, which should require allocation of the central processor's limited, general-purpose, storage capacity, apparently are able to proceed concurrently with little mutual interference.

Monsell (1984) has expressed an alternative viewpoint to Baddeley's which will now be considered. He suggests that 'working memory' "is merely a label for heterogeneous storage capacities intrinsic to diverse domain-specific subsystems"

(Monsell, 1984, p327). He makes the point that language processing, for instance, requires the interplay of various subsystems, some of which "deal with modality-specific, relatively peripheral input or output processes, some with lexical mappings between modality-specific representations and meaning, some with analysis or generation of supralexical - syntactic, conceptual and prosodic - structure" (Monsell, 1984, p330). It is argued that distributed processing is most compatible with distributed storage - "capacities for temporary storage specific to and intrinsic to each processing module" (Monsell, 1984, p331). In this case working memory is the summation of specific storage capacities and no 'general purpose' storage capacity need be assumed.

Multiple short-term storage capacities are argued to be associated with various representational domains, each specific to the various characteristics of language units (eg. auditory, phonological, articulatory, visual, imaginal, lexical, syntactic, conceptual). In this case dual task interference will be expected when competition for at least one <u>specific</u> capacity is engendered by tasks which might otherwise be quite different. Monsell suggests that each of the domains of processing could be organised with respect to various temporal, spatial, serial-order, syntactic and semantic relationships together with several kinds of 'process-management' information. On the other hand interference might be the result of competition for executive controlling processes.

The Working Memory theory of Baddeley and Hitch has proved extremely fruitful in that it has generated a wealth of research activity. As a result of this, it has been successful in generating new knowledge and fresh insights. The approach that they, and their colleagues, have adopted is typical of what Lakatos (1971) defines, in his paper on the history and philosophy of science, as a 'Research Programme'. This consists of a "conventionally accepted hard core with a positive heuristic which defines problems, outlines the construction of a belt of auxiliary hypotheses, foresees anomalies and turns them victoriously into examples, all according to a preconceived plan" (Lakatos, 1971, p99). It is the positive heuristic that guides research activity. Falsification, in the Popperian sense, does not imply rejection of this sort of theory. Anomalies need to be recorded and publicly displayed but they need not be acted upon, at least until they become sufficiently numerous or important to herald the demise of the original Research Programme and the accession of a more plausible rival account.

Lakatos makes a distinction by which Research Programmes can be evaluated. He distinguishes between advancing problem shifts, which occur "as long as its theoretical growth anticipates its empirical growth" (Lakatos, 1971, p100), and stagnating problem shifts, which occur when its "theoretical growth lags behind its empirical growth, that is, as long as it gives only post hoc explanations either of chance discoveries or of facts anticipated by, and discovered in, a rival programme" (Lakatos, 1971, p100). Baddeley's admission that the Working Memory model was not intended to be predictive is disadvantageous in that it characterises a stagnating problem shift. Hyland (1981) suggests that these are typical of psychological research which tends to be "data or methodology oriented" (Hyland, 1981, p13).

The data obtained via the experiments described in the last three chapters do not appear to be compatible with the Working Memory theory in its original formulation. However, the current account that Baddeley (1983) gives of Working Memory is substantially modified from its previous form. As a last resort, the number of discrete stores proposed has been increased in the light of anomalous data. A hint has been given by Baddeley that even the notion of a general, limited capacity, central executive store might prove unnecessary (Baddeley, 1983). Eventually it is possible that Baddeley's description of Working Memory will not differ substantially from rival accounts such as that of Monsell (1984).

CHAPTER 8

INTERPRETATION OF THE EXPERIMENTS IN RELATION TO EVANS' DEVELOPED DUAL PROCESS THEORY OF REASONING.

The initial object of this research project was to investigate the idea that reasoning data reflect the operation of dual thought processes. This hypothesis was originally suggested by the theory of Wason and Evans (1975). In chapter 1 it was shown how Wason and Evans' theory underwent modification as a result of subsequent research. In the light of this, Evans (1980a; 1980b) hypothesised specifically that the logical component of performance reflects a verbal-rational Type 2 process, and the non-logical component a non-verbal and non-rational Type 1 thought process.

In chapter 2 some possible connections between the revised Dual Process theory of reasoning and other theories of high-level cognitive performance were considered. Several parallels were drawn between the former theory and the account of Paivio (1971; 1975). If the Type 2 process corresponds to Paivio's verbal system, then it was considered feasible that the Type 1 process corresponds to his visual system. This would imply, perhaps, that 'Matching Bias' has an imagery component.

'Matching Bias' is clearly demonstrated in conditional reasoning when negatives are introduced into the rules. The bias has been demonstrated in various paradigms such as the Wason Selection Task and studies where Truth Tables are either constructed or evaluated. The logical relevance of an item appears to be subjectively determined by the extent to which the combinations of values mentioned in the rule and the instance correspond (see, for example, Table 1.10). Although, Evans (1983a, discussed in Chapter 1) has shown that 'Matching Bias' can be reduced in a Truth Table evaluation task when the items mentioned in conditional rules are explicitly negated (eg. 'not a 3') rather than implicitly negated as is usually the case, it is plausible that explicit negation merely induces subjects to favour a verbal strategy thus leading to reduced 'Matching Bias'. It could be that, normally, subjects are more inclined to use an imagery strategy which might be expected to increase the subjective relevance of values which are perceptually present, irrespective of their logical significance.

In order to test the revised Dual Process theory, it was decided to adopt selective interference methods. Since the notion that the Type 1 process may be imagery related is rather tentative and without direct empirical support, it was decided to concentrate initially on the verbal/non-verbal distinction. A technique was required which might be expected to interfere selectively with a verbal process. In view of the extensive literature reviewed in chapter 3, it was thought that the methodology known as articulatory suppression might prove appropriate in the present case. If Evans is correct then suppression might be expected to disrupt the logical component of reasoning performance relative to the non-logical component. However, before Evans' approach is assessed in the light of the data obtained from the various experiments presented in chapters 4 to 6, it is worthwhile considering how the truth table evaluation task might be solved under ideal circumstances.

A COMPETENCE MODEL FOR TRUTH TABLE EVALUATION.

In this section a competence model for truth table evaluation, based on an unpublished account of Evans, will be described. It is anticipated that, with the aid of this model, the locus of various experimental effects will be more easily ascertained. The model is presented in Table 8.1 .

Operations:

1) Represent the rule.

2) Represent the instance.

3) Compare the representations of rule and instance.

   a) Note whether the attributes of the instance match or mismatch the named attributes of the rule, ignoring the presence of negative components. Store the following information in the truth index:

      TT if both values match

      TF if only antecedent matches

      FT if only consequent matches

      FF if neither match.

   b) Inspect polarity of each component. When a negative is found, reverse the sign of the corresponding part of the truth index.

4) Compare value of truth index with stored truth table of the rule.

5) Output response in accordance with truth table.

Table 8.1 . Competence model for solution of a truth table evaluation task.

The model illustrated in Table 8.1 is used to describe an ideal strategy, to account for the logical component of performance. Deviations of observed behaviour from this competence model will be explained with reference to non-logical processes. The model assumes representation and processing stages. The representation of the rule may occur in a propositional form, in which the presence of negatives is preserved. However, it is also plausible to assume that items could be encoded in some other manner (eg. phonologically or visually) depending upon the strategy occasioned by specific task demands. As we shall see, the representation of the instance appears to depend upon the manner in which it is presented, but with verbal instances a propositional form might normally be the case. In accordance with propositional (eg. Clark and Chase, 1972) and mixed imagery/propositional (eg. Beech, 1980) models of sentence-picture verification, it is assumed that a truth index is set by a match or mismatch of named values (operation 3.a) and subsequently reversed due to the presence of negatives in the rule (operation 3.b). The difference is that on conditional rule evaluation, the subject must determine two truth values (antecedent and consequent), and must refer the combination to a truth table (operation 4) in order to determine his response.

Let us consider an example of how the model would work. Suppose the following problem, using a verbal instance, is presented:

IF THE COLOUR IS RED THEN THE SHAPE IS NOT A SQUARE.

BLUE SQUARE

The subject encodes the rule in two parts, antecedent and consequent, eg. ANT (RED), CON (Not(SQUARE)). When the encoded instance is compared to this, the colour (BLUE) mismatches, but the shape (SQUARE) matches, so the result of operation 3.a is that FT is stored in the truth index.

At operation 3.b it is observed that the consequent is negative, so its stored truth value is reversed leading to the truth index value of FF. Finally, the subject looks up FF in his stored truth table and responds accordingly (operations 4 and 5). Assuming a truth table for defective implication, he responds 'irrelevant'.

Deviations from the competence model can be explained by assuming that the logical process competes for control with a non-logical tendency to disregard negatives and to consider mismatching items as irrelevant. This leads to the 'Matching Bias' effect. Also, the various stages are error prone but it is more likely that errors will occur at the comparison stage (3.a). These possibilities will be considered further in due course.

It is also necessary to explain why errors tend to be fewest with TT and TF logical cases, which lead to determinate ('True' or 'False') responses, and substantially increased with FT and FF logical cases, which lead to an indeterminate ('Irrelevant') response. It is assumed that determinate responses are most reliably and quickly available from the stored truth table. However, when no determinate response is prompted on an initial run-through, there is an increased tendency to cycle back to the comparison stage (3.a) for a rapid check and more weighting is given to the non-logical matching process on this second run-through.

How does this view correspond to the revised Dual Process theory? Essentially it is saying that the logical process has greater control with the TT case, and usually for the TF case, and thus more correct responding is achieved with these logical cases. It is in the 'defective' region of the truth table where the process tends to fail most often (Evans and Newstead, 1977), and here that non-logical processes take over. 'Matching bias' is a plausible idea for a subject who is 'stuck' on the task.

VERBAL INTERFERENCE AND THE DUAL PROCESS THEORY.

The results of the experiments reported in Section 2 of this thesis give modest support to the hypothesis that articulatory suppression selectively disrupts reasoning performance in the manner expected according to the revised Dual Process theory of reasoning. In Experiment II there was significant evidence of disruption of logical performance by suppression. Also, in support Evans' predictions, the 'Matching Bias' effect - attributed to a non-verbal process - was not affected by presence of verbal interference tasks. The logically correct classifications of TT and TF cases were significantly reduced under concurrent articulation (with and without memory load). Also the modal 'irrelevant' response to FT and FF was reduced in the Articulation group. Although significant, the decrement was of modest proportions. Whilst the direction of this effect was similar in Experiment VI, the effect was far from significant in the analyses of variance. In Experiment III, the effect disappeared altogether.

The results of Experiment I appear to be even less encouraging. For example, Groups did not interact with Inference rate. However, if we look more closely at the two inferences which are logically necessary regardless of interpretation, a slight trend in the expected direction is observed. MP dropped from 97% in the Control group to 92% in the Articulation group and MT dropped from 60% to 51%. The awkward group to explain is Memory. When the concurrent articulation task incorporated a memory load this, quite contrary to expectations, affected the non-logical component of performance on the inferential task. However, it is concurrent articulation which is essentially verbal, and the presence of this in itself (without memory load) does not alter 'Conclusion Bias', while it is associated with slightly more logical errors. There was not a comparable effect of articulatory suppression on the non-logical factor ('Matching Bias') in any of the Truth Table Evaluation studies (Experiments II, III and VI).

It was found in Experiment II, in line with previous research (eg. Evans and Newstead, 1977), that negative components increased latencies in both CT and VT. The effect of negatives in either component was additive. This result fits the model (Table 8.1) well if it is assumed that CT is the time taken to complete operation 1, and VT measures the time taken for the other operations. Thus, in line with previous models of sentence-picture verification (eg. Clark and Chase, 1972; Beech, 1980), encoding negatives takes longer and accounts for the CT effect. Also reversing the value of the truth index takes time and this accounts for the VT effect.

Let us suppose, in view of the finding that suppression tends to speed up conditional reasoning, that a subject habitually encodes items phonologically (or in an articulatory manner). If concurrent articulation is introduced then such encoding is prevented and so the subject switches to an alternative faster code which may be visual or semantic. The speeding up effect is most marked on the VT period because more operations involving the code occur at this stage (Encoding of the instance and comparison with the sentence value).

There was a facilitating effect on comprehension and verification latencies in the visual condition of Experiment III when compared with conditions emphasising phonological or semantic characteristics. This effect will now be considered with reference to the competence model described above. Only operation 1 can be completed in the CT period. Since faster latencies were observed for the visual condition in both CT and VT, this effect cannot be explained simply in terms of the time taken to represent the instance. Also if the effect were due to instance recoding time, one would expect a comparable effect of pictorial versus verbal presentation in the VT analysis of Experiment II, where no significant difference was observed. It is quite possible that the Visual condition of Experiment III induces a visual mode of processing that operates throughout the various reasoning stages. It is perfectly reasonable to suppose that the need to consider phonological or semantic characteristics includes more complex cognitive mechanisms than those involved in colour discriminating, and are consequently slower to operate.

In conclusion, the hypothesis that reasoning data reflect distinct types of thought has received some support in these experiments but it is only of modest proportions. The disruptive effect of the competing verbal task was not as large as might have been expected in view of Evans' (1980a; 1980b) theoretical position. However, in view of support given to Evans by the hemispheric studies of Golding (Golding, Reich and Wason, 1974; Golding, 1980; 1981) which were considered in chapter 1, the Dual Process hypothesis cannot be dismissed as incorrect. On balance, the Dual Process hypothesis might be correct, but the verbal interference task used in the present studies could have been chosen inappropriately. Articulatory suppression might be expected to inhibit verbal thinking, if the latter results from 'inner speech' based on implicit articulation, an approach that was discussed in chapter 7. However, it is possible that the logical component of reasoning data is the result of verbal processes which are not exlusively articulatory in nature.

THE EFFECT OF PICTORIAL VERSUS VERBAL INSTANCES.

After considering various links between the revised Dual Process theory of reasoning and the theory of Paivio (see Chapter 2), it was decided to investigate the possible influence of visual imagery in the reasoning process. In line with Paivio's general approach, visual imagery has been implicated in problem solving. For instance, in chapter 2, it was shown how the data from transitive inference and sentence-picture verification tasks are explained parsimoniously in terms of theories incorporating visual imagery processes.

Until the present studies, the possible effects of visual imagery in conditional reasoning have not been explored.

However, this was done in Experiments II, VI and VII. It was suggested in the review (Chapters 1 and 2) that the non-verbal Type 1 process could involve visual imagery processes. If this were the case then it might be expected that any manipulation which encourages the use of a visual strategy would enhance non-logical tendencies such as 'Matching Bias'.

This hypothesis was first tested in Experiment II by manipulation of the verbal versus pictorial display of the instance to be evaluated in a truth table task. Results doubly confounded expectations, since the manipulation (i) had no effect on 'Matching Bias' and (ii) led to significantly better logical performance with pictorial presentation. Furthermore, in Experiment III the design necessitated the use of verbal instances, but in one condition subjects were required to attend to only a visual (colour) characteristic. Whilst this did not affect the frequency of correct decisions, a facilitatory effect was found in that both comprehension and verification times were faster in this condition.

Compared to the effects of articulatory suppression and the effects of coding manipulations of Experiment III, it is harder to explain the increase in logically correct responding in the Pictorial conditions of Experiment II. In terms of the model, this result suggests either that instances are encoded more accurately (operation 2) or else compared more accurately at operation 3.a . Once the match/mismatch decision has been achieved, it is hard to see how the mode of instance can have any further effect upon the chances of a logical decision.

In Experiment II, there was a possibility of encoding errors being increased in the Verbal condition due to a confounding factor. The order of the shape and colour values was inconsistent in the rules and instances. As a consequence, the subject had to reverse the attributes of the verbal instance in order to encode it in a comparable form to the rule. With pictorial instances features could be extracted in an optional order. If this was the cause of the effect observed, then it need have nothing to do with verbal and visual mechanisms as such. However, if this was the explanation, it would be surprising that the reversal involved in encoding the Verbal instance did not result in increased VT's. In any case, the possibility that the facilitative effect of pictures might be due to this confounding factor was eliminated in Experiment VI. Once again, pictorial presentation of the instance led to superior responding and no effect on 'Matching Bias' was observed. It appears that the effect is genuine and it is unlikely to be the result of errors during the encoding stage. It appears more likely that the effect arises in the comparison stage.

In Experiment VII a possible reason why pictures should lead to better performance than verbal descriptions was tested. This relates to the combination of features into a 'gestalt' with the picturial instances. This can be compared to the discrete naming of features which is typical of verbal descriptions. It was found that when the values mentioned in pictorial instances were presented in a discrete manner, performance was, if anything, inferior to that with verbal information. It appears that the effect is in some way dependent on the 'gestalt' nature of pictorial instances.

It is worth considering whether the facilitating effect of pictures, relates to realism effects in reasoning. In chapter 1 several studies were mentioned which were originally taken to indicate that the more plausible the context, and the more realistic the reference of the materials, the better people perform (see. Wason and Johnson-Laird, 1972; Evans, 1982; Griggs, 1983). In Experiments II, VI and VII the sentences refer to coloured shapes. In the verbal conditions subjects are given descriptions of such shapes, whereas in the conventional pictorial condition they are given the actual coloured shapes themselves. Thus the verbal condition is relatively hypothetical, and the pictorial relatively concrete. However, it will be remembered that Griggs (1983) argues that the facilitating effects of realism, for instance when they occur in the Wason Selection Task, are most likely caused by the cueing effects of thematic materials from long-term memory, instructional effects or changes in the problem context. These explanations are not plausible in the present case where identical instructions and problem contexts were used for all instance conditions.

The facilitation of reasoning with the pictorial instance is hard to reconcile with a unitary propositional approach. One could only explain it in such terms by assuming that subjects made additional errors in encoding a verbal instance. For, once encoded, pictures and words are treated as semantically identical according to a pure propositional approach. It is implausible, however, to suppose that any significant process of error should influence the encoding of two well defined attributes. If the error difference arises at the processing stage, however, then the two types of problem are evidently, <u>not</u> semantically equivalent.

On the other hand, the results do not appear to fit very well to the notion of dual coding such as that of Paivio, for instance. If the pictorial instance induces use of a visual code it would be expected to inhibit rather than facilitate verbal reasoning. The point is that images are associated with concrete rather than abstract logical thought (cf. Paivio, 1975). For example how does one represent a negative - such as 'not red' in a visual image? Also why should a 'gestalt' representation of the instance lead to improved performance in conditional reasoning when the very nature of the task demands an analysis of that instance into its constituent parts?

The most plausible answer to this is that a mixed model is appropriate and that the particular strategy that a subject uses is dependent upon the apparent demands of the task. These kinds of explanation have been considered in chapter 2. On those occasions, applications were to other problem solving tasks such as those involving transitive inference or simple sentence-picture verification. An explanation in similar terms, appropriate to a more complex sentence-picture verification task, is not wholly unprecedented. The basis of the present explanation is a model proposed by Snodgrass (1980; 1984) for picture-word processing. It is illustrated in Figure 8.1 . This model predicts both amodal and modality specific effects dependent upon particular task demands. It was originally developed to characterise similarities and differences between pictures and words (Snodgrass, 1980) and has recently been extended to characterise similarities and differences between words in one language and their translations into another (Snodgrass, 1984).

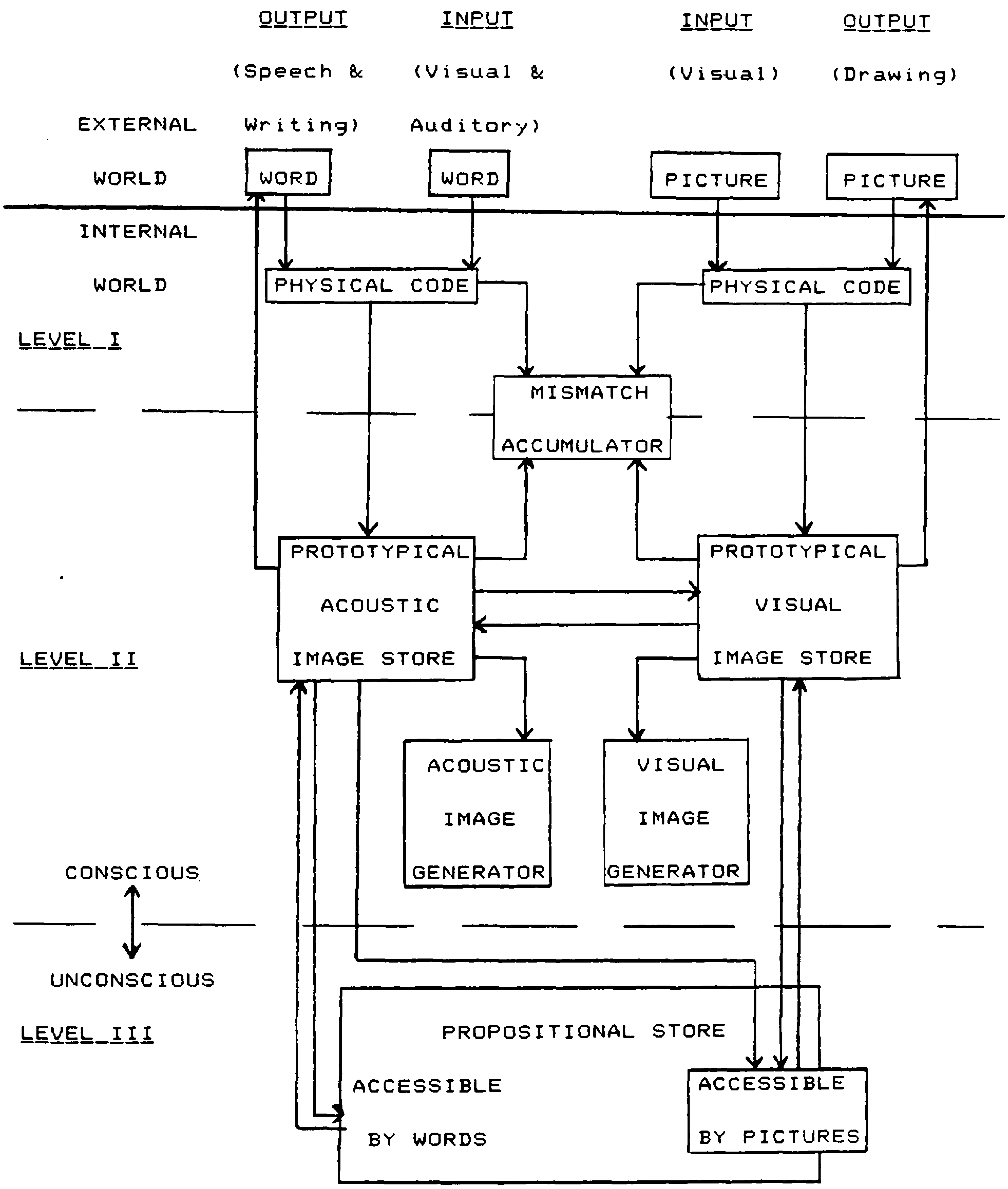


Figure 8.1 . A schematic diagram of a model for picture- and word-processing. From Snodgrass (1984) figure 1.

The model consists of three levels which are described in the following manner. Level I consists of the raw codes which result from low-level processing of words presented either visually or auditorily or for pictures presented visually. Physical attributes are extracted and stored at this level. These include characteristics such as voice and intonation, or typeface and colour for words presented in the auditory or visual modality respectively. For pictures, orientation, amount of detail and other simple physical properties are extracted and stored.

Level II consists of prototypical information about words and pictures generated, or potentially available, in the form of acoustic or visual images (eg. the sound of words, or how objects look). The stores at this level are prototypical in that they represent only the basic characteristics leaving out non-essential details. Snodgrass suggests that these images are available to introspection. It is also suggested that the acoustic image prototype store corresponds to the products of inner speech and is accessed during verbal thinking. The visual image prototype store is assumed to correspond to the products of visual imagery and is accessed during visual imaging and visual thinking. However, since images can be in a potential form within these stores, one does not necessarily experience an image. Also, it is possible to experience partial information as in 'tip-of-the-tongue' states. The visual and acoustic image generators are assumed to be limited in capacity. Each can perhaps produce only a single image at a time. However, both genertors can be used in parallel so that simultaneous visual and acoustic images can be experienced but not two images within a single modality.

In recognising speech, written language or pictures, information is assumed to be accumulated about the degree of mismatch between the physical image store at Level I and the prototypical image store at Level II. More mismatch information is accumulated when the written or spoken word does not correspond to the prototypical visual or acoustic image stored at Level II, such as when a word is written in an unusual typeface or spoken in a strange accent. Similarly, if the picture of an object differs in some way from the image that has been generated, more mismatch information will be accumulated.

Level III is the propositional or semantic store. This is viewed as an abstract set of nodes and interconnections. It is unavailable to introspection and, as a consequence, is labelled at the unconscious level. However access to this store is available in both directions via the image stores. In the case of the visual image store, only concepts and relationships which can be pictured are available. It is assumed that words can access more nodes in the propositional store than pictures although there is a degree of overlap in the case of concrete objects.

In Figure 8.1 direct connections are drawn between the image stores. This allows for the possibility that picture-word matches can be made without accessing the propositional store, although it is suggested that this is not usually the case. It is more usual to make access between these stores via the propositional route. This because it has proved useful in the past, such as when verbal thoughts are ambiguous and input from propositional memory is called upon to disambiguate them.

The model outlined above can be seen to include aspects of dual coding in that two imagery systems are described at Level II. But, unlike in Paivio's account, the meaning of images is contained in the propositional system. It was stated that, typically, access between the two systems is made via the propositional system rather than directly. However Snodgrass (1980, p575) claims that comparisons among entries in the two image systems can be made but "only on the basis of shared feature values; thus phonemic and visual similarity judgements can be made within Level II but conceptual similarity judgements must be made by accessing Level III". Snodgrass suggests that the propositional level corresponds to the propositional memory of unitary propositional theorists. It also includes Pylyshyn's (1973) assumption that the operation of this system is not available to introspection.

Let us see how this model can help to explain the facilitating effect of picturial over verbal instances in the present studies of conditional reasoning. When rules and instances are both presented verbally, each is processed via the same (left-hand) route in Snodgrass's model. Some confusion regarding the features attributed to rules might be occasioned when verbal instances are simultaneously processed for comparison at Level II. There may be a tendency to forget or over-write relevant features at this stage particularly in view of the limited capacity of this store. This disruption would lead to the generation of a faulty truth index in terms of the truth table evaluation model presented in Table 8.1 with a consequent reduction in correct responding.

When the rule is presented verbally and the instance pictorially, each is likely to be processed via different routes in terms of Snodgrass's model. Whilst verbal and pictorial features could be compared for identity at Level II, there would be little likelihood of confusion between features extracted from the rule and from the instance due to their association with distinct modalities. Furthermore, when the rule and instance are presented in modality specific manners, there is a markedly reduced likelihood of overflowing the limited storage capacity available in specific stores at Level II. Thus performance is facilitated with pictorial instances.

In the case of 'Split-Pictorial' instances, used in Experiment VII, perfomance on the truth table evaluation task was poorer than with either conventional pictorial or verbal instances. It is possible that subjects were unable to combine the discrete features into an effective visual image in this condition or that they were induced, to their disadvantage, to employ a verbal strategy. Paivio (1971) suggests that both imaginal and verbal coding systems are available in certain circumstances (see Figure 2.2). However, the verbal system is more specialised for dealing with information that is sequential in nature. The features of the 'Split-Pictorial' instance, being presented separately, might be translated into a format more suited to a sequential system of thinking.

Contrary to expectations, the use of conventional pictorial rather than verbal instances did <u>not</u> increase

'Matching Bias'. This finding does not, however, refute Evans' revised account of the Dual Process theory. This is because Evans' (1980a; 1980b) attributes the non-logical factor to a non-verbal but not necessarily a visual thought process.

A REASSESSMENT OF THE REVISED DUAL PROCESS THEORY.

In spite of the relative lack of support which the present series of experiments have given to the revised version of the Dual Process theory, it has received some support from Golding's hemispheric studies (see Chapter 1). As a result of links drawn between the theory and Paivio's dual coding account (see Chapter 2), it seemed feasible that the non-verbal process might, in fact, be visual in nature.

A final attempt to interfere selectively with possible visual and verbal processes was made in another experiment, not reported in the main experimental Section, which will be described briefly. Subjects attempted a version of the Wason Selection Task (see Chapter 1) presented on a mini-computer. Eight trials were given and both AA and AN rules were incorporated so that logical and non-logical aspects of performance could be distinguished. Prior to each reasoning trial, a verbal or pictorial item was presented on the V.D.U. for memorisation. In a Control condition no additional memory load was imposed. The verbal item consisted of a randomly generated five-digit number. The visual item consisted of a design incorporating five rectangles stacked one on top of another but distributed randomly to the left or to the right of a vertical axis. It was felt that the visual design could not easily be memorised in a verbal manner. The memorised item was to be identified from amongst four similar items presented simultaneously after the reasoning trial was completed. It was hypothesised that the influence of the logical factor ('Truth Value') would be diminished by the verbal interference relative to the Control group, whereas the influence of the non-logical factor ('Matching Bias') would be diminished by the visual inteference relative to the Control group. The percentage of items selected by each Group, broken down by Logical Case (and Matching Case) for both AA and AN rules, is shown in Table 8.2.

Interference Group	Rule AA				Rule AN			
	Logical Case				Logical Case			
	TA	FA	TC	FC	TA	FA	TC	FC
Control	66	21	54	30	77	34	48	59
Verbal	57	30	36	39	70	32	39	57
Visual	61	29	68	27	70	25	52	54
Mean	61	27	53	32	72	30	46	57
(Matching Case)	(p)	($\bar{p}$)	(q)	($\bar{q}$)	(p)	($\bar{p}$)	($\bar{q}$)	(q)

TA=True Antecedent FA=False Antecedent

TC=True Consequent FC=False Consequent

Table 8.2 . The percentage of items selected by Control, Verbal memory load and Visual memory load groups. Each point is based on 14 subjects. Total N=42.

In fact, neither of the experimental hypotheses were confirmed. The results shown in Table 8.2 indicate that reasoning performance is consistent with previous research for all three Groups. Neither the main effect nor any interactions involving Groups approached significance. As regards the memory task, errors were at about 10% overall for both the Verbal and Visual interference groups.

Unfortunately, the present research has been unsuccessful in its attempt to selectively influence the non-logical tendency known as 'Matching Bias'. Whether or not it is the result of a visual thought process remains a mystery. However, no evidence has been derived from the present series of experiments to support this hypothesis.

The extent of disruption of the verbal-logical process caused by articulatory suppression was not of the magnitude expected, in view of Evans' developed Dual Process theory, but this could be due to an inappropriate choice of verbal interference. After all, the suggestion that a process is verbal in nature does not necessarily tie it to an articulatory process.

Despite its lack of success in terms of the hypotheses tested, an interesting finding did emerge from the Selection Task experiment just described. Account was taken, by the computer, of the order in which cards were selected or rejected by subjects. Obviously, the relative (left to right) positions of the four 'cards' presented for consideration were randomised independently for each trial. The mean rank order of the decisions for each card, designated by Matching Case, is shown in Table 8.2 .

It was found that the order of selection or rejection of the cards was dominated by the cards matching value rather than its logical status. p was selected before $\bar{p}$ and q before $\bar{q}$ for <u>both</u> rules, although the logical significance of q and $\bar{q}$ are reversed in the two rules. This effect was significant for both rules combined. Evans (1983b) has interpreted this finding in support of his recent suggestion that 'Matching Bias' is the result of selective processing. He supposes that "the subject is more likely to <u>attend</u> to matching values and consider their logical significance" (Evans, 1983b, p139).

Rule	p	$\bar{p}$	q	$\bar{q}$
If p then q.	2.15	2.80	2.38	2.66
If p then not q.	2.08	2.81	2.41	2.69

Table 8.2 . Mean rank decision order for selections on two rules in the Wason Selection Task. From Evans (1983b) table 5.1.

The increased attention given to values which are present is suggested to result from a 'bias to positivity' which is pervasive in human thought. Various studies in the literature, and the present experiments, confirm that negatives lead to increased difficulty in comprehension. It is also stressed that negatives are employed to deny <u>affirmative</u> statements. In Evans' (1983b, p141) words "linguistically, negatives make statements about affirmatives". In rules such as 'If the colour is red then the shape is not a square', attention is still directed to the features actually present (ie. 'red' and 'square') rather than alternative possibilities. Evans (1983b, p141) concludes that "accuracy of performance is then a consequence of whether subjects' attention is directed to the logically important, rather than psychologically salient, aspects of the problem".

In chapter 1, a recent experiment by Evans (1983a) was discussed which demonstrated the influence of linguistic factors on the non-logical 'Matching Bias' phenomenon. The extent of 'Matching Bias' was significantly reduced in a truth table evaluation task when instances explicitly, rather than implicitly, negated the features present in the rule. This demonstrates the influence of linguistic factors on the processing rather than just the initial representation of problem information. The possibility that linguistic effects do not wholly explain 'Matching Bias' is suggested by the marked extent of it remaining even in the explicitly negated group.

In conclusion, only modest support is derived from the present project for the revised Dual Process theory of reasoning (Evans, 1980a; 1980b). There is a little evidence that interfering tasks of a verbal nature selectively disrupted the logical component of reasoning performance. The magnitude of interference caused by articulatory suppression is modest and achieved significance in only one of the experiments. However, it is possible that Evans' verbal-logical process is not articulatory in nature. Attempts to increase the influence of the non-logical process proved unsuccessful. It is surmised that Evans' characterisation of the non-logical process as non-verbal does not necessarily link it to a visual process.

However, a number of surprising and novel effects of relevance to current issues in cognitive psychology have emerged. In chapter 2 the debate between dual coding and unitary propositional theorists was considered. The present research has produced data which cannot easily be explained in terms of either of these approaches. The facilitatory effects of manipulations likely to encourage visual modes of processing are best explained in terms of Snodgrass' (1980; 1984) tricoding model of picture and word processing. Whilst it is, perhaps, premature to suggest specific practical implications for this result, it is plainly important to determine whether pictorial (rather than verbal) presentation of information also leads to improved performance with other tasks where presentation mode can be varied. In the interests of effective communication and education it is necessary to determine the most efficient means of presenting different types of complex information. The influence of presentation factors on statistical inference are now under investigation by Evans who is the holder of a current research grant from the Economic and Social Research Council with this as one of its primary aims. The second interesting finding is that the articulatory suppression technique tends to accelerate solution latencies in conditional reasoning whilst interfering very little with reasoning performance. Much of chapter 7 discussed the relevance of this and other results to the theory of Working Memory.

The field of conditional reasoning has been somewhat insular in the past. Much attention has been paid to the role of logic, and arguments for and against rationalistic explanations of behaviour. The present project has not concentrated on such matters. It has explored the nature of cognitive mechanisms underlying reasoning performance and has been concerned with areas of interest more central to cognitive psychology. The results achieved by various experimental manipulations were often surprising and were shown to be of relevance to general issues in cognition. Although little support has been found for the revised theory of Dual Processes which instigated the present research, the investigations have, nevertheless, proved worthwhile.

REFERENCES.

ALLPORT, D.A. (1980) Attention and performance. In G. Claxton (Ed) Cognitive psychology: New directions. London: Routledge and Kegan Paul.

ANDERSON, J.R. (1978) Arguments concerning representations for mental imagery. Psychological Review, 85:4, 249 - 277.

ANDERSON, J.R. & BOWER, G.H. (1973) Human associative memory. Washington, D.C.: Hemisphere Press.

ATKINSON, R.C. & SHIFFRIN, R.M. (1968) Human memory: A proposed system and its control processes. In K.W. Spence & J.T. Spence (Eds): The psychology of learning and motivation: Advances in research and theory, Volume 2. New York: Academic Press.

ATKINSON, R.C. & SHIFFRIN, R.M. (1971) The control of short-term memory. Scientific American, 225, 82 - 90.

ATWOOD, G. (1971) An experimental study of visual imagination and memory. Cognitive Psychology, 2, 290 - 299.

BADDELEY, A.D. (1966) Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. Quarterly Journal of Experimental Psychology, 18, 362 - 365.

BADDELEY, A.D. (1968) A 3 min reasoning test based on grammatical transformation. Psychonomic Science, 10:10, 341 - 342.

BADDELEY, A.D. (1976) The psychology of memory. New York: Basic Books.

BADDELEY, A.D. (1978) The trouble with levels: A reexamination of Craik and Lockhart's framework for memory research. Psychological Review, 85:3, 139 - 152.

BADDELEY, A.D. (1979) Working memory and reading. In P.A. Kolers, M.E. Wrolstad & M. Bouma (Eds): Processing of visible language, Volume 1. New York: Plenum Press.

BADDELEY, A.D. (1981) The concept of working memory: A view of its current state and possible future development. Cognition, 10, 17 - 23.

BADDELEY, A.D. (1983) Working memory. Philosophical Transactions of the Royal Society - London, B 302, 311 - 324.

BADDELEY, A.D., ELDRIDGE, M. & LEWIS, V. (1981) The role of subvocalisation in reading. Quarterly Journal of Experimental Psychology, 33A, 439 - 454.

BADDELEY, A.D., GRANT, S., WIGHT, E. & THOMSON, N. (1975) Imagery and visual working memory. In P.M.A. Rabbitt & S. Dornic (Eds): Attention and performance V. London: Academic Press.

BADDELEY, A.D. & HITCH, G.J. (1974) Working memory. The Psychology of Learning and Motivation, 8, 47 - 89. Reprinted in G.H. Bower (Ed): Human memory: Basic processes. London: Academic Press (1977, p199 - 241).

BADDELEY, A.D. & HITCH, G.J. (1977) Commentary on working memory. In G.H. Bower (Ed): Human memory: Basic processes. London: Academic Press.

BADDELEY, A.D. & LEWIS, V.J. (1981) Inner active processes in reading: The inner voice, the inner ear and the inner eye. In A.M. Lesgold & C.A. Perfetti (Eds): Interactive processes in reading. Hillsdale, N.J.: Erlbaum.

BADDELEY, A.D., LEWIS, V. & VALLAR, G. (1984) Exploring the articulatory loop. Quarterly Journal of Experimental Psychology, 36A, 233 - 252.

BADDELEY, A.D. & LIEBERMAN, K. (1980) Spatial working memory. In R. Nickerson (Ed): Attention and Performance VIII. Hillsdale, N.J.: Erlbaum.

BADDELEY, A.D., THOMSON, N. & BUCHANAN, M. (1975) Word length and the structure of short-term memory. Journal of Verbal Learning and Verbal Behaviour, 14, 575 - 589.

BEECH, J.R. (1980) An alternative model to account for the Clark and Chase picture verification experiments. Journal of Mental Imagery, 4, 1 - 11.

BESNER, D., DAVIES, J. & DANIELS, S. (1981) Reading for meaning: The effects of concurrent articulation. Quarterly Journal of Experimental Psychology, 33A, 415 - 437.

BRACEWELL, R.J. (1974) Interpretation factors in the four card selection task. Paper presented at the Trento conference on the selection task.

BROADBENT, D.E. (1958) Perception and communication. London and New York: Pergamon.

BROOKS, L.R. (1967) The suppression of visualisation by reading. Quarterly Journal of Experimental Psychology, 19, 289 - 299.

BROOKS, L.R. (1968) Spatial and verbal components of the act of recall. Canadian Journal of Psychology, 22, 349 - 368.

BUGELSKI, B.R. (1977) Imagery and verbal behaviour. Journal of Mental Imagery, 1, 39 - 52.

CARPENTER, P.A. & JUST, M.A. (1975) Sentence comprehension: A psycholinguistic processing model of verification. Psychological Review, 82, 45 - 73.

CHASE, W.G. & CLARK, H.H. (1971) Semantics in the perception of verticality. British Journal of Psychology, 62, 311 - 316.

CHMIEL, N. (1984) Phonological recoding for reading: The effect of concurrent articulation in a Stroop task. British Journal of Psychology, 75, 213 - 220.

CHOMSKY, N. (1965) Aspects of the theory of syntax. Cambridge, Massachusetts: The M.I.T. Press.

CLARK, H.H. (1969) Linguistic processes in deductive reasoning. Psychological Review, 76:4, 387 - 404.

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

COHEN, G. (1983) The psychology of cognition. (2nd edition). London: Academic Press.

COLTHEART, M. (1978) Lexical access in simple reading tasks. In G. Underwood (Ed): Strategies of information processing. London: Academic Press.

COLTHEART, M. (1983) Iconic memory. Philosophical Transactions of the Royal Society - London, B 302, 283 - 294.

CONRAD, R. (1962) An association between memory errors and errors due to acoustic masking of speech. Nature, 193, 1314 - 1315.

CONRAD, R. (1964) Acoustic confusions in immediate memory. British Journal of Psychology, 55, 75 - 84.

CONRAD, R. (1970) Short-term memory processes in the deaf. British Journal of Psychology, 61, 179 - 195.

CONRAD, R. & HULL, A.J. (1964) Information, acoustic confusion and memory span. British Journal of Psychology, 55, 429 - 432.

COOPER, L.A. & SHEPARD, P.N. (1973) Chronometric studies of rotation of mental images. In W.G. Chase (Ed): Visual information processing. New York: Academic Press.

COPI, I.M. (1982) Introduction to logic. (6th edition). London: Collier Macmillan Publishers.

CRAIK, F.I.M. & LOCKHART, R.S. (1972) Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behaviour, 11, 671 - 684.

CRAIK, F.I.M. & TULVING, E. (1975) Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology, 104, 268 - 294.

CROWDER, R.G. & MORTON, J. (1969) Precategorical acoustic storage (PAS). Perception and Psychophysics, 5, 365 - 373.

DE SOTO, L.B., LONDON, M. & HANDEL, L.S. (1965) Social reasoning and spatial paralogic. Journal of Personality and Social Psychology, 2, 513 - 521.

DI VESTA, F.J., INGERSOLL, G. & SUNSHINE, P. (1971) A factor analysis of imagery tests. Journal of Verbal Learning and Verbal Behaviour, 10, 471 - 479.

DONALDSON, M. (1963) A study of children's thinking. London: Tavistock.

DYER, F.N. (1973) The Stroop phenomenon and its use in the study of perceptual, cognitive and response processes. Memory and Cognition, 1, 106 - 120.

EDWARDS, A.L. (1967) Statistical methods. New York: Holt, Rinehart and Winston.

ERNEST, C.H. (1977) Imagery ability and cognition: A critical review. Journal of Mental Imagery, 1, 181 - 216.

EVANS, J.St.B.T. (1972a) On the problems of interpreting reasoning data: Logical and psychological approaches. Cognition, 1, 373 - 384.

EVANS, J.St.B.T. (1972b) Interpretation and matching bias in a reasoning task. Quarterly Journal of Experimental Psychology, 24, 193 - 199.

EVANS, J.St.B.T. (1972c) Reasoning with negatives. British Journal of Psychology, 63, 213 - 219.

EVANS, J.St.B.T. (1975) On interpreting reasoning data: A reply to Van Duyne. Cognition, 3, 387 - 390.

EVANS, J.St.B.T. (1976) A critical note on Quinton and Fellows' observation of reasoning strategies. British Journal of Psychology, 67, 517 - 518.

EVANS, J.St.B.T. (1977a) Linguistic factors in reasoning. Quarterly Journal of Experimental Psychology, 29, 297 - 306.

EVANS, J.St.B.T. (1977b) Toward a statistical theory of reasoning. Quarterly Journal of Experimental Psychology, 29, 621 - 635.

EVANS, J.St.B.T. (1978) The psychology of deductive reasoning: Logic. In A. Burton & J. Radford (Eds): Thinking in Perspective: Critical essays in the study of thought processes. London: Methuen.

EVANS, J.St.B.T. (1980a) Current issues in the psychology of reasoning. British Journal of Psychology, 71, 227 - 239.

EVANS, J.St.B.T. (1980b) Thinking: Experiential and information processing approaches. In G. Claxton (Ed): Cognitive Psychology: New directions. London: Routledge and Kegan Paul.

EVANS, J.St.B.T. (1982) The Psychology of Deductive Reasoning. London: Routledge and Kegan Paul.

EVANS, J.St.B.T. (1983a) Linguistic determinants of bias in conditional reasoning. Quarterly Journal of Experimental Psychology, 35A, 635 - 644.

EVANS, J.St.B.T. (1983b) Selective processes in reasoning. In J.St.B.T. Evans (Ed): Thinking and reasoning: Psychological approaches. London: Routledge and Kegan Paul.

EVANS, J.St.B.T. & LYNCH, J.S. (1973) Matching bias in the selection task. British Journal of Psychology, 64, 391 - 397.

EVANS, J.St.B.T. & NEWSTEAD, S.E. (1977) Language and reasoning: A study of temporal factors. Cognition, 5, 265 - 283.

EVANS, J.St.B.T. & WASON, P.C. (1976) Rationalisation in a reasoning task. British Journal of Psychology, 67, 479 - 486.

FILLENBAUM, S. (1978) How to do some things with IF. In J.L. Cotton & R.C. Klatzky (Eds): Semantic factors in cognition. Hillsdale, N.J.: Erlbaum.

GLANZER, M. (1972) Storage mechanisms in recall. In G.H. Bower (Ed) The psychology of learning and motivation: Advances in research and theory, Volume 5. New York: Academic Press.

GLUCKSBERG, S. (1984) Commentary: The functional equivalence of common and multiple codes. Journal of Verbal Learning and Verbal Behaviour, 23, 100 - 104.

GOLDING, E. (1980) Non-dominant hemisphere ECT and reasoning. Paper presented to the annual conference of the British Psychological Society, Aberdeen.

GOLDING, E. (1981) The effect of unilateral brain lesion on reasoning. Cortex, 17, 31 - 40.

GOLDING, E., REICH, S.S. & WASON, P.C. (1974) Inter-hemispheric differences in problem solving. Perception, 3, 231 - 235.

GOODWIN, R.Q. & WASON, P.C. (1972) Degrees of insight. British Journal of Psychology, 63, 205 - 212.

GRIGGS, R.A. (1983) The role of problem content in the selection task and in the THOG problem. In J.St.B.T. Evans (Ed): Thinking and reasoning: Psychological approaches. London: Routledge.and Kegan Paul.

HAMMERTON, M. (1969) Interference between low information verbal output and a cognitive task. Nature, 222, 196.

HAYES, J.R. (1973) On the function of visual imagery in elementary mathematics. In W.G. Chase (Ed): Visual information processing. New York: Academic Press.

HENLE, M. (1962) On the relation between logic and thinking. Psychological Review, 69, 366 -378.

HILES, D.R. (1973) Visual encoding in short-term memory. Unpublished Ph.D. Thesis. McGill University, Canada.

HINTZMAN, D.L. (1967) Articulatory coding in short-term memory. Journal of Verbal Learning and Verbal Behaviour, 6, 312 - 316.

HITCH, G.J. (1978) The role of short-term working memory in mental arithmetic. Cognitive Psychology, 10, 302 - 323.

HITCH, G.J. (1980) Developing the concept of working memory. In G. Claxton (Ed): Cognitive psychology: New directions. London: Routledge and Kegan Paul.

HITCH, G.J. & BADDELEY, A.D. (1976) Verbal reasoning and verbal memory. Quarterly Journal of Experimental Psychology, 28, 603 - 621.

HITCH, G.J. & BADDELEY, A.D. (1977) Unit 15 Working Memory. In: Cognitive psychology, Block 3, Memory (part 1). Milton Keynes: The Open University Press.

HUTTENLOCHER, J. (1968) Constructing spatial images: A strategy in reasoning. Psychological Review, 75, 550 - 560.

HUTTENLOCHER, J., HIGGINS, E.T., MILLIGAN, C. & KAUFFMAN, B. (1970) The mystery of the 'negative equative' construction. Journal of Verbal Learning and Verbal Behaviour, 9, 334 - 341.

HUTTENLOCHER, J. & STRAUSS, S. (1968) Comprehension and a statement's relation to the situation it describes. Journal of Verbal Learning and Verbal Behaviour, 7, 527 - 530.

HYLAND, M. (1981) Introduction to theoretical psychology. London: Macmillan.

JANSSEN, W.H. (1976) Selective interference in paired-associate and free recall learning: Messing up the image. Acta Psychologica, 40, 35 - 48.

JOHNSON-LAIRD, P.N. (1972) The three-term series problem. Cognition, 1:1, 57 - 82.

JOHNSON-LAIRD, P.N., LEGRENZI, P. & LEGRENZI, M.S. (1972) Reasoning and a sense of reality. British Journal of Psychology, 63, 395 - 400.

JOHNSON-LAIRD P.N. & STEEDMAN, M. (1978) The psychology of syllogisms. Cognitive Psychology, 10, 64 - 98.

JOHNSON-LAIRD P.N. & TAGART, J. (1969) How implication is understood. American Journal of Psychology, 82, 367 - 373.

JOHNSON-LAIRD, P.N. & WASON, P.C. (1970) A theoretical analysis of insight in a reasoning task. Cognitive Psychology, 1, 134 - 148.

KAHNEMAN, D. (1973) Attention and effort. Englewood Cliffs, N.J.: Prentice-Hall.

KIERAS, D. (1978) Beyond pictures and words: Alternative information-processing models for imagery effects in verbal memory. *Psychological Bulletin*, *85:3*, 532 - 554.

KLAPP, S.T., MARSHBURN, E.A. & LESTER, P.T. (1983) Short-term memory does not involve the "working memory" of information processing: The demise of a common assumption. *Journal of Experimental Psychology: General*, *112*, 240 - 264.

KLEE, H. & EYSENCK, M.W. (1973) Comprehension of abstract and concrete sentences. *Journal of Verbal Learning and Verbal Behaviour*, *12*, 522 - 529.

KLEIMAN, G.M. (1975) Speech recoding in reading. *Journal of Verbal Learning and Verbal Behaviour*, *14*, 323 - 339.

KNEALE, W. & KNEALE, M. (1962) *The Development of Logic*. London: Oxford University Press.

KOSSLYN, S.M. (1976) Can imagery be distinguished from other forms of internal representation? Evidence from studies of information retrieval time. *Memory and Cognition*, *4*, 291 - 297.

KOSSLYN, S.M. (1980) *Image and mind*. Cambridge, Massachusetts: Harvard University Press.

KOSSLYN, S.M. (1981) The medium and the message in mental imagery: A theory. *Psychological Review*, *88:1*, 46 - 66.

KOSSLYN, S.M., PINKER, S., SMITH, G.E. & SHWARTZ, S.P. (1979) On the demystification of mental imagery. *The Behavioral and Brain Sciences*, *2*, 535 - 581.

KOSSLYN, S.M. & POMERANTZ, J.R. (1977) Imagery, propositions, and the form of internal representations. *Cognitive Psychology*, *9*, 52 - 76.

KOSSLYN, S.M. & SHWARTZ, S.P. (1977) A simulation of visual imagery. *Cognitive Science*, *1*, 265 - 295.

KROLL, N.E.A., PARKS, T., PARKINSON, S.R., BIEBER, S.L. & JOHNSON, A.L. (1970) Short-term memory while shadowing: Recall of visually and aurally presented letters. *Journal of Experimental Psychology*, *85*, 220 - 224.

LAKATOS, I. (1971) History of science and its rational reconstructions. In R.C. Buck & R.S. Cohen (Eds): *Boston Studies in the Philosophy of Science VIII*. Dordrecht, The Netherlands: Reidel.

LAWSON, R. (1977) Representation of individual sentences and holistic ideas. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 1 - 9.

LEVY, B.A. (1971) Role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning and Verbal Behaviour*, *10*, 123 - 132.

LEVY, B.A. (1975) Vocalisation and suppression effects in sentence memory. Journal of Verbal Learning and Verbal Behaviour, 14, 304 - 316.

LEVY, B.A. (1978) Speech analysis during sentence processing: Reading and listening. Visible Language, 12, 81 - 102.

LURIA, A.R. (1959) The directive function of speech in developments and dissolution, part 1. Word, 15, 341 - 352.

MACLEOD, C.M., HUNT, E.B. & MATHEWS, N.N. (1978) Individual differences in the verification of sentence-picture relationships. Journal of Verbal Learning and Verbal Behaviour, 17, 493 - 507.

MADIGAN, S. (1983) Picture memory. In J.C. Yuille (Ed): Imagery, memory and cognition: Essays in honor of Allan Paivio. Hillsdale, N.J.: Lawrence Erlbaum Associates.

MADIGAN, S. & LAWRENCE, V. (1980) Factors affecting item recovery and reminiscence in free recall. American Journal of Psychology, 93, 489 - 504.

MANKTELOW, K.I. (1978) Realistic reasoning: It's not that easy. Paper presented at the London conference of the British Psychological Society.

MANKTELOW, K.I. & EVANS, J.St.B.T. (1979) Facilitation of reasoning by realism: Effect or non-effect. British Journal of Psychology, 70, 477 - 488.

MARCEL, A.J. (1983) Conscious and unconscious perception: Experiments on visual masking and word recognition. Cognitive Psychology, 15, 197 - 237.

MARCUS, S.L. & RIPS, L.J. (1979) Conditional reasoning. Journal of Verbal Learning and Verbal Behaviour, 18, 199 - 223.

MARTIN, M. (1978) Speech recoding in silent reading. Memory and Cognition, 6:2, 108 - 114.

MATHEWS, N.N., HUNT, E.B. & MACLEOD, C.M. (1980) Strategy choice and strategy training in sentence-picture verification. Journal of Verbal Learning and Verbal Behaviour, 19, 531 - 548.

MATHEWS, W.A. (1983) The effects of concurrent secondary tasks on the use of imagery in a free recall task. Acta Psychologica, 53, 231 - 241.

MATTINGLY, I.G. (1972) Reading, the linguistic process and linguistic awareness. In J.F. Kavanaugh & I.G. Mattingly (Eds): Language by ear and by eye. Cambridge, Massachusetts: The M.I.T. Press.

MAYER, R.E. (1981) The promise of cognitive psychology. San Francisco: W.H. Freeman & Company.

MCGUIGAN, F.J. (1966) Thinking: Studies of covert language language processes. New York: Appleton Century.

MCGUIGAN, F.J. (1970) Covert oral behaviour during the silent performance of language tasks. Psychological Bulletin, 74, 309 - 326.

MILLER, G.A. & NICELY, P.E. (1961) An analysis of perceptual confusions among some English consonants. In S. Saporta (Ed): Psycholinguistics. New York: Holt, Rinehart and Winston.

MONSELL, S. (1984) Components of working memory underlying verbal skills: A 'distributed capacities' view - A tutorial review. In H. Bouma & D.G. Bouwhuis (Eds): Attention and performance X. london: Lawrence Erlbaum Associates.

MORTON, J. (1970) A functional model for memory. In D.A. Norman (Ed): Models of human memory. New York: Academic Press.

MURDOCK, B.B.Jr. & WALKER, K.D. (1969) Modality effects in free recall. Journal of Verbal Learning and Verbal Behaviour, 8, 665 - 676.

MURRAY, D.J. (1967) The role of speech responses in short-term memory. Canadian Journal of Psychology, 21, 263 - 276.

MURRAY, D.J. (1968) Articulation and acoustic confusibility in short-term memory. Journal of Experimental Psychology, 78:4, 679 - 684.

NEISSER, U. (1963) The multiplicity of thought. British Journal of Psychology, 54, 1 - 14.

NEISSER, U. (1967) Cognitive psychology. New York: Meredith Publishing Company.

NELSON, D.L. (1979) Remembering pictures and words: Appearance, significance, and name. In L.C. Cermak & F.I.M. Craik (Eds): Levels of processing in human memory. Hillsdale, N.J.: Lawrence Erlbaum Associates.

NELSON, D.L., REED, V.S. & WALLING, J.R. (1976) The pictorial superiority effect. Journal of Experimental Psychology: Human Learning and Memory, 2, 523 - 528.

NEWSTEAD, S.E. (1979) Student subjects in psychological experiments. Bulletin of the British Psychological Society, 32, 384 - 386.

NEWSTEAD, S.E., MANKTELOW, K.I. & EVANS, J.St.B.T. (1982) The role of imagery in the representation of linear orderings. Current Psychological Research, 2:1, 21 - 32.

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

NISBETT, R.E. & WILSON, T.D. (1977) Telling more than one can know: Verbal reports on mental processes. Psychological Review, 84, 231 - 259.

OAKHILL, J.V. & JOHNSON-LAIRD, P.N. (1984) Representation of spatial descriptions in working memory. Current Psychological Research and Reviews, 52 - 62.

PAIVIO, A. (1969) Mental imagery in associative learning and memory. Psychological Review, 76:3, 241 - 263.

PAIVIO, A. (1971) Imagery and verbal processes. New York: Holt, Rinehart and Winston.

PAIVIO, A. (1975) Imagery and synchronic thinking. Canadian Psychological Review, 16, 147 -163.

PAIVIO, A. (1976) Images, propositions and knowledge. In J.M. Nicholas (Ed): Images, perception and knowledge (The Western Ontario series in the philosophy of science). Dordrecht, The Netherlands: Reidel.

PAIVIO, A. (1983) The empirical case for dual coding. In J.C. Yuille (Ed): Imagery, memory and cognition: Essays in honor of Allan Paivio. Hillsdale, N.J.: Lawrence Erlbaum Associates.

PAIVIO, A. & LAMBERT, W. (1981) Dual coding and bilingual memory. Journal of Verbal Learning and Verbal Behaviour, 20, 532 - 539.

PAIVIO, A. & OKOVITA, H.W. (1971) Word imagery modalities and associative learning in blind and sighted subjects. Journal of Verbal Learning and Verbal Behaviour, 10, 506 - 510.

PETERSON, L.R. (1969) Concurrent verbal activity. Psychological Review, 76:4, 376 - 386.

PETERSON, L.R. & JOHNSON, S.T. (1971) Some effects of minimising articulation on short-term retention. Journal of Verbal Learning and Verbal Behaviour, 10, 346 - 354.

PETERSON, L.R. & PETERSON, M.J. (1959) Short-term retention of individual items. Journal of Experimental Psychology, 58, 193 - 198.

PHILLIPS, W.A. (1983) Short-term visual memory. Philosophical Transactions of the Royal Society London, B 302, 295 - 309.

PHILLIPS, W.A. & CHRISTIE, D.F.M. (1977) Interference with visualization. Quarterly Journal of Experimental Psychology, 29, 637 - 650.

PIAGET, J. (1972) Intellectual development from adolescence to adulthood. Human Development, 15, 1 - 12.

PINTNER, R. (1913) Inner speech during silent reading. Psychological Review, 20, 129 - 153.

POLLARD, P. (1979) Human reasoning: Logical and non-logical explanations. Unpublished PhD Thesis, Plymouth Polytechnic.

POLLARD, P. (1981) The effect of thematic content on the 'Wason selection task'. Current Psychological Research, 1, 21 - 29.

POLLARD, P. & EVANS, J.St.B.T. (1980) The influence of logic on conditional reasoning performance. Quarterly Journal of Experimental Psychology, 32, 605 - 624.

POPPER, K.R. (1968) The logic of scientific discovery. (3rd edition). London: Hutchinson.

POSNER, M.I. (1973) Cognition: An introduction. Glenview, Illinois: Scott, Foresman and Company.

POSNER, M.I. & ROSSMAN, E. (1965) Effect of size and location of informational transforms on short-term memory. Journal of Experimental Psychology, 70, 496 - 505.

POSTMAN, L. & PHILLIPS, L.W. (1965) Short-term temporal changes in free recall. Quarterly Journal of Experimental Psychology, 17, 132 - 138.

POTTS, G.R. (1972) Information processing strategies used in the encoding of linear orderings. Journal of Verbal Learning and Verbal Behaviour, 11, 727 - 740.

POTTS, G.R. (1974) Storing and retrieving information about ordered relationships. Journal of Experimental Psychology, 103, 431 - 439.

PYLYSHYN, Z.W. (1973) What the mind's eye tells the mind's brain: A critique of mental imagery. Psychological Bulletin, 80:1, 1 - 24.

PYLYSHYN, Z.W. (1981) The imagery debate: Analogue media versus tacit knowledge. Psychological Review, 88:1, 16 - 45.

QUINTON, G. & FELLOWS, B.J. (1975) 'Perceptual' strategies in the solving of three-term series problems. British Journal of Psychology, 66, 69 - 78.

REED, S.K., HOCK, H.S. & LOCKHEAD, G.R. (1983) Tacit knowledge and the effect of pattern configuration on mental scanning. Memory & Cognition, 11:2, 137 - 143.

RICHARDSON, J.T.E. (1978) Mental imagery and memory: Coding ability or coding preference? Journal of Mental Imagery, 2, 101 - 116.

RICHARDSON, J.T.E. (1980a) Mental imagery and human memory. London: Macmillan.

RICHARDSON, J.T.E. (1980b) Concreteness, imagery, and semantic categorization. Journal of Mental Imagery, 4, 51 - 58.

RICHARDSON, J.T.E. (1980c) Mental imagery and stimulus concreteness. Journal of Mental Imagery, 4, 87 - 97.

RICHARDSON, J.T.E. (1981) Working memory: A critical evaluation. Paper presented to the British Psychological Society, Cognitive Psychology Section. Plymouth Polytechnic, 9th - 11th September.

RICHARDSON, J.T.E. (1984) Developing the theory of working memory. Memory and Cognition, 12:1, 71 - 83.

RICHARDSON, J.T.E., GREAVES, D.E. & SMITH, M.M.C. (1980) Does articulatory suppression eliminate the phonemic similarity effect in short-term recall? Bulletin of the Psychonomic Society, 16:6, 417 - 420.

ROBERGE, J.J. (1976) Reasoning with exclusive disjunction arguments. Quarterly Journal of Experimental Psychology, 28, 419 - 427.

ROBERGE, J.J. (1978) Linguistic and psychometric factors in propositional reasoning. Quarterly Journal of Experimental Psychology, 30, 705 -716.

ROTHKOPF, E.Z. (1971) Incidental memory for location of information in text. Journal of Verbal Learning and Verbal Behaviour, 10, 608 - 613.

SALAMÉ, P. & BADDELEY, A.D. (1982) Disruption of short-term memory by unattended speech: Implications for the structure of working memory. Journal of Verbal Learning and Verbal Behaviour, 21, 150 -164.

SEGAL, S.J. & FUSELLA, V. (1969) Effects of imaging on signal-to-noise ratio with varying signal conditions. British Journal of Psychology, 60, 459 - 464.

SEYMOUR, P.H.K. (1979) Human visual cognition: A study in experimental cognitive psychology. London: Collier Macmillan.

SHALLICE, T. & WARRINGTON, E.K. (1970) Independent functioning of verbal memory stores: A neuropsychological study. Quarterly Journal of Experimental Psychology, 22, 261 - 273.

SHAVER, P., PIERSON, L. & LANG, S. (1975) Converging evidence for the functional significance of imagery in problem solving. Cognition, 3:4, 359 - 375.

SMALLEY, N.S. (1974) Evaluating a rule against possible instances. British Journal of Psychology, 65, 293 - 304.

SNODGRASS, J.G. (1980) Towards a model for picture and word processing. In P.A. Kolers, M.E. Wrolstad & H. Bouma (Eds): Processing of visible language, Volume 2. New York: Plenum Press.

SNODGRASS, J.G. (1984) Concepts and their surface representations. Journal of Verbal Learning and Verbal Behaviour, 23, 3 - 22.

SOKOLOV, A.N. (1972) Inner speech and thought. London: Plenum Press.

STANDING, L. (1973) Learning 10,000 pictures. Quarterly Journal of Experimental Psychology, 25, 207 - 222.

STAUDENMAYER, H. (1975) Understanding conditional reasoning with meaningful propositions. In R.J. Falmagne (Ed): Reasoning: Representation and process. New York: Wiley.

STERNBERG, R.J. (1980) Representation and process in linear syllogistic reasoning. Journal of Experimental Psychology: General, 109:2, 119 - 159.

STERNBERG, R.J. & WEIL, E.M. (1980) An Aptitude X Strategy interaction in linear syllogistic reasoning. Journal of Educational Psychology, 72:2, 226 - 239.

STROOP, J.R. (1935) Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 18, 643 - 662.

TAPLIN, J.E. (1971) Reasoning with conditional sentences. Journal of Verbal Learning and Verbal Behaviour, 10, 219 - 225.

TAPLIN, J.E. & STAUDENMAYER, H. (1973) Interpretation of abstract conditional sentences in deductive reasoning. Journal of Verbal Learning and Verbal Behaviour, 12, 530 -542.

TRABASSO, T. (1972) Mental operations in language comprehension. In R.O. Freedle & J.B. Carroll (Eds): Language comprehension and the acquisition of knowledge. Washington, D.C.: Winston.

UNDERWOOD, G. & HOLT, P. O'B. (1979) Cognitive skills in the reading process: A review. Journal of Research in Reading, 2, 82 - 94.

VALLAR, G. & BADDELEY, A.D. (1982) Short-term forgetting and the articulatory loop. Quarterly Journal of Experimental Psychology, 34A, 53 - 60.

VAN DUYNE, P.C. (1973) A short note on Evans' criticism of reasoning experiments and his matching bias hypothesis. Cognition, 2, 239 -242.

VYGOTSKY, L.S. (1962) Thought and language. Cambridge, Massachusetts: The M.I.T. Press.

WASON, P.C. (1960) On the failure to eliminate hypotheses in a conceptual task. Quarterly Journal of Experimental Psychology, 12, 129 - 140.

WASON, P.C. (1966) Reasoning. In B.M. Foss (Ed): New horizons in psychology 1. Harmondsworth: Penguin.

WASON, P.C. (1968) Reasoning about a rule. Quarterly Journal of Experimental Psychology, 20, 273 - 281.

WASON, P.C. (1969) Regression in reasoning. British Journal of Psychology, 60, 471 - 480.

WASON, P.C. (1983) Realism and rationality in the selection task. In J.St.B.T. Evans (Ed): Thinking and reasoning: Psychological approaches. London: Routledge and Kegan Paul.

WASON, P.C. & EVANS, J.St.B.T. (1975) Dual processes in reasoning? _Cognition_, _3_, 141 - 154.

WASON, P.C. & JOHNSON-LAIRD, P.N. (1970) A conflict between selecting and evaluating information in an inferential task. _British Journal of Psychology_, _61_, 509 - 515.

WASON, P.C. & JOHNSON-LAIRD, P.N. (1972) _Psychology of reasoning: Structure and content_. London: Batsford.

WASON, P.C. & SHAPIRO, D. (1971) Natural and contrived experience in a reasoning problem. _Quarterly Journal of Experimental Psychology_, _23_, 63 - 71.

WATSON, J.B. (1930) _Behaviourism_. Chicago: University of Chicago Press.

WAUGH, N.C. & NORMAN, D.A. (1965) Primary memory. _Psychological Review_, _72_, 89 -104.

WHITE, K., SHEEHAN, P.W. & ASHTON, R. (1977) Imagery assessment: A survey of self-report measures. _Journal of Mental Imagery_, _1_, 145 - 170.

WICKELGREN, W.A. (1965) Acoustic similarity and intrusion errors in short-term memory. _Journal of Experimental_ Psychology, _70_, 102 - 108.

YUILLE, J.C. (1983) _Imagery, memory and cognition: Essays in honor of Allan Paivio (Edited)_. Hillsdale, N.J.: Lawrence Erlbaum Associates.

ZIMLER, J. & KEENAN, J.M. (1983) Imagery in the congenitally blind: How visual are visual images? _Journal of Experimental Psychology: Learning, Memory and Cognition_. _9:2_, 269 - 282.

APPENDICES A - J.

APPENDIX A.

Current Psychological Research (1981), 1:2, 139–147

Competing with Reasoning: A Test of the Working Memory Hypothesis

J. St B. T. EVANS & P. G. BROOKS

Plymouth Polytechnic

An experiment is reported in which subjects attempted conditional reasoning problems while concurrently articulating a series of digits, with or without memory load. Logical performance was not impaired by the competing tasks and the latency of responding was actually faster under concurrent articulation, without memory load, than in a control group. The results are discussed with reference to the Baddeley and Hitch (1974) model of working memory.

Baddeley and Hitch (1974) suggested that the short-term memory store acts as a working memory underlying various cognitive activities such as reading and reasoning. Their model of working memory has two components: a central executive and an articulatory loop. The latter not only permits rehearsal of items in short-term store, but is also supposed to assist in the processing of word-order information (see also Baddeley, 1979). Both rehearsal and word-order functions would seem to be required for the solution of verbal reasoning problems, in which separate pieces of information need to be compared in order for deductions to be made.

The working memory model may be assessed by use of competing task methodology. The logic of such methods is essentially as follows: if task A and B each require the use of a common mechanism, then they will compete for its use, with consequent interference. A famous example is Brooks' (1967, 1968) demonstration that tasks requiring mental imagery are subject to interference by concurrent perceptual tasks utilizing the same modality. In the case of the articulatory loop, the accepted interference task is that of concurrent speaking aloud of irrelevant verbal material. This technique, often referred to as articulatory suppression, has been used in a number of recent studies of memory and reading tasks (see Baddeley, 1976, 1979). It appears that the effect of suppression is to deny access of visually presented words to the loop. A central executive is assumed to be partially occupied by the requirement to hold a short-term memory load (cf. Baddeley & Hitch, 1974).

0143-3887/81/0102-139 $02.00

An investigation of working memory and reasoning was reported by Hitch and Baddeley (1976). For example, in their Experiment III, subjects were asked to solve a reasoning problem under one of three conditions:

1. Control. No accompanying task.
2. Articulation. Subject said either 'the-the-the' or '1-2-3-4-5-6' repeatedly while performing the task.
3. Memory. Subjects were presented with a different six-digit number at the start of each trial and were required to repeat it aloud. This adds a short-term memory load to the concurrent articulation task.

Reasoning accuracy was not significantly disrupted by the interference task, although a significant main effect on response latency was observed. Control was the fastest condition, Memory the slowest and Articulation intermediate. However, the reasoning task adopted by Hitch and Baddeley was relatively simple compared with those normally used in the study of logical reasoning. We therefore decided to repeat their experiment with conditional reasoning problems, which are known to produce systematic errors under normal conditions (see Evans, 1982). There is consequently no risk of a 'floor' effect, and we might expect an increase in logical errors under either interference condition. Since Hitch and Baddeley observed interference on the latency scores, however, a similar measure was also employed in the present study.

The experiment employed a set of reasoning problems in which subjects were invited to make each of four inferences for each of four rules, thus producing 16 distinct problems (see Table 1). In formal logic only two of these inferences, MP and MT, are considered valid. However, in natural language the conditional sentence

Table 1. The different logical forms used in the experiment

	Inference							
	MP		DA		AC		MT	
Rule	Given	Conclusion	Given	Conclusion	Given	Conclusion	Given	Conclusion
If p then q	p	q	not p	not q	q	p	not q	not p
If p then not q	p	not q	not p	q	not q	p	q	not p
If not p then q	not p	q	p	not q	q	not p	not q	p
If not p then not q	not p	not q	p	q	not q	not p	q	p

MP Modus Ponens
DA Denial of the Antecedent
AC Affirmation of the Consequent
MT Modus Tollens

is sometimes used to express equivalence, rather than implication, in which case all four inferences are valid.

It can be seen from Table 1 that each of the four inferences results in an affirmative conclusion on two rules, and a negative conclusion on two rules. Previous research has shown that, all else being equal, subjects prefer to accept negative conclusions (e.g. Evans, 1972, 1977; Roberge, 1978; Pollard & Evans, 1980), which has been interpreted as a non-logical response bias.

Performance of subjects on the 16 problems was assessed under similar conditions to those of Hitch and Baddeley (1976) except that one of the Articulation groups (the-the-the) was dropped.

METHOD

Design

Three experimental groups, each consisting of six male and six female subjects, were tested on an inference task using conditional rules.

In the Control group subjects were required to remain silent during the task. In the Articulation group subjects were instructed to repeat the heavily overlearned counting sequence 'one-two-three-four-five-six' repeatedly at a rate of about four words per second. In the Memory group the subjects heard a spoken sequence of six random digits at the start of each trial. They were required to speak the sequence repeatedly at a rate of about four words per second. In this condition alone the sequence to be articulated was changed on each trial.

Within groups, subjects were required to evaluate each of the 16 types of inference shown in Table 1. To assess the effects of practice, three blocks of the 16 problems using different lexical content were constructed. A different randomized order of presentation of problems within each block was used during each session. Also, the order of presentation of the three blocks was varied systematically within groups. Each subject thus received a total of 48 problems.

Subjects

Thirty-six students at Plymouth Polytechnic, having no previous experience with this type of task, served as subjects on a paid volunteer basis. They were tested individually.

Task and Materials

Subjects were presented with the two premises of each argument, together with the appropriate valid or fallacious conclusion (cf. Table 1). They were required to decide whether or not the conclusion necessarily followed logically from the premises.

All the arguments concerned shape-colour relationships. One of four shapes (triangle, circle, square or diamond), together with one of four colours (red, blue, yellow or green), were named in systematically randomized combinations for each problem.

The materials may be illustrated with the following sample problem which uses an AC inference with an 'If not p then not q' rule.

Given:

1. If it is not a triangle then it is not red.
2. It is not red.

Conclusion:

It is not a triangle.

Problems were presented on a two-field tachistoscope whose onset and offset were synchronized with an automatic timer.

The subject's task was to decide whether or not the conclusion necessarily followed logically and to signal his response by pressing a toggle switch to indicate 'yes' or 'no'.

Procedure

Each subject was read a standard set of instructions in which the nature of the logical task was explained. It was emphasized that subjects should press the 'yes' key if they considered that the conclusion necessarily followed from the premises. If not, they should press the 'no' key. They were told that they would be timed, but that accuracy was more important than speed.

Subjects were then given eight practice trials which differed from the test problems in that disjunctive rules were used in place of conditionals. Following the practice trials, the instructions to the three groups were as follows.

Control group Do you understand what you have to do? We will now start on the main problems. I will give the signal 'Ready, Start' as in the practice session ...

Articulation group Do you understand what you have to do? We will now start on the main problems. I would like you to carry out an additional task while solving these problems. Please say aloud the sequence of numbers '1, 2, 3, 4, 5, 6'. Speak them repeatedly at an even pace, like this (DEMONSTRATE), when I give you the signal 'Ready, Start'. You may stop counting as soon as you have indicated your answer to the problem by pushing the key. I cannot tell you the purpose of this procedure at this stage, but I will be happy to discuss it with you after the experiment. Do you understand what you have to do? I will give you the signal, 'Ready, Start' as in the practice session ...

Memory group Do you understand what you have to do? We will now start on the main problems. I would like you to carry out an additional task while solving these problems. Please say aloud the sequence of numbers which I will give you after the 'Ready' signal. Speak them repeatedly at an even pace, like this (DEMONSTRATE), when I give you the signal 'Ready, Number'. You may stop counting as soon as you have indicated your answer to the problem by pushing the key. I cannot tell you the purpose of this procedure at this stage, but I will be happy to discuss it with you after the experiment. Do you understand what you have to do? I will give you the signal 'Ready, Number' as in the practice session ...

In the Articulation and Memory groups the problems were presented immediately after the subject commenced articulation. The experimenter was vigilant for any perceptible drop in the articulation rate throughout the experiment, and prompted subjects to maintain the required rate as and when necessary.

RESULTS

The percentage frequency with which subjects accepted each inference is shown in Table 2, together with a summary of the significant effects in the analysis of variance. In order to provide reasonable numbers for the analysis, the subjects'

Table 2

(a) The percentage of 'yes' responses (arguments accepted) in each condition broken down by polarity of conclusion (n = 12 in each group)

Group	Polarity of conclusion	Inference				$\bar{X}$
		MP	DA	AC	MT	
Control	Affirmative	97	46	76	43	66
	Negative	97	63	86	77	81
Articulation	Affirmative	90	46	69	35	60
	Negative	93	68	69	67	74
Memory	Affirmative	90	67	88	58	76
	Negative	85	58	82	77	75
	$\bar{X}$	92	58	79	59	

(b) Significant effects in the ANOVA

Within subjects	d.f.	F	P
Inference	1,33	40.34	<0.001
Conclusion	1,33	21.42	<0.001
Conclusion x Inference	1,33	9.91	<0.01
Conclusion x Groups	2,33	5.76	<0.01

three attempts at each problem were combined in the frequency scores. Since the effect of negatives is known to operate through the polarity (affirmative/negative) of the conclusion evaluated, the data were organized in this way for the analysis of variance. The factors were thus Groups (three levels), Inference (four levels), and Conclusion (two levels), the last two being within subject factors.

As would be expected from previous work, the rate of acceptance of inferences was significantly influenced by their logical classification (MP, DA, AC or MT). If the competing verbal tasks interfered with subjects' ability to make such logical discriminations, then a Groups x Inference interaction would be expected. No such interaction occurred. Owing to the ambiguity of the conditional referred to previously we can only clearly recognize responses as 'correct' in the cases of MP and MT, which should be accepted. The percentage correct on these inferences combined are: Control, 79 per cent; Articulation, 71 per cent; and Memory, 78 per cent. There is, then, little support for the hypothesis that competing articulation tasks (with or without memory load) should disrupt logical reasoning.

The results, as a whole, show high incidence of logical errors, as is characteristic of conditional reasoning tasks. The normal systematic bias to prefer negative conclusions was observed and its interaction with Inference is consistent with previous work. Conclusion bias was not, however, expected to interact with Groups. It can be seen that the bias is completely absent in the Memory group.

Hitch and Baddeley (1976) also found little disruption of performance in terms of errors, but reconciled the data with their working memory model with reference to the latency data. In their study, interference conditions significantly increased latencies, and also interacted with the complexity of the problems. More complex problems - with greater demands on working memory - were slowed down more under interference. Our latency analysis (Table 3) tests for replication of these effects. The analysis of variance differed from that of the frequency data by inclusion of a Blocks factor, and by the inclusion of a Rules rather than Conclusion factor to test for the effect of negatives. Previous research (Evans, 1977; Evans & Newstead, 1977) indicates that the presence of a negative in either component slows negatives in an additive manner.

Although the Groups factor was significant in the analysis of variance, the direction of effect was unexpected. The order of latency was Articulation fastest and Memory slowest, with Control in between. Two-tailed t tests revealed that the acceleration of Articulation responses relative to Control was significant ($t_{22} = 2.27$, $P < 0.05$), as was the slowing of Memory relative to Control ($t_{22} = 2.14$, $P < 0.05$). Groups did not significantly interact with any other factor.

DISCUSSION AND CONCLUSIONS

The present study, like that of Hitch and Baddeley (1976), fails to find evidence that imposing loads on working memory increases the frequency of reasoning errors, despite the much greater complexity of the reasoning task used. The latency analysis, however, has shown a quite different pattern from theirs, and one that is more difficult to reconcile with the Baddeley and Hitch (1974) model.

Although articulation with memory load significantly increased latencies, the effect was not interactive with variables affecting problem complexity, such as the presence or absence of negative components. It is hard, therefore, to interpret the slowing of response times in terms of a competing load on the central executive. A

Table 3

(a) The mean latency in seconds for each condition (n = 12 in each group)

Group	Rule	Inference MP	DA	AC	MT	$\bar{X}$
Control	If p then q	6.03	8.30	5.94	7.33	6.90
	If p then not q	6.67	10.00	9.85	8.71	8.81
$\bar{X}$ = 8.58	If not p then q	6.96	9.87	8.28	10.08	8.80
	If not p then not q	7.64	10.01	9.85	11.85	9.84
Articulation	If p then q	4.23	5.70	4.29	6.08	5.07
	If p then not q	5.36	6.95	6.73	5.71	6.19
$\bar{X}$ = 6.13	If not p then q	4.83	6.96	5.98	7.94	6.43
	If not p then not q	5.61	7.22	7.41	7.11	6.84
Memory	If p then q	9.39	11.70	9.37	10.26	10.18
	If p then not q	9.29	11.92	11.36	13.22	11.45
$\bar{X}$ = 11.62	If not p then q	9.34	13.83	12.83	14.64	12.63
	If not p then not q	9.53	12.67	13.53	13.13	12.22
	$\bar{X}$	7.07	9.59	8.77	9.67	

(b) Significant effects in the ANOVA

	d.f.	F	P
Between subjects:			
Groups	2,30	9.66	<0.001
Within subjects:			
Blocks	1,30	18.51	<0.001
Rules	1,30	14.97	<0.001
Inference	1,30	24.89	<0.001

much more plausible explanation is that subjects did not attempt to solve the reasoning tasks until they had said the novel digits several times. Thus, in effect, the digits are rehearsed and committed to long-term memory before they attempt the problems. Consequently, the lack of increase in logical errors under memory load cannot be taken as evidence against the working memory model. It rather points to a weakness of the method which does not impose a concurrent load on working memory, as intended. This account, however, does not explain the suppression of conclusion bias under memory load, discussed later.

The significant acceleration of response times under straight concurrent articulation is hard to reconcile with the Baddeley and Hitch model. In the light of experimental evidence, Baddeley (1979) has weakened the claim for the articulatory loop, regarding it as an optional control strategy. For reasons given in the introduction his view still seems to imply that it should be helpful to conditional reasoning, but the very least he could expect is for it to have no effect. An explanation of why suppression of the loop should actually facilitate reasoning speed is beyond the scope of the model.

There are two lines of explanation for the acceleration effect. It could be specifically related to speech mechanisms. It is possible that on verbal reasoning tasks subjects habitually subvocalize. Certainly, electromyographic studies find evidence of micromovements and electrical potentials in speech organs during problem solving, especially when the problems are complex or novel (see McGuigan, 1966; Sokolov, 1972). Such subvocalization might slow down thinking, as it is thought to slow reading. Hence concurrent articulation, which prevents subvocalization of the problem content, may speed up the thought process. This view implies that, contrary to traditional theories of 'inner speech' (cf. McGuigan, 1966), such a process has little functional value. (It should, however, be noted that there was a non-significant increase in errors in the Articulation group). A second line of explanation is that the repetition of an overlearned sequence has a very general effect of increasing concentration or level of arousal. The effect may not be specifically related to speech at all. However, there is no reason to suppose a general facilitatory effect of concurrent articulation on problem solving. Sokolov (1972) reviews a number of studies of mental arithmetic under various techniques of articulatory suppression. Usually the competing task slowed performance, although an acceleration effect for exceptionally skilled and practised subjects sometimes occurred. There is also no acceleration of anagram solution speed under concurrent articulation in experiments by Peterson (1969) and an unpublished study by the present authors.

An unexpected finding in the frequency analysis was the apparent disappearance of the 'negative conclusion bias' under memory load. From previous research we would expect the bias to be strongly marked on DA and MT, weak but present in AC, and absent on MP, where inference rates are very high on all rules. Inspection of the Control group data in Table 2 reveals precisely this pattern. The anomalous results in the Memory group are the high rates of affirming DA and AC when the conclusion was affirmative (against the bias); MT actually showed the normal trend. Since the correctness of DA and AC is ambiguous, depending on how the subject interprets the conditional, one cannot say that memory load is necessarily interfering with reasoning. The result is also hard to interpret, since reasoning research has not established whether conclusion bias reflects a preference for negatives or an aversion to affirmatives. Only if the latter were the case could one regard the effect of memory load as 'releasing' the subject from the conclusion bias.

Allport (1980 a,b) has recently attacked the notion of 'general purpose limited capacity central processors' of the sort entailed by the Baddeley and Hitch model, and also questioned the value of competing task methodology. The present experiment lends some support to his views. Certainly the effects of the 'competing task' of concurrent articulation seem to be specific to the type of problem content used, and the present findings are not encouraging to those wishing to pursue a general-purpose concept of working memory.

ACKNOWLEDGEMENTS

The authors are grateful to Ian Dennis for a critical reading of an earlier draft of this paper. The research reported in this paper was supported by a grant from the Science Research Council.

REFERENCES

Allport, D.A. (1980a). Patterns and actions: cognitive mechanisms are content specific. In G. Claxton (ed.), Cognitive Psychology, New Directions. London: Routledge & Kegan Paul.

Allport, D.A. (1980b). Attention and performance. In G. Claxton (ed.), Cognitive Psychology, New Directions. London: Routledge & Kegan Paul.

Baddeley, A.D. (1976). The Psychology of Memory. New York: Basic Books.

Baddeley, A.D. (1979). Working memory and reading. In P.A. Kolers, M.E. Wrolstad & M. Bourne (eds.), The Proceedings of the Conference on the Processing of Visible Language, Eindhoven.

Baddeley, A.D. & Hitch, G.J. (1974). Working memory. In G.A. Bower (ed.), The Psychology of Learning and Motivation, 8, 47-90.

Brookes, L.R. (1967). The suppression of visualisation by reading. Quarterly Journal of Experimental Psychology, 19, 289-299.

Brookes, L.R. (1968). Spatial and verbal components of the act of recall. Canadian Journal of Psychology, 22, 349-368.

Evans, J. St B.T. (1972). Reasoning with negatives. British Journal of Psychology, 63, 213-219.

Evans, J. St B.T. (1977). Linguistic factors in reasoning. Quarterly Journal of Experimental Psychology, 29, 297-306.

Evans, J. St B.T. (1982). The Psychology of Deductive Reasoning. London: Routledge & Kegan Paul (in press).

Evans, J. St B.T. & Newstead, S.E. (1977). Language and reasoning: a study of temporal factors. Cognition, 5, 265-283.

Hitch, G.J. & Baddeley, A.D. (1976). Verbal reasoning and working memory. Quarterly Journal of Experimental Psychology, 28, 603-621.

McGuigan, F.J. (1966). Thinking: Studies of Covert Language Processes. New York: Appleton Century.

Peterson, L.R. (1969). Concurrent verbal activity. Psychological Review, 76, 376-386.

Pollard, P. & Evans, J. St B.T. (1980). The influence of logic on conditional reasoning performances. Quarterly Journal of Experimental Psychology, 32, 605-624.

Roberge, J.J. (1978). Linguistic and psychometric factors in propositional reasoning. Quarterly Journal of Psychology, 30, 705-716.

Sokolov, A.N. (1972). Inner Speech and Thought. London: Plenum Press.

J. ST B.T. EVANS . Department of Psychology, Plymouth Polytechnic, Plymouth PL4 8AA.

Date of acceptance for publication: April 1981

APPENDIX B.

THE ANALYSIS OF VARIANCE TABLES FOR EXPERIMENT I.

Significance levels (F PR) for all repeated measures factors use conservative degrees of freedom (see Edwards, 1967).

(i) Response Frequencies.

Factors are: G (Groups), S (Sex), I (Inference), C (Conclusion Type), B (Blocks).

Source	DF	SS	MS	VR	F PR
Between S's.					
G	2	4.2315	2.116	0.68	ns
S	1	1.9456	1.946	0.63	ns
G.S	2	2.0093	1.005	0.32	ns
RESID	30	93.1042	3.104		
Within S's.					
I	3	69.9016	23.301	37.65	<.001
G.I	6	5.8796	0.980	1.58	ns
S.I	3	0.7627	0.254	0.41	ns
G.S.I	6	0.7130	0.119	0.19	ns
RESID	90	55.7014	0.619		
C	1	7.9734	7.973	20.76	<.001
G.C	2	4.2870	2.144	5.58	<.01
S.C	1	0.0104	0.010	0.03	ns
G.S.C	2	0.7500	0.375	0.98	ns
RESID	30	11.5208	0.384		
B	2	1.1551	0.578	1.11	ns
G.B	4	1.1019	0.276	0.53	ns
S.B	2	0.2245	0.112	0.22	ns
G.S.B	4	2.6019	0.651	1.25	ns
RESID	60	31.2500	0.521		
I.C	3	10.9942	3.665	10.25	<.01
G.I.C	6	1.6759	0.279	0.78	ns
S.I.C	3	1.2905	0.430	1.20	ns
G.S.I.C	6	3.1574	0.526	1.47	ns
RESID	90	32.1736	0.358		
I.B	6	2.0671	0.345	1.58	ns
G.I.B	12	2.3704	0.198	0.91	ns
S.I.B	6	2.1088	0.352	1.61	ns
G.S.I.B	12	3.9259	0.327	1.50	ns
RESID	180	39.1944	0.218		
C.B	2	0.0718	0.036	0.14	ns
G.C.B	4	0.1574	0.039	0.16	ns
S.C.B	2	0.0486	0.024	0.10	ns
G.S.C.B	4	0.3889	0.097	0.39	ns
RESID	60	15.1667	0.253		
I.C.B	6	1.1690	0.195	0.79	ns
G.I.C.B	12	4.2963	0.358	1.45	ns
S.I.C.B	6	1.3032	0.217	0.88	ns
G.S.I.C.B	12	1.6759	0.140	0.57	ns
RESID	180	44.3889	0.247		

(ii) Response Latencies (logarithmically transformed).

Factors are: G (Groups), S (Sex), I (Inference), R (Rules), B (Blocks).

Source	DF	SS	MS	VR	F PR
Between S's.					
G	2	23.0176	11.509	9.66	<.001
S	1	3.1141	3.114	2.61	ns
G.S	2	0.1997	0.100	0.08	ns
RESID	30	35.7423	1.19		
Within S's.					
I	3	4.1325	1.377	24.89	<.001
G.I	6	0.0650	0.011	0.20	ns
S.I	3	0.0468	0.016	0.28	ns
G.S.I	6	0.1333	0.022	0.40	ns
RESID	90	4.9806	0.06		
R	3	1.9476	0.649	14.97	<.001
G.R	6	0.3171	0.053	1.22	ns
S.R	3	0.1088	0.036	0.84	ns
G.S.R	6	0.2214	0.037	0.85	ns
RESID	90	3.9041	0.04		
B	2	3.5436	1.772	18.51	<.001
G.B	4	0.5883	0.147	1.54	ns
S.B	2	0.0289	0.014	0.15	ns
G.S.B	4	0.7295	0.182	1.91	ns
RESID	60	5.7428	0.10		
I.R	9	0.7849	0.087	3.11	ns
G.I.R	18	0.1954	0.011	0.39	ns
S.I.R	9	0.1937	0.022	0.77	ns
G.S.I.R	18	0.5427	0.030	1.07	ns
RESID	270	7.5791	0.03		
I.B	6	0.2732	0.046	1.70	ns
G.I.B	12	0.3171	0.026	0.99	ns
S.I.B	6	0.2504	0.042	1.56	ns
G.S.I.B	12	0.5376	0.045	1.68	ns
RESID	180	4.8080	0.03		
R.B	6	0.0342	0.006	0.24	ns
G.R.B	12	0.3351	0.028	1.19	ns
S.R.B	6	0.2641	0.044	1.88	ns
G.S.R.B	12	0.1749	0.015	0.62	ns
RESID	180	4.2122	0.02		
I.R.B	18	0.6451	0.036	1.26	ns
G.I.R.B	36	0.6518	0.018	0.64	ns
S.I.R.B	18	0.4032	0.022	0.79	ns
G.S.I.R.B	36	0.8229	0.023	0.80	ns
RESID	540	15.3704	0.03		

APPENDIX C.

THE ANALYSIS OF VARIANCE TABLES FOR EXPERIMENT II.

Significance levels (F PR) for all repeated measures factors use conservative degrees of freedom (see Edwards, 1967).

(i) <u>Response Frequencies.</u>

Factors are: G (Groups), IN (Instance), M (Matching Case), B (Blocks).

(a) <u>TT as 'Conforms'.</u>

Source	DF	SS	MS	VR	F PR
<u>Between S's.</u>					
G	2	1.Ø139	Ø.5Ø7	3.74	<.Ø5
IN	1	Ø.84Ø3	Ø.84Ø	6.21	<.Ø25
G.IN	2	Ø.18Ø6	Ø.Ø9Ø	Ø.67	ns
RESID	42	5.6875	Ø.135		
<u>Within S's.</u>					
M	3	4.9444	1.648	15.8Ø	<.ØØ1
G.M	6	1.Ø972	Ø.183	1.75	ns
IN.M	3	Ø.52Ø8	Ø.174	1.66	ns
G.IN.M	6	Ø.125Ø	Ø.Ø21	Ø.2Ø	ns
RESID	126	13.1458	Ø.1Ø4		
B	2	Ø.7743	Ø.387	5.66	<.Ø25
G.B	4	Ø.6944	Ø.174	2.54	ns
IN.B	2	Ø.17Ø1	Ø.Ø85	1.24	ns
G.IN.B	4	Ø.2778	Ø.Ø69	1.Ø1	ns
RESID	84	5.75ØØ	Ø.Ø68		
M.B	6	Ø.42Ø1	Ø.Ø7Ø	Ø.97	ns
G.M.B	12	1.5694	Ø.131	1.81	ns
IN.M.B	6	Ø.8Ø21	Ø.134	1.85	ns
G.IN.M.B	12	Ø.7Ø83	Ø.Ø59	Ø.82	ns
RESID	252	18.1667	Ø.Ø72		

(b) <u>TF as 'Conflicts'.</u>

Source	DF	SS	MS	VR	F PR
<u>Between S's.</u>					
G	2	1.7222	0.861	4.82	<.025
IN	1	2.1267	2.127	11.89	<.01
G.IN	2	0.8472	0.424	2.37	ns
RESID	42	7.5104	0.179		
<u>Within S's.</u>					
M	3	7.5330	2.511	16.77	<.001
G.M	6	1.1389	0.190	1.27	ns
IN.M	3	0.6441	0.215	1.43	ns
G.IN.M	6	0.7361	0.123	0.82	ns
RESID	126	18.8646	0.150		
B	2	2.0035	1.002	6.56	<.025
G.B	4	0.0694	0.017	0.11	ns
IN.B	2	1.1910	0.596	3.90	ns
G.IN.B	4	1.0694	0.267	1.75	ns
RESID	84	12.8333	0.153		
M.B	6	0.4410	0.074	0.72	ns
G.M.B	12	1.1528	0.096	0.93	ns
IN.M.B	6	0.2257	0.038	0.37	ns
G.IN.M.B	12	1.0972	0.091	0.89	ns
RESID	252	25.9167	0.103		

(c)FT as 'Irrelevant'.

Source	DF	SS	MS	VR	F PR
Between S's.					
G	2	7.4618	3.731	3.82	<.Ø5
IN	1	Ø.84Ø3	Ø.84Ø	Ø.86	ns
G.IN	2	2.3993	1.2ØØ	1.23	ns
RESID	42	41.Ø2Ø8	Ø.977		
Within S's.					
M	3	9.125Ø	3.Ø42	14.38	<.ØØ1
G.M	6	2.26Ø4	Ø.377	1.78	ns
IN.M	3	Ø.Ø347	Ø.Ø12	Ø.Ø6	ns
G.IN.M	6	Ø.434Ø	Ø.Ø72	Ø.34	ns
RESID	126	26.6458	Ø.212		
B	2	2.191Ø	1.Ø96	7.34	<.Ø1
G.B	4	1.5694	Ø.392	2.63	ns
IN.B	2	Ø.Ø243	Ø.Ø12	Ø.Ø8	ns
G.IN.B	4	Ø.6736	Ø.168	1.13	ns
RESID	84	12.5417	Ø.149		
M.B	6	Ø.4Ø63	Ø.Ø68	Ø.54	ns
G.M.B	12	Ø.7Ø83	Ø.Ø59	Ø.47	ns
IN.M.B	6	Ø.684Ø	Ø.114	Ø.91	ns
G.IN.M.B	12	1.6597	Ø.138	1.11	ns
RESID	252	31.5417	Ø.125		

(d) <u>FF as 'Irrelevant'.</u>

Source	DF	SS	MS	VR	F PR
<u>Between S's.</u>					
G	2	7.0868	3.543	3.53	<.05
IN	1	1.1736	1.174	1.17	ns
G.IN	2	1.6285	0.814	0.81	ns
RESID	42	42.1667	1.004		
<u>Within S's.</u>					
M	3	8.9028	2.968	16.81	<.001
G.M	6	2.4410	0.407	2.30	ns
IN.M	3	0.4792	0.160	0.90	ns
G.IN.M	6	0.4271	0.071	0.40	ns
RESID	126	22.2500	0.177		
B	2	4.1076	2.054	14.74	<.001
G.B	4	0.6736	0.168	1.21	ns
IN.B	2	0.3576	0.179	1.28	ns
G.IN.B	4	0.1528	0.038	0.27	ns
RESID	84	11.7083	0.139		
M.B	6	0.8785	0.146	1.26	ns
G.M.B	12	3.0903	0.258	2.21	ns
IN.M.B	6	0.4896	0.082	0.70	ns
G.IN.M.B	12	1.1667	0.097	0.83	ns
RESID	252	29.3750	0.117		

(ii) Response Latencies (logarithmically transformed).

(a) Comprehension Times.

Factors are: G (Groups), R (Rules), B (Blocks).
Logical Case was included as a dummy factor. This did not affect the analysis and is not included in the Table.

Source	DF	SS	MS	VR	F PR
Between S's.					
G	2	15.3879	7.694	3.95	<.05
RESID	45	87.6957	1.949		
Within S's.					
R	3	5.0241	1.675	42.85	<.001
G.R	6	0.9041	0.151	3.86	<.05
RESID	135	5.2768	0.039		
B	2	1.4770	0.738	6.88	<.025
G.B	4	0.0911	0.023	0.21	ns
RESID	90	9.6535	0.107		
R.B	6	0.3158	0.053	2.19	ns
G.R.B	12	0.2128	0.018	0.74	ns
RESID	270	6.5341	0.02		

(b) Verification Times.

Factors are: G (Groups), IN (Instance), LC (Logical Case), R (Rules), B (Blocks).

Source	DF	SS	MS	VR	F PR
Between S's.					
G	2	7.6798	3.840	4.50	<.025
IN	1	1.1138	1.114	1.30	ns
G.IN	2	0.1682	0.084	0.10	ns
RESID	42	35.8680	0.85		
Within S's.					
LC	3	6.7867	2.262	39.46	<.001
G.LC	6	1.6026	0.267	4.66	<.025
IN.LC	3	0.1224	0.041	0.71	ns
G.IN.LC	6	0.1911	0.032	0.56	ns
RESID	126	7.2230	0.06		
R	3	9.8287	3.276	68.25	<.001
G.R	6	0.5684	0.095	1.97	ns
IN.R	3	0.0999	0.033	0.69	ns
G.IN.R	6	0.3417	0.057	1.19	ns
RESID	126	6.0482	0.05		
B	2	9.1642	4.582	58.16	<.001
G.B	4	0.3017	0.075	0.96	ns
IN.B	2	0.0548	0.027	0.35	ns
G.IN.B	4	0.3441	0.086	1.09	ns
RESID	84	6.6174	0.08		
LC.R	9	1.6172	0.180	3.05	ns
G.LC.R	18	1.1402	0.063	1.07	ns
IN.LC.R	9	0.7498	0.083	1.41	ns
G.IN.LC.R	18	0.6583	0.037	0.62	ns
RESID	378	22.2922	0.06		
LC.B	6	0.3254	0.054	1.31	ns
G.LC.B	12	0.5535	0.046	1.11	ns
IN.LC.B	6	0.1825	0.030	0.73	ns
G.IN.LC.B	12	0.4668	0.039	0.94	ns
RESID	252	10.4619	0.04		
R.B	6	0.2297	0.038	1.04	ns
G.R.B	12	0.9044	0.075	2.05	ns
IN.R.B	6	0.2479	0.041	1.12	ns
G.IN.R.B	12	0.3861	0.032	0.87	ns
RESID	252	9.2846	0.04		
LC.R.B	18	0.5407	0.030	0.84	ns
G.LC.R.B	36	1.5691	0.044	1.21	ns
IN.LC.R.B	18	0.6293	0.035	0.97	ns
G.IN.LC.R.B	36	1.8552	0.052	1.44	ns
RESID	756	27.1367	0.04		

APPENDIX D.

THE STIMULUS WORDS USED IN EXPERIMENTS III AND V.

(i) Classified by the colour of the ink in which they were printed:

Ink Colour			
Red	Blue	Yellow*	Green
Pram	Joe	Goat	Brain
Doe	Throat	Train	Sam
Tram	Flo	Stoat	Vein
Jane	Ram	Toe	Boat

*Gold in Experiment V.

(ii) Classified by their rhyming characteristics:

Rhyme With			
Feign	Gramme	Glow	Vote
Brain	Pram	Doe	Throat
Vein	Ram	Flo	Goat
Train	Tram	Joe	Stoat
Jane	Sam	Toe	Boat

(iii) Classified by their semantic category:

Christian Name	Part of the Body	Means of Transport	Type of Animal
Jane	Throat	Pram	Doe
Joe	Toe	Tram	Ram
Flo	Brain	Train	Goat
Sam	Vein	Boat	Stoat

APPENDIX E.

THE ANALYSIS OF VARIANCE TABLES FOR EXPERIMENT III.

Significance levels (F PR) for all repeated measures factors use conservative degrees of freedom (see Edwards, 1967).

(i) Response Frequencies.

Factors are: G (Groups), M (Matching Case), C (Conditions).

(a) TT as 'Conforms'.

Source	DF	SS	MS	VR	F PR
Between S's.					
G	2	Ø.5Ø46	Ø.252	Ø.55	ns
RESID	33	15.2639	Ø.463		
Within S's.					
M	3	2.75ØØ	Ø.917	8.59	<.Ø1
G.M	6	Ø.68Ø6	Ø.113	1.Ø6	ns
RESID	99	1Ø.5695	Ø.1Ø7		
C	2	Ø.Ø185	Ø.ØØ9	Ø.1Ø	ns
G.C	4	1.Ø787	Ø.27Ø	2.85	ns
RESID	66	6.2631	Ø.Ø94		
M.C	6	Ø.5ØØØ	Ø.Ø83	1.14	ns
G.M.C	12	1.Ø694	Ø.Ø89	1.22	ns
RESID	198	14.43Ø6	Ø.Ø73		

(b) TF as 'Conflicts'.

Source	DF	SS	MS	VR	F PR
Between S's.					
G	2	Ø.Ø88Ø	Ø.Ø44	Ø.12	ns
RESID	33	12.1528	Ø.368		
Within S's.					
M	3	1.7963	Ø.599	2.93	ns
G.M	6	Ø.4676	Ø.Ø78	Ø.38	ns
RESID	99	2Ø.2361	Ø.2Ø4		
C	2	Ø.24Ø7	Ø.12Ø	1.ØØ	ns
G.C	4	1.12Ø4	Ø.28Ø	2.32	ns
RESID	66	7.9722	Ø.121		
M.C	6	1.4259	Ø.238	2.Ø2	ns
G.M.C	12	1.2685	Ø.1Ø6	Ø.9Ø	ns
RESID	198	23.3Ø56	Ø.118		

(c) FT as 'Irrelevant'.

Source	DF	SS	MS	VR	F PR
Between S's.					
G	2	3.8519	1.926	1.95	ns
RESID	33	32.5278	Ø.986		
Within S's.					
M	3	4.6Ø19	1.534	8.78	<.ØØ1
G.M	6	1.2593	Ø.21Ø	1.2Ø	ns
RESID	99	17.3Ø56	Ø.175		
C	2	Ø.56Ø2	Ø.28Ø	1.14	ns
G.C	4	Ø.5Ø93	Ø.127	Ø.52	ns
RESID	66	16.2639	Ø.246		
M.C	6	Ø.7176	Ø.12Ø	1.ØØ	ns
G.M.C	12	1.Ø463	Ø.Ø87	Ø.73	ns
RESID	198	23.5695	Ø.119		

(d) FF as 'Irrelevant'.

Source	DF	SS	MS	VR	F PR
Between S's.					
G	2	3.9074	1.954	1.76	ns
RESID	33	36.7292	1.113		
Within S's.					
M	3	3.3588	1.120	5.90	<.025
G.M	6	0.4259	0.071	0.37	ns
RESID	99	18.7986	0.190		
C	2	0.7269	0.363	1.50	ns
G.C	4	0.3982	0.100	0.41	ns
RESID	66	16.0417	0.243		
M.C	6	0.8843	0.147	1.22	ns
G.M.C	12	1.2685	0.106	0.87	ns
RESID	198	24.0139	0.121		

(ii) Response Latencies (logarithmically transformed).

(a) Comprehension Times.

Factors are: G (Groups), R (Rules), C (Conditions).
Logical Case was included as a dummy factor. This did not affect the analysis and is not included in the Table.

Source	DF	SS	MS	VR	F PR
Between S's.					
G	2	4.1427	2.071	1.96	ns
RESID	33	34.7928	1.05		
Within S's.					
R	3	0.3743	0.125	4.95	<.05
G.R	6	0.5398	0.090	3.57	<.05
RESID	99	2.4941	0.03		
C	2	2.5850	1.292	11.59	<.01
G.C	4	0.1117	0.028	0.25	ns
RESID	66	7.3618	0.11		
R.C	6	0.1063	0.018	0.68	ns
G.R.C	12	0.2150	0.018	0.69	ns
RESID	198	5.1769	0.03		

(b) <u>Verification Times.</u>

Factors are: G (Groups), LC (Logical Case), R (Rules), C (Conditions).

Source	DF	SS	MS	VR	F PR
<u>Between S's.</u>					
G	2	3.3207	1.660	0.88	ns
RESID	33	62.4911	1.89		
<u>Within S's.</u>					
LC	3	4.9331	1.644	19.13	<.001
G.LC	6	1.1394	0.190	2.21	ns
RESID	99	8.5116	0.09		
R	3	3.4747	1.158	22.72	<.001
G.R	6	0.2690	0.045	0.88	ns
RESID	99	5.0471	0.05		
C	2	3.4196	1.710	13.24	<.01
G.C	4	1.0385	0.260	2.01	ns
RESID	66	8.5261	0.13		
LC.R	9	0.8517	0.095	2.29	ns
G.LC.R	18	0.9149	0.051	1.23	ns
RESID	297	12.2834	0.04		
LC.C	6	0.4750	0.079	1.83	ns
G.LC.C	12	0.4146	0.035	0.80	ns
RESID	198	8.5672	0.04		
R.C	6	0.2515	0.042	1.26	ns
G.R.C	12	0.7505	0.063	1.88	ns
RESID	198	6.5735	0.03		
LC.R.C	18	1.6553	0.092	2.73	ns
G.LC.R.C	36	0.9401	0.026	0.78	ns
RESID	594	19.9960	0.03		

APPENDIX F.

THE STIMULUS MATERIALS USED IN EXPERIMENT IV.

(i) Letters

R	A	Z	K
X	V	U	F
G	H	C	M
L	S	Y	O

(ii) Words

Rose	Coal	Girl	Door
Book	Wall	Path	King
Lake	Ship	Fish	Neck
Hill	Moon	Tree	Bird

APPENDIX G.

THE ANALYSIS OF VARIANCE TABLES FOR EXPERIMENT IV.

Significance levels (F PR) for all repeated measures factors use conservative degrees of freedom (see Edwards, 1967).

(i) Response Latencies (logarithmically transformed).

Factors are: G (Groups), LW (Letters/Words), T (Truth Value), V (Sentence Voice), P (Polarity).

Source	DF	SS	MS	VR	F PR
Between S's.					
G	2	1.6660	0.833	4.73	<.025
RESID	21	3.6953	0.18		
Within S's.					
LW	1	0.0613	0.061	2.33	ns
G.LW	2	0.0069	0.003	0.13	ns
RESID	21	0.5522	0.03		
T	1	0.0408	0.041	3.16	ns
G.T	2	0.0140	0.007	0.54	ns
RESID	21	0.2708	0.01		
V	1	0.1876	0.188	16.82	<.001
G.V	2	0.0131	0.007	0.59	ns
RESID	21	0.2342	0.01		
P	1	2.3090	2.309	158.07	<.001
G.P	2	0.1807	0.090	6.18	<.01
RESID	21	0.3067	0.01		
LW.T	1	0.0000	0.000	0.00	ns
G.LW.T	2	0.0030	0.001	0.12	ns
RESID	21	8.2637	0.01		
LW.V	1	0.0103	0.010	1.53	ns
G.LW.V	2	0.0068	0.003	0.51	ns
RESID	21	0.1405	0.01		
LW.P	1	0.0024	0.002	0.17	ns
G.LW.P	2	0.0005	0.000	0.02	ns
RESID	21	0.2997	0.01		
T.V	1	0.0009	0.001	0.15	ns
G.T.V	2	0.0087	0.004	0.69	ns
RESID	21	0.1324	0.01		
T.P	1	0.1294	0.129	12.35	<.01
G.T.P	2	0.0083	0.004	0.40	ns
RESID	21	0.2201	0.01		
V.P	1	0.1622	0.162	10.75	<.01
G.V.P	2	0.0316	0.016	1.05	ns
RESID	21	0.3170	0.02		
LW.T.V	1	0.0068	0.007	0.51	ns
G.LW.T.V	2	0.0023	0.001	0.09	ns
RESID	21	0.2808	0.01		

LW.T.P	1	0.0070	0.007	0.61	ns
G.LW.T.P	2	0.0054	0.003	0.23	ns
RESID	21	0.2426	0.01		
LW.V.P	1	0.0057	0.006	0.47	ns
G.LW.V.P	2	0.0028	0.001	0.11	ns
RESID	21	0.2591	0.01		
T.V.P	1	0.0042	0.004	0.39	ns
G.T.V.P	2	0.0823	0.041	3.77	<.05
RESID	21	0.2291	0.01		
LW.T.V.P	1	0.0008	0.001	0.09	ns
G.LW.T.V.P	2	0.0290	0.015	1.55	ns
RESID	21	0.1960	0.01		

(ii) Response Frequencies.

Factors are: G (Groups), LW (Letters/Words), T (Truth Value), V (Sentence Voice), P (Polarity).

Source	DF	SS	MS	VR	F PR
Between S's.					
G	2	0.8763	0.438	2.27	ns
RESID	21	4.0605	0.19		
Within S's.					
LW	1	0.0007	0.001	0.01	ns
G.LW	2	0.2201	0.110	1.20	ns
RESID	21	1.9199	0.09		
T	1	0.0163	0.016	0.13	ns
G.T	2	0.3685	0.184	1.51	ns
RESID	21	2.5684	0.12		
V	1	0.5475	0.548	9.33	<.01
G.V	2	0.1732	0.087	1.48	ns
RESID	21	1.2324	0.06		
P	1	1.2038	1.204	18.84	<.001
G.P	2	0.1576	0.079	1.23	ns
RESID	21	1.3418	0.06		
LW.T	1	0.1100	0.110	1.46	ns
G.LW.T	2	0.0091	0.005	0.06	ns
RESID	21	1.5840	0.08		
LW.V	1	0.1882	0.188	2.51	ns
G.LW.V	2	0.4388	0.219	2.92	ns
RESID	21	1.5762	0.08		
LW.P	1	0.0059	0.006	0.10	ns
G.LW.P	2	0.0742	0.037	0.62	ns
RESID	21	1.2480	0.06		
T.V	1	0.1465	0.146	4.65	<.05
G.T.V	2	0.0820	0.041	1.30	ns
RESID	21	0.6621	0.03		
T.P	1	0.8913	0.891	7.66	<.025
G.T.P	2	0.0560	0.028	0.24	ns
RESID	21	2.4434	0.12		
V.P	1	0.0007	0.001	0.02	ns
G.V.P	2	0.1419	0.071	1.99	ns
RESID	21	0.7480	0.04		
LW.T.V	1	0.0007	0.001	0.01	ns
G.LW.T.V	2	0.0091	0.005	0.08	ns
RESID	21	1.2559	0.06		
LW.T.P	1	0.0319	0.032	0.79	ns
G.LW.T.P	2	0.0091	0.005	0.11	ns
RESID	21	0.8496	0.04		

LW.V.P	1	0.0007	0.001	0.01	ns
G.LW.V.P	2	0.0013	0.001	0.01	ns
RESID	21	1.6387	0.08		
T.V.P	1	0.0788	0.079	1.67	ns
G.T.V.P	2	0.0091	0.005	0.10	ns
RESID	21	0.9902	0.05		
LW.T.V.P	1	0.0007	0.001	0.01	ns
G.LW.T.V.P	2	0.3529	0.176	2.75	ns
RESID	21	1.3496	0.06		

APPENDIX H.

THE ANALYSIS OF VARIANCE TABLES FOR EXPERIMENT V.

Significance levels (F PR) for all repeated measures factors use conservative degrees of freedom (see Edwards, 1967).

(i) Response Latencies (logarithmically transformed).

Factors are: G (Groups), C (Conditions), T (Truth Value), V (Sentence Voice), P (Polarity).

Source	DF	SS	MS	VR	F PR
Between S's.					
G	2	0.5448	0.272	0.60	ns
RESID	33	14.9916	0.45		
Within S's.					
C	3	9.1128	3.038	88.67	<.001
G.C	6	0.1134	0.019	0.55	ns
RESID	99	3.3913	0.03		
T	1	0.0054	0.005	0.38	ns
G.T	2	0.0201	0.010	0.70	ns
RESID	33	0.4749	0.01		
V	1	0.2328	0.233	21.62	<.001
G.V	2	0.0116	0.006	0.54	ns
RESID	33	0.3553	0.01		
P	1	3.7359	3.736	151.63	<.001
G.P	2	0.2386	0.119	4.84	<.025
RESID	33	0.8130	0.02		
C.T	3	0.0453	0.015	1.32	ns
G.C.T	6	0.1387	0.023	2.01	ns
RESID	99	1.1373	0.01		
C.V	3	0.2574	0.086	7.01	<.025
G.C.V	6	0.0874	0.015	1.19	ns
RESID	99	1.2116	0.01		
C.P	3	0.0734	0.024	2.01	ns
G.C.P	6	0.0852	0.014	1.16	ns
RESID	99	1.2078	0.01		
T.V	1	0.0000	0.000	0.00	ns
G.T.V	2	0.0055	0.003	0.29	ns
RESID	33	0.3082	0.01		
T.P	1	0.1378	0.138	12.84	<.01
G.T.P	2	0.0069	0.003	0.32	ns
RESID	33	0.3542	0.01		
V.P	1	0.0034	0.003	0.33	ns
G.V.P	2	0.0567	0.028	2.72	ns
RESID	33	0.3439	0.01		
C.T.V	3	0.0695	0.023	2.56	ns
G.C.T.V	6	0.0526	0.009	0.97	ns
RESID	99	0.8970	0.01		

C.T.P	3	0.0446	0.015	2.01	ns
G.C.T.P	6	0.0657	0.011	1.48	ns
RESID	99	0.7335	0.01		
C.V.P	3	0.1246	0.042	3.58	ns
G.C.V.P	6	0.0419	0.007	0.60	ns
RESID	99	1.1488	0.01		
T.V.P	1	0.0389	0.039	2.89	ns
G.T.V.P	2	0.0167	0.008	0.62	ns
RESID	33	0.4450	0.01		
C.T.V.P	3	0.0745	0.025	2.21	ns
G.C.T.V.P	6	0.0810	0.013	1.20	ns
RESID	99	1.1121	0.01		

(ii)Response Frequencies.

Factors are: G (Groups), C (Conditions), T (Truth Value), V (Sentence Voice), P (Polarity).

Source	DF	SS	MS	VR	F PR
Between S's.					
G	2	315.4080	157.704	4.98	<.025
RESID	33	1044.1406	31.64		
Within S's.					
C	3	57.1181	19.039	1.77	ns
G.C	6	22.4392	3.740	0.35	ns
RESID	99	1062.6302	10.73		
T	1	68.0556	68.056	5.63	<.025
G.T	2	37.8038	18.902	1.56	ns
RESID	33	398.8281	12.09		
V	1	100.3472	100.347	11.51	<.01
G.V	2	7.3351	3.668	0.42	ns
RESID	33	287.6302	8.72		
P	1	145.9201	145.920	13.86	<.001
G.P	2	17.4913	8.746	0.83	ns
RESID	33	347.5260	10.53		
C.T	3	6.4236	2.141	0.32	ns
G.C.T	6	40.3212	6.720	1.00	ns
RESID	99	667.3177	6.74		
C.V	3	7.8125	2.604	0.44	ns
G.C.V	6	42.3177	7.053	1.19	ns
RESID	99	585.8073	5.92		
C.P	3	20.9201	6.973	1.37	ns
G.C.P	6	57.8559	9.643	1.89	ns
RESID	99	504.0365	5.09		
T.V	1	0.3472	0.347	0.06	ns
G.T.V	2	6.8142	3.407	0.63	ns
RESID	33	178.7760	5.42		
T.P	1	83.4201	83.420	9.80	<.01
G.T.P	2	24.7830	12.391	1.46	ns
RESID	33	280.8594	8.51		
V.P	1	7.0313	7.031	1.27	ns
G.V.P	2	2.4740	1.237	0.22	ns
RESID	33	182.6823	5.54		
C.T.V	3	47.0486	15.683	2.53	ns
G.C.T.V	6	33.8108	5.635	0.91	ns
RESID	99	614.4531	6.21		
C.T.P	3	5.9896	1.997	0.26	ns
G.C.T.P	6	70.7031	11.784	1.55	ns
RESID	99	752.9948	7.61		

C.V.P	3	13.2813	4.427	Ø.95	ns
G.C.V.P	6	15.2344	2.539	Ø.55	ns
RESID	99	46Ø.5469	4.65		
T.V.P	1	Ø.Ø868	Ø.Ø87	Ø.Ø1	ns
G.T.V.P	2	5.2517	2.626	Ø.34	ns
RESID	33	252.474Ø	7.65		
C.T.V.P	3	5.2951	1.765	Ø.33	ns
G.C.T.V.P	6	8Ø.5122	13.419	2.47	ns
RESID	99	537.63Ø2	5.43		

APPENDIX I.

THE ANALYSIS OF VARIANCE TABLES FOR EXPERIMENT VI.

Significance levels (F PR) for all repeated measures factors use conservative degrees of freedom (see Edwards, 1967).

(i) Response Frequencies.

Factors are: G (Groups), IN (Instance), O (Feature Order in Rule), M (Matching Case).

(a) TT as 'Conforms'.

Source	DF	SS	MS	VR	F PR
Between S's.					
IN	2	Ø.2569	Ø.129	Ø.74	ns
G	1	Ø.Ø556	Ø.Ø56	Ø.32	ns
IN.G	2	Ø.1319	Ø.Ø66	Ø.38	ns
RESID	3Ø	5.2Ø83	Ø.174		
Within S's.					
O	1	Ø.Ø556	Ø.Ø56	Ø.4Ø	ns
IN.O	2	Ø.Ø486	Ø.Ø24	Ø.17	ns
G.O	1	Ø.Ø139	Ø.Ø14	Ø.1Ø	ns
IN.G.O	2	Ø.1736	Ø.Ø87	Ø.62	ns
RESID	3Ø	4.2Ø83	Ø.14Ø		
M	3	3.8472	1.282	11.79	<.Ø1
IN.M	6	Ø.9653	Ø.161	1.48	ns
G.M	3	Ø.1389	Ø.Ø46	Ø.43	ns
IN.G.M	6	Ø.2569	Ø.Ø43	Ø.39	ns
RESID	9Ø	9.7917	Ø.1Ø9		
O.M	3	Ø.4167	Ø.139	1.Ø3	ns
IN.O.M	6	Ø.7292	Ø.122	Ø.9Ø	ns
G.O.M	3	Ø.Ø139	Ø.ØØ5	Ø.Ø3	ns
IN.G.O.M	6	Ø.2153	Ø.Ø36	Ø.27	ns
RESID	9Ø	12.125Ø	Ø.135		

(b) TF as 'Conflicts'.

Source	DF	SS	MS	VR	F PR
Between S's.					
IN	2	2.5833	1.292	10.11	<.001
G	1	0.3472	0.347	2.72	ns
IN.G	2	0.1111	0.056	0.44	ns
RESID	30	3.8333	0.128		
Within S's.					
O	1	0.1250	0.125	1.29	ns
IN.O	2	0.0833	0.042	0.43	ns
G.O	1	0.0139	0.014	0.14	ns
IN.G.O	2	0.3611	0.181	1.86	ns
RESID	30	2.9167	0.097		
M	3	8.4583	2.819	16.55	<.001
IN.M	6	1.5833	0.264	1.55	ns
G.M	3	0.9028	0.301	1.77	ns
IN.G.M	6	1.7222	0.287	1.69	ns
RESID	90	15.3333	0.170		
O.M	3	0.4028	0.134	1.22	ns
IN.O.M	6	0.8056	0.134	1.22	ns
G.O.M	3	0.8472	0.282	2.56	ns
IN.G.O.M	6	0.5278	0.088	0.80	ns
RESID	90	9.9167	0.110		

(c) FT as 'Irrelevant'.

Source	DF	SS	MS	VR	F PR
Between S's.					
IN	2	0.0625	0.031	0.07	ns
G	1	0.1250	0.125	0.29	ns
IN.G	2	1.2708	0.635	1.49	ns
RESID	30	12.7917	0.426		
Within S's.					
O	1	0.3472	0.347	5.32	<.05
IN.O	2	0.1319	0.066	1.01	ns
G.O	1	0.0556	0.056	0.85	ns
IN.G.O	2	0.2569	0.128	1.97	ns
RESID	30	1.9583	0.065		
M	3	18.4722	6.157	41.43	<.001
IN.M	6	1.4653	0.244	1.64	ns
G.M	3	0.0694	0.023	0.16	ns
IN.G.M	6	0.8681	0.145	0.97	ns
RESID	90	13.3750	0.149		
O.M	3	0.2917	0.097	1.41	ns
IN.O.M	6	0.5625	0.094	1.36	ns
G.O.M	3	0.4722	0.157	2.28	ns
IN.G.O.M	6	0.7153	0.119	1.73	ns
RESID	90	6.2083	0.069		

(d)FF as 'Irrelevant'.

Source	DF	SS	MS	VR	F PR
Between S's.					
IN	2	Ø.34Ø3	Ø.17Ø	Ø.37	ns
G	1	Ø.42Ø1	Ø.42Ø	Ø.92	ns
IN.G	2	5.2986	2.649	5.79	<.Ø1
RESID	3Ø	13.7292	Ø.458		
Within S's.					
O	1	Ø.17Ø1	Ø.17Ø	1.58	ns
IN.O	2	Ø.Ø486	Ø.Ø24	Ø.23	ns
G.O	1	Ø.Ø313	Ø.Ø31	Ø.29	ns
IN.G.O	2	Ø.1458	Ø.Ø73	Ø.68	ns
RESID	3Ø	3.2292	Ø.1Ø8		
M	3	14.51Ø4	4.837	42.Ø4	<.ØØ1
IN.M	6	2.4375	Ø.4Ø6	3.53	<.Ø5
G.M	3	Ø.2Ø49	Ø.Ø68	Ø.59	ns
IN.G.M	6	1.3681	Ø.228	1.98	ns
RESID	9Ø	1Ø.3542	Ø.115		
O.M	3	1.Ø66Ø	Ø.355	2.95	ns
IN.O.M	6	1.ØØ69	Ø.168	1.39	ns
G.O.M	3	Ø.1493	Ø.Ø5Ø	Ø.41	ns
IN.G.O.M	6	Ø.7986	Ø.133	1.1Ø	ns
RESID	9Ø	1Ø.8542	Ø.121		

(ii) Response Latencies (logarithmically transformed).

(a) Comprehension Times.

Factors are: G (Groups), O (Feature Order in Rule), R (Rules). Logical Case was included as a dummy factor. This did not affect the analysis and is not included in the Table.

Source	DF	SS	MS	VR	F PR
Between S's.					
G	1	0.0246	0.025	0.04	ns
RESID	34	19.0037	0.559		
Within S's.					
R	3	2.7822	0.927	29.71	<.001
G.R	3	0.0629	0.021	0.67	ns
RESID	102	3.1823	0.031		
O	1	0.0145	0.015	0.38	ns
G.O	1	0.0114	0.011	0.28	ns
RESID	34	1.3274	0.039		
R.O	3	0.0300	0.010	0.63	ns
G.R.O	3	0.0161	0.005	0.32	ns
RESID	102	1.6170	0.016		

(b) Verification Times.

Factors are: G (Groups), IN (Instance), LC (Logical Case), R (Rules), O (Feature Order in Rule).

Source	DF	SS	MS	VR	F PR
Between S's.					
IN	2	1.4973	0.749	1.16	ns
G	1	1.1134	1.113	1.72	ns
IN.G	2	0.2659	0.133	0.21	ns
RESID	30	19.4036	0.65		
Within S's.					
LC	3	3.6452	1.215	24.73	<.001
IN.LC	6	0.0730	0.012	0.25	ns
G.LC	3	0.4264	0.142	2.89	ns
IN.G.LC	6	0.2817	0.047	0.96	ns
RESID	90	4.4222	0.05		
R	3	2.3095	0.770	20.32	<.001
IN.R	6	0.1838	0.031	0.81	ns
G.R	3	0.0272	0.009	0.24	ns
IN.G.R	6	0.1926	0.032	0.85	ns
RESID	90	3.4106	0.04		
O	1	0.0667	0.067	0.69	ns
IN.O	2	0.1334	0.067	0.69	ns
G.O	1	0.0097	0.010	0.10	ns
IN.G.O	2	0.0154	0.008	0.08	ns
RESID	30	2.9024	0.10		
LC.R	9	1.6496	0.183	6.91	<.025
IN.LC.R	18	0.4264	0.024	0.89	ns
G.LC.R	9	0.2820	0.031	1.18	ns
IN.G.LC.R	18	0.3390	0.019	0.71	ns
RESID	270	7.1619	0.03		
LC.O	3	0.0651	0.022	0.73	ns
IN.LC.O	6	0.2227	0.037	1.25	ns
G.LC.O	3	0.0906	0.030	1.02	ns
IN.G.LC.O	6	0.0328	0.005	0.18	ns
RESID	90	2.6758	0.03		
R.O	3	0.0253	0.008	0.34	ns
IN.R.O	6	0.3804	0.063	2.58	ns
G.R.O	3	0.0251	0.008	0.34	ns
IN.G.R.O	6	0.2883	0.048	1.96	ns
RESID	90	2.2078	0.02		
LC.R.O	9	0.4975	0.055	2.23	ns
IN.LC.R.O	18	0.3047	0.017	0.68	ns
G.LC.R.O	9	0.1119	0.012	0.50	ns
IN.G.LC.R.O	18	0.5702	0.032	1.28	ns
RESID	270	6.6968	0.02		

APPENDIX J.

THE ANALYSIS OF VARIANCE TABLES FOR EXPERIMENT VII.

Significance levels (F PR) for all repeated measures factors use conservative degrees of freedom (see Edwards, 1967).

(i) Response Frequencies.

Factors are: IN (Instance), M (Matching Case), B (Blocks).

(a) TT as 'Conforms'.

Source	DF	SS	MS	VR	F PR
Between S's.					
IN	2	1.9740	0.987	5.01	<.025
RESID	45	8.8594	0.197		
Within S's.					
M	3	7.5000	2.500	19.27	<.001
IN.M	6	0.9844	0.164	1.26	ns
RESID	135	17.5156	0.130		
B	1	0.0000	0.000	0.00	ns
IN.B	2	0.1406	0.070	1.34	ns
RESID	45	2.3594	0.052		
M.B	3	0.0833	0.028	0.48	ns
IN.M.B	6	0.5260	0.088	1.50	ns
RESID	135	7.8906	0.058		

(b) TF as 'Conflicts'.

Source	DF	SS	MS	VR	F PR
Between S's.					
IN	2	Ø.3281	Ø.164	Ø.75	ns
RESID	45	9.7969	Ø.218		
Within S's.					
M	3	5.1458	1.715	13.32	<.ØØ1
IN.M	6	Ø.9635	Ø.161	1.25	ns
RESID	135	17.39Ø6	Ø.129		
B	1	Ø.Ø1Ø4	Ø.Ø1Ø	Ø.14	ns
IN.B	2	Ø.Ø677	Ø.Ø34	Ø.45	ns
RESID	45	3.4219	Ø.Ø76		
M.B	3	Ø.5521	Ø.184	1.95	ns
IN.M.B	6	Ø.1823	Ø.Ø3Ø	Ø.32	ns
RESID	135	12.7656	Ø.Ø95		

(c) FT as 'Irrelevant'.

Source	DF	SS	MS	VR	F PR
Between S's.					
IN	2	Ø.27Ø8	Ø.135	Ø.22	ns
RESID	45	27.4688	Ø.61Ø		
Within S's.					
M	3	22.3854	7.462	45.59	<.ØØ1
IN.M	6	1.27Ø8	Ø.212	1.29	ns
RESID	135	22.Ø938	Ø.164		
B	1	Ø.Ø938	Ø.Ø94	1.ØØ	ns
IN.B	2	Ø.4375	Ø.219	2.33	ns
RESID	45	4.2188	Ø.Ø94		
M.B	3	Ø.26Ø4	Ø.Ø87	1.Ø9	ns
IN.M.B	6	Ø.27Ø8	Ø.Ø45	Ø.57	ns
RESID	135	1Ø.7188	Ø.Ø79		

(d) FF as 'Irrelevant'.

Source	DF	SS	MS	VR	F PR
Between S's.					
IN	2	Ø.8177	Ø.4Ø9	Ø.64	ns
RESID	45	28.9297	Ø.643		
Within S's.					
M	3	12.57Ø3	4.19Ø	21.93	<.ØØ1
IN.M	6	2.Ø156	Ø.336	1.76	ns
RESID	135	25.7891	Ø.191		
B	1	Ø.44Ø1	Ø.44Ø	2.45	ns
IN.B	2	Ø.Ø99Ø	Ø.Ø5Ø	Ø.28	ns
RESID	45	8.Ø859	Ø.18Ø		
M.B	3	Ø.Ø495	Ø.Ø17	Ø.15	ns
IN.M.B	6	Ø.9427	Ø.157	1.43	ns
RESID	135	14.8828	Ø.11Ø		

(ii) <u>Response Latencies (logarithmically transformed).</u>

Factors are: IN (Instance), LC (Logical Case), R (Rules), B (Blocks).

Source	DF	SS	MS	VR	F PR
<u>Between S's.</u>					
IN	2	3.9098	1.955	2.71	ns
RESID	45	32.4742	0.722		
<u>Within S's.</u>					
LC	3	6.1288	2.043	36.83	<.001
IN.LC	6	0.6186	0.103	1.86	ns
RESID	135	7.4877	0.055		
R	3	7.2183	2.406	79.71	<.001
IN.R	6	0.2201	0.037	1.22	ns
RESID	135	4.0749	0.030		
B	1	2.0630	2.063	36.44	<.001
IN.B	2	0.3295	0.165	2.91	ns
RESID	45	2.5475	0.057		
LC.R	9	1.4554	0.162	6.77	<.025
IN.LC.R	18	0.4446	0.025	1.04	ns
RESID	405	9.6694	0.024		
LC.B	3	0.0273	0.009	0.31	ns
IN.LC.B	6	0.2633	0.044	1.51	ns
RESID	135	3.9373	0.029		
R.B	3	0.0388	0.013	0.68	ns
IN.R.B	6	0.1896	0.032	1.67	ns
RESID	135	2.5581	0.019		
LC.R.B	9	0.1176	0.013	0.60	ns
IN.LC.R.B	18	0.3690	0.021	0.94	ns
RESID	405	8.7971	0.022		