City Research Online

City St George’s, University of London

Citation: Goodman, C. (1978). The representation and monitoring of recall. (Unpublished Doctoral thesis, The City University)

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/37829/

THE REPRESENTATION AND MONITORING OF RECALL

CATHERINE GOODMAN

Thesis submitted for the degree of Doctor of Philosophy.

Department of Social Science and Humanities,
The City University, St. John Street, London, EC1V 4PB

July 1978

ABSTRACT

The main objective of the thesis was to investigate how an act of recall is represented in memory and monitored. The thesis attempted to understand the effects recall might have on the underlying representation of an item in terms of (i) episodic theory, and (ii) the pre-theoretical assumption that a memory trace constitutes a collection of attributes. A further aim of the thesis was to trace the development of the ability to assess previous recall performances.

These aims were pursued in three sets of studies. In the first it was proposed that performance features derived from the articulatory and sensory feedback of a response during recall become established as part of the attribute ensemble defining an item in memory. It was supposed that these changes would at least partly mediate memory for past recall and also influence learning in a multi-trial free recall situation. These proposals were supported in that both memory for remembered events and the rate of learning varied directly as a function of the number of performance features available during recall.

The development of the ability to assess past recall was investigated in a second set of studies. Contrary to recent findings, young children were found to be relatively poor in assessing past recall. This was true even where the memory task had been otherwise modified to suit the vocabulary and estimated memory spans of the specific age groups involved.

The notion that recall might result in the formation of a unique trace in episodic memory was tested in a third set of studies, by means of the paradigm developed by Murdock and Anderson, (1975). Little support for this hypothesis or for Murdock and Anderson's model was found, however. A second experiment undertook to investigate their 'conveyor belt' model of memory further but again the results were not completely definitive. The question of whether an interpolated recall task influences recognition

of items which have been presented but not recalled, was pursued in a final study. Whilst initial testing made no difference to overall recognition scores a closer analysis seemed to show that recognition of non recalled items was worse following recall, than recognition of comparable items was where no recall trial had intervened.

It was concluded that recall has effects upon the underlying representation of an item which are different to those of presentation and constitute more than a simple increment in trace strength. It was also suggested that this influence was unlikely to be uniform and constant but will vary, for instance, with trials in multi-trial recall. It remains uncertain as to whether or not the influence of recall extends to the formation of a trace separate to that formed at presentation. Young children's relative inability to appraise their recall was speculated to perhaps underlie their well documented failure to make spontaneous use of mnemonic strategies.

ACKNOWLEDGEMENTS

I would like to thank the following people for their help and advice in the completion of this thesis.

My supervisor, Dr. J.M. Gardiner, for freely availing me of his time and consideration.

[redacted] for his encouragement and valuable criticisms of draft manuscripts.

[redacted] [redacted] for her assistance in designing and analysing Experiment 1.

The technicians of the Psychology Department, The City University.

My fellow students, [redacted], [redacted] and [redacted].

[redacted] [redacted] for typing the manuscript.

My husband, [redacted], for his encouragement.

[redacted] [redacted] for allowing me time whilst at work to write this thesis.

CONTENTS

Page

Chapter 1 Introduction to the Thesis

1.1 Introduction to the Chapter

1.2 The nature of the memory trace

1.3 Memory monitoring processes

d. The development of knowledge of past recall performance

1.4 Episodic memory

Chapter 2 The Role of Performance Features in Mediating Memory for Past Recall and in Learning

LIST OF TABLES

LIST OF TABLES IN APPENDIX

Experiment 1

Experiment 2

Experiment 3

Experiment 4

Experiment 5

Experiment 6

Experiment 7

LIST OF FIGURES

Page

Experiment 1

Experiment 2

Experiment 3

Experiment 4

Experiment 5

Experiment 6

Experiment 7

List of Figures in Appendix

Experiment 3

Experiment 4

CHAPTER 1

Introduction to the Thesis

1.1 Introduction to the Chapter

a) General Introduction

Unlike the playing of a record, the 'playing' of a memory does not leave the record unchanged. An act of recall has consequences which are at least comparable to those of a presentation trial, (Tulving, 1967) and in general one retrieval will facilitate a subsequent retrieval, though this effect may vary with the retrieval effort (Bjork, 1975).

For the purposes of the thesis the effects recall has upon the memory system are conceptualised as occurring at two levels. At a specific level, recall is contended to affect the underlying memory trace of an item, which effects in turn are supposed to enable monitoring of output to take place and assessments of past recall to be made. Memory monitoring was a term originally applied to the knowledge possessed of the internal storage states of memory (Hart, 1967), but has since been extended to include knowledge of the processes and performances of memory, (Robinson and Kulp, 1970; Gardiner and Klee, 1976).

Such monitoring is presumed to mediate the second, more general, type of effect of recall. If assessments of the level of recall can be made then various mnemonic strategies can be appraised, with reference to task requirements and the nature of the stimuli. These strategies can then be used appropriately in future. Thus memory is perhaps well conceived of as a closed system whereby any information generated both as a product or a by-product of the process of memorising, re-enters the system and transforms the contents of the store.

This thesis is concerned with how an act of recall is represented in memory and monitored. The thesis also aims to trace the development of the ability to monitor and appraise memory. An attempt to accomplish these aims is made from within two compatible frameworks. 1) Throughout the thesis the pre-theoretical assumption is made that the memory trace is multi-componential (Bower, 1967; Wickens, 1970; Underwood, 1969). No single unit representing a target item is present but instead a collection of characteristics or attributes which together define the target are stored in memory. From this perspective the act of recall is viewed as resulting in performance attributes, derived from response feedback, becoming established as part of the attribute ensemble defining the recalled item. It is further proposed that the presence of these performance attributes mediates, in part, the proven ability to discriminate recalled from nonrecalled items, (Gardiner and Klee, 1976; Klee and Gardiner, 1976) 2) An attempt is then made to understand these proposed changes in the attribute ensemble as a result of recall, in terms of episodic theory (Tulving, 1972; 1976). Basically episodic theory assumes that word events are stored independently of their semantic representations in memory. It is proposed here that the act of recall constitutes a unique event in terms of temporal,performance, and autobiographical attributes, and as such is represented by the formation of a unique trace in episodic memory.

The value of these notions is explored in 3 sets of related studies. In the second chapter the role of response produced feedback in mediating memory for past recall, and in promoting learning, is investigated. In the third chapter the development of the ability to assess past recall performances is traced. The well documented, (e.g. Flavell, Friedrichs and Hoyt, 1970) improvement in memory with age is usually attributed to the growing skill and awareness children show in using mnemonic strategies (Flavell, 1977). However, if it were the case that young children are poor at appraising previous memory performances then they may well remain unaware of the need to improve memory and might anyway be incapable of using mnemonic strategies to any good effect.

The proposal that a unique trace is formed in episodic memory as a result of recall is investigated in Chapter 4 by means of a model developed by Murdock and Anderson, (1975). The nature of the search process which takes place when episodic information is retrieved from memory is also considered, as well as the effect an interpolated recall trial has upon recognition performance.

The implications of the experimental findings are discussed in a final chapter.

b) Introduction to the remainder of the chapter

The remainder of this chapter is divided into two main sections, the first of which is concerned with memory monitoring processes, and the second with episodic theory. Each major section is further divided into subsections.

In the first main section the conceptualisation of the memory trace as a bundle of attributes or features is outlined, and the identity of these encoding dimensions considered. Memory monitoring processes are next reviewed, where monitoring is both of memory content and memory performances. Finally, the development of memory appraisal skills is traced and the relation of these to the use of mnemonic strategies and to memory performance per se considered.

Episodic theory, and the episodic-semantic distinction upon which it rests, are described in the second major section. Two lines of research which support this theoretical approach are briefly reviewed in further sections. One line of research involves the investigation of the effects of changing the contextual cues present from input to output for both recall and recognition performances, whilst the second involves the study of response latencies in the retrieval of episodic information from memory.

1.2 The Nature of the Memory Trace

a) Introduction

Throughout the thesis the widely held pre-theoretical assumption is made, that a stimulus item is evaluated and encoded in terms of a number of constituent attributes, (Underwood, 1969), components (Bower, 1967), dimensions, (Wickens, 1970) or elements (Tulving and Watkins, 1975). This assumption simply provides a point of reference within which the aims of the thesis might be discussed, and its validity is not seen as being essential for the acceptance of the tenets and conclusions of the thesis.

The conceptualisation of the memory trace as an ensemble of attributes, is outlined in the following section and the identity of the possible encoding dimensions considered.

b) An attribute view of representation in memory

A memory trace is popularly conceived to be multidimensional in nature, no single corpus or literal copy of the target being stored, (Bower, 1967; Underwood, 1969; Wickens, 1970; Morton, 1970; Tulving and Watkins, 1975). An idea,word or relationship is represented by a unique collection of features or attributes from which the target is reconstructed at retrieval. The success of retrieval depends on whether sufficient features have been retained to allow the unequivocal identification of the target.

Whilst Underwood and Wickens are more concerned with the identification of the various encoding dimensions used to represent items in memory, Bower considers the nature of encoding and forgetting given a memory trace which is multidimensional. He assumes that some kind of feature analysis takes place at input which produces an ordered list of attributes with their corresponding values. Such a listing constitutes the primary code of the target, as distinct from the secondary code which is the verbal label elicited by the primary code, and is itself represented by a listing of vocal attributes. Either the primary code or the secondary code or both are stored in memory. Others, (e.g. Underwood) make no such distinction and simply consider acoustic attributes to comprise part of the total attribute ensemble defining a target.

For a particular situation, (the need for this qualification will become apparent later), an item is assumed to be encoded along a fixed number of components, each of which have a fixed and equivalent number of possible values, one of these being assigned to the input. Component information can be lost, the loss from a single component dimension being an all or none event. Since information is not lost from all components at once, forgetting of target information appears gradual.

The encoding of items into a variety of attributes is not a conscious process and the subject is unaware of their identity. Two methods popularly used to arrive at the identity of the constituents of the memory trace are the release from proactive inhibition paradigm, (Wickens, 1970; 1972), and the false recognition paradigm, (Underwood, 1965). In the false recognition paradigm distractor items in the recognition test share a common attribute with various target items. A distractor and a target might for instance share the same meaning or sound. To the extent that the distractor is mistaken for the target it is assumed that the shared feature forms a part of the memory trace representing the target. The release from PI paradigm derives from the finding that proactive inhibition builds up in a short term Peterson and Peterson distractor task, so that retention decreases from the first to the fourth trial. It is assumed that (Wickens, 1970, p. 3), "the process of perceiving a word involves encoding that word into positions within categories; if a series of items comes from the same set of categories they will interfere with each other and depress retention performance'. If the encoding category is altered or shifted and retention improves and returns to initial levels then it can be assumed that there was a common way of encoding within the first category. In reviewing evidence of this kind, Underwood (1969) was able to conclude that stimuli can be encoded in terms of temporal, spatial, frequency, modality, orthographic, associative verbal and associative nonverbal attributes.

As well as permitting the identification of encoding dimensions, the release from PI paradigm enables the comparison of their relative salience. The amount of release which occurs upon shifting categories is taken as indicative of the relative importance of the first category in the encoding and representation of an item. Some of the largest release effects have been found where the shift has involved semantic dimensions such as taxonomic class, where for example names of birds might be presented for 3 trials and trees on the fourth. Shifting physical or grammatical features of the stimuli results in relatively little release from PI.

This paradigm provides some elucidation of a question raised by Bower. The problem is whether or not attributes are arranged in a hierarchical fashion and are forgotten at a rate inversely proportional to their importance in the hierarchy. From the release from PI studies it certainly seems that particular types of attributes (semantic), are favoured as encoding dimensions, and from incidental learning studies (e.g. Hyde and Jenkins, 1973) that using semantic attributes to encode stimuli results in more durable memories. Clearly however all stimuli are not encoded along all possible dimensions, and the pattern and relative importance of attributes used in encoding will vary with the nature of the task, (Jarvella, 1971), the nature of the stimuli (Gardiner, Klee, Redman and Ball, 1976) and the age of the subject, (Klastorin Naron, 1978). For example if the colour in which stimuli are presented is shifted on the fourth trial, little release effect is found where the stimuli are words. However, where the stimuli are nonsense syllables and do not have as many possible encoding dimensions as words, a colour shift brings about a large release effect (Gardiner et al., 1976).

Very little has so far been said about evidence which offers general support to an attribute conceptualisation of the memory trace. The most convincing evidence comes from the tip-of-the tongue (TOT) or feeling-of-knowing studies (Brown and McNeill, 1966; Hart, 1967), in which the subject can identify various features of the target item without actually being able to provide the target. The more features known, the more imminent recall is. The behaviour of a person in a TOT state clearly reflects a process of reconstructing a target from a stored attribute ensemble. Some evidence regarding the TOT or feeling of knowing state is considered in the next section, though this is cited specifically for the purpose of illustrating one type of memory monitoring process.

1.3 Memory Monitoring Processes

a) Introduction to the Section

Memory monitoring processes are examples of control processes, which 'are not permanent features of memory, but are instead transient phenomena under the control of the subject', and 'include any schemes, coding techniques or mnemonics used by the subject in his effort to remember', (Atkinson and Shiffrin, 1968, p. 106). The importance of this distinction is perhaps most apparent in the developmental rather than adult literature. When considering the improvement in memory which occurs with age,the issue has become one of whether this reflects a growth in physiological capacity or changes in the way in which the task is controlled by the child, the latter view being currently favoured (e.g. Flavell, Friedrichs and Hoyt, 1970). Section 1.3d considers the evidence which suggests that improvement in memory with age can be attributed to the increasing use made of mnemonic strategies. It will be proposed that in turn, the increasing use made of strategies could be attributed to a growing capacity to monitor past memory performances.

Two types of memory monitoring processes can broadly be distinguished;

those of memory content, and those of memory performance. The evidence describing these two types of monitoring and the factors underlying their use will be reviewed in the following sections (1.3b and c).

Whilst the relevance of these monitoring processes to actual memory performance is briefly discussed here, an attempt to understand this relationship more fully is made in Section 1.3d.

b) Tip-of-the-tongue or Feeling of knowing states

Failure to recall something does not necessarily mean that there is no representation of it stored in memory. As Tulving and Pearlstone (1966) showed an item can be available in memory but inaccessible. A capacity to assess when a sought after target which cannot be retrieved is actually stored in memory is a necessity for a fallible retrieval system such as that of human memory, (Hart, 1967). If an item is judged to be represented in memory, then continued retrieval efforts have a chance of success.

That the memory is capable of making such an assessment has been shown by Hart (1967), Brown and McNeill (1966) and Blake (1973). The subjective state experienced when an item which cannot be retrieved is felt to be imminent such that it will be subsequently recognised or recalled is referred to as a tip-of-the-tongue (TOT) state, (Brown and McNeill, 1966), or a feeling of knowing (Fk) state, (Hart, 1967). There is every indication that these two terms refer to the same phenomenon.

Brown and McNeill read out dictionary definitions of fairly rarely occurring words and asked subjects who were unable to supply the word so defined if they were in a TOT state. When a TOT state was experienced there was a high probability that the target item would be subsequently recalled or recognised. Prior to recall or recognition many of the features of the target could be provided. Often the initial letter, final letter, number of syllables, and location of the primary stress in the target word were known. Words similar in sound or meaning, ranked according to their degree of similarity to the target could also be supplied. Similarly accurate knowledge is possessed of targets which cannot be retrieved from episodic memory (Yarmey, 1973).

Feelings of knowing and tip-of-the-tongue states are accurate predictors of subsequent recognition, and in general the more features of the target known the more likely subsequent recognition is. To determine whether the reporting of a feeling of knowing reflected the experience of a particular state of mind rather than an inference of the kind 'I know some features of the target therefore I should be able to recognise it', Blake made partial knowledge of the target effectively useless in the recognition task. Reports of feeling of knowing states did diminish in relation to a task in wich partial attribute knowledge was of use, but were still reported nevertheless. Blake was thus able to conclude that Fks were genuine reports of the experiencing of a certain state of mind. Reports of Fks increased as a function of the number of features of the target available to the subject regardless of whether or not partial recall was free to influence recognition performance.

That certain features of an elusive target can be retrieved before the target itself is accessed is strong evidence in favour of the conceptualisation of a memory trace as a bundle of attributes which collectively define the target, and of retrieval as a process of reconstruction from this; (Underwood, 1969; Wickens, 1970; Bower, 1967; Morton, 1970).

c) The monitoring of recall

Not only can the contents of memory be accurately appraised when retrieval is unsuccessful, but also the results of retrieval attempts can be monitored to exclude errors in recall. That this kind of monitoring is both necessary and accomplished is demonstrated by Bousfield and Rosner

(1970). When a group of subjects was given 'uninhibited' free recall instructions, such that they were to recall any and every word which came into their head during the recall session, errors, predominantly of repetition, increased approximately sixfold relative to those made by a group given standard free recall instructions. Obviously target words can be retrieved more than once from memory, necessitating that retrieval is monitored to exclude possible repetitions in recall.

Again monitoring of output is necessary according to a model described by Murdock (1974) since whilst the standard free recall task implicitly requires sampling-without-replacement, the human retrieval system is capable only of sampling-with-replacement. 'After list presentation there is a store of available items. To retrieve, the central processor samples randomly with replacement from this store. To avoid repetition, each recalled item is saved in an output pool, and every sampled item is compared against this output pool before recall. If a match is found, the item is not recalled; if not, it is. As recall proceeds, the number of unrecalled items in the store decreases while the size of the pool increases,' (p. 221).

This model can account for the longer inter-response intervals commonly found as the recall trial proceeds (Murdock and Okada, 1970). As the number of target words not already recalled decreases in the pool of available items, successive samples will yield no output. In addition, the number of items saved in the output pool will have increased over the trial so that, assuming the search made of these is serial and self terminating, the time taken to check this will also increase.

Output monitoring partly mediates memory for remembered events (MRE) (Gardiner and Klee, 1976; Klee and Gardiner, 1976), a term which refers to the knowledge held of past memory performances. Memory for past recall was originally investigated by Robinson and Kulp (1970). A single free recall trial was followed by a recognition task in which the original stimuli were re-presented and subjects required to discriminate words they had recalled from words they had not recalled. This task was performed accurately, even after some irrelevant activity had intervened between the recall trial and the recognition task.

This paradigm was extended by Gardiner and Klee, (1976) who gave a series of free recall trials followed by a final unexpected recognition test in which recalled and nonrecalled items were to be discriminated. Whilst MRE was accurate for items from primacy and asymptotic portions of the serial position curve, it was relatively poor for items from the recency portion of the curve.

This negative recency effect (Craik, 1970), may have been due to the fact that subjects were instructed to recall the last few items presented to them, first. Such an instruction increases the likelihood that recency items were recalled from primary memory rather than secondary memory and were thus probably less well registered in memory than pre recency items, (Craik, 1970). Another possible effect of this instruction could have been to subject the recency items which were output first in the recall trial, to more retroactive interference from the subsequent recall of other items than the pre recency items had received.

In order to assess the contribution of output interference and depth of initial registration to the negative recency effect in MRE, Klee and Gardiner, (1976), carried out two further experiments. In the first of these, serial recall, (the recalling of items in the same order as they were presented), was required. This meant that recency items would occupy final output positions in recall and would not only have minimal output interference, but should also be retrieved from secondary rather than primary memory.

However, a negative recency effect was still found following serial recall, though this was significantly smaller than the negative recency effect found in the experiment (Gardiner and Klee, 1976), using modified free recall instructions. Registration differences and/or output interference effects cannot then fully account for the negative recency effect in MRE. Neither does the manner in which these two variables influence MRE seem to be straightforward, for whilst order of recalling affected the recency portions of the curve, it had no effect upon the pre recency portions.

The second experiment involved recognition of previous recognition performances. The order in which items from the primacy, asymptotic, and recency portions of the list were encountered in the initial recognition test was manipulated so that if A, B and C represent these three portions of the list, they were tested in the sequence ABC, CBA, CAB, ACB, BCA and BAC. There was no influence of testing order on MRE performance, with respect to the serial position curve. Recognition of previously recognised items was poor whatever portion of the presentation list was tested first. The finding that recognition of recall produces curves which differ to those of recognition favours the contention of Gardiner and Klee, (1976), that MRE does at least in part reflect the operation of memory monitoring processes at output.

Memory monitoring can be expected to differ substantially between recognition and recall tasks, for whilst recall involves a continuous monitoring of output to ensure that no errors of repetition or intrusion are made, recognition tests do not necessitate such a process. Thus it seems reasonable to assume that the 'processes arising from, and peculiar to, the act of recall' (Klee and Gardiner, 1976, p. 478) are influential factors in MRE performance, where this involves recognition of previous recall. That an act of retrieval does have important consequences for the underlying storage state of an item is well attested to by Bjork, (1975, p. 123) who states that 'an item can seldom, if ever, be retrieved from memory without modifying the representation of that item in memory in significant ways'.

The present thesis includes an experiment designed to determine whether attributes associated with the act of recall become represented in memory, during recall, and are used in a subsequent MRE task to differentiate previously recalled from non recalled items. An investigation into this is described in Chapter Two.* The following sections consider the development of the ability to assess past recall performance.

*(Experiment 1, P.44, has subsequently been published: Gardiner, J.M., Passmore, C., Herriot,P.,and Klee,H. "Memory for remembered events: effects of response mode and response-produced feedback." J.V.L.V.B. 16,45-54,1977).

d) The development of knowledge of past recall

(i) Introduction

The developmental literature on memory shows more concern with the role of memory appraisal skills in memory performance, than the adult literature does. The reason for this emphasis lies primarily in the findings which suggest that the well documented increase in memory span with age, (Flavell, Friedrichs and Hoyt, 1970; Huttenlocher and Burke, 1976) is due, not to any physiological changes in the structure of the brain, but to psychological changes occurring within the child which alter his approach to the memory task, (Moely, Olson, Halwes and Flavell, 1969; Flavell et al., 1970; Flavell, 1977). Instead of basic memory capacity altering with age, the child increasingly learns how to make the most effective use of a limited cognitive space. The child is able to do this by means of mnemonic strategies or plans of action specifically undertaken for the purpose of memorising, (Section 1.3d, ii).

At about the same time that the child comes spontaneously to bring strategies to bear on the memory task, he also becomes much more skilful in assessing the storage states and processes of his memory (Wellman, 1977), and in estimating its performance both past and future, (Flavell, et al., 1970). It would seem to be rather obvious that all of these aspects of memory are linked. As recall span increases so does the use of strategies, and the skill with which memory can be appraised. Intuitively certain strategies could not be implemented to improve the memory performance unless certain memory appraisals had first been carried out, and conversely the employment of some strategies would enable better appraisals of memory to be made.

The experiments reported in this thesis (Chapter 3) were concerned with what is felt to be a basic aspect of memory appraisal - the knowledge possessed of past recall performance. It is hoped, in the following sections, to place this knowledge within the context of the literature and in addition to outline possible interrelations between memory, memory appraisal skills and mnemonic strategies.

ii) Mnemonic strategies

A mnemonic strategy can be defined as a plan of action drawn up for the purpose of memorising, given an otherwise limited capacity for so doing. A strategy may be broadly of acquisition or of retrieval, and consists of a specific means of treating the relevant information in order that it may be better retained and retrieved. Rehearsal, imaging, naming and organising the input according to semantic categories are all instances of mnemonic strategies.

A given strategy may be more appropriate to one type of task or material than another (Tversky and Teiffer, 1976), and similarly may be differentially effective depending on the item's position in the serial position curve, (Hagen, 1971). Adults clearly use a wide variety of strategies, combined in a complex manner, in any memory task (Kelly, Scholnick, Travers and Johnson, 1976).

iii) The child as a memoriser

There is now a large body of evidence which shows that children become increasingly strategic and planful in their approach to memory tasks as they grow older. They increasingly tend to rehearse (Flavell, Beach and Chinsky, 1966), categorise items at input (Moely, Olson, Halwes and Flavell, 1969), use indirect retrieval strategies, (Ritter, Kaprove, Fitch and Flavell, 1973), and use effective study period behaviours, such as anticipating and naming items (Flavell, Friedrichs and Hoyt, 1970). However, it seems possible that young children who do not use such strategies are not incapable of doing so. They appear to possess much of the information necessary to the performance of certain study strategy behaviours, (Kelly, Scholnick, Travers and Johnson, 1976), and when instructed to do so, can effectively rehearse (Keeney, Cannizzo and Flavell, 1967), and categorise items on input (Moely et al., 1969).

The notion that there is a gap between the emergence of the necessary skills for employing certain strategies and their spontaneous use is known as the production deficiency hypothesis (Flavell et al., 1966). This is to be contrasted with the mediation deficiency hypothesis of Reese, (1962), which states that even if various strategies were used by the child they would not benefit his memory performance. Clearly (Keeney et al., 1969; Moely et al., 1969), the child is simply failing to produce strategies which he is capable of using to effect.

The reasons for this failure are currently of major research interest. Most of the explanations considered, (to be discussed below), assume some lack of metamemorial skills. The term metamemory, (Flavell, 1970) designates the introspective knowledge a person has of the workings of his memory. Wellman (1977) has further distinguished two types of metamemory, a) the knowledge of "certain timeless facts that a person could know about memory", and b), "certain ongoing transient assessments a person could make about items in his memory". This second type of metamemory, Wellman states, can be termed memory monitoring, and would include knowledge of the current storage states of items in memory and knowledge held of past recall performance.

Possible reasons why children fail to make use of mnemonic strategies will now be considered. A pre-requisite to the use of strategies would seem to be the awareness on the part of the subject that he is able actively to influence the memorisation process. Appel, Cooper, McCarrell, Sims-Knight, Yussen and Flavell (1972), found that young children's behaviour and recall levels were unaffected by whether they had been instructed to memorise or to simply perceive the stimuli. Clearly the young child fails to differentiate memory from perception and does not appreciate that memorisation may require an active effort on his part. Once he has re-assessed his role in the task, he presumably searches for the means to influence the memory process and finds it in mnemonic strategies.

The degree to which the skill underlying the use of a strategy is possessed may be another factor determining the spontaneous emergence of strategies in the child's approach to the memory task. Moely et al., (1969) suggest that the greater the degree of mastery of the requisite skill, the more likely it is that it will emerge spontaneously.

It seems obvious that a strategy could not be used unless its availability were known to the child (Wellman, 1978). Yet even where children have experienced the beneficial effects of a strategy, (i.e. their recall level has improved) they do not necessarily adopt the strategy spontaneously (Flavell et al., 1966; Moynahan, 1973). Possibly the child fails to realise that he has recalled any better when using that strategy. If children are unable to assess their memory performance accurately then they will remain ignorant of which strategy choice will yield the better retention of stimulus material.

Again if the child was unaware of the limits of his memory performance then it is difficult to understand why he should want to utilise strategies to improve it. Also if his memory appraisal skills were lacking he would not be able to use strategies to any good effect, being unable to assess the effectiveness and appropriateness of any given strategy. Whether the child is actually able to appraise memory accurately and whether the above postulated dependencies between the use of strategies and appraisal skills are operative will now be considered.

iv) Memory appraisal skills and their relation to mnemonic strategies and to memory performance

Whilst there is relatively little research evidence concerning metamemorial skills, what there is, is also somewhat conflicting. Flavell, Friedrichs and Hoyt (1970), found that young children were prone to overestimating their memory span when predicting their recall levels. They were also poor at estimating when they had studied sufficiently to be ready to recall. Since they were also deficient in using certain appropriate study period behaviours, such as naming the to-be-remembered items to themselves, (see later in this section), Flavell et al., suggested that poor memory appraisal skills might be linked to the failure to use appropriate mnemonic strategies.

However there is little additional evidence that this is indeed the case. Moynahan (1976) found that children of all ages were equally accurate in assessing their memory performance after recall and were capable of adjusting their estimates according to whether stimulus items belonged to a single category or were drawn from several different categories. Kelly et al. (1976) also report that children are as accurate as adults in estimating their memory performance, both prior to and following the memory task. Thus their additional finding, that children do not make use of a study strategy of reviewing previously missed items, could not be attributed to poor memory appraisal skills. Also those who were best at estimating memory did not necessarily have higher recall levels or make more use of the above study strategy.

Several possibilities could account for a failure to find a relationship between memory appraisal skills, actual retention, and the use of mnemonic strategies. First of all,

whilst certain memory appraisal skills may yield the information necessary to the use of a given study strategy, there are many other equally expedient strategies available for the subject to choose from, which presumably do not all depend on appraisal skills, (Kelly et al., 1976).

Secondly a kind of mediation deficiency may be operative. Study strategy behaviour is rather complex. The oldest group of children studied by Flavell et al.,(1970) for instance, engaged in approximately three different kinds of strategic behaviour which were variously employed depending on the amount of the study period which had elapsed. In the first quarter of the allocated study period fourth grade children engaged in naming the stimulus items to themselves. This activity dropped off over the study period with rehearsal taking precedence over other strategic behaviours in the final quarter. Whilst a child may be able to implement one particular strategy when instructed, it is clearly doubtful whether he would be able to combine a number of strategies in an effective study sequence of his own accord.

A further factor which might obscure possible relationships between memory appraisal skills and the use of strategies, is the fact that the child appears to be more or less planful depending on the degree of the abstractness of the strategy being investigated. For instance, Kreutzer, Leonard and Flavell (1975) found that young children were quite planful when the constraints of the laboratory were lifted and the memory task made closer to a real life situation. If a child was asked something like 'Where did you leave the pen?', he would get up and search for it in the last place he had been. An adult is more likely to perform this search mentally. Similarly children presented with pictures of people occupying familiar roles paired with appropriate toy objects, (e.g. king - crown), were more likely to think of using one member of the pair to aid the retrieval of the hidden other if the task was more meaningful, and if more physical reminders of the availability of the retrieval cues were present (Ritter, Kaprove, Fitch and Flavell, 1973).

Finally, perhaps more specific information than that necessary for accurate estimates of memory is required for the implementation of study strategies. If a child does not know precisely which items he has or has not recalled then he will be unable to differentially rehearse or allocate more study time to previously non recalled items. Whilst Masur, McIntyre and Flavell (1973) have presented evidence that the child's failure to use a strategy of giving additional rehearsal to items previously not recalled was not due to their being unable to discriminate recalled and non recalled items, this ability was assessed following only a single recall trial using the eight subjects who had failed to employ this strategy. The experiments reported in Chapter 3 seek to extend this procedure using the paradigm developed by Gardiner and Klee (1976).

e) Conclusions to Section 1.3

The MRE function describing recognition of previous recognition is quite different to that which describes recognition of previous recall (Klee and Gardiner, 1976). Since recall involves output monitoring whereas recognition presumably does not, it was concluded that memory for previous recall must be a function of the processes occurring at output, as well as those taking place at input. It is one of the contentions of this thesis that performance features are encoded during recall and aid the subsequent discrimination of recalled and nonrecalled words in an MRE task. This proposal is tested in Chapter Two.

Children's failure to use appropriate mnemonic strategies, even when they are capable of doing so, was concluded to perhaps be due to an incapacity to assess accurately their own memory performances. Such an incapacity would mean that children would remain unaware of both the need to make use of strategies and of the improvement in performance the use of these brings. The question of whether children can appraise their past recall performance accurately is investigated in Chapter 3.

1.4 Episodic Memory

a) Introduction

A major controversy which has developed in recent years has been the issue of whether episodic information is represented by the marking or tagging of semantic structures which existed pre experimentally, (Anderson and Bower, 1972, 1974; Bahrick, 1970; Martin, 1975), or as episodic theory asserts is represented independently of these, (Tulving, 1972, 1976; Tulving and Thomson, 1971, 1973; Watkins, Ho and Tulving 1976). The present section seeks to extend the conceptualisation of the effects of recall on the memory system developed previously (Section 1.3), by placing it within the perspective of episodic theory. It will be proposed that an act of recall constitutes a unique event, and as such is represented by the formation of a trace in episodic memory.

Episodic theory, and the episodic-semantic distinction which is basic to it, will be outlined in the following section. 'Tagging' theories of memory will be briefly reviewed in a further section, and evidence which contra indicates these considered. A final section will look at a different line of evidence which is also supportive of episodic theory, and which will be used in the thesis as a basis from which to explore the hypotheses made regarding recall and episodic theory.

b) Episodic memory and episodic theory

According to Tulving (1972), two major types of memory can be distinguished, episodic and semantic memory. Semantic memory is necessary for the making of inferences and the comprehension of language. 'It is a mental thesaurus (the) organised knowledge a person possesses about words and other verbal symbols, their meaning and referents, about relations among them, and about rules, formulas, and algorithms for the manipulation of these symbols, concepts and relations'. (Tulving, 1972, p. 386). Many theorists (e.g. Anderson and Bower, 1972), conceptualise semantic memory as a vast network of interconnected nodes, where a node is a basic unit of meaning, defined by its position in the semantic space, and might be a word, a concept, or an image.

Information is stored in semantic memory in terms of its relations to the already existing cognitive structure, rather than in terms of its perceptual properties. It is possible to retrieve information from the store which has not been entered into it directly, by drawing inferences from the existant data base. The contents of semantic memory are relatively immune to loss and transformation, and an act of retrieval from the store leaves it unchanged.

In contrast to this highly organised structure, episodic memory stores only personally experienced events or episodes in terms of their spatio-temporal relations to one another. Whilst the relations among the nodes in semantic memory are multidirectional and associative, the relations among events in episodic memory are unidirectional in that storage characterises whether one event preceded or followed another event. Memories held in the episodic store are thus accompanied by the subjective experience of self in time and space.

Due to the nature of storage, events in episodic memory are more likely to undergo transformation and loss. Since each event is unique in its circumstances and in time, there is no distinction made in episodic memory between the occurrence of two separate events and two occurrences of the 'same' nominal event, a unique trace being formed in each case. An act of retrieval from either episodic or semantic memory constitutes a unique input into episodic memory and alters the contents of the store.

While semantic and episodic memory are able to operate independently of one another, the information in one can be used in the encoding and retrieval of information in the other. The use of semantic information in the encoding and retrieval of items from episodic memory is better documented and understood than the role of episodic information in the encoding and retrieval of items from semantic memory. The formation of a trace in episodic memory can be thought of as the taking of a nominal copy of information retained in semantic memory and placing it with its context in a file of events ordered according to the time of their occurrence. Tulving (1976) likens this to the distinction between a photograph and the person the photograph is of; one does not exactly constitute the other, though they have certain features in common.

The nature of the memory trace within episodic memory is conceived to be that of an ensemble of attributes, Tulving and Thomson (1973, p. 354) agreeing with the 'widely held pre-theoretical assumption that central representation of the to-be-remembered event, the memory trace, is a multidimensional collection of elements, features or attributes'. An important component of the trace in episodic memory is temporal information, since search in episodic memory is guided by the temporal date of an event's occurrence.

Atkinson, Hermann and Wescourt, (1974), similarly distinguish between an event knowledge store (EKS) which encodes events along spatial and temporal dimensions, and a conceptual store (CS), in which traces are located according to their content. They go further than Tulving in proposing that structures evolve in the CS as a result of the repeated encounter and representation of an item in the EKS.

From the episodic-semantic distinction, and the encoding specificity principle, which broadly 'asserts that only that can be retrieved that has been stored, and that how it can be retrieved depends on how it was stored', (Tulving and Thomson, 1973, p. 359) episodic theory has developed (Tulving, 1976; Tulving and Thomson, 1971, 1973; Watkins and Tulving, 1975). This holds that the occurrence of an event, such as a word, results in the formation of a unique trace in episodic memory. Whilst the nature of this trace is largely governed by the knowledge retained about the word in semantic memory, it is stored independently of this knowledge. An integral part of the episodic trace is the context prevailing at the time of encoding, and retrieval cues will be effective only to the extent that they reinstate this context. The very essence of episodic theory is that every item is unique in terms of its context, its autobiographical referents, time of occurrence, etc., and an item plus its context constitutes the storage unit of episodic memory - an episode. Since every episode is unique and can never be experienced again, it is necessarily represented by the formation of a new trace in episodic memory.

Retrieval from episodic memory is a matching process known as ecphory, (Tulving, 1976). This depends on information from two sources, the memory trace and the retrieval cue, and successful retrieval is a function of the degree of overlap of information contained in, and extracted from, the two. Semantic properties of a word will serve as effective retrieval cues only in so far as these formed part of the context at input and were encoded in the episodic trace. Within the framework of episodic-ecphory, recall and recognition involve the same process and differ only in the nature of the retrieval cues given, a nominal copy of the target item being present in recognition.

The main empirical thrust of episodic theory has been addressed to the issue of whether episodic information is represented by modifications to a permanent semantic structure as 'tagging' theorists propose, (Anderson and Bower, 1972; 1974; Martin, 1975; Bahrick, 1970), or is in some sense independent of this. This issue has major implications for the understanding of the processes of recall and recognition.

The following section briefly considers how episodic information might be represented within semantic memory, and reviews some of the evidence which is inconsistent with such a view.

c) The representation of episodic information in memory

In contrast to episodic theory the so-called 'tagging' theories of memory, (Anderson and Bower, 1972; 1974; Bahrick, 1970) assert that the occurrence of a word event results in the marking of pathways within a permanent semantic structure. In general, the semantic structure is considered to consist of a network of nodes, (where a node is a word, a concept or a cluster of information), interconnected by associative relationships, the position of a node within the network being determined by its content and meaning. Anderson and Bower (1972) assume that when a word is presented, the node corresponding to this is activated. At the same time the nodes corresponding to the contextual elements of the word, such as implicit associations to it, the current mood of the subject, etc., are also activated and become associated to the node of the word via a list maker. This marker is a sort of label, serving to identify the contextual elements belonging to a word. A recognition task is considered to be essentially one of deciding whether an item was drawn from the list under test. For any list there is a pool of elements which uniquely define that list. In recognition the subject gains automatic access to the node corresponding to the probed word. He then simply has to decide whether a sufficient number of the contextual elements marked as associated to that node belong to the pool of elements defining that particular list. In recall possible candidates for retrieval have to be generated by searching back along associative pathways which have been recently marked. This generation process is followed by a recognition stage in which the subject has to decide whether any of the candidates so retrieved are appropriately marked. Thus recognition consists of a one stage decision process whereas recall involves a retrieval phase as well as a decision stage. Recognition failure can only be attributed to the incorrect marking of a node or to the loss of an appropriate marker.

Whilst there are various versions of tagging theory they all make two related assumptions, (Tulving and Thomson, 1973); 1) that a word has only one representation in memory, or at the most a few corresponding to each of its possible meanings, and 2) that access to the underlying representation of a word is automatically guaranteed by the presence of a nominal copy of the target - only recall has a retrieval problem.

Thus tagging theory makes no allowance for context effects within recognition. As long as a nominal copy of the target is present, the context within which this is embedded can have no influence on recognition performance. This contrasts with episodic theory which assumes that recognition will always be context specific since it is the item plus its context which is stored and recollected in episodic memory. Again according to tagging theory the amount recalled can never be greater than the amount recognised. Since recall involves recognition as a substage successful recall implies that recognition has also been successful. Recall can be higher than recognition within the framework of episodic theory, since they both have a retrieval problem and involve the same processes of ecphory.

Quite a large body of evidence has now accumulated which contradicts the basic tenets of tagging theories, (Watkins and Tulving, 1975; Light and Carter-Sobell, 1970; Watkins, Ho and Tulving, 1976). This research falls into two broad groups: 1) that which considers context change effects in recognition and 2) that which considers the effects of extra list cueing on recall. In the experiments known as context change here, a word is typically presented either uncued or with a strong or weak associate, and recognition tested under the same, different and no cue conditions. Recognition is usually impaired under conditions where the cue has been altered from input to output, relative to conditions where cues have remained the same (Tulving and Osler, 1968; Tulving and Thomson, 1971).

Tagging theories have, however, been able to accommodate such findings by making the additional assumption that a word may have more than one representation in memory corresponding to each of its possible meanings, (Anderson and Bower, 1972; Martin, 1975). Thus pairing a target word with a different cue at output to that with which it was presented means that a different, incorrect node was accessed. But even these revised versions of tagging theory are unable to account for findings of contextual influences on recognition for stimuli which can be presumed to have had no prior representation in memory. Watkins, Ho and Tulving (1976), presented subjects with picture pairs of unfamiliar faces and tested recognition for right hand members where the left hand member was changed or unchanged from input to output. Recognition performance was superior where the context had remained unchanged. In a similar experiment, (Winograd and Rivers- Bulkeley, 1977) recognition of faces was worse where the original presentation context had been deleted or substituted. Since unfamiliar faces could only conceivably have one node, (formed at presentation), representing them in memory, it is difficult to explain these kind of context change effects in terms of accessing of the incorrect node. Tagging theory also has difficulties in accounting for false positive errors, (the mistaking of a distractor for a target), in recognition of unfamiliar faces. Since there can be no underlying representation for a picture of an unfamiliar face which is acting as a distractor, then there can be no inappropriate marking of semantic structures. A false positive error could only result in this instance from accessing the wrong node, thus recognition too has a retrieval problem.

The facilitation of recall by the presence of strong pre-experimental associates at retrieval, (Hudson and Austin, 1970; Bilodeau, 1967; Bahrick, 1970) has typically been ascribed by tagging theorists to a facilitation of the generation phase of recall. The strong extra list associates are presumed to aid the retrieval stage of recall by reducing the probability that the relevant information is not found. Episodic theory on the other hand assumes that strong pre-experimental associates will only serve as effective retrieval cues in so far as they were encoded as part of an item's context at input. Thomson and Tulving (1970) manipulated encoding by pairing to-be-remembered words either with strong, or weak cues or no cues at input and asking subjects to recall in the presence of strong, weak or no cues. Strong cues at output facilitated recall only where these were present at input or where the to-be-remembered word had been presented alone. Strong extra-list associates failed to facilitate recall where weak associates had been present at input.

Similarly, Tulving and Thomson (1973) found that the presence of strong extra list cues at retrieval provided no facilitation of recall over uncued conditions, where weak associates were present at input. In addition they found that recognition was poorer than recall, where recall was cued with items present at input and recognition tested in the presence of strong extra-list cues. Since the success of retrieval from episodic memory depends on the degree to which information contained in the memory trace and in the retrieval cue overlap, and since the context of an item is stored with it, then it follows that recall may well be superior to recognition under these circumstances.

Where a word has several meanings recourse has again been made to the assumption that there will be several representations in memory corresponding to each of these. It has been argued (Reder, Anderson and Bjork, 1974) that in the above experiment the context present during recognition predisposes accessing of the incorrect node, whilst the context at recall induces the accessing of the correct node, so that recall will naturally be higher than recognition. This too has been shown to be untenable. Recognition failure of recallable words has been demonstrated for words which have only one possible meaning, and thus can only have one representation in memory (Tulving and Watkins, 1977).

In the following section, evidence (Murdock and Anderson, 1975;

Carroll and White, 1973b),which implies that the retrieval of episodic information from memory involves a 'search along something corresponding to the temporal dimension in a system different from semantic memory', (Tulving, 1976, p. 71) is considered.

d) A conveyor belt model of memory

Episodic theory is essentially a multiple trace theory, every event being sufficiently distinct to warrant a unique representation in memory. The basic assumptions and postulates of the theory require that there can be no accumulation of episodic traces at any one point in memory on the basis of semantic or other features. Any imposition of organisation on events at encoding must therefore be at a minimum. (This does not preclude the possibility that similarities may be sought amongst traces after encoding. If events are evoked together this may result in their further representation in episodic memory as a single unit, any act of retrieval from episodic memory constituting a unique input into the store (Tulving, 1972). This is basically the proposal derived and tested in the present thesis).

Temporal information is considered to form an important component of the episodic trace, search within episodic memory being guided by this, and episodic memory is accompanied by a sense of personal continuity through time. It thus seems most probable that events are stored directly in the temporal sequence of their occurrence, the nature of the temporal code being, as Tulving (1976) suggests, a qualification in terms of other events, whether preceded, followed by or simultaneous with these. The temporal code would thus be highly subjective.

That such a conceptualisation of the storage of episodic information may be quite appropriate is attested to by the work of Murdock and his associates (Murdock and Anderson, 1975; Murdock, 1974; Murdock and Dufty, 1972). Their results support a model of memory which is best illustrated by an analogy with a conveyor belt; 'Imagine a conveyor belt moving at a constant speed. Little packets or globs of material are dropped onto it, much as suitcases on a loading ramp at an airport. However, it is an endless conveyor belt, receding ever further into the distance. As a glob or packet gets further and further away it becomes more amorphous; it loses its attributes, the contents of the suitcase if you will. But the glob does not disappear; it simply becomes less distinctive. Retrieval of item information involves looking over this conveyor belt to find if a particular packet (or suitcase) is on the belt. If it was loaded it will be; but the further back it is, the less distinct it will be', (Murdock, 1974, p. 266). Incoming events are directly encoded in the store in the same temporal sequence as they occurred and retrieval from the store takes place by means of a backward, serial self terminating scan. Whilst such a system may seem, intuitively, inappropriate for the retrieval of childhood memories, latencies in naming objects have been shown to be a function of the age at which these words were acquired (Carroll and White, 1973b). Similarly latencies of adults in supplying appropriate members of categories were governed to a large extent by the frequency of the category in children's vocabulary and the frequency of the most likely response in children's vocabulary, (Loftus and Suppes, 1972).

The evidence favouring a conveyor belt model of memory was derived from a series of experiments, (Murdock and Anderson, 1975) in which lists of target words were typically presented for study followed by a recognition task in which response latencies were recorded. Half the recognition test items were targets, and half distractors or new items. Reaction times were found to increase steadily as a linear function of testing position in the case of new items, and of lag in the case of targets. Lag is the number of items intervening between study and test and is therefore a composite measure of both presentation and test positions.

Distractor items of necessity have no lag. These results were interpreted in terms of a backward, serial, self terminating scan taking place over both study and test items. For new items, the search process was presumed to return to the beginning of the presentation list. Low confidence responses were slower than fast confidence responses and this was assumed to reflect directly the number of interrogations of memory which took place. Recognition accuracy decreased as a function of lag and test position for targets and new items respectively, and was attributed to deterioration in the traces over time, due to the loss of features or attributes.

The linear functions relating reaction times to test position and lag cannot be explained simply in terms of a direct access process whereby latency reflects the strength of the underlying trace. A strength model could not account for the smooth linear relationship between test position and new items, since the strength of the underlying representation of new items could not be presumed to be related to test position. Neither can increases in reaction times as a function of lag or test position be explained in terms of the effects of the course of testing, otherwise reaction times to targets would bear no relationship to lag. Any given lag measure represents a mixture of presentation and test positions and might be derived from an item that was tested first or tested last.

The present thesis aims to utilise the model and experimental paradigm developed by Murdock to test a proposal derived from episodic theory. An act of recall is here proposed to constitute a unique event in terms of performance features (Section 1.3c), autobiographical and temporal referents, and as such is assumed to be represented by a unique trace in episodic memory. Given that the temporal sequence of events is directly represented in the store it is predicted that response latencies to recalled items will be faster than those to non recalled items, and in addition will show a linear relationship to lag, where lag is taken as the number of items intervening between recall and test.

This proposal is tested in Chapter 4. An additional experiment considers the nature of the search process which might take place within an episodic system. A final experiment considers the effects an interpolated recall trial has on recognition.

e) Conclusions to Section 1.4

Much evidence has accrued (Watkins and Tulving, 1975; Tulving and Thomson, 1971, 1973; Thomson and Tulving, 1970; Watkins, Ho and Tulving, 1976) that episodic information is represented in a system which is in some sense independent of, and separate to semantic memory.

Generation-recognition models (Anderson and Bower, 1972, 1974; Bahrick, 1970) which propose that episodic information is stored in the form of marking of semantic structures, have been shown to be untenable in two different lines of studies. In the first, context cues are either changed from input to output or remain the same, and recall or recognition tested. The results of these studies contra indicate the two related and basic assumptions of generation recognition models, namely that a nominal copy of a target guarantees access to the underlying representation of the target and that contextual changes from input to output can have no influence on recognition performance. The second set of studies, briefly reviewed in this section, show that retrieval of episodic information involves a search, the duration of which is related to the time that the information was encoded. This strongly suggests that episodic information is represented in a system organised in a spatial temporal fashion, and is not represented merely as ancillary to semantic structures.

The model and paradigm developed in the second set of studies by Murdock and his associates are used in the thesis (Chapter 4), to investigate the proposal, derived from the assumptions of episodic theory, that an act of recall results in the formation of a unique trace in episodic memory.

CHAPTER 2

The Role of Performance Features in Mediating Memory for Past Recall and in Learning

2.1 Introduction to the Chapter

This chapter sought to understand the effect recalling an item has upon its underlying representation in memory, where a memory trace was assumed to consist of a collection of features or attributes, (e.g. Underwood, 1969; Bower, 1967). It was speculated that performance features available during recall are encoded to become part of the attribute ensemble defining an item. These changes were proposed to mediate, at least in part, memory for previous recall performance.

These proposals were investigated in two ways in Experiment 1, (i) by varying the mode of recall so that recall was either written or spoken or both, and (ii) by attenuating feedback of recall on half of the recall trials.

Experiment 2 considered the role performance features might have in learning. In a multi-trial-free-recall learning paradigm response mode was again varied so that recall was either spoken or both written and spoken.

2.2 Introduction to Experiment 1

Memory for past recall is accurate in the case of secondary memory items but a negative recency effect, (Craik, 1970), is typically found for primary memory items, (Gardiner and Klee, 1976). Whilst primary memory items are generally agreed to be poorly registered in memory, (Craik, 1970), memory for past recall is not solely a function of initial item registration effects. Memory for serial recall, in which any primary memory items recalled can be assumed to have been well registered and are subject to minimal output interference, also exhibits a negative recency effect, though this is less marked than that occurring under free recall conditions, (Klee and Gardiner, 1976).

It seems that the events occurring at output, as well as those of input, govern memory for past recall, since the function describing memory for previous recognition performance differs markedly from that describing memory for previous recall, (Klee and Gardiner, 1976). Whilst standard recall instructions require that output is monitored to avoid repetition errors, a recognition task would involve no such process.

Making the pre-theoretical assumption that a memory trace can be conceptualised as a collection of characteristics or attributes, (Underwood, 1969; Bower, 1967; Wickens, 1970), it was proposed here that performance features derived from the articulatory and sensory feedback of a response during recall, become established as part of the underlying attribute ensemble representing an item. Such changes in the memory trace as a result of recall might, in turn, be used to aid the discrimination of previously recalled from non recalled items in an MRE task. These notions were tested here using the experimental paradigm developed by Gardiner and Klee (1976), such that a series of recall trials were followed by a final unexpected recognition test of what had been recalled. Response feedback was attenuated on half of the recall trials, and response mode varied so that recall was written, spoken, or both written and spoken for three groups of subjects. In so far as performance features available during recall are important in mediating memory for past recall it was predicted that this would be worse where feedback of response had been attenuated. Similarly performance was expected to vary as a function of response mode, being better where recall was in two modes, and more performance features available for encoding.

2.3 Method

Design

Subjects were divided at random into 3 groups according to whether recall was written, spoken or both written and spoken. There were twelve recall trials, and on half of these feedback of response was impaired, so that S could not see what he was writing, and/or hear what he was saying.

Following the recall trials all of the words initially read out by E were re-presented, and S was required to discriminate between words previously recalled, and words previously not recalled.

Subjects

Thirty subjects were assigned to each group. They were male and female undergraduates of the City University. All subjects were tested individually. One subject was omitted from the study since he appeared to be responding in an entirely random fashion. Another subject was recruited in his place.

Materials

360 words were chosen, at random, from the Toronto word pool of common two syllable nouns, and these were divided into 4 sets each of 90 words, here, for descriptive purposes, labelled A,B,C and D. Each of these sets was coupled with another in 5 of the possible 6 combinations, (AB, BC etc.), to form 5 relatively disparate sets of 180 words. Six subjects within each of the 3 experimental groups, received one such set of words. The words were organised into 12 lists of 15 words, the order of words within lists, and the order of lists being randomised for each subject.

For the recognition of recall task, the words were removed from list order and typed on one sheet of paper, in two different random orders for each 180-word set.

For the written mode, impairment of feedback was achieved by having the subjects write their recall on a cellophane sheet, beneath which was a carbon sheet, and in turn this covered a plain piece of paper.

For the spoken mode S spoke into a microphone and under normal recall conditions heard himself through headphones. E addressed S via another microphone. For impaired feedback conditions both microphones were switched off, and white noise, at a tolerable level, played over the headphones. Whilst this manipulation did not completely eliminate auditory response feedback, it did substantially degrade it.

Procedure

The subject was instructed as to the unusual conditions under which he would be required to perform. Then E read out a list of 15 words at a rate of one word every two seconds. S was then asked to recall either by speaking or writing or both, and was asked to try and recall the last few words of the list first. This latter instruction was given in an attempt to equate recall strategies across subjects. On a random basis, on half of the trials, normal feedback of response was allowed, and on the other trials feedback of response was attenuated. There was one constraint however, viz. that no more than 3 consecutive trials were of the same kind. The subject did not know whether feedback of recall would be normal or impaired until the list had been read out. For written recall E indicated whether recall was to be on ordinary paper or on the cellophane arrangement. In the case of spoken recall the subject was unaware of the conditions under which recall was to be attempted until E switched on white noise and turned the microphones off. For the spoken mode, the subjects' recall was recorded by E. A maximum period of 1½ minutes was allowed for recall. This procedure was repeated for twelve trials.

For the recognition of recall test, S was given a sheet with all 180 initially presented words typed on it in a new random order, and was asked to tick the words he had recalled, and to cross those words which had not been recalled. In addition, confidence ratings along a 3-point scale,were asked for these responses, where 1 = unsure, 2 = fairly sure, and 3 = certain. The recognition task was paced by E, who read out each word on the sheet at a rate of one every 3 seconds.

2.4 Results

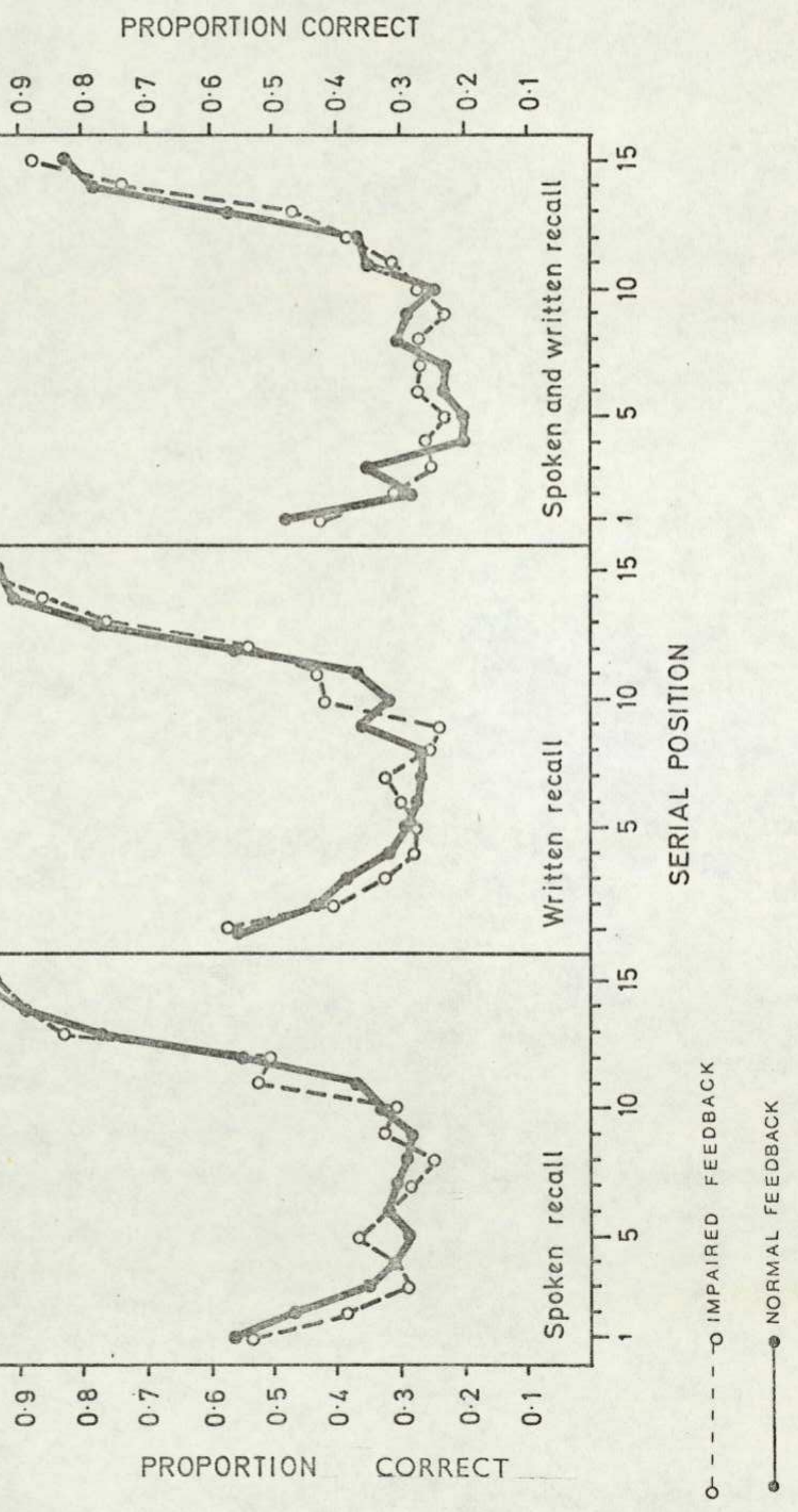
The recall data will be considered initially.

The proportion of words recalled over serial position, as a function of the mode of recall, and whether feedback of response was normal or impaired, was calculated, and these data are shown in Figure 2.1. Typical primacy and recency effects were found.Whilst the feedback manipulation appears to have had no effect upon the level of recall, recall levels seem to be lower where in two modes than where in a single mode.

For purposes of statistical analysis serial position was collapsed into 3 points, (serial positions 1-5, 6-10 and 11-15 inclusive), retrieval of items from the last 5 serial positions being assumed to represent retrieval from primary memory. Such an assumption was felt to be justified since, (i) the instruction given to recall the terminal list items first maximises the identity between recency items and primary memory, and (ii) the identification of the last 5 items in the list with primary memory is likely to represent an overestimate of primary memory capacity, (Tulving and Colotla, 1970). Any differences found between primary and secondary memory components should then be conservative rather than exaggerated estimates. A 3-way ANOVA with repeated measures on two factors was performed on these data and revealed that only the observed effects of serial position and response mode were reliable, (serial position $F(2,174) = 436.49$. $p < 0.001$; Response mode, ($F(2,87) = 10.35$, $p < 0.001$), / Post hoc comparisons using a Newmans Keul procedure were performed. These showed that whilst recall in two modes was significantly lower than recall in either single mode over all portions of the serial position curve, written and

Figure 2.1. Experiment 1

PROPORTION OF WORDS RECALLED AS A FUNCTION OF MODE OF RESPONSE, FEEDBACK CONDITION AND SERIAL POSITION



spoken recall levels did not differ substantially.

Though the feedback manipulation had no effect upon the level of initial recall, repetition errors increased under impaired feedback conditions, particularly over the recency portion of the curve, and in the written and written plus spoken modes (Table 2A). The greater increase in repetition errors from normal to impaired feedback conditions, where recall involved the written mode, was not surprising since, under normal recall conditions, S had his recall protocol available for checking. A χ^2 analysis on these data collapsed over serial position, revealed the significance of the observed differences (χ^2 (2) = 52.66, p 0.01).

No comparable increase in the number of intrusion errors, occurring under impaired feedback conditions, was observed. Whilst there was a 300% increase in the total number of repetition errors from normal to impaired feedback conditions, there was only a 13% increase in the number of intrusion errors.

Table 2A Total repetition errors occurring as a function of serial position, response mode, and feedback conditions.

	Feedback			Impaired feedback		
Serial Position	1-5	6-10	11-15	1-5	6-10	11-15
Spoken recall	5	8	39	8	8	44
Written recall	1	0	4	10	13	58
Spoken and written recall	0	0	1	1	6	25

Signal detection theory methods were used to analyse the results of the recognition of recall phase of the experiment. Since confidence ratings were largely extreme, performance was collapsed over these and estimates of d' values calculated on the basis of a binary classification of performance into hit rate and false alarm rate. Hit rate was considered to be the proportion of the words recalled which were correctly identified as such, and false alarm rate as the proportion of words not recalled which were incorrectly identified as having been recalled. Mean d' values as a function of response mode, feedback conditions and serial position, where serial position was collapsed into 3 points as above, are shown in Figure 2.2. Recognition of recall seems to be worse where feedback of response was impaired and also over the recency portions of the curve. Performance is clearly better where recall had been written than where spoken, and in turn better when recall had been both written and spoken than where written alone.

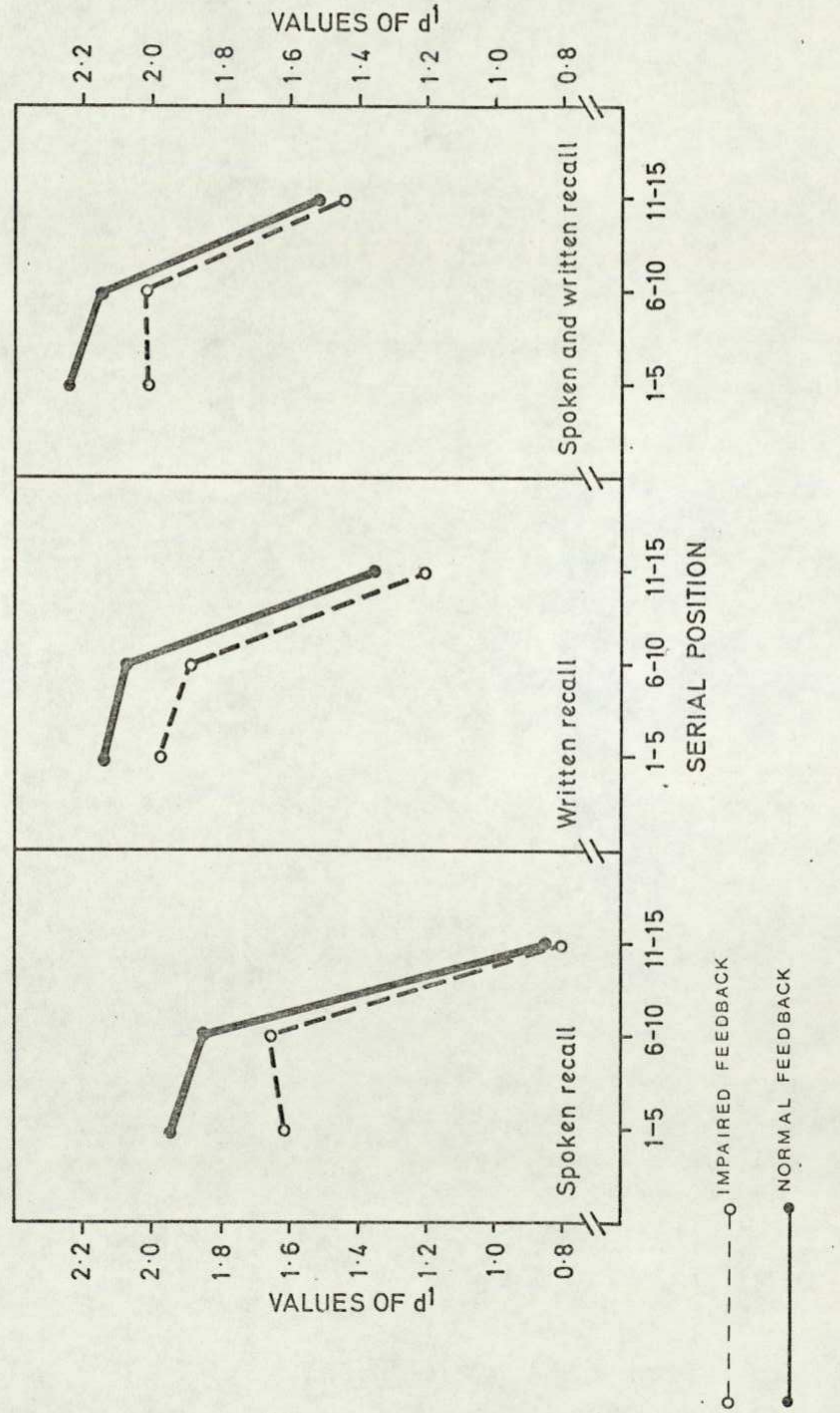
A 3-way ANOVA with repeated measures on 2 factors, showed all of these observed effects to be reliable, (Feedback condition, $F(1,87) = 10.13$, $p < 0.01$; Serial position, $F(2,174) = 79.26$ $p < 0.001$; Response mode, $F(2,87) = 6.23$ $p < 0.01$). There were no significant interaction effects.

In order to determine the location of these effects, three further 2-way ANOVAs compared performance over feedback conditions and response mode, separately for the primacy, asymptotic and recency portions of the curve. These revealed a significant effect of response mode only for the recency portion of the curve, ($F(2,87) = 11.08$, $p < 0.01$), and a significant effect of the feedback manipulation only for the primacy portion of the curve, ($F(1,87) = 7.57$, $p < 0.01$).

An analysis of false positive rates was undertaken to assess whether the estimates of d' scores obtained, were good estimates of the underlying 'true' d' values. Response bias is independent of sensitivity to the signal, or in this case of sensitivity to the memory trace, only when the underlying

Figure 2.2. Experiment 1

MEAN d^1 VALUES AS A FUNCTION OF RESPONSE MODE, FEEDBACK CONDITION AND SERIAL POSITION



receiver operating characteristic (ROC) curves are symmetrical about the negative axis. The estimation of d' values derived from the use of Z scores in analysing the results of a binary classification of data into false alarm and hit rates, assumes that the underlying ROC curves are symmetrical. This assumption does not matter however, provided response bias is equivalent across conditions. A comparison of false alarm rates reveals whether response bias might be equivalent or not, across conditions.

A 3-way ANOVA, with repeated measures on 2 factors, showed that response bias varied as a function of response mode, ($F(2,87) = 7.17$, $p < 0.01$). The differences in false positive rates were in the opposite direction to the obtained values of d' scores, with most false positives occurring in the spoken mode, and the least in the both written and spoken mode. These findings indicate that differences in sensitivity, with respect to response mode, may be somewhat greater than was indicated by the estimates of d' scores obtained here.

2.5 Discussion

The feedback manipulation was successful in that whilst it had no influence upon the level of initial recall the desired effect of attenuating response feedback was obviously achieved, evidenced by the increase in repetition errors from normal to impaired feedback conditions. The increase in repetition errors also provides further evidence that monitoring of output takes place normally prior to recall.

Whereas Craik (1970), found written recall to be slightly higher than spoken recall over the recency portion of the serial position curve, no differences were observed here between written and spoken recall. The explanation for the discrepancy between these results possibly lies in the fact that Craik used either auditory or visual presentation, and collapsed over these to consider mode of recall effects.

However, recall was found to be lower, in the present experiment, when in two modes, than when either spoken or written alone. Conversely, memory for past recall was better when recall had been in two modes than when in one mode alone. A sampling-with-replacement model of memory such as those of Murdock (1974) and Rundus (1973), can accommodate these findings.

According to Murdock, a sample of items is retrieved from a pool of items available in memory, and searched for possible targets. Any target found is output and then saved in an 'output pool'. Any items retrieved from memory are checked against those in the output pool, before being recalled. The probability of an item being retrieved is directly proportional to its strength of representation in memory.

Rundus (1973), similarly conceives of retrieval from memory as being a process of sampling-with-replacement. Since his model represents an expansion of Shiffrin's (1970), it has come to be known as the Shiffrin-Rundus model. When a list of words is presented the subject is considered to derive an organising theme, which later serves as a retrieval cue for the list (RQ list). To this is attached in varying strengths of association, a number of units which correspond roughly to category superordinates, and which serve as retrieval cues (rqs) for the specific list items. The list items are attached to their appropriate rqs, again in varying strengths of association, the sum of which represents the degree of association of that rq to the RQ list.

At retrieval an rq is selected and within it an item retrieved. The probability of any given rq being retrieved is governed by a ratio rule, such that the probability of retrieval equals the ratio of the strength of association between an rq and RQ list, to the sum of the associative strengths between every rq and RQ list. This can be summarised by the following formula, where k represents the number of rqs available for a given list, and $\underset{\longleftrightarrow}{S}$ represents the associative strength existing between two units,

a) $$p\,(rq_x) = \frac{rq_x \overset{S}{\longleftrightarrow} \text{RQ list}}{\sum_{1}^{k} (rq \overset{S}{\longleftrightarrow} \text{RQ list})}$$

The probability of a given list item being retrieved is similarly determined by the ratio of the strength of association between a given item (I_x), and its rq, (rq_x), to the sum of the strengths of association between rq_x and every list item connected to it. This can be summarised,

b) $$P(I_x) = \frac{I_x \overset{S}{\longleftrightarrow} rq_x}{\sum_{1}^{n} (I \overset{S}{\longleftrightarrow} rq_x)}$$

where n represents the number of items attached to a given rq. Note that the bottom line of formula (b) is equivalent to the top line of formula (a) for a given retrieval cue.

A major problem in retrieval lies in the competition between items which have been recalled and those which have not. An act of retrieval strengthens the association between a list item and its rq, and thus in line with the ratio rule, lessens the probability of other non recalled items being retrieved. As recall progresses, there is less and less chance of a previously non recalled item being retrieved and after a certain number of unsuccessful draws S decides to terminate the trial.

Applying this model to the current experiment it could be assumed that recall was lower and memory for past recall greater, where recall was in two modes, since this had a greater 'strengthening' effect upon the memory trace. Strength in the context of this experiment is perhaps best conceived of as 'a convenient summary term, an overall measure of the amount of information available about an item', (Wells, 1974, p. 378). This point is discussed further below. Whilst more information added during recall could be expected to result in better memory for recall, it would also increase the competition between recalled and non recalled items and would lead to the earlier termination of recall. Experiment 2 considers some predictions of this model with respect to a multi-trial free recall learning situation, where recall is either spoken or both written and spoken.

Memory for past recall was better where recall had been written than where spoken. This was probably simply an effect of the recall protocol being available for inspection throughout the recall trial during written recall. That subjects availed themselves of this opportunity to inspect their protocols is shown by the increase in repetition errors from normal to impaired feedback conditions.

The influence of response mode and the feedback manipulation upon memory for past recall indicates that the effects of recall cannot be conceived of as simply an increment in the underlying trace strength of an item. Strength is traditionally taken to be a cumulative measure of how frequently or recently an event has occurred. Clearly, however, performance features are encoded during recall and are used in the discrimination of previously recalled and non recalled items. If the number of features available at recall is decreased, as under impaired feedback conditions, or increased as in the two modes of recall condition, memory for past recall varies accordingly.

Item registration effects, as well as effects arising at recall, are responsible for variations in memory for past recall, (Klee and Gardiner, 1976). It may well be that these two factors are interdependent, the extent to which performance features are encoded at recall and utilised in a subsequent recognition of recall task being a function of the initial depth of registration. Some evidence for such a relationship was found in the present experiment, in that impairing feedback of response appeared to be most detrimental over the primacy portion of the curve, and response mode of greatest importance over the recency portions of the curve.

No significance can be attached to the value of the decrement in recognition performance where feedback of response had been impaired. Total elimination of response produced feedback was not possible for either written or spoken recall. In spoken recall feedback of response was still available by means of bone conduction, and articulatory features of the response were still present. Similarly kinaesthetic feedback could not be eliminated in written recall.

The following experiment considers the role of performance features generated during recall in promoting learning, and discusses further implications for sampling-with-replacement models of memory.

2.6 Introduction to Experiment 2

Where recall was in two modes in Experiment 1, recall levels were depressed relative to those of single mode conditions, whilst conversely memory for remembered events was better. The Shiffrin-Rundus sampling-with-replacement model of memory, (Shiffrin, 1970; Rundus, 1973), seemed to fit these results quite well. This model assumes that retrieval from memory is governed by a ratio rule. Since recall is assumed to strengthen the association between an item and its retrieval cue, recall lessens the probability of other items being retrieved from memory in two ways; (i) non recalled items are relatively weakened and in line with the ratio rule their chances of retrieval lessened, and (ii) the chances of re-retrieving an already recalled item are increased, and since recall is terminated after a certain number of unsuccessful draws, the end of the recall trial is hastened.

It follows from this that the more strengthening a certain act of recall is the more the chances of retrieving other non recalled items from the store are decreased. The fact that recall in two modes was terminated sooner than recall in a single mode in Experiment 1 suggests that recall in two modes was more strengthening. In accord with the findings of Experiment 1 more strengthening is here again taken to mean that more performance features were available and encoded during recall to become part of the memory trace. If recall in two modes does result in more information being represented in memory than recall in a single mode does, then in a multi-trial free recall learning paradigm where recall is either in one or two modes, a cross-over effect should occur over trials. Whilst recall levels should initially be lower where recall is in two modes, learning should eventually be greater. Whereas less items will be recalled on initial trials relative to single mode conditions, inter-trial retention should be superior. This prediction was tested in the present experiment by having one group of subjects recall in the spoken mode and a second group of subjects recall by both speaking and writing the response, within a multi-trial free recall learning paradigm. The spoken mode was chosen for the single mode of recall group since, in Experiment 1, memory for past recall where recall had been both written and spoken, differed most from memory for past recall where recall had been spoken. Any effects found in the present experiment should thus be more pronounced than if the written mode had been utilised.

2.7 Method

Design

In a multi-trial free recall learning paradigm S was required to learn a list of words. For one group recall was spoken, and for another group recall was both written and spoken.

Subjects

Twelve subjects were allocated at random to each group. They were male and female undergraduates of the City University. All subjects were tested individually.

Materials

432 words were selected at random from the Toronto word pool, and these were arranged by means of random draws into 12 sets each of 36 words. These words were written on cards so that they could be shuffled after each trial. One subject in each group received one of the 12 word-sets to learn.

Procedure

S was told that a list of words would be read out. Then for the one mode of recall group, he was instructed to call out as many words as he could remember. For the both modes of recall group, Ss were told to both call out and write down as many of the words as they could remember. S was told that the same list of words would be read out 8 times, each time in a different order, and that he should aim to learn the list over these trials.

E read the list at a rate of one word every 2 seconds, then asked S to recall. If recall was spoken E recorded it. After each trial the order of words was randomised, and the procedure repeated.

2.8 Results

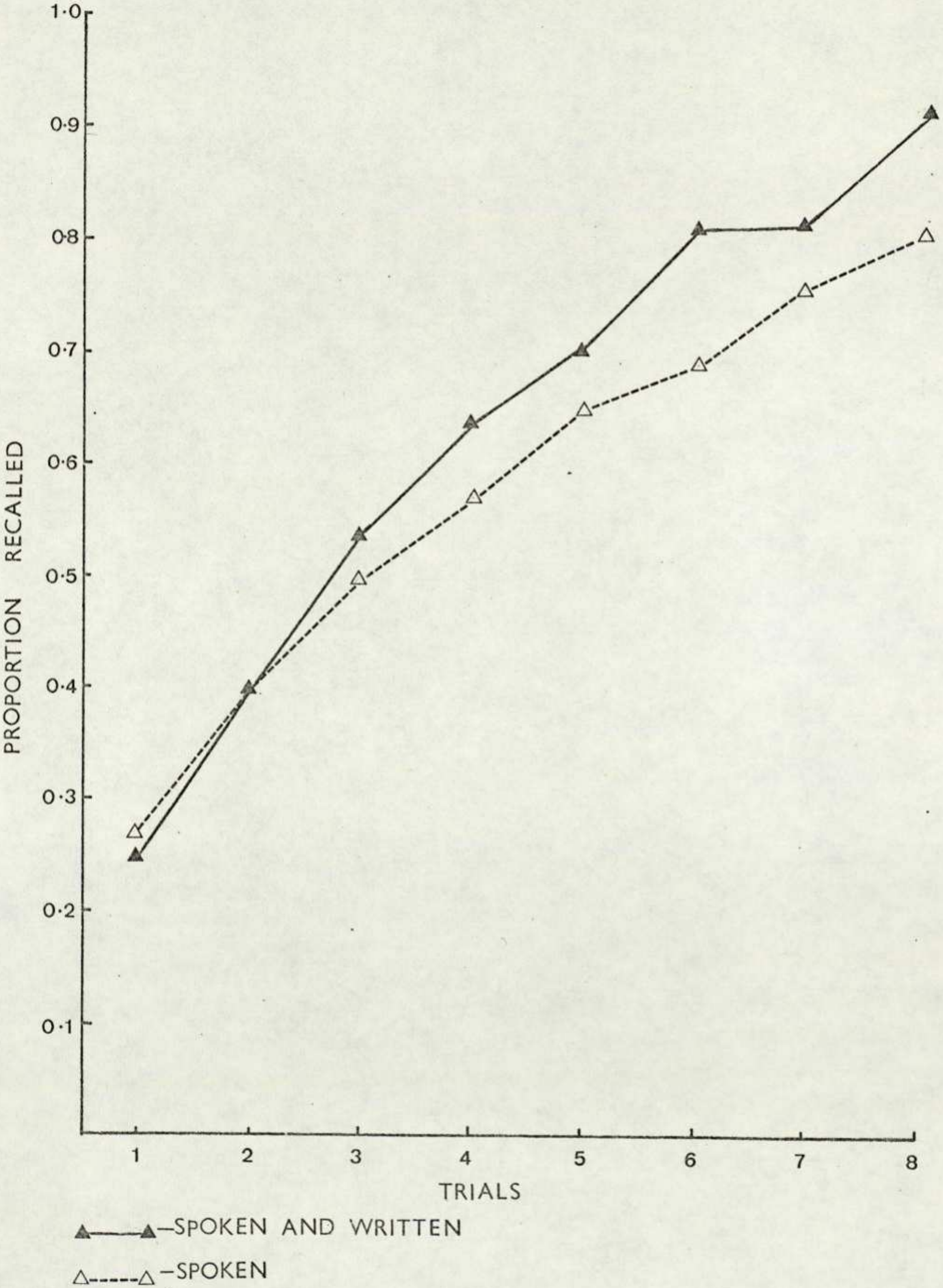
The proportion of words recalled on each trial across subjects was calculated, and plotted (Figure 2.3). Only a very slight, clearly non-significant cross-over effect occurred, which took place on the second trial. Typical multi-trial-free recall learning curves were found (Tulving, 1967).

The number of words recalled by each subject, on each trial, was calculated, and a 2-way ANOVA performed on the data. This revealed a significant main effect of trials only, ($F(7,154) = 182.61$, $p < 0.01$), and a significant interaction between trials and mode of recall ($F(7,154) = 2.54$, $p < 0.05$). Learning thus increased at a faster rate over trials when recall was both written and spoken.

The mean number of repetition errors made by each subject, over all

FIGURE 2.3. EXPERIMENT 2.

PROPORTION RECALLED AS A FUNCTION OF MODE OF RECALL AND TRIALS.



of the 8 trials was 9.8 words, where recall was spoken, and 0.583 words where recall was both written and spoken. That many more repetitions occurred when recall was spoken, than where it was in two modes, is not surprising since in the two modes of recall condition, the recall protocol was available for inspection. No such difference was observed for intrusion errors, which averaged 1, over all trials, per subject, for the spoken mode of recall group, and 0.583 for the two modes of recall group.

Repetition errors increased steadily over trials, from an average of 0.083 per subject on trial 1, to an average of 1.833 on trial 8, where recall was spoken. There were too few repetition errors for calculations to be performed for the two modes of recall group. The increase in repetition errors over trials can probably be attributed to the increased difficulty of monitoring output as the number of words recalled on each trial increased.

A post hoc examination was made of PRNI (the priority in recall of newly recalled items) effects, within each of the recall conditions. The PRNI effect, originally observed by Battig, Allen and Jensen, (1965), refers to a general tendency on the part of the subject to recall items with a poor recall history early in the recall trial. Whilst Battig et al. suggested that this effect might arise from a strategy of giving additional attention during study to previously non recalled items, and/or postponing the recall of strong items until weaker items had been recalled, Rundus (1974), explains this effect as arising from a strategy of giving additional rehearsal to previously non recalled items. Whatever PRNI effects are attributable to, it seems clear that an active discrimination of previously recalled from non recalled items might well underlie the implementation of any such attentional or rehearsal strategy. If this is the case then PRNI effects could be expected to be more pronounced where recall was in two modes, since from Experiment 1 the discrimination of previously recalled from non recalled items was shown to be better where recall had been both written and spoken, than when either written or spoken alone.

To this end recall protocols were vincentised, following Rundus (1974) into pentiles, excluding any primary memory items, which were identified using the procedure of Tulving and Colotla (1970). Primary memory items were omitted from the analysis to ensure that any PRNI effects obtained were not simply a result of subjects recalling from short-term memory. The proportion of words recalled, ($R_{n-1}\ R_n$), or not recalled ($\overline{R}_{n-1}\ R_n$) on the previous trial was calculated within each pentile across subjects, from the total number of $R_{n-1}R_n$ and $\overline{R}_{n-1}R_n$ words respectively occurring on that trial.

Though PRNI effects were present, as shown by the greater proportion of $\overline{R}_{n-1}R_n$ words recalled within the first pentiles on each trial, relative to the proportion of $R_{n-1}R_n$ words, (Table 2.B), no difference in the magnitude of this effect was observed between the one and two modes of recall conditions.

However, totalling the number of $\overline{R}_{n-1}R_n$ and $R_{n-1}R_n$ words for each trial, did reveal that the superior learning exhibited by the two modes of recall group was entirely attributable to their superiority in re-recalling items which had been recalled on the previous trial (Table 2B). There was no difference between the two grpups in the number of items recalled on any trial which had not been recalled on the previous trial.

2.9 Discusssion

Learning was superior where recall had been both written and spoken than where only spoken. However, recall levels did not differ over initial trials between the two groups, (recall in two modes was only slightly lower than recall in a single mode on the first trial), and no cross-over effect occurred.

Within the sampling-with-replacement model of Rundus (1973), the point at which recall levels become equivalent and cross, where recall is

Table 2B

The words recalled on any trial n, can be grouped into 2 categories, (i) those which were also recalled on the previous trial, designated $R_{n-1}R_n$, and (ii) those which were not recalled on the previous trial, designated $\overline{R}_{n-1}R_n$. The proportion of words from each category, falling within the pentiles of each trial, is shown below, as a function of mode of recall.

The total number of words in each category, on every trial is shown in brackets.

Mode of recall / Pentiles	Spoken and Written		Spoken	
	$\overline{R}_1R_2$	R_1R_2	$\overline{R}_1R_2$	R_1R_2
1	0.25	0.14	0.23	0.15
2	0.22	0.20	0.16	0.22
3	0.12	0.28	0.20	0.21
4	0.18	0.21	0.18	0.26
5	0.24	0.18	0.23	0.15
	(93)	(80)	(95)	(72)
	$\overline{R}_2R_3$	R_2R_3	$\overline{R}_2R_3$	R_2R_3
1	0.22	0.16	0.27	0.14
2	0.22	0.16	0.14	0.26
3	0.22	0.19	0.17	0.26
4	0.17	0.24	0.20	0.19
5	0.16	0.25	0.23	0.16
	(116)	(116)	(101)	(109)
	$\overline{R}_3R_4$	R_3R_4	$\overline{R}_3R_4$	R_3R_4
1	0.33	0.12	0.24	0.15
2	0.15	0.22	0.18	0.20
3	0.17	0.22	0.20	0.22
4	0.20	0.20	0.15	0.23
5	0.15	0.24	0.24	0.20
	(92)	(181)	(96)	(144)

Table 2B continued

	Spoken and Written		Spoken	
	$\overline{R}_4R_5$	R_4R_5	$\overline{R}_4R_5$	R_4R_5
1	0.24	0.18	0.25	0.20
2	0.24	0.19	0.20	0.22
3	0.13	0.23	0.13	0.26
4	0.15	0.22	0.20	0.17
5	0.26	0.17	0.23	0.15
	(88)	(213)	(87)	(158)
	$\overline{R}_5R_6$	R_5R_6	$\overline{R}_5R_6$	R_5R_6
1	0.29	0.17	0.23	0.18
2	0.14	0.23	0.19	0.19
3	0.12	0.22	0.12	0.25
4	0.23	0.19	0.21	0.20
5	0.22	0.19	0.25	0.18
	(73)	(244)	(73)	(191)
	$\overline{R}_6R_7$	R_6R_7	$\overline{R}_6R_7$	R_6R_7
1	0.46	0.15	0.24	0.17
2	0.11	0.21	0.22	0.19
3	0.16	0.21	0.11	0.23
4	0.11	0.22	0.19	0.21
5	0.18	0.21	0.24	0.19
	(57)	(294)	(79)	(226)
	$\overline{R}_7R_8$	R_7R_8	$\overline{R}_7R_8$	R_7R_8
1	0.28	0.19	0.35	0.16
2	0.17	0.20	0.13	0.22
3	0.15	0.21	0.10	0.22
4	0.15	0.21	0.14	0.21
5	0.25	0.19	0.29	0.18
	(65)	(313)	(63)	(263)

attempted under two different conditions in multi-trial recall, will obviously depend on how differentially strengthening the two conditions are. However, whilst it is unquestionably true that recall in two modes represents an optimal alternative to recall in a single mode, in that it is sufficiently more strengthening to influence learning, yet not so much more strengthening that recall levels are seriously affected, the full pattern of results is not entirely consistent with Rundus' model. Whilst this is well able to account for the termination of recall in single trial recall, it may represent an oversimplification of the processes influencing recall in a multi-trial situation.

In line with Rundus' model which assumes (i) that retrieval is governed by a ratio rule and (ii) that recall strengthens the association of an item to its retrieval cue (Section 2.5), the recall trial will be terminated sooner under conditions where recall is more strengthening. Strengthening is again taken to refer to the amount of information contained in the trace. In a multi-trial situation on the first trial less items can be expected to be recalled where recall is in two modes, rather than in a single mode. However, the probability of recalling these items on the next trial should be greater than where recall was in single mode, since recall was more strengthening. On any trial after the first, recall is made up of two types of items; those which were recalled on the previous trial (RR), and those which were not recalled on the previous trial ($\overline{R}$R). As a recall trial progresses the influence of recall on the probability of retrieving items remaining in the store will not apply so much to RR items, where recall was both spoken and written. Recall in two modes can only be expected to decrease the chances, relative to single mode recall, of retrieving as yet unrecalled items during the current recall trial, where the items remaining are $\overline{R}$R items. However, it was found here that whilst the two modes of recall group were superior, as expected, at recalling items which they had recalled on the previous trial, there was no difference between the two groups on any trial, in the number of items recalled which were not recalled on the previous trial. The increased probability of inter-trial retention and the equivalent probability of retrieving items not recalled on the previous trial, meant that the two modes of recall group quickly overcame the initial slight disadvantage they were at.

Conditions under which recall is attempted are, of course, not the only factor influencing the amount of information stored in a memory trace. Possibly the tendency to allocate additional attention (Battig et al., 1965) or rehearsal (Rundus, 1974), during presentation, and to give priority in output (Rundus, 1974) to items which were not recalled on the previous trial, served to protect these items from any additional depressing effect recalling in two modes might have had, over and above those of single mode recall, on subsequent retrieval efforts. Neither do apparently similar acts of recall necessarily increment the amount of information in the trace by a constant amount. The effect retrieval has on the underlying representation of an item can vary with the retrieval effort, (Bjork, 1975) such that the more complex and difficult retrieval is, the greater the long term benefits are, (Gardiner, Craik and Bleasdale, 1973). It seems reasonable to suppose that as learning progresses and the number of items consistently recalled increases, that less and less effort will be involved in retrieving previously recalled items from the store. As performance attributes become established as part of the underlying memory trace, less and less information will be added as a consequence of recall. The influence of recall on the retrieval of other, as yet unrecalled, items from the store is thus probably not as straightforward as Rundus' model might suggest.

Since memory for previous recall was more accurate in Experiment 1, where recall was both spoken and written, it was suggested that PRNI effects, which intuitively ought to involve the discrimination of previously recalled and non recalled items, might be greater for the two modes of recall condition in Experiment 2. Although PRNI effects were present no difference in the magnitude of these according to whether recall was in one or two modes was observed. That such a relationship should exist was only a post hoc speculation and the experiment was not designed specifically to test for this. If the ability to discriminate previously recalled and nonrecalled items does mediate PRNI effects the relationship between them may well be indirect. Rundus (1974) suggests that additional rehearsal is only given to some of the items which were not recalled on the previous trial, during a study period, and these items are then given priority in recall. He suggests that 'there appears to be no consistent strategy for emphasising previously unrecalled items' during a study period. Many other diverse strategies are open to the subject to engage in during study.

Finally, the fact that the rate of learning over trials directly reflected the number of performance features which were available during recall, corroborates the findings and conclusions of Experiment 1. Clearly performance features do become established as part of the trace representing an item in memory.

2.10 General Conclusions

Performance features available during recall do apparently become established as part of the underlying representation of an item in memory and mediate, at least to some extent, memory for past recall. Such performance features are utilised in learning and the more features there are available during recall, as for instance where recall is both written and spoken, the faster learning is.

A sampling-with-replacement model of memory (Rundus, 1973), appeared to fit the results of Experiments 1 and 2 quite well, although it was suggested that the concept of strength within this context be re-defined. Since recall obviously has qualitative effects upon the memory trace of an item, conceptualising recall as simply resulting in an increment in trace strength is inadequate. In addition it was suggested that Rundus' model might represent an oversimplification of the processes occurring during recall in a multi-trial recall situation. The influence of recall upon the underlying representation of an item cannot be assumed to be constant but will vary with the nature of the retrieval effort (Bjork, 1975).

CHAPTER 3

The Development of the Ability to Assess Past Recall Performance

3.1 Introduction to Experiments 3 and 4

Both of the experiments reported here, were designed to investigate the development of the ability to discriminate previously recalled from non recalled items.

The paradigm used was the same as that employed in Experiment 1, in that a series of recall trials were followed by a recognition test, in which recalled items were to be identified as such. Signal detection methods were used to analyze the data.

Experiment 4 was essentially a replication of Experiment 3 except that it was designed to provide a task at a more appropriate level for each group.

3.2 Introduction to Experiment 3

Young children often fail to make use of a mnemonic strategy, appropriate to a given task, even when they are capable of doing so (Flavell, Beach and Chinsky, 1966; Keeney, Cannizzo and Flavell, 1967). One important factor, in the emergence of the spontaneous use of strategies, would appear to be the ability to assess memory performance. It seems unlikely that the child would use a given strategy if he was unaware that his memory performance was in need of improvement and could be improved, or that he would use a strategy appropriately, unless he was able to monitor the effects it had upon memory. However, as yet, no such relationship between memory appraisal skills, and the use of strategies has been demonstrated. In fact most research shows that young children are as skilled as older children, and adults, in appraising their memory (Kelly, Scholnick, Travers and Johnson, 1976; Moynahan, 1976).

The present experiment attempts to clarify further the development of the ability to assess past memory performance, and is aimed in particular at overcoming one or two objections which can be raised to the procedures of Kelly et al. (1976), and Moynahan (1976). Both of these experiments required that an estimate of recall performance be given after each recall trial, a procedure which would prime subjects, and hence provide an inflated measure of the knowledge children normally hold of their past memory performances. The present experiment overcomes this by giving only one final unexpected test of what was recalled. Moynahan, and Kelly et al., also asked respectively simply for estimates of either how many items had been recalled, or whether the placement of items in an array was correct or not. Here it is required, more specifically, that recalled and non recalled items be identified as such.

Masur, McIntyre and Flavell (1973), did ask for the identification of recalled and non recalled items, and found children to be 90% accurate in this. However they asked this of only eight children, (Mean age 7.2 years), who had used a strategy of studying previously non recalled items the least, following a single study-test trial. In the present experiment this procedure was expanded to include a wider age range of children, and more recall trials. Further, performance was measured by means of the more sensitive d' scores, derived from a signal detection theory analysis.

3.3 Method

Subjects

Seventeen children within each of the following age ranges, 5 to 7 years, 8 to 10 years, and 11 to 15 years, were tested. The mean age of each group, where the age of each child was taken to the nearest year, was 6.8, 9.0 and 12.9 years respectively. Subjects were tested in small groups of mixed ages.

Design

The children were presented with a series of pictures plus their labels, and were required to free recall, where recall was written. This procedure was repeated for ten trials, then the pictures plus labels were re-presented, and the subjects asked to indicate which of these they had recalled.

Materials

A pool of 126 words, naming objects highly likely to be encountered in the first few years of life, was drawn up. These objects were depicted by simple line drawings on cards with their labels printed underneath. The cards were then randomly arranged into 14 sets each of 9 cards. 10 sets were chosen at a time for presentation to one group of subjects, the order of words within lists, and the order of lists being randomized after each presentation.

For the recognition test the same pictures plus labels, though on a smaller scale, were printed on sheets of paper, 18 items to a sheet. Each subject thus received 5 of these sheets, stapled together in a random order.

Procedure

The subjects were told that they would be shown some pictures and that the names of the depicted objects would be called out. Then, when instructed to do so, they were to write down as many of these names as possible. After an example had been given E held up 9 cards, one at a time, each for a duration of two seconds, and read out their labels. Two Es were present to aid the youngest members of the group in recording their recall.

A maximum of 1½ minutes was allowed for recall, then the subjects were told to turn the page of the recall booklet and the procedure was repeated for ten trials.

Finally, recall booklets were collected, and the recognition sheets with all 90 initially presented items, both depicted and labelled printed on them, handed around. Ss were instructed to look at every item on these sheets, and to place a tick beside those they believed they had written down in the recall booklet, and a cross beside those they had not.

3.4 Results

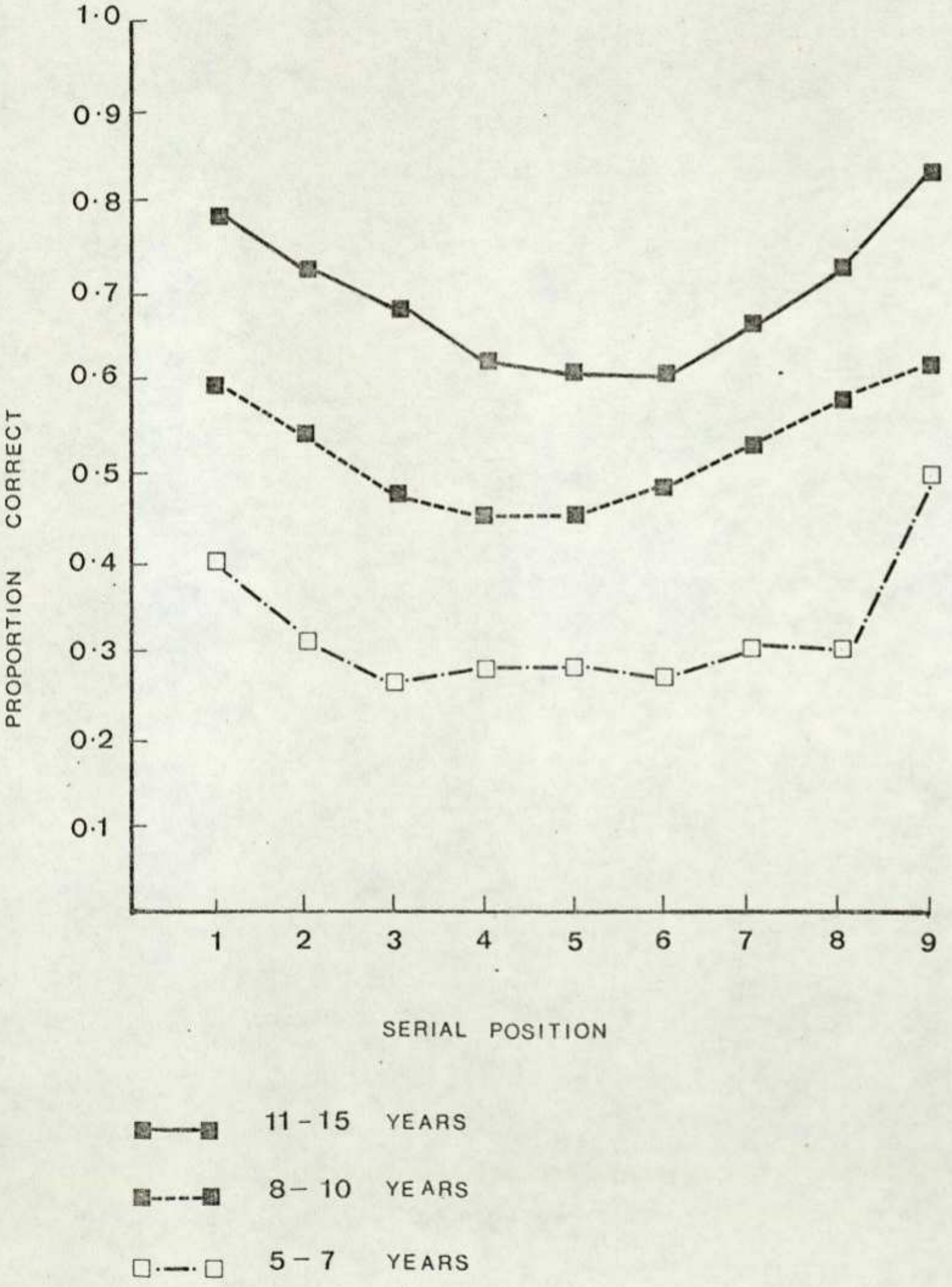
The recall data will be considered initially, then the recognition of recall data.

The proportion of words recalled, over all subjects, is shown as a function of age group and serial position in figure 3.1. Typical primacy and recency effects are present, and there is a steady increase in the level of recall with age. A 2-way ANOVA performed on these data revealed the reliability of these effects, (serial position, F (8,384) = 14.66, $p<0.01$; age, F (2,48) = 24.168 $p<0.01$).

Repetition errors in recall totalled 7 for the 5 to 7 year olds, 3 for the 8 to 10 year olds, and 12 for the 11 to 15 year olds. Intrusion errors were virtually non existent, totalling 3 over all 3 age groups. Thus there are no apparent age related differences in the monitoring of output, as evidenced by the error data, though written recall may have masked true age differences in an internal checking process (see section 1.3.c.), since an external protocol was available for inspection. This possibility is examined in Experiment 4, where childrens' recall is spoken.

FIGURE 3.1. EXPERIMENT 3.

PROPORTION RECALLED AS A FUNCTION OF AGE AND SERIAL POSITION.



For the recognition of recall data, a signal detection analysis was used, and d' scores were derived from a classification of responses into hit and false alarm rates. Hit rate was calculated as the proportion of items recalled which were correctly identified as such, and false alarm rate as the proportion of items not recalled, which were incorrectly identified as recalled.

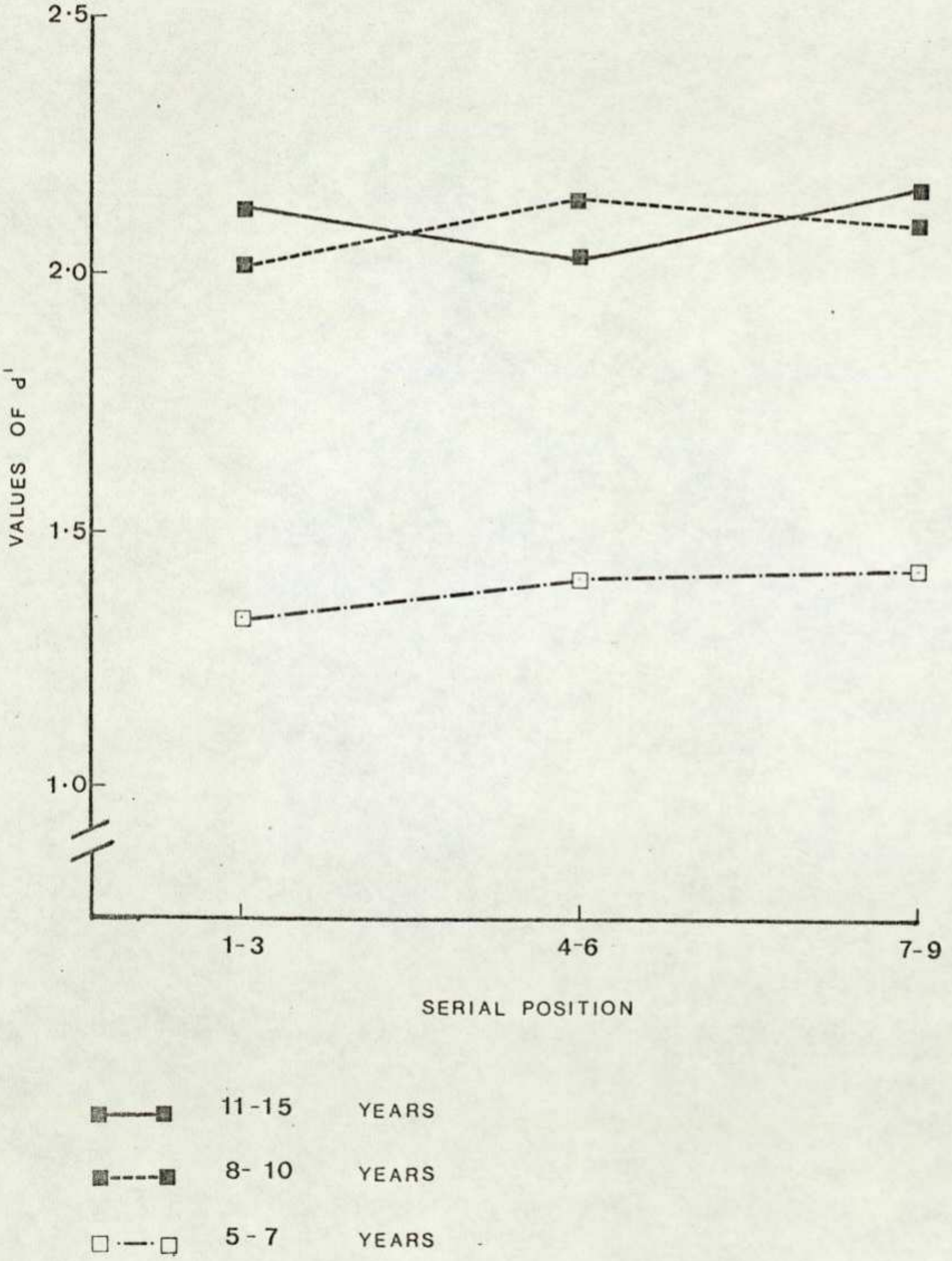
Serial position was collapsed into 3 points, positions 1-3, 4-6 and 7-9, inclusive, and an aggregate d' score derived for each age group, (figure 3.2). Aggregate scores were derived since in some cases there were too few data points for reliable, individual estimates of d' scores to be made. It is clear from figure 3.2 that the negative recency effects found by Gardiner and Klee (1976), and Klee and Gardiner (1976), are not present.

Still considering figure 3.2, it is clear that whilst the two older groups have performed at much the same level as each other on the recognition of recall task, the youngest group performed much worse. Individual d' scores were calculated, ignoring serial position, and a 1-way ANOVA performed on the data. This confirmed that age was a significant factor ($F(2,48) = 4.34$, $p < 0.05$).

A separate analysis of false positive rates was undertaken (for rationale see section 2.4), revealing that there were no significant differences in response bias, according to age or serial position.

To assess the contribution of each of the ten years of age represented in the sample tested, to the development of the ability to recognize previously recalled items, a scatterplot of age against the mean d' value obtained by the subjects within each year was constructed (figure 3.3. appendix). This showed that there was no untoward contribution of any year of age to the d' value of their respective groups; performance increased steadily as a function of age, up to 8 years, where it levelled off.

FIGURE 3.2. EXPERIMENT 3.
d' VALUES AS A FUNCTION OF AGE AND SERIAL POSITION.



3.5 Discussion

Both level of recall and memory for past recall were found to improve with age. That memory span increases with age is a well documented finding (Flavell et al., 1970; Tversky and Teiffer, 1976), and Huttenlocher and Burke (1976), report that span increases from 2 to 8 items, between the ages of 2½ and 16 years. In contrast, though, whilst recognition of previously recalled items was found here to improve over the ages of 5, 6 and 7 years, levelling off at about age 8, most evidence in the literature points to a lack of age related differences in the ability to assess past memory performance (Masur, et al, 1973; Moynahan, 1976; Kelly et al., 1976).

However there are certain factors in this experiment, which might have biased the results in favour of the older groups. First of all recall was written, a fact which must have caused some difficulty for the youngest group with their relative lack of writing skills.

Secondly, the use of both pictorial and verbal stimuli may have made the task somewhat more difficult for the youngest group. Pictures were used as stimuli, from a commonly made, but mistaken assumption, that pictorial representation precedes verbal representation in the developmental sequence. Rohwer (1970), presents evidence however, that children derive maximum benefit from pictorial stimuli only to the extent that they are able to store a verbal label simultaneously. He states that a 'preference for and a capacity to make effective use of visual representation and storage develops later than is the case for verbal modes of representing and storing information', P(401). Whilst labels were always given in the present experiment with the pictures, children were left to their own devices on the recognition test. Perhaps they could not read, or supply, the appropriate label themselves, or simply chose to consider the pictures alone. Whatever the effects of this, it seems clear that the memory task would have been fairer to all of the subjects, if verbal stimuli alone had been used, and presented auditorily throughout the experiment.

The use of pictorial stimuli may also underlie the absence of negative recency effects in the present experiment, since Madigan, McCabe and Itatani (1972) found poorer memory for terminal list items in delayed recall, where pictures, rather than words, were used. Further discussion of this will be postponed, until the results of Experiment 4 are considered.

Finally the task of recall, and of recognizing recall, appeared to be simply too easy for the older groups, the materials and list lengths being designed to suit the capabilities of the youngest children. That developmental effects may be due to the ease with which stimuli are identified, is supported by Huttenlocher and Burke (1976) who suggest that developmental differences in recall span are attributable to differences in the ease with which items are identified at encoding.

Experiment 4 was designed then, to equate the difficulty of the task presented to different age groups, in all respects except in the ability to assess past recall.

3.6 Introduction to Experiment 4

In order to overcome the possible confounding factors present in Experiment 3, Experiment 4 was designed to eliminate any differences in the task presented to each age group, other than those due to the ability to recognize previously recalled items. To this end the stimuli were chosen to represent a suitable difficulty level for each group, and were arranged into lists of an appropriate length. The words were chosen from the age of acquisition norms of Carroll and White (1973 a), such that the age of acquisition of each set of words was just below that of the group it was to be presented to. Each set of words was then arranged into lists, the lengths of which varied according to the age group they were presented to, the youngest group receiving the smallest list lengths. So that each age group was presented with approximately the same total number of items, the number of lists presented to each group was also varied. Thus the youngest group received more lists of fewer words, which were all acquired by a mean age of four years.

In addition to this, recall was spoken. Spoken recall permitted a more accurate assessment of the extent to which children monitor their output at recall to be made, than did the written recall of Experiment 3.

3.7 Method

Design

Experiment 4 was essentially identical to Experiment 3, in that a series of recall trials were followed by a final recognition test of which items had been recalled, and differed only in the mode of recall used, and the nature and organization of the stimulus materials.

Subjects

42 children were chosen to form three groups, the mean ages of which were, (A), 6.02 years; (B), 7.88 years, and (C) 10.02 years. Subjects were tested individually.

Stimuli

The stimuli were selected from the age of acquisition norms presented by Carroll and White (1973, a). For group A, (five and six year olds), the stimuli were 72 words acquired by both males and females, according to the norms by a mean age of four years. For group B (seven and eight year olds), a random 50 of group A's words, plus 40 words acquired between the ages of 4 and 6, by males and females, were used. Similarly for group C, (9 and 10 year olds), a random 20 of group A's words plus the 40 words used for group B, which were acquired between 4 and 6, plus another 30 acquired between the ages of 6 and 8 were used. In this way it was hoped that the material would be of approximately equal difficulty level to each of the 3 groups.

These words were arranged at random in 8 lists of 9 words for group A, 7 lists of 12 words for group B, and 6 lists of 15 words for group C, the order of words within lists, and the order of lists being randomised for each subject. It was expected from the data of Experiment 3 that subjects would be able to recall approximately 50% of the lists of the respective length. Thus on the recognition test no group would have the advantage that recalling most or very few of the items initially presented to them, might give.

For the recognition test the words were removed from list order and printed, for each group, in large capitals on to 3 sheets of paper. These 3 sheets were then stapled together in a random order.

Procedure

A list of words was read out at a rate of one word every 2 seconds. S was instructed to call out as many of these as he could remember, saying the last few read by E, first. This latter instruction was given to try and equate recall strategies across subjects. A period of $1\frac{1}{2}$ minutes was allowed for recall before the next list was read. E recorded Ss recall.

After all of the lists had been presented the recognition test sheets were placed in front of S. A 'Yes' response was to be given if that word had been called out previously and a 'No' response if it had not.

E controlled the rate at which the words were considered by reading each word aloud at a rate of one every 3 seconds.

3.8 Results

The proportion of words recalled over all subjects, as a function of age and serial position, is shown in figure 3.4. The proportion of words recalled of the total number presented to each group, appears to be approximately equal across the 3 groups.

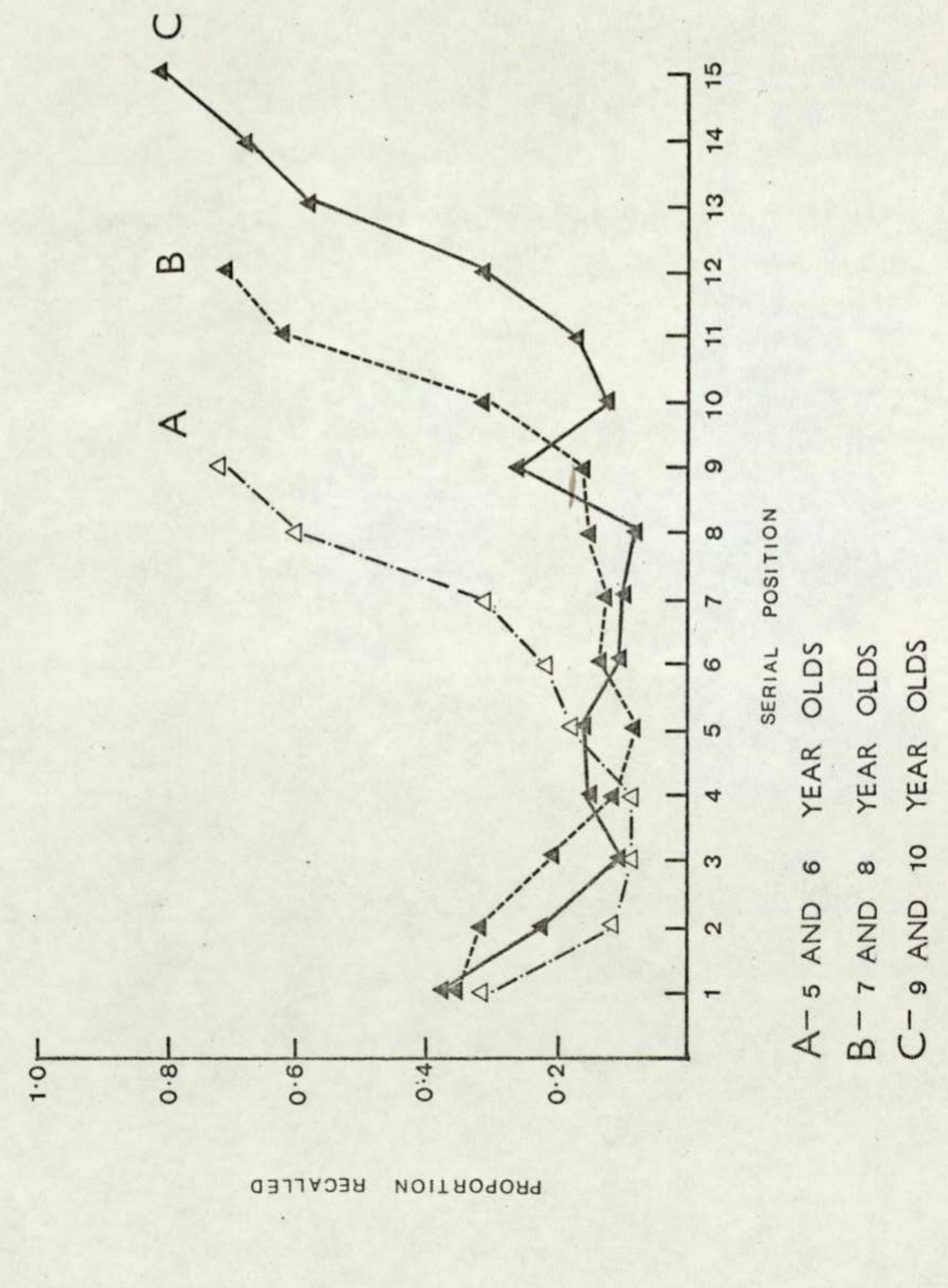
To enable direct comparisons of the recall levels of the 3 groups, in relation to serial position to be made, serial position was collapsed into 3 points for each group; group A, positions 1-3, 4-6, and 7-9 inclusive, group B, positions 1-4, 5-8, 9-12 inclusive, and group C, positions 1-5, 6-10, and 11-15 inclusive. Whilst these divisions did not represent an equal number of items on the serial position curve, it was felt that in view of the approximate memory span of the respective age groups, they represented roughly comparable portions of the serial position curve. The total number of items, over trials, falling within each of these divisions, was approximately equal across the 3 groups. The proportion of words recalled within each of these divisions was calculated for each subject, and a 2-way ANOVA, with repeated measures on one factor, performed on the data. The only significant effect was that of serial position, ($F(2,78) = 14.24$, $p<0.01$), indicating that typical primacy and recency effects were present. In figure 3.4. the recency effect is somewhat pronounced probably due to the modified free recall instructions used in this experiment.

Whilst the manipulation of list length, number and composition, were successful, in that equivalent proportions of words were recalled by each group, (A = 0.29, B = 0.28, C = 0.29), it should be borne in mind that these were calculated from different totals. Thus typical age related effects in the level of recall were found, the mean number of words recalled per list being 2.6, 3.9 and 4.3, for groups A, B and C, respectively.

Both repetition and other kinds of error were higher in this experiment than in Experiment 3, totalling 33 repetitions and 72 other

FIGURE 3.4. EXPERIMENT 4 .

PROPORTION RECALLED AS A FUNCTION OF AGE AND SERIAL POSITION.



errors over all groups, as opposed to 22 repetitions and 3 errors in Experiment 3. This increase in errors can probably be attributed to the absence of the recall protocol during spoken recall, (see section 2.5.). The total number of repetition and other errors, did not differ substantially across the 3 groups in this experiment; (repetitions, A = 5, B = 16, C = 8; intra-list intrusions, A = 18, B = 10, C = 14; extra list intrusions; A = 13, B = 9, C = 6). While the relative lack of errors in recall by the youngest group would suggest that no age differences in the operation of an internal monitoring process exist, (see section 1. 3. c), it may well be the case that there are age differences in the extent to which items are retrieved prior to their monitoring.

A signal detection analysis, as in Experiment 3, (see section 3.4), was used to analyze the recognition of recall data. Mean d' values are shown in figure 3.5, as a function of age and serial position, where serial position was collapsed into 3 points as above. As in Experiment 3, recognition of recall seems to be age related, but in contrast to Experiment 3, negative recency effects are present, probably due to the use of verbal stimuli.

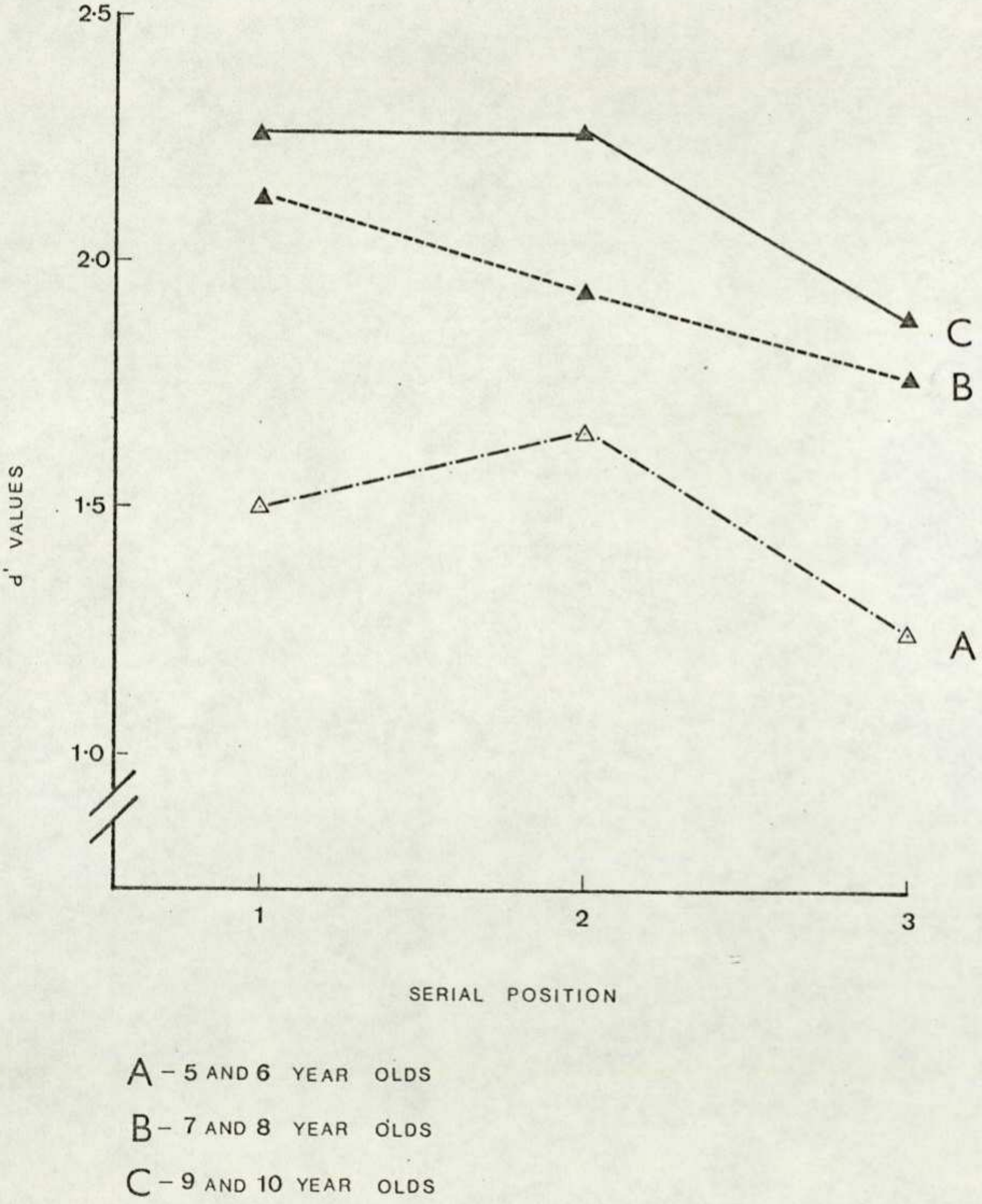
A 2-way ANOVA, with repeated measures on one factor, showed the negative recency effect and the age effects to be reliable (age $F(2,39) = 3.40$, $p<0.05$; serial position, $F(2, 78) = 6.68$, $p<0.01$).

Post hoc comparisons showed that the significant effects of age were largely attributable to differences in performance between the youngest and the two oldest groups, on the recency portion of the curve, (Between groups A and C, $F(1,39) = 6.51$, $p<0.01$; between groups A and B, $F(1,39) = 5.15$, $p<0.01$), and between the youngest and the oldest group, on the primacy portion of the curve, (between A and C, $F(1,39) = 4.9$, $p < 0.01$).

An analysis of false positive rates, (for rationale see section 2.4.), showed no significant differences in response bias, due to age or

FIGURE 3.5. EXPERIMENT 4.

MEAN d' VALUES AS A FUNCTION OF AGE AND SERIAL POSITION.



serial position.

A scatterplot of year of age against the mean d' value, scored by the children falling within that year, indicated, as in Experiment 3, that there was no untoward contribution of any particular year to the recognition of recall performance (figure 3.6. appendix).

3.9 Discussion

The results of Experiment 3 were replicated, in that despite the task and materials being of a more appropriate level for each age group involved, memory for past recall varied as a function of age. Whilst children of 5 and 6 years were poor in discriminating previously recalled from non recalled items, performance improved significantly with age.

However, in contrast to Experiment 3, negative recency effects in memory for past recall performance were found. The discrepancy in results can probably be attributed to the nature of the stimuli used as suggested earlier, (section 3.5), negative recency effects typically being found with verbal stimuli in a recognition of recall task (Gardiner and Klee, 1976), but not on a delayed recall task where pictorial stimuli are used (Madigan, McCabe and Itatani, 1972). The differential effects of verbal and pictorial stimuli, upon the recency portion of the recognition of recall function, may well be due to the differential depth of registration each gives rise to. Whereas a long term representation of an item may be established immediately, where this is depicted (Madigan et al., 1972), a long term representation of a verbal item is established only after a certain duration of stay in a short term store. Typically terminal list items are not in the short term store, for a period long enough to ensure their representation in long term memory, (Craik, 1970).

Whilst age related differences in memory appraisal skills were found in Experiments 3 and 4, others have found no such effects (Moynahan, 1976; Kelly et al., 1976; Masur et al., 1973). However, it may be that the discrepancy between these results and those of the present experiments, lies principally within the point at which the ability to appraise past memory performance was tested. The lack of errors in recall, made by the youngest group, in the present experiment, seems to suggest, in accord with the findings of Masur et al., that young children are able to discriminate effectively, recalled from non recalled items, during and immediately following recall. This could equally well be due to age differences in the extent to which items are sampled from memory prior to their monitoring, in which case the monitoring of output might be somewhat redundant for the youngest group. Whether the lack of errors in recall, in the youngest group, can be attributed to effective monitoring of output or to a relative lack of retrievals from memory, could be tested by means of Bousfield and Rosners, (1970), free versus uninhibited recall paradigm, (section 1.3.c.).

The results of Experiment 1, and the work of Gardiner and Klee (1976), and Klee and Gardiner (1976), show that memory for past recall reflects the monitoring of output, and varies also as a function of initial item registration effects. Thus the poor memory for past recall exhibited by the younger children can probably be attributed to either one or both of these factors. If age related differences in the extent to which an available pool of memory traces are sampled (section 2.5.) exist, then it would be reasonable to expect parallel differences in the monitoring of output as a consequence, and thus memory for past recall would be affected. If on the other hand there are no age related differences, in the extent to which recall is monitored, then there may still be age differences in the manner in which it is accomplished. Young children possibly do not encode performance features arising from the act of recall, which Experiment 1 showed to be important in mediating memory for past recall. This could be tested in the same way as in Experiment 1 by attenuating the feedback of the response in recall.

Considering initial item registration effects, there is some evidence that the encoding of items by young children may be somewhat less elaborate and therefore less durable than that of adults and older children. The pattern and richness of encoding varies with age, the emergence of various encoding dimensions reflecting the salience of these to adults so that the more salient emerge earlier developmentally, (Hagen, Jongeward and Kail, 1975). Bach and Underwood (1974) and Means and Rohwer (1976) present evidence that acoustic encoding is relatively more important to the child than the adult, the adult using it in conjunction with a greater number of other encoding dimensions. However, if registration effects did underlie young children's poorer memory for past recall, differences in the pattern of results in relation to serial position, rather than simple main effects, might have been expected to occur (Brown, 1975).

3.10 Conclusions

The findings of the present experiment suggest that memory appraisal skills are not as well developed in the young, as previous research might have indicated. In so far as memory for past recall reflects the monitoring of output, and initial item registration effects, then either or both of these factors may underlie the poor memory for past recall exhibited here by younger children. The possibility that the spontaneous use of strategies depends upon the initial development of memory appraisal skills thus remains open.

CHAPTER 4

Episodic Theory: The Representation of an Act of Recall and the Nature of the Search Process Involved in the Retrieval of Episodic Information from Memory

4.1 Introduction to the Chapter

The main issue to which episodic theory, (e.g. Tulving, 1972, 1976; Tulving and Thomson, 1973) has been directed is the question of whether episodic information is represented by modifications or taggings of semantic structures, or whether it is represented independently of these. There is evidence to suggest that the retrieval of episodic information from memory involves a search along what appears to be a temporal dimension within a system which is other than semantic, (Murdock and Anderson, 1975; Carroll and White, 1973b). Using the reaction time paradigm of Murdock and Anderson, Experiment 5 tests the proposal, derived here from the basic assumptions of episodic theory, that an act of recall results in the formation of a unique trace in episodic memory.

The nature of the episodic system and of the search process which takes place within this when episodic memory is interrogated is examined further in Experiment 6.

The question of whether an interpolated recall task influences a subsequent recognition task arose from Experiment 5 and was tested in Experiment 7.

4.2 Introduction to Experiment 5

Semantic memory is generally conceived of as a network of nodes, where a node might represent a word, a concept, or an image, (Rumelhart, Lindsay and Norman, 1972; Anderson and Bower, 1972, 1974). The nodes are interconnected by associative pathways and the location of a node within the semantic space thus defines its meaning. Typically episodic information is assumed to be represented in the form of markings or taggings of semantic structures. In a recognition task the test item allows direct access to its representation and the subject simply has to decide whether or not the node so accessed is appropriately marked. Only recall involves a search, along semantic pathways, which can be aided by the presence of semantic associates to the target.

Much evidence has now accumulated, however, to suggest that episodic information is represented in a system which is separate to and independent of semantic memory, (Tulving, 1972, 1976; Watkins, Ho and Tulving, 1976; Tulving and Watkins, 1977). Tulving distinguishes two kinds of memory, episodic and semantic memory. Whilst semantic memory stores the meaning and referents of items in the fashion of an internal lexicon, episodic memory stores only personally experienced events in terms of their spatio-temporal relations to one another. The episodic-semantic distinction is basic to episodic theory, (Tulving, 1976; Tulving and Thompson, 1971, 1973). This assumes that the unit of storage in episodic memory constitutes an item with its context. Since every event is unique in terms of its context, it is necessarily represented by a unique trace, which is retained independently of semantic structures.

Only one line of evidence favouring episodic theory was pursued further in this thesis. Murdock and Anderson (1975), found that retrieval of episodic information appeared to involve a search within a system in which the temporal sequence of events was directly preserved in the format of the store, rather in the fashion of a tape recorder. Reaction times in a recognition task were found to increase steadily as a linear function of an item's testing position, where the item was a distractor, and as a linear function of the number of items intervening between presentation and test (lag), where the item was a target. These results were interpreted in terms of a backward, serial, self-terminating scan taking place over both study and test items within a system in which events are organised according to their sequence of occurrence. In the case of distractor items the search was presumed to return to the beginning of the study list. Reaction times increased with response ratings of uncertainty, and more than one interrogation of memory was presumed to be taking place. The traces stored in this system were assumed to become degraded over time so that recognition accuracy decreased as a function of test position and lag.

Since Tulving (1972, 1976), proposed that (i) retrieval from either episodic or semantic memory constitutes an input into episodic, and (ii) that nominally identical events separated in time are unique with respect to their autobiographical referents and would thus have more than one representation in memory, it was here deduced that an act of recall might also be represented by a unique trace in episodic memory. This trace might differ from earlier traces of the same nominal item in terms of performance features (from Experiment 1 of this thesis) as well as other contextual elements.

The notion that an act of recall results in the formation of a unique trace in episodic memory was tested here by means of the paradigm and model of Murdock and Anderson. A list of words was presented for free recall and followed by a recognition task in which the subject was required to discriminate recalled, old but not recalled, and new distractor items. Latencies for these responses were recorded. Given an episodic system in which the temporal sequence of events is preserved in the store, and from which retrieval takes place by means of a backward serial self-terminating scan from the point of test, it was predicted that reaction times to items which had been recalled should be faster than those to items which had been presented but not recalled, and in turn these should be faster than reaction times to distractor items. Such a finding could however only be taken as supportive of the experimental hypothesis and interpreted in terms of episodic theory, if response latencies within each response category also showed linear increases as a function of test position or lag, where lag for recalled items was measured as the number of items intervening between recall and test. Differences in reaction times between the three response categories could otherwise simply be interpreted as a function of the same factors which resulted in those items being recalled or not recalled initially.

4.3 Method

Subjects

The subjects were fifteen male and female undergraduates of the City University. All subjects were tested individually.

Design

22 words were presented for immediate spoken free recall. Then these words plus another 22 new words were shown in a tachistoscope and the subjects asked to respond 'new', if the word _had not_ been presented, 'old' if the word _had_ been presented, and 'recalled' if the word had been recalled. Confidence ratings of 'sure', or 'unsure' were required for these responses, and response latencies were recorded.

This procedure was repeated four times.

Materials

(i) _Stimuli._ 176 6-letter words, chosen at random from the Toronto word pool, were stencilled in black letters about the centre point of white cards. The cards were then randomly divided into 4 sets of 44 words, 22 of which served as targets, and 22 as distractors. Whereas one subject received one set of 22 words as targets, the next subject received these as distractors.
(ii) _Apparatus_. One field of a 3-field tachistoscope was used in the recognition phase of the experiment. This was fitted with an automatic

card changer, to which a digital timer was wired and onset whenever a card was displayed. This timer was stopped by a manual key press from the subject, and reaction times were automatically recorded by a print-out machine.

Procedure

E read out a list of 22 words, at a rate of one word every second. On completion S was asked to free recall, for which a maximum period of 1½ minutes was allowed. E recorded S's recall.

Then the subject was asked to look in the tachistoscope, where the 22 words read to him ('old'), some of which he had recalled, ('recalled'), plus 22 other words, ('new'), were shown, to which he was to respond as indicated. In addition, a confidence rating of 'sure' or 'unsure' was required for each response. The order of presentation of the 44 words was randomised for each subject, and the order of presentation of the 4 word sets was also varied.

The subjects were instructed to rest their hands close to the manual key press and to depress this when making a response. It was explained that the time taken to respond would be recorded, so that responses should be as fast but as accurate as possible.

Each card was automatically displayed for 3.5 seconds, and the field was off for 5 seconds. Four practice cards were inserted at the beginning of the first trial.

This procedure was repeated four times.

4.4 Results

Though it should be noted that recall levels were fairly low, (of the 88 words presented over 4 trials to each subject, only 14 on average were recalled), the initial recall data is not of major interest here, and will not be considered further.

For the recognition of recall data all reaction times are reported in milliseconds. In order to exclude any extraneous influences of uncertainty and incorrectness, only the response latencies where the response was correct and rated as sure were included in the following analyses.

For items which had been recalled lag was calculated as the number of items intervening between the recall of that item and its test. For old items, which had not been recalled, lag was calculated as the number of items intervening between presentation and test, inclusive of those items which were recalled. Test position alone was considered for new items.

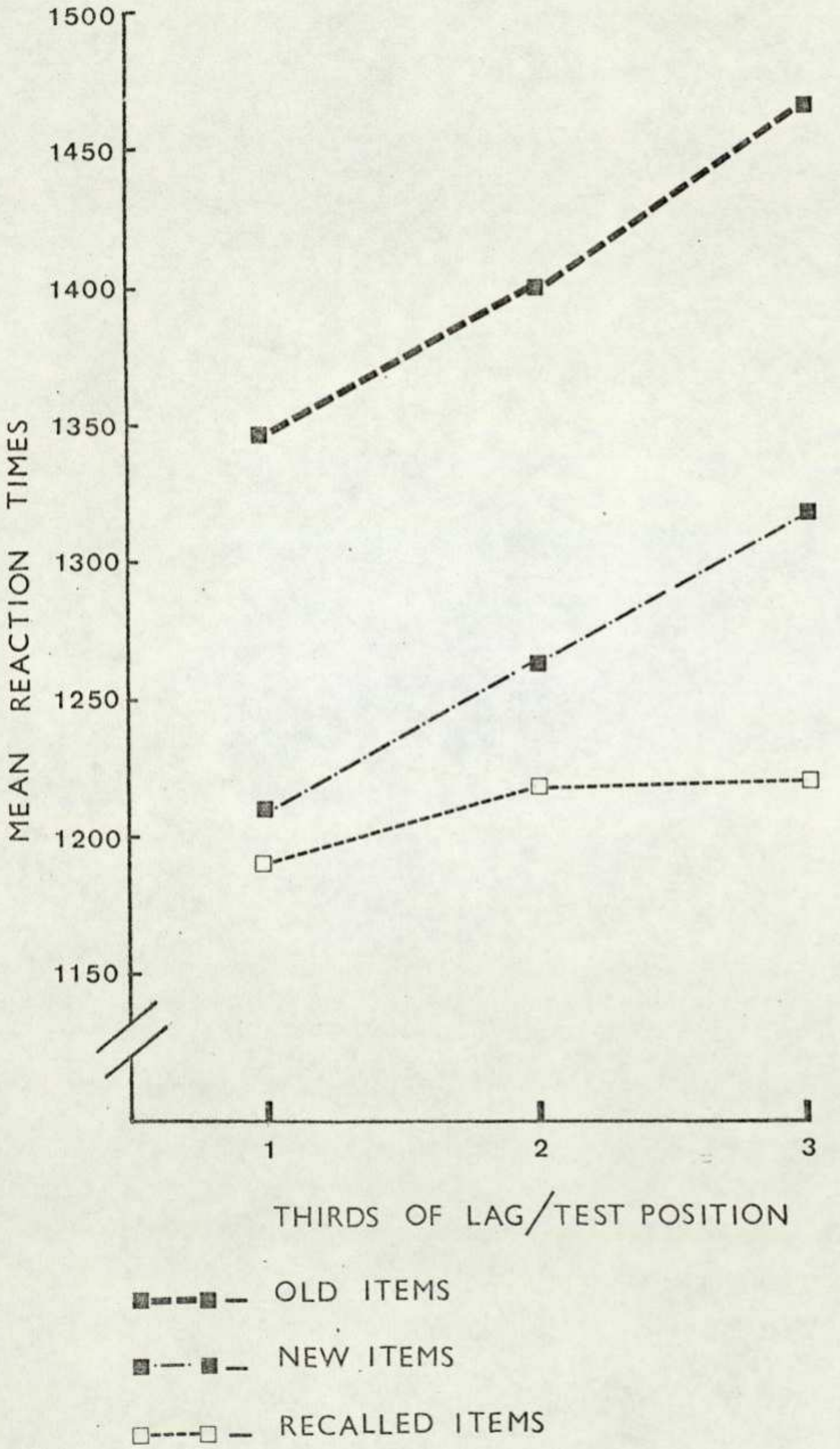
Since lag was not systematically controlled the number of responses occurring within a given range of lags and test positions were not equivalent. To overcome this, response latencies were vincentised into thirds in order of increasing lag or test position. The mean reaction time occurring within each vincentised third was calculated for each subject, as a function of response type, and these means are shown in Figure 4.1. From this graph it is apparent that reaction times to items which had been recalled were faster than those to items which were simply old, or new. However, response latencies to items which had been recalled show no apparent linear increase as a function of lag.

A 2-way ANOVA, with repeated measures on both factors, compared response latencies within each response type and across lag or test position. This revealed a significant effect of response type only, ($F(2,28) = 5.59$, $p < 0.01$). Trend analyses showed that the functions graphed in Figure 4.1 had no significant linear or higher order components.

In order to compare response latencies for the three types of response, these data were collapsed over the vincentised thirds of test position or lag. Whilst vincentising had allowed the influence of test position or lag upon response latencies to be assessed, post hoc comparisons within each vincentised third would have been uninformative and probably invalid,

FIGURE 4.1. EXPERIMENT 5.

MEAN REACTION TIMES AS A FUNCTION OF VINCENTIZED THIRDS OF LAG AND TEST POSITION.



since reaction times had been considered as a function of test position for new items, and of lag for old and recalled items. Post hoc comparisons on the collapsed data revealed a significant difference only between response latencies to recalled and old items, $(F,(1\ 28) = 7{\cdot}35,\ P < 0{\cdot}01)$.

A 3x3 probability matrix was constructed over all subjects to consider the accuracy of recognition performance, (Table 4A). The probability of scoring a hit is shown in the top left to right diagonal for recalled, new and old items. Whilst the probability of scoring a hit where the item had been recalled was rather high, the probability of recognising an old, non recalled item was low by comparison.

Table 4A

Subjects' Response

		'Recalled'	'Old'	'New'
Test Item	Recalled	0.89	0.08	0.03
	Old	0.12	0.52	0.36
	New	0.03	0.22	0.75

Probability Matrix. Experiment 5.

The probability of correctly recognising a recalled, a new and an old item as such (hits) is shown in the top left to right diagonal.

4.5 Discussion

Whilst response latencies were fastest for recalled items, these showed no linear increase with lag. An interpretation of this result in terms of the same factors influencing recognition reaction times as those which initially led to certain items being recalled and others not, seems untenable, or at least premature for the following two reasons.
1) Reaction times within the other two response categories did show evidence of linear increases as a function of lag and test position though these did not reach significance. 2) The data for recalled items is probably unreliable since only a small proportion of the total words presented were recalled. Of the 88 words presented to each subject, only 14 were on average recalled, and again only some of these correctly recognised as such.

Lack of conclusive support for the proposal that an act of recall results in the formation of a unique trace in episodic memory may also be attributable to the way in which lag was calculated for recalled items. According to the experimental hypothesis items which have been recalled would have two representations in episodic memory, one formed at presentation and another upon recall. Thus recalled items would also have two lags, that between presentation and recall, and that between recall and test, only the second of these measures being considered in this experiment. Whilst it was tacitly assumed that for recalled items to be identified as such only the second of the two traces would need to be accessed it is not inconceivable that the first trace, and thus the first lag was also influential.

Though linear trends relating reaction times for non recalled and new items to lag and test position respectively were discernible, they failed to reach significance. It seems that the difficulty of the task may have interfered to some extent with these functions. Subjects appeared to find recognition of old non recalled items particularly difficult and it is possible that the interpolated recall task was partly responsible for this, (Kay and Skemp, 1956; Brown and Packham, 1967; Packham, 1968). Experiment 7 was designed to follow up this suggestion.

On the whole relatively little support for the notion that recall results in the formation of a unique trace in episodic memory was found. Neither was there any strong evidence in favour of Murdock and Anderson's model. The following experiment was designed to investigate reaction time-lag and reaction time-test position functions further, as well as to consider the nature of the search process which takes place when episodic information is retrieved from memory.

4.6 Introduction to Experiment 6

To account for linear increases in reaction times found in a recognition task as a function of lag, in the case of target items, and of test position in the case of distractors, Murdock and Anderson (1975) proposed that a backward, serial, self-terminating scan encompassing both study and test items was taking place within a memory system, organised in a spatio-temporal fashion. That test items should also be included in this search seems a somewhat redundant procedure. It would make better sense, for instance, if the search process by-passed test items and sought a match for the probed item amongst the stored representations of target items. The following experiment was consequently designed to assess whether the search process does indeed proceed backwards from the point of test. A further aim was to investigate whether semantic cues are utilised in the search as Tulving (1976, p. 71) suggested; 'since the scanning speed in episodic memory may be very high - Murdock has obtained rates close to 1 msec per item, and higher rates cannot be ruled out - the data also make more reasonable the proposition that semantic and other kinds of specific cues can be effectively and quickly used to locate information in episodic memory'.

To accomplish these aims two groups of subjects were tested in a standard recognition task, in which half of the items were targets and half were distractors, and to which they were to respond 'Yes' and 'No' respectively. For the experimental group, also known as the 'related' group, the distractors were pairs of semantically associated words. Each pair was associated more strongly in one direction than in the reverse direction, so that if A is likely to evoke the response B, B is much less likely to evoke A. These distractor pairs were inserted into the recognition test sequence so that the first member of each pair to be encountered was that least likely to evoke the other. For the control group, (the 'unrelated' group), other non-associated words were substituted for the first members of each distractor pair.

Since the distractor pairs for the experimental group were associated strongly in a direction which was the reverse of that of the recognition test sequence, it was assumed that their presence could only influence the speed and correctness of responding if a backward search utilising semantic cues took place over test as well as study items. More specifically the experimental group were expected to make more false alarms (mistaking distractors for targets) than the control group. When presented with the second distractors of the pairs in the recognition test, a backward search through the test items would yield, for the related group, a word closely matching the probe in meaning. Given that semantic features form part of the attribute ensemble in terms of which a match is sought, this close match could be mistaken for an actual match, and a false positive error made. No apparent match for the probe would of course be found by the control group. Since the second distractors of the pairs were unlikely to occur as implicit associative responses, (Underwood, 1965) to the first distractors, false alarms would presumably not be made without a backward search, on the basis of that same item having been 'encountered' earlier. Where second members of the distractor pairs were correctly identified as distractors by the experimental group, it was expected that response latencies would be slower than those of the control group. Extra response time would presumably be required to consider and then correctly reject the word closely matching the probe in meaning.

The present experiment was also designed to re-assess reaction time-lag and reaction time-test position functions, since, contrary to predictions, no significant linear trends describing these were found in Experiment 5. It was aimed to discover whether any more convincing evidence in favour of Murdock and Anderson's model would be obtained in a task where no recall trial intervened between presentation and test, and in which a standard discrimination of targets and distractors was required. In addition, lag was systematically controlled, and distractors and targets distributed evenly across the recognition test sequence. Lag was manipulated so that targets presented in certain study positions were tested at given positions in the test sequence. This was achieved by the simplest means available; targets were tested either in the same order as that in which they were presented, or in the reverse order. This meant that lag increased at two different rates with test position, increasing at a greater rate across test position where the study order had been reversed. Where study order was reversed lag would be smaller over initial test positions and larger over final test positions, in relation to where study order was preserved in the test sequence. Thus in so far as lag is an important variable in influencing response latencies, an interaction effect between these two conditions was expected.

4.7 Method

Design

6 lists of 16 words were presented. After each list these words plus 16 distractors were shown one at a time, in a tachistoscope. The subject was required to respond 'Yes' if the word had been presented, and 'No', if the word had not been presented, and to provide confidence ratings for

these responses. Response latencies were recorded.

The subjects were divided into two groups, designated here the related and the unrelated groups. For the related group the distractors for each list consisted of 8 pairs of words associated in one direction much more strongly than in the reverse direction, so that if A is quite likely to evoke response B, B is much less likely to evoke response A. These pairs were inserted in the recognition test sequence at a distance of 1 to 5 items apart, the first member of each pair to be tested being the one least likely to evoke the other.

For the unrelated group, the distractors were unassociated pairs of words. Whereas the second members of each of these pairs were the same words used for the related group and were placed in identical positions in the test sequence, the first pair members were other unassociated words.

Subjects

The subjects were 32 male and female undergraduates of the City University, half of whom were assigned to the related group, and half to the unrelated group, on a random basis. All subjects were tested individually.

Materials

(i) Apparatus

One field of a 2-field tachistoscope was used. A digital timer was wired up to this which onset automatically whenever a word display was initiated. Depression of a manual response key stopped the timer, and response latencies were recorded automatically.

(ii) Stimuli

The distractor pairs for the related group were chosen from the norms presented by Keppel and Strand in Postman and Keppel's (1970), word association norms. 48 pairs of words were chosen so that they were related in the strongest direction from 20 to 70% and in the weakest direction not more than 15%. The 96 target words, and the 48 distractor words used as first pair members for the unrelated group, were all chosen at random from the tables presented in Postman and Keppel. This was done to provide a pool of words relatively indistinguishable from the associated word pairs in terms of word frequency, word length, and part of speech. They were chosen so that there were no obvious associations amongst them.

The target words and distractors were divided randomly into 6 sets each of 32 words, with the constraint that each set contained 16 targets and 16 distractors. In the case of the related group there were 8 associated pairs of distractors, and for the unrelated group 8 non-associated pairs of distractors. Halfway through the experiment this process was repeated and list membership reshuffled to control for any possible unplanned associations which might have existed between the targets and distractors within any one list.

For presentation in the tachistoscope, all of the words were stencilled in black letters on white cards.

(iii) The recognition test

Lag was manipulated for both groups so that it increased at two different rates with test position. To achieve this presentation order was varied, and targets were either presented in the same or the reverse order to that in which they were tested. For this purpose the presentation list was treated as 4 blocks (ABCD) each of 4 target words and the recognition test similarly as 4 corresponding blocks, (ABCD) each of 4 targets and 2 pairs of distractors. The 4 targets tested in block A were those which were also presented in block A, and similarly for blocks B, C and D. Whereas the testing sequence was always ABCD, presentation order was varied so that on three randomly selected trials the list was presented in order ABCD, and on the remaining three trials in order DCBA. Where a list was both presented and tested in order ABCD, lag increased only slightly as a function of test position, the minimum and maximum lags possible being respectively, 12 at the beginning of the test sequence, and 31 at the end of the test sequence. Where a list was presented DCBA and tested ABCD, however, lag increased at a much greater rate over test position, the minimum and maximum lags possible here ranging from 0 at the beginning of the test to 46 at the end.

Whereas 2 subjects, one in each group, received a certain list of targets in the order ABCD, another two subjects, again one from each group, were presented with this list in the reverse order, DCBA.

Two targets and two distractors were placed within each half of the four blocks of items comprising the recognition test. The first members of two distractor pairs occurred in the first half of each block, and the second members of these pairs in the second half of the blocks. These pairs were semantically associated in the case of the related group, and unassociated in the case of the unrelated group. The sequence of occurrence of the two pairs was preserved across the two halves of the block so that if the first member of pair X occurred before the first member of pair Y in the first half of the block, then member 2 of pair X would also occur before the second member of pair Y in the second half of the block. The order of words within blocks, and the order of the four blocks was varied as much as was possible for every two subjects within each group. The number of items intervening between the two members of a distractor pair could thus vary from 1 to 5 words.

Every four subjects received an identical recognition test sequence except that two of these subjects belonged to the unrelated group, and thus had other unassociated distractors substituted for the first members of the associated pairs presented to the related group.

Procedure

Each target list was presented auditorily at a rate of one word every two seconds. Then the test cards were shown in a tachistoscope and the subject required to respond 'yes' or 'no' depending on whether or not the word had been presented. Confidence ratings on a 3-point scale were required for these responses, where 3 = certain, 2 = fairly certain, and 1 = unsure. The subject was asked to depress a manual response key, at the same time that he initiated a response, and his reaction time was automatically recorded. The subjects were asked to rest their hands beside the manual response key throughout the experiment, and were instructed to be as fast but as accurate as possible in making their responses.

Each card was displayed for 3 seconds and the field was off for 4 seconds. Two buffer cards were inserted at the beginning of each test sequence, the responses to which were not included in any analyses, following Murdock and Anderson (1975).

This procedure was repeated six times.

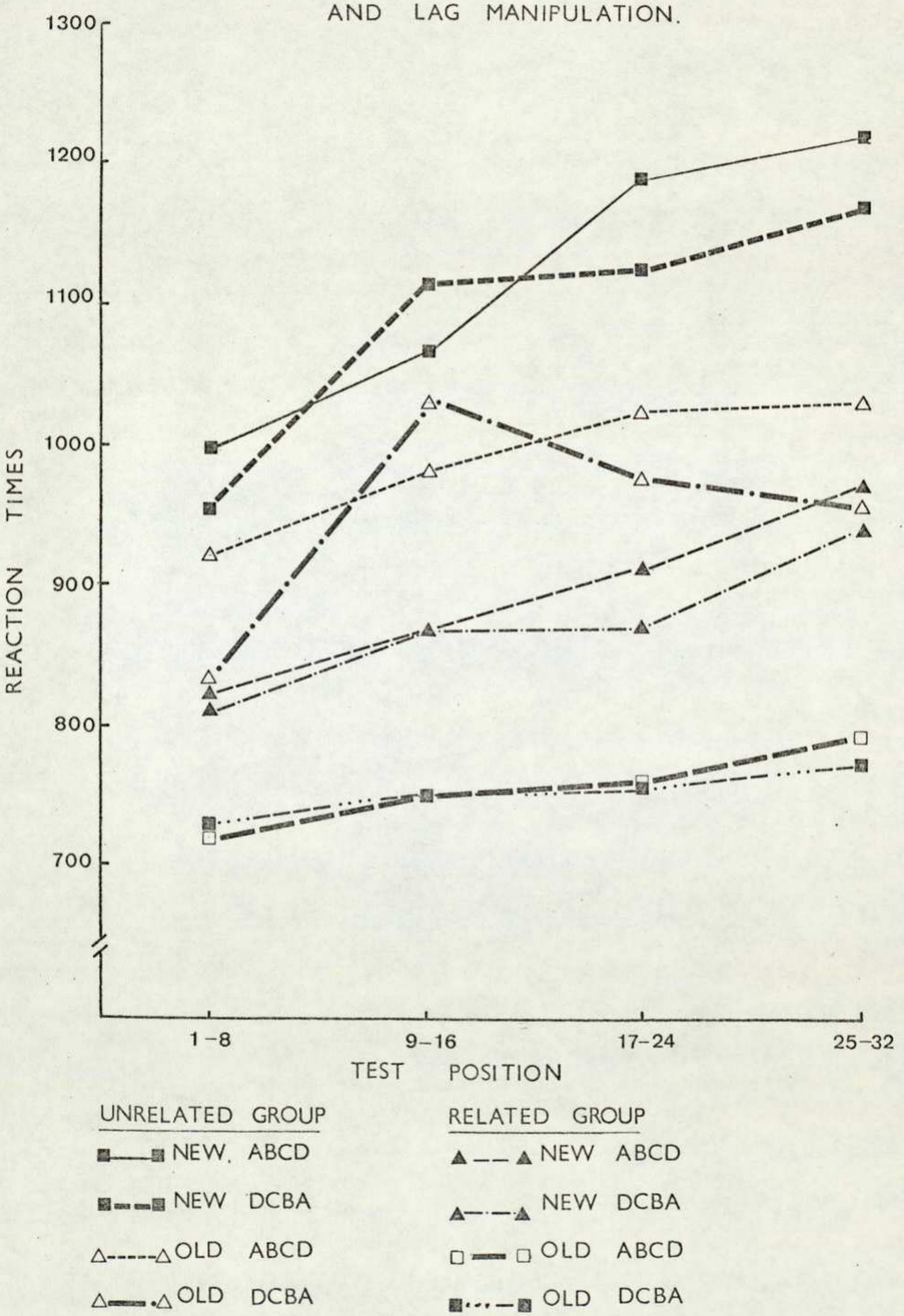
4.8 Results

Considering the reaction time data first of all, only those response latencies where responses were correct and rated as certain were included in the following analyses, (after Murdock and Anderson, 1975). All reaction times are reported in milliseconds.

For purposes of analysis, test position was collapsed into 4 blocks, (positions 1-8, 9-16, 17-24, 25-32 inclusive), and the mean reaction time found for each subject within the two groups as a function of test position, response type and lag manipulation. Mean reaction times calculated over all subjects are shown in Figure 4.2. In general reaction times were slower to both targets and distractors within the related group, and slower within both groups to distractors than targets. Reaction times clearly increase as a function of test position. The influence of the lag manipulation seems to be somewhat less straightforward than was predicted.

A 4-way ANOVA performed on these data showed that the differences between the two groups were significant, and that the effects of response

FIGURE 4.2 EXPERIMENT 6.
REACTION TIMES WITHIN EACH GROUP AS A FUNCTION OF TEST POSITION RESPONSE TYPE AND LAG MANIPULATION.



type and test position were reliable, (related/unrelated group, $F(1,30) = 4.70$ $p < 0.05$; response type, $F(1,30) = 24.91$, $p < 0.01$; test position, $F(3,90) = 24.08$, $p < 0.01$), whilst those of the lag manipulation were unreliable. Only one 2-way interaction reached significance, that of test position and response type ($F,(3,90) = 4.19$, $p < 0.01$).

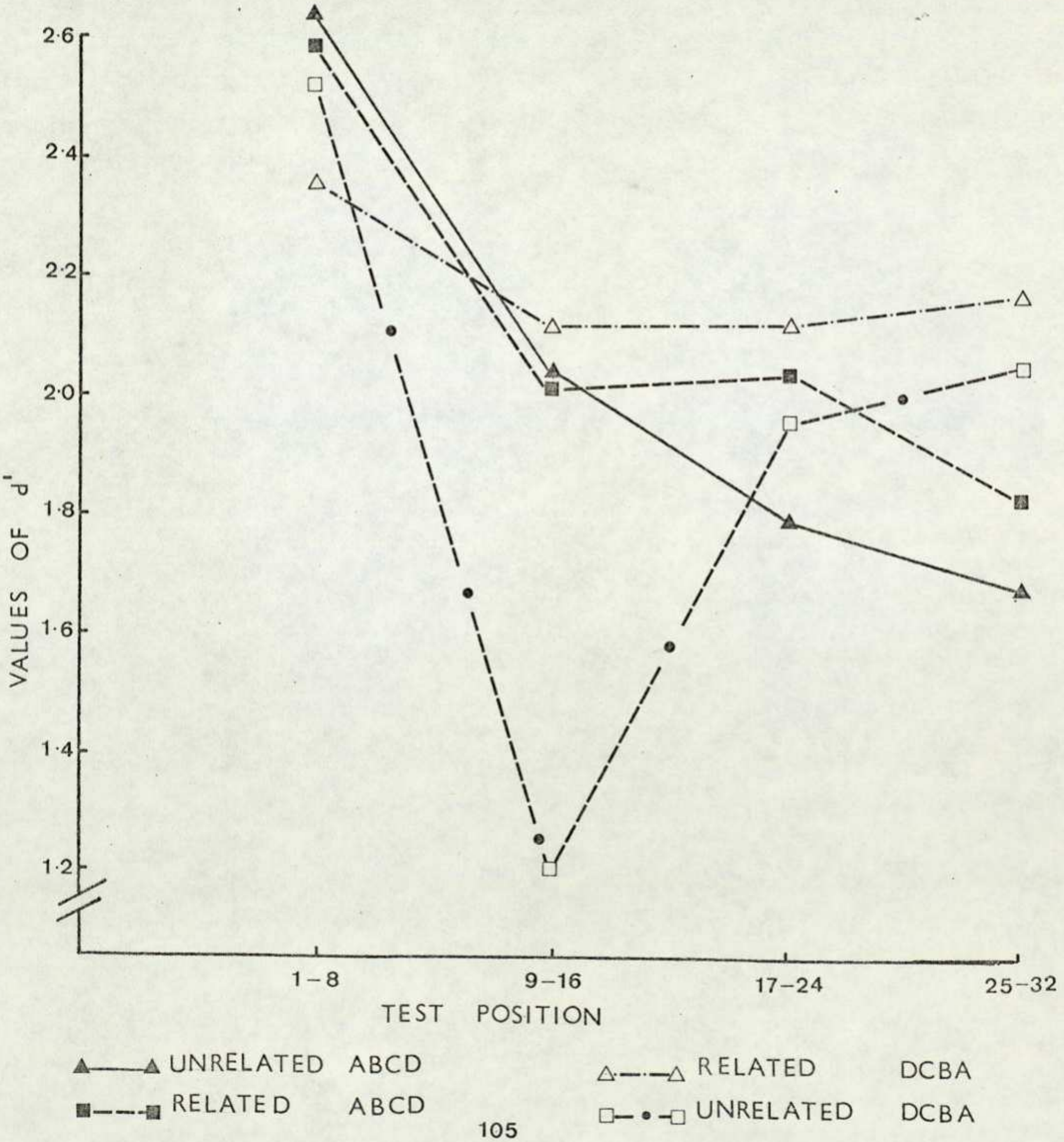
Since the lag manipulation had no significant effects the data were collapsed over these and analysed for trend components. These analyses revealed significant linear components in the functions relating reaction times to targets and distractors for both the related and unrelated groups. In addition, for the unrelated groups the function describing reaction times to targets also had a significant quadratic component, (targets, unrelated group, $F_{lin}(1,45) = 10.53$, $p < 0.01$; $F_{quad}(1,45) = 6.47$ $p < 0.05$; distractors, unrelated group, $F_{lin}(1,45) = 35.93$, $p < 0.01$; targets, related group, $F_{lin}(1,45) = 17.84$, $p < 0.01$; distractors related group, $F_{lin}(1,45) = 27.8$, $p < 0.01$).

The data were collapsed further over test position, (since post hoc comparisons which took this factor into account would have been relatively uninformative), and post hoc comparisons, using a Newman Keuls procedure performed. These compared total reaction times within and between each group for targets and distractors. Of all the comparisons made only that between reaction times to targets for the unrelated group and reaction times to distractors in the related group, failed to reach significance at the 0.01% level.

Accuracy of recognition was assessed using a signal detection theory analysis. Performance was classified into hit rate, (the proportion of targets correctly identified as such), and false alarm rate, (the proportion of distractors incorrectly identified as targets). A d' score was derived for each subject within both groups as a function of test position and the lag manipulation. Test position was collapsed into four points, as above. Mean d' values are shown in Figure 4.3. As expected, recognition performance

FIGURE 4.3 EXPERIMENT 6.

MEAN d' VALUES AS A FUNCTION OF LAG MANIPULATION, TEST POSITION, AND ASSOCIATIVENESS OF DISTRACTORS.



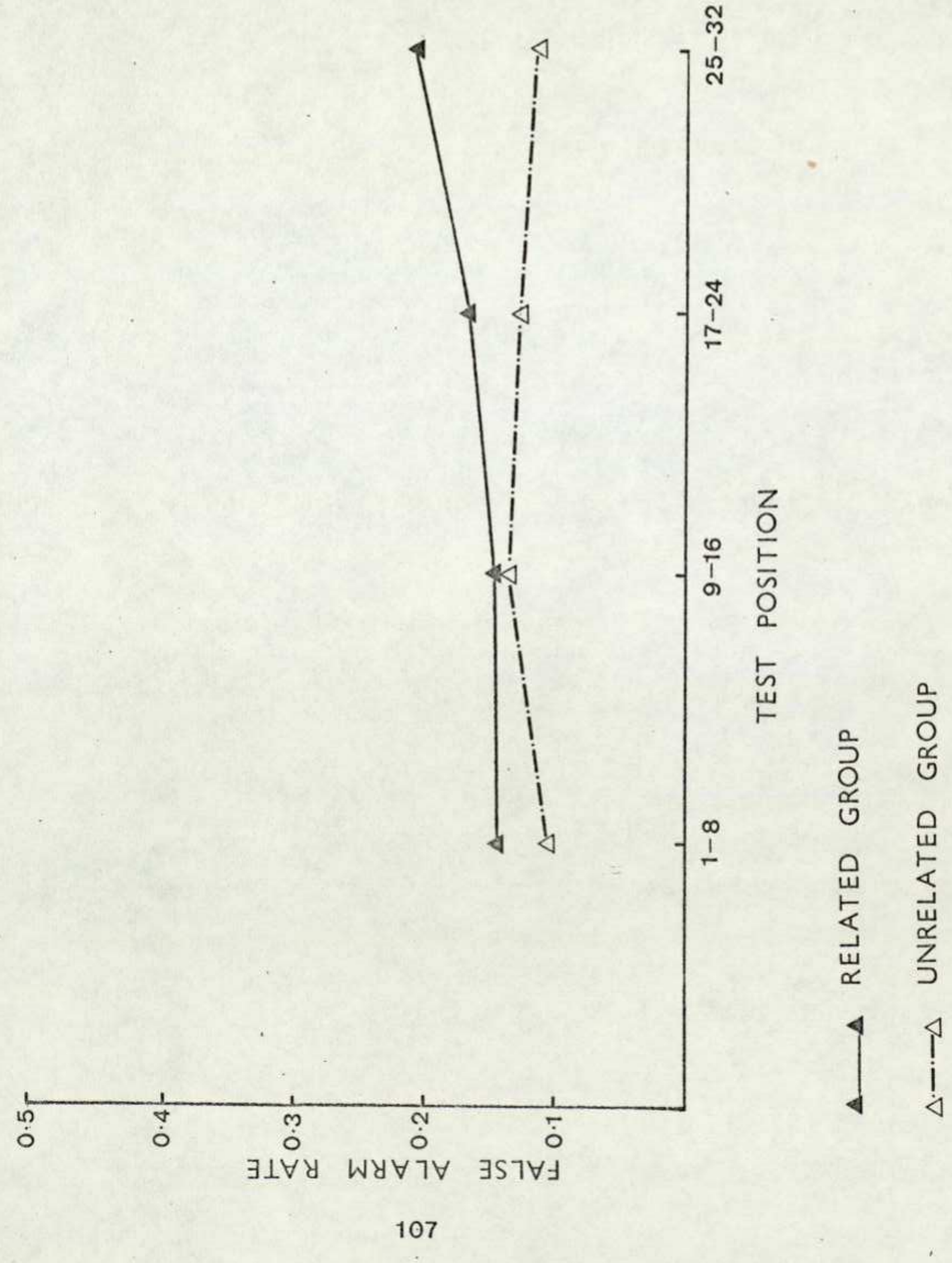
appears to decline with test position. A 3-way ANOVA confirmed that this effect was reliable, ($F(3,90) = 17.8$, $p < 0.01$). Recognition performance was unaffected by whether distractor pairs were associated or not, and by the lag manipulation.Only one two-way interaction effect reached significance, that of test position and distractor associativeness, ($F(3,90) = 3.5$ $p < 0.05$).

The possibility remained that considering recognition accuracy for both targets and distractors, and for both the first and second pair members in the latter case, had obscured a trend towards an increased false alarm rate to the second critical members of the distractor pairs in the related group. A further analysis was thus undertaken to compare false alarm rates to only the second members of the distractor pairs in the related and unrelated groups. Mean false alarm rates to these second members of distractor pairs as a function of test position are shown in Figure 4.4. The data were collapsed over the lag manipulation here since this had proved to be of no significance above. Although false alarm rates do seem to be somewhat greater where distractor words were preceded by an associated word in the test sequence, a 2-way ANOVA performed on the data revealed no main effects of test position or associativeness of distractor pairs, nor were there any significant interaction effects.

4.9 Discussion

Regardless of whether the test item was a distractor or a target the presence of semantically associated pairs of distractors in the recognition test meant that a search of memory was accomplished much more swiftly than where non-associated distractor pairs were present. Two possibilities might account for this result. 1) The presence of associated words in the recognition test may simply have had the general effect of inducing subjects in the related group to encode test items more deeply (Craik and Lockhart, 1972), and to make a search based to a larger extent on semantic features. 2) A more attractive possibility is that the trap posed by distractors closely

FIGURE 4.4. EXPERIMENT 6.
MEAN FALSE ALARM RATE TO SECOND MEMBERS
OF DISTRACTOR PAIRS.



matching probe items in meaning may have been realised. To avoid this a finer discrimination in terms of semantic features would have to be made. This may well have led to all probe items being encoded in a deeper, more elaborative fashion (Craik and Lockhart, 1972),than would otherwise have been the case, and a match sought along these dimensions. Such a search would presumably be faster than one which relied to a lesser extent on semantic features (Tulving, 1976). If either of these possibilities is justified then the unexpected direction of the results is not necessarily contrary to Murdock and Anderson's model. Since the associations between distractor pairs were in the reverse direction to that of the test sequence these should not have influenced the speed and possibly the nature of the search process, unless a backward search using semantic cues took place. However this remains to be confirmed.

Whilst significant linear relationships were found between reaction times to new items and test position, the relationship between reaction times to old items and lag did not emerge as clearly as anticipated in this experiment. Though an interaction effect seemed to be occurring between the two lag conditions, this was insignificant and was not of the simple crossover nature predicted. However, the lag manipulation was inadvertently confounded here with serial position effects, which have been shown to influence reaction times in investigations of short term memory, (Clifton and Birenbaum, 1970; Corballis, 1967). Where the presentation order of targets was reversed relative to testing order, recency items were tested first and primacy items tested last. Where presentation and testing order corresponded, recency items were tested last, and primacy items,tested first
The fact that recency items,
/ which are quite poorly registered in memory, (Craik, 1970), were tested last in the corresponding presentation-test order conditions may well account for the converging rather than the diverging, as expected, of reaction times within the two conditions over the final test positions.

Whilst more false alarms did occur to second members of distractor pairs, where these were semantically associated to first pair members, this difference did not reach significance.

Although some aspects of these data were consistent with Murdock and Anderson's model of memory, the full pattern of results was not wholly supportive of it. No conclusive evidence as to whether or not a backward search utilising semantic cues takes place over test and study items when episodic memory is questioned was found here. The following experiment considers a question which arose from Experiment 5, namely whether an interpolated recall task influences recognition of presented but non recalled items.

4.10 Introduction to Experiment 7

It was previously suggested (Section 4.5) that the relatively poor discrimination of items which the subject had failed to recall in Experiment 5, may have been due to the recall task interpolated between presentation and recognition testing. Whilst intuitively 'recall should strengthen the retention of items recalled and leave unaffected the retention of items not recalled so that, on average, performance in a recognition test will be facilitated', (Brown and Packham, 1967, p.356), recognition has been shown to be worse following recall than after an interval filled with an irrelevant activity, (Brown and Packham, 1967; Kay and Skemp, 1956). Kay and Skemp showed subjects a scene of a riverback and asked one group of subjects to recall items from it. A control group engaged in some other activity in the interim. Recognition of 30 items compiled from the scene was 64.2% for the control group as compared to only 43.5% for the experimental group. Whilst recognition performance was high in both cases for the 10 most frequently recalled items, the experimental group recognised only about half as many of the remaining 20 less popularly recalled items as the control group did.

The locus of this depressing influence of recall upon recognition has been attributed to a shift in recognition thresholds brought about by the presence of stronger items in the recognition test. Where stronger items are present the subject is 'trying to detect a relatively weak signal amongst many stronger signals', (Kay and Skemp, 1956, p. 161), so that weaker signals which might ordinarily have been detected go unnoticed. Packham (1968) attributes the influence of recall on recognition to alterations in the subject's set which the presence of familiar items in the recognition test brings about. Two groups of subjects both received a recall test prior to recognition. The recognition task consisted of two stages. In one the test contained no previously recalled items, (test B) and in the other the test contained recalled items, (test A). Recognition performance was superior where recognition test B was completed first. Packham suggests that 'The initial presence of highly familiar recalled words could convey the impression that crude judgements were required for the experimental task, whereas in fact fine ones were needed to separate the non-recalled words from distractors', (p. 291-292).

In contrast to the above experiments, Darley and Murdock (1971) found that an intervening recall task was inconsequential for subsequent recognition performance. A series of word lists were followed either by recall or by an irrelevant activity. Whilst final recall levels benefited from an initial testing, final recognition performance was unaffected. A similar paradigm to Darley and Murdock's was employed in the present experiment, such that the presentation of word lists was followed on half of the trials by recall and on the remaining trials by an anagram solving task. However, instead of a final forced choice recognition task a free, unconstrained (in that subjects could select as many items as they wished as targets) recognition test was given after <u>every</u> trial. Performance was analysed by means of a signal detection theory analysis. This procedure more closely approximated that of Experiment 5. It was speculated that if the influence of recall upon non recalled items arises from the contrasting of better and worse known items at the testing stage, any effects of recall on recognition would not have been so apparent where a forced choice procedure involving the comparison of only three items at a time was used. Whilst Brown and Packham found an effect of recall on recognition using a forced choice procedure, all of the test items were presented together and the subjects required to rank these.

4.11 Method

Design

Six lists each of 20 words were presented. On a random basis free spoken recall took place after half of these, otherwise subjects engaged in solving anagrams. A recognition test followed every trial regardless of whether or not recall had been required. A list of 40 words, half of which were targets and half distractors, was presented and subjects required to discriminate between these.

Subjects

Twenty male and female undergraduates of the City University took part in the experiment. They were all tested individually.

Materials

480 words were drawn at random from the Toronto word pool and divided into two sets of 240 words. Each set was presented to 10 subjects, five of whom received 120 words as targets and the remaining 120 as distractors, whilst another 5 subjects received these as distractors and targets respectively.

Each 120 word set was organised into 6 x 20 word lists. The order of list presentation and of words within lists was randomised for every subject.

For the recognition test the 20 target words were removed from their list order and typed, together with 20 distractors, in a random order on a sheet of paper. For the trials where no recall was required 200 x 3 letter anagrams were printed on a sheet of paper.

Procedure

A list of words was read out by E at a rate of one word every two seconds. After each list had been read the subject was told whether to free recall or to solve anagrams. E recorded Ss recall, whilst for the anagram solving trials S wrote the solution beside each anagram. One minute was allowed for each of these activities.

Finally the sheet with the 20 targets and 20 distractors printed on it was given out and the subject instructed to tick the words which E had read out, and to cross those which had not been read out. This procedure was repeated six times.

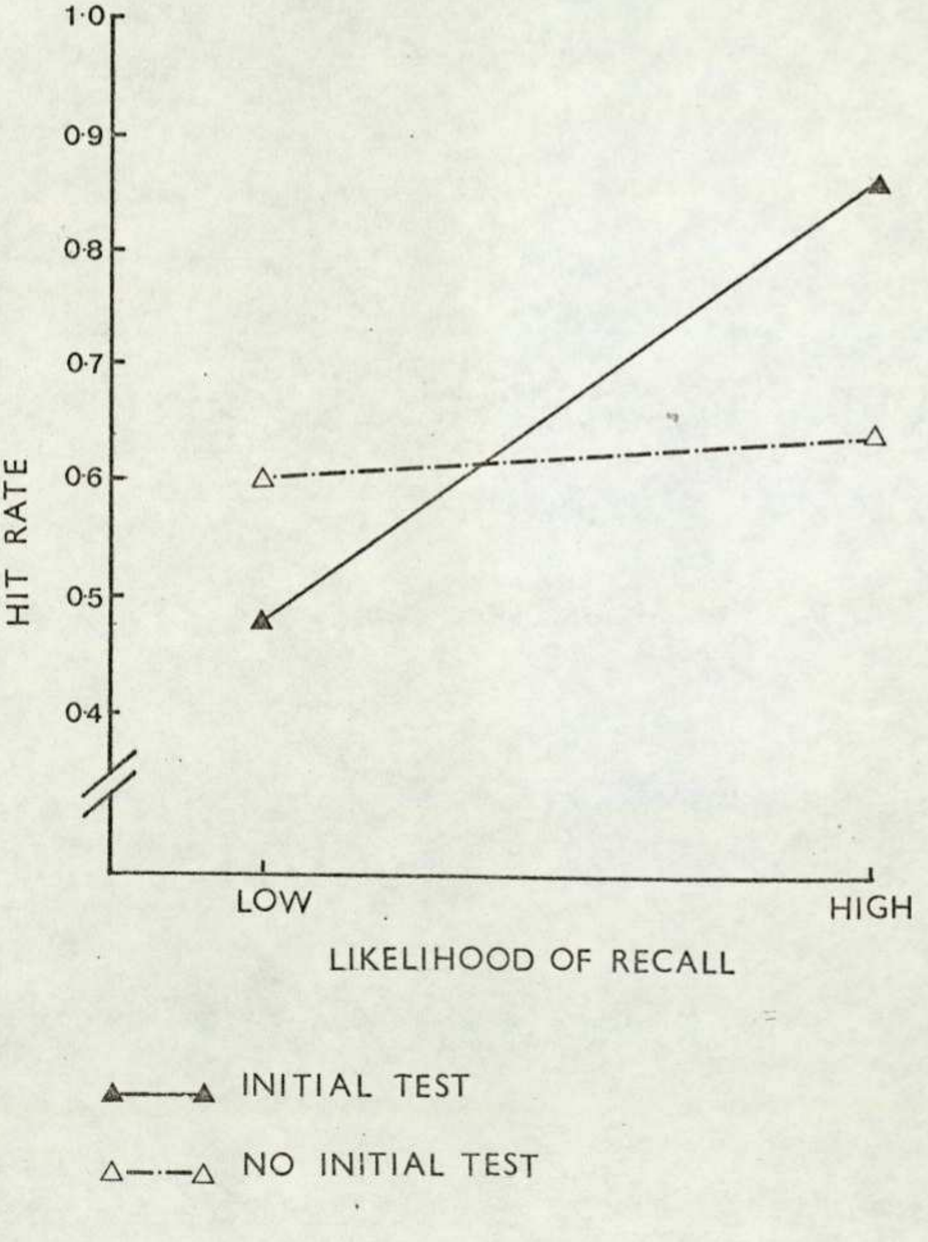
4.12 Results

A signal detection theory analysis was used to analyse recognition performance. A d' score was derived for each subject for both recall and no recall trials. Hit rate was defined as the probability of correctly identifying a target item as such, and false alarm rate as the probability of incorrectly identifying a distractor as a target. Where recall had intervened between presentation and test the mean d' value obtained was 1.57, and where no recall intervened, 1.55. A t test for dependent means confirmed that these means did not differ significantly ($t(19) < 1$).

A more detailed analysis of recognition performance was achieved by considering the target population as two groups of words, those 6 (since an average of 6 words was recalled on each trial) presented in the serial positions from which recall was most frequent, (Positions 1, and 16-20 inclusive), and the 14 words presented in the remaining serial positions, from which recall least frequently occurred. Recognition performance was analysed as a function of this division as well as of whether or not a recall trial had intervened between presentation and test. The mean hit rate achieved for these two groups of targets following recall and anagram solving are shown in Figure 4.5. A 2-way ANOVA with repeated measures

FIGURE 4.5. EXPERIMENT 7.

MEAN HIT RATE AS A FUNCTION OF WHETHER OR NOT RECALL WAS TESTED AND LIKELIHOOD OF RECALL.



on both factors was performed on these data. This revealed main effects of both interpolated recall, ($F(1,19) = 4.71$; $p < 0.05$) and high or low probability of recall, ($F(1,19) = 45.97$, $p < 0.01$) as well as a significant interaction effect, ($F(1,19) = 87.23$, $p < 0.01$). Post hoc comparisons using a Newman Keuls procedure showed that only hit rate for recalled items differed significantly from that for non recalled items, ($q_4(19) = 5{\cdot}38$, $p < 0{\cdot}05$).

4.13 Discussion

Whilst overall recognition scores were unaffected by whether or not a recall trial intervened between presentation and test, recognition performance did nonetheless seem to be affected by this manipulation. Recognition performance was uniform across the target population where no recall intervened, but differed significantly between previously recalled and non recalled items where recall had been required. Though care in interpreting these results must be taken due to possible item selection effects it does seem as though recognition of non recalled items is worse following recall than recognition of comparable items where no recall intervened. The conclusions of Darley and Murdock (1971) that an initial recall test has no influence upon subsequent recognition would then not be completely justified. An initial recall test appears to facilitate the recognition of recalled items and to depress the recognition of other non recalled items.

The effects of recall upon recognition seem to arise from the contrasting of better and less well known items in the recognition test, (Packham, 1968), but it is possible that the effects arise from within the recall trial. There seems to be no reason why the influence recall has upon the subsequent retrieval of non recalled items remaining in the store described by Rundus (1973) (Section 2.6 of this thesis), should not extend to their recognition. Such a possibility becomes more plausible if Tulving's (1976) episodic-ecphory viewpoint (Section 1.4b) is accepted. This suggests that recall and recognition involve basically the same process and differ only in the nature of the retrieval cues present. Recognition also has a retrieval problem and access to the underlying representation of an item is not necessarily guaranteed by the presence of that item in a recognition test. It is interesting to note that if Tulving's theoretical approach is accepted then Darley and Murdock's conclusion that recall affects item accessibility but not item availability would also be untenable. They assume that recall is a function of an item's availability and accessibility (Tulving and Pearlstone, 1966) in memory whereas a recognition test allows automatic access to the underlying trace of an item and performance thus depends only on an item's availability.

4.14 General conclusions

No support was found for the contention that an act of recall results in the formation of a new trace in Episodic Memory, in Experiment 5. While reaction times were fastest for items which had been recalled there was no evidence of any linear relationship between these and measures of lag. Neither was much support for Murdock and Anderson's model of memory found, and this was investigated further in Experiment 6.

The nature of the search which takes place when episodic memory is questioned was considered. The speed of the search as well as the accuracy, to some extent, was affected by the presence of semantically associated pairs of distractors in the recognition test sequence. The search was faster, rather than slower as expected, and this in response to targets as well as distractors. However, since the distractors were semantically associated in a direction which was the reverse of the test sequence, it was suggested that they were unlikely to have had any influence upon the search unless this proceeded backwards, from the point of test and utilised semantic cues. These results did not then appear entirely contradictory to Murdock and Anderson's model. The relationship of reaction times to lag measures did not emerge very clearly and this was attributed to a possible confounding influence of serial position effects.

The influence of initial recall upon subsequent recognition was considered in Experiment 7. Whilst, as Darley and Murdock (1971), found an initial recall test had no effect upon overall recognition scores, there was some evidence that initial recall depressed the recognition of non recalled items and enhanced the recognition of recalled items.

CHAPTER 5

Concluding Discussion

5.1 Introduction

The principal objective of the thesis was to explore how an act of retrieval from memory is represented in the system and monitored. Given Tulving's (1972) distinction between episodic and semantic memory, it was argued that it is useful to assume that each recall event constitutes a new event in episodic memory. The thesis also adopted the generally accepted pre-theoretical position that where the events concern memory for word lists, it is useful to assume that the underlying memory traces comprise an ensemble of the characteristics, attributes or features of the words, (Wickens, 1970; Underwood, 1969; Bower, 1967; Rumelhart, Lindsay and Norman, 1972; Morton, 1970). Within this general framework three specific lines of research have been described in the experimental chapters of the thesis. These were (i) whether the subject's knowledge of his previous recall performance is mediated by features or attributes that only become available at the time of recall, (Chapter Two); (ii) whether there are any age-related differences in the accuracy with which subjects can distinguish previously recalled from previously non recalled items, (Chapter Three); and (iii) whether it might be appropriate to view recall as resulting in the formation of a unique trace in episodic memory, which is functionally separate to other representations of that item (Chapter Four).

This chapter summarises and discusses the principal findings reported in the previous chapters, in the light of the objectives of the thesis. The possible direction of theoretical developments arising from the present findings are briefly mentioned. Finally, the main conclusions drawn from the thesis are presented.

5.2 The Principal Findings of the Thesis

a) The representation and the monitoring of an act of recall

Control processes are transient phenomena brought to bear on the permanent structures of memory at the will of the subject, and include such processes as rehearsal, and various means of coding and retrieving stimuli. That control processes form a 'pervasive and integral component of human memory' (Atkinson and Shiffrin, 1968, p. 191) has only just come to be recognised for one class of control processes, that of memory monitoring, (Tulving and Madigan, 1970). Whilst the role of monitoring the contents of memory, (Hart, 1967; Blake, 1973; Brown and McNeill, 1966) seems to be unquestionably that of informing the subject of whether continued retrieval efforts might be worthwhile, (Hart, 1967), the role of assessing past memory performances is less well defined. The immediate relevance of monitoring output during recall is in the avoidance of repetition errors, (Gardiner and Klee, 1976), but monitoring past recall ought, intuitively, to have more important implications than this. It might, for instance, be involved in the modulation and regulation of other control processes, (Kelly, Scholnick, Travers and Johnson, 1976; Klee and Gardiner, 1976). Presumably assessing and comparing recall levels under particular task conditions enables different mnemonic behaviours to be evaluated and allows for their appropriate selection and deployment in future.

That past recall performances are monitored and remembered has been shown by Gardiner and Klee (1976) and Robinson and Kulp, (1970). Memory for items which have previously been recalled is only accurate however, where these were drawn from secondary memory. In the case of primary memory items a negative recency effect (Craik, 1970) is found. Since primary memory items are generally believed to be less well registered initially in memory (Craik, 1970), it might be assumed that memory for past recall simply reflects the processes occurring at input. However memory for serial recall, in which any primary memory items recalled can be assumed to have been well registered and in addition have been subject to minimal output interference, still shows a negative recency effect, (Klee and Gardiner, 1976). Registration effects have at least some influence in memory for past recall however, since the negative recency effect found here was less pronounced than that found under modified free recall conditions. In a second experiment (Klee and Gardiner, 1976) recognition of previous recognition performance was tested, and this produced a function differing substantially from that describing recognition of recall. Since recognition involves no overt production of the target words, and presumably no monitoring of output it thus seems that memory for past recall is a function of the events occurring at output as well as of those at input.

Previously the effects that an act of recall has upon the underlying representation of an item in memory have been conceptualised either in terms of quantitative increments in the strength value of the trace, (Rundus, 1973) or else compared to the effects of a presentation trial, (Tulving, 1967). In the present thesis the pre-theoretical assumption that a memory trace comprises an attribute bundle which defines the target, (Underwood, 1969; Bower, 1967; Wickens, 1970) was adopted, and it was proposed that performance features derived from feedback of the response during recall become established as part of the underlying attribute ensemble. It was further proposed that these changes in the memory trace would serve in part at least, to mediate memory for past recall. To test this, response feedback was attenuated on half of the recall trials in Experiment 1, where a series of recall trials was followed by a recognition test in which the discrimination of previously recalled and non recalled items was required. It was predicted that memory for past recall would be correspondingly attenuated under conditions where less performance features were available for encoding during recall.

In addition, response mode was varied so that recall for three groups of subjects was either written, spoken or both written and spoken. To the extent that performance features are important in mediating memory for past recall it was argued that this would vary as a function of response mode, and in particular would be best where recall was in two modes, since more performance features would be available for encoding. Both of these predictions were borne out and it was concluded that performance features are encoded during recall and serve in part to mediate memory for past recall. Memory for previous recall performances thus seems to be a joint function of registration effects and the events taking place at recall. Conceivably these two factors operate in such a fashion that the importance of output events in determining memory for remembered events varies as a function of how well an item was initially registered in memory.

Where recall had been both spoken and written, recall levels were lower and memory for past recall higher than where recall was either spoken or written. The Shiffrin-Rundus model of memory (Shiffrin, 1970; Rundus, 1973), appeared to fit these results quite well. The probability of retrieving an item from memory is assumed to be a function of its strength of association to a retrieval cue in relation to the sum of the strengths of all the associations existing between that retrieval cue and other items. Sampling from the store is with replacement. Recall is assumed to increment the strength of an item and this decreases the chances of retrieving as yet unrecalled items from the store in two ways, (i) the strength values of non recalled items are relatively weakened, and (ii) since the chances of re-retrieving an already recalled item are increased, successive retrieval attempts are unlikely to yield previously unrecalled items, and the end of the recall trial is hastened. It follows from this that the more strengthening recall is, the sooner a recall trial will be terminated. Recall in two modes was thus apparently more strengthening than recall in a single mode in Experiment 1. Strength was here taken to refer to the amount of information contained in the memory trace, (Wells, 1974), and in this instance to the number of performance features available and encoded during recall.

If recall in two modes does add more information to the underlying trace than recall in a single mode, then in a multi-trial free recall learning situation, where recall is attempted in either one or two modes, a crossover effect should occur. Whilst recall levels will be lower over initial trials, where this is both written and spoken, inter-trial retention should be superior and recall levels eventually surpass those of single mode recall. This prediction was tested in Experiment 2, where recall was spoken for one group of subjects and both written and spoken for a second group of subjects, within a multi-trial-free recall learning paradigm.

Although learning increased at a greater rate over trials where recall was in two modes, recall levels over initial trials were not depressed and no crossover effect occurred. The additional performance features available during both written and spoken recall did apparently become established as part of the underlying memory trace and infuenced learning. However, whilst the greater amount of information added to the trace, as a result of both speaking and writing the response, had the expected positive influence of increasing the probability of recalling that item on a subsequent trial, it did not have the expected negative influence on subsequent retrieval efforts during the current recall trial. The negative influence of recall on the retrieval of items remaining in the store could be expected to fall to a greater extent on items which had not been recalled on the previous trial, where recall was in two modes. Recall in two modes means that that item has a better chance of retrieval on the subsequent trial, than is the case for items recalled in one mode. However, whilst the two modes of recall group were better at re-recalling items which had been recalled on the previous trial, there was no difference between the two groups on any trial in the number of items recalled which had not been recalled on the previous trial. The small difference in recall levels between the two groups on the first trial was quickly overcome by the two modes of recall group. It was concluded that the sampling-with-replacement model (Rundus, 1973) might represent an oversimplification of the processes occurring during recall in a multi-trial situation. Possibly the tendency to give additional rehearsal or attention during presentation, and priority in output (Rundus, 1974) to items which were not recalled on the previous trial served to offset the greater negative influences on subsequent retrieval efforts resulting from recall in two modes. In addition to this it seems reasonable to assume that the amount of information added to an underlying trace as a result of recall will not be uniform (Bjork, 1975) but will vary as a function of progressive recall trials. Once the attributes of an item become established as part of the memory trace little additional information will be added by further presentation and recall trials.

Clearly recall is an important memorial event which does not serve simply to increment the strength value of a trace that is already qualitatively complete. Nor is recall a simple reading out of the contents of memory, and neither does it constitute a re-presentation of the original stimulus. There is obviously a need for current theoretical models of memory to recognise the importance of recall as an input process, as well as a means of assessing what was stored initially. It then becomes necessary to understand how the monitoring and encoding which occurs at recall interacts with the encoding which took place at presentation.

Given that recall influences the underlying memory trace of an item and that this influence is likely to vary over trials in multi-trial recall, it becomes of interest to explore the variety of outcomes possible where recall is attempted under two different conditions.

A situation (which is plausible from a direct application of Rundus' model to multi-trial recall), where recall adds so much information to the trace that only re-recall can occur now seems an impossibility. The more information recall adds to the trace the more quickly, presumably, the attributes defining that item will become established. Less effort will then be involved in retrieving these items from memory and so the sooner their influence upon non recalled items remaining in the store will be diminished. Conceivably the probability of recalling previously non recalled items may even be higher for a group of subjects attempting recall under the more informative conditions.

b) (i) The development of memory appraisal skills

Memory performance improves with age, (Flavell, Friedrichs and Hoyt, 1970; Huttenlocher and Burke, 1976), yet this is not simply a function of maturation at the physiological level, (Hagen, 1971, p. 267). Psychological changes occur so that the child becomes increasingly planful and strategic in his approach to the memory task, (Flavell et al., 1970; Masur, McIntyre, and Flavell, 1973). In general two different hypotheses have been proposed to account for the younger child's failure to use mnemonic strategies, where a strategy can be defined as an intention or plan to treat input in a certain manner specifically for the purpose of memorising. These hypotheses are, (i) the mediation deficiency hypothesis, (e.g. Reese, 1962) which suggests that even if young children were to employ strategies, this would not benefit memory performance, and (ii) the production deficiency hypothesis, (e.g. Flavell, Beach and Chinsky, 1966), which proposes that children are simply failing to produce strategies which they are capable of using to good effect. The experimental evidence currently favours the production deficiency hypothesis since where children are instructed for instance to categorise input, (Moely, Olson, Halwes and Flavell, 1969) or rehearse items (Keeney, Cannizzo and Flavell, 1967), improvements in recall levels result.

Why children should fail to produce strategies that they are capable of using effectively is rather puzzling, and the issue in the study of the development of memory has become one of which factors mediate the transition from non-strategic to strategic behaviour (Hagen, 1971). Two factors which seem important are the degree of mastery of the skill in question, (Moely et al., 1969) and the extent to which the experimental environment suggests the use of a particular strategy, (Ritter, Kaprove, Fitch and Flavell, 1973; Ryan, Hegion and Flavell, 1970).

An essential pre-requisite to the use of cognitive skills which are part of a child's repertoire, appears to be the realisation of self as an active agent in the environment, capable of influencing the memorising process. That memorising and perceiving are initially functionally and conceptually undifferentiated is shown by the lack of differences in recall levels where young children have been instructed either to simply look at stimuli, or to memorise them, (Appel, Cooper, McCarrell, Sims-Knight, Yussen and Flavell, 1972).

A further factor which intuitively should underlie the spontaneous use of strategies is memory appraisal skills. Experiencing the beneficial effects of strategic behaviour on recall levels is alone insufficient to induce their continued use once experimenter pressure to employ them is removed, (Hagen, 1971; Moynahan, 1973). This suggests that young children are unaware of the benefits the use of a given strategy brings. This situation would occur if young children were unable to assess past recall levels. The ability to assess previous recall levels is seemingly essential 'if the subject is to use and evaluate different memory strategies; in order to determine if a memory strategy is effective, one must know how well one has remembered, while using and not using, the strategy' (Moynahan, 1976, p. 94).

While there is little evidence as yet that memory appraisal skills and the use of memory strategies are directly related, such a relationship would undoubtedly be of a complex nature, (Kelly, Scholnick, Travers and Johnson, 1976). An investigation of this is further complicated by the fact that there is little agreement as to whether or not appraisal skills are lacking in the young. Whereas there is some evidence that children's estimates of their recall capacities are unrealistic, (Flavell, Friedrichs and Hoyt, 1970; Levin, Yussen, De Rose, and Pressley, 1977), other evidence has shown children to be reasonably accurate in predicting their recall spans, (Kelly et al., 1976), as well as in assessing past recall levels, (Moynahan, 1976; Kelly et al., 1976).

There are some objections which can be raised against the procedures used to investigate children's ability to appraise past recall levels. Moynahan (1976) considered the accuracy with which children could estimate how many items they had recalled previously, whilst Kelly et al. (1976) asked subjects to judge whether their placement of a duplicate set of stimulus items in an array was correct or not. In both of these experiments these judgements were required following each of several recall trials, a procedure which would be likely to prime subjects to attend primarily to the level and accuracy of their recall. Experiments 3 and 4 in the present thesis assessed whether children can actually discriminate previously recalled from non recalled items by giving a final unexpected recognition test. Whilst Masur, McIntyre and Flavell (1973) did consider children's ability to discriminate recalled and non recalled items this was considered only after the experiment as a possible reason underlying the children's failure to employ a certain strategy. Their ability to assess past recall was investigated using 8 of the youngest subjects following only a single recall trial.

In Experiment 3, children ranging in age from 5 to 15 years were given a series of free recall trials, followed by a final recognition test in which they were required to discriminate previously recalled and previously non recalled items. The accuracy with which this task was performed increased steadily with age up to about ten years, children of 5 to 7 years being relatively poor at identifying previously recalled and non recalled items.

It was felt however that Experiment 3 had placed the older children at an unfair advantage in that the task and the choice of stimuli were designed to suit the vocabulary and estimated memory span of the youngest subjects. Experiment 4 was thus designed to replicate Experiment 3 except that the task difficulty was manipulated to be of an appropriate level for each age group involved. List length was varied according to the age of the subjects, and similarly the stimuli varied in the normative age, (Carroll and White, 1973a),at which they become established as part of a child's vocabulary. In addition, the stimuli were verbal rather than pictorial, as they were in Experiment 3, since there is evidence to suggest that young children may find pictorial stimuli more difficult to cope with than verbal stimuli (Rohwer, 1970).

However, despite these modifications, children of approximately 5 to 7 years were again found to be poor in discriminating previously recalled from non recalled items. Item registration effects and the events occurring at recall have been shown, (Gardiner and Klee, 1976; Klee and Gardiner, 1976; Experiment 1 of the present thesis) to govern memory for past recall. The poorer performance of the youngest children might then be assumed to be due to age differences in the monitoring of output and to differences in the way in which items are registered initially. Whilst the number of encoding dimensions used and the preference for these does vary with age (Bach and Underwood, 1970; Klastarin Naron, 1978; Means and Rohwer, 1976) it seems unlikely that registration differences could account for the young child's poorer memory for past recall. If this was the case differences in the pattern of results in relation to serial position, rather than simple main effects of age, should have been found, (Brown, 1975).

It initially appeared that the young child's poor memory for past recall could not be attributed to a failure to monitor output at the time of recall, since repetition errors in spoken recall (Experiment 4) were unrelated to age. However, it is possible that young children do not make as many retrieval attempts as older children and thus have less need to monitor output for the purpose of avoiding repetition errors in recall. It may then still be that age differences in appraising past recall are attributable to differences in the extent and nature of output monitoring at recall.

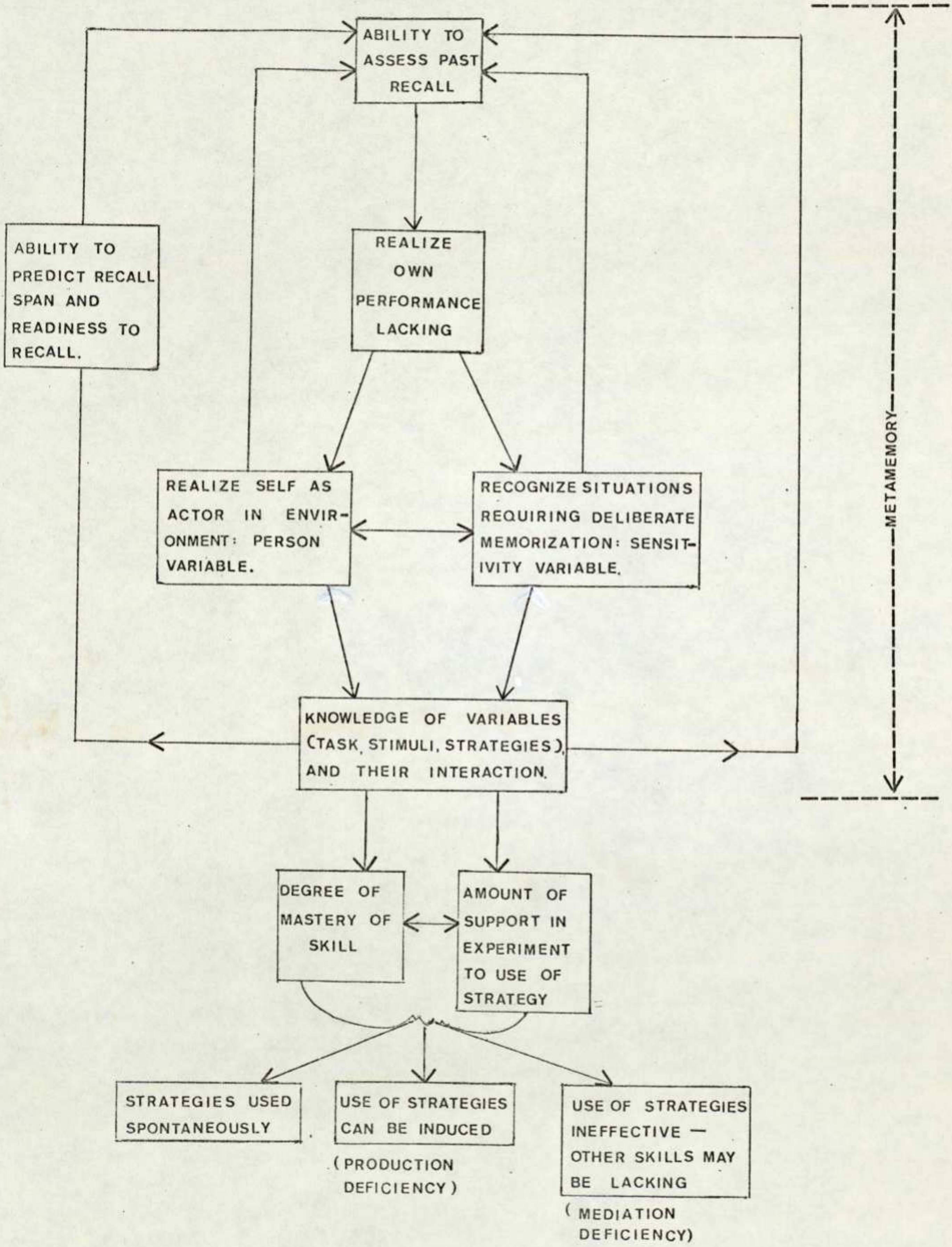
There is growing conviction that poor memory performance in the young may be as much attributable to a lack of metamemorial skills as to the absence of strategic behaviour, and that in fact an inability to accurately appraise memory might underlie the failure to make spontaneous use of strategies, (Wellman, 1978; Kelly et al., 1976; Brown, 1975). It then becomes essential to (i) adequately document the whole range of appraisal skills and to better understand the apparent inconsistencies in the literature, and (ii) to specify the role appraisal skills play in the transition from non-strategic to strategic behaviour, and in the development of other metamemorial skills. A framework which should permit the formulation of useful questions regarding the development of appraisal skills and their role in the development of other mnemonic skills is outlined in the following section. This conceptualisation summarises many of the points made in earlier sections.

(ii) The role of metamemorial skills in the development of strategic behaviour

The following formulation (Figure 5.1) is intended to clarify the role of metamemorial skills in the development of competence in memory tasks. The probable chronological sequence of achievement of various

FIGURE 5.1.

THE ROLE OF APPRAISALS OF PAST RECALL IN THE DEVELOPMENT OF COMPETENCE IN MEMORY TASKS.



skills and concepts, which are generally recognised as necessary to the transition from non-strategic to strategic behaviour, is loosely traced. It is suggested that the ability to appraise past recall may be of fundamental importance in the development of other metamemorial skills, and possible relationships between these are outlined. Finally it is suggested that a mediation-production distinction, such as has been applied to the use of mnemonic strategies, might be of value when considering the development of appraisal skills. It is hoped that this framework will be of some value in formulating useful research questions, and in resolving some of the inconsistencies in the literature.

Metamemorial knowledge can be classified into two categories according to Wellman (1978) (i) sensitivity and (ii) knowledge of variables. Sensitivity refers to the awareness of whether a certain memory task necessitates a deliberate effort or whether remembering will be relatively automatic. Brown (1975) suggests that remembering will be automatic where the task is meaningful and of a reconstructive as opposed to reproductive nature. Knowledge of variables includes knowledge of stimuli, tasks, strategies and person, and in addition knowledge of how these variables interact. Knowledge of person is achieved when the child comes to recognise himself as an actor in the environment, capable of influencing and even controlling the memorising process, (Appel et al., 1972). The task variable refers to learning that a certain task will be easy or difficult. Serial recall would for instance prove more difficult than free recall. Knowledge of stimuli involves learning to recognise which stimulus characteristics will make memorising easy, and similarly knowledge of strategies involves learning the various ways stimuli can be treated in order to make the task easier. An understanding of how these variables interact must also be gained so that certain strategies can be applied to the appropriate stimuli under certain task conditions, etc.

Within the present framework it is suggested that the achievement of these different types of knowledge probably depends on, or leads to, the concurrent achievement of the other types of knowledge. More importantly it is suggested that the attainment of all these, and knowledge of their interaction ought, intuitively, to hinge upon the development of the ability to appraise past recall. The child will not come to understand that certain stimuli, various task constraints and particular ways of treating the material, render a task more or less difficult unless he can appreciate that his recall levels are lower in some instances and higher in others. Once able to assess past recall the child may then be able to relate his behaviour during presentation to resultant recall levels, and begin to realise that he is able to control the memorisation process. Again it is difficult to see how a child could come to recognise that certain tasks will require a deliberate effort unless it is known that recall levels were poor under similar circumstances previously. Similarly, by assessing and comparing recall levels attained under particular conditions in relation to his own behaviour the child should reach an understanding of which strategies are appropriate to which tasks and which stimuli. Again a child would presumably be unable to predict his recall span unless he knows how well he performed under similar conditions previously. Predictions of readiness to recall would remain inaccurate until the child had learnt to associate the outcome of his recall effort to internal states of 'preparedness'. When accurate predictions of recall span can be made under a variety of task constraints this would mean that the child is capable of understanding the interaction of various task requirements, stimuli and various strategies, (Wellman, 1978). It is also speculated here that the child must realise that his own memory performance is lacking before any deliberate attempt to improve it is made, (Kelly et al., 1976).

Two other factors which have been shown to govern the spontaneous use of strategies are the degree of mastery over the requisite mnemonic skill, (Moely, et al., 1969) and the degree to which the use of a relevant strategy is suggested by the experimental environment, (Ritter et al., 1973). Conceivably the more skilled the child is in using a certain strategy, the less important the support to its use given in the experiment becomes. Whether the above concepts and skills are attained will determine whether the subject will use strategies spontaneously in a memory task. Three outcomes are possible, (i) strategies will be used spontaneously, (ii) strategies will not be used spontaneously, but if specifically instructed to do so, will be used effectively, (the production deficiency hypothesis), and (iii) the basic skills are possessed but using the strategy has no effect upon the child's memory performance, (the mediation deficiency hypothesis, Reese, 1962).

A mediation-production deficiency distinction might prove of value when considering the development of appraisal skills. Perhaps the necessary skills are within the child's capabilities but he simply doesnot realise the need for their use, or alternatively the child does make appraisals of memory, but other skills are lacking so that these are either inaccurate or else of no benefit to memory performance.

c) Episodic Memory and the Act of Recall

Tulving (1972) considered that there are two kinds of memory, episodic and semantic memory. Whilst semantic memory is rather like an internal lexicon which stores the meaning and referents of words and concepts, episodic meory stores personally experienced events or episodes in terms of their spatio-temporal relations to one another. The two stores operate independently though semantic information is frequently used in the encoding and retrieval of information from episodic memory. The episodic-semantic distinction is basic to episodic theory, (e.g. Tulving, 1976; Watkins, Ho and Tulving, 1976; Tulving and Watkins, 1977), the essence of which is that any input cannot be understood aside from its context, where this includes autobiographical and temporal referents, and is thus necessarily unique. Every event is then represented by the formation of a unique trace in episodic memory. Since what is stored in episodic memory is not an item, but an item plus its context, the success of retrieval depends on the extent to which the encoding context is reinstated at the time of retrieval. Semantic associates will aid retrieval only if these were stored as part of the context at the time of encoding.

The approach of episodic theory to the representation of episodic information in memory, contrasts with the so-called 'tagging' theories of memory, (Anderson and Bower, 1972, 1974; Bahrick, 1970; Martin, 1975). These typically assume that episodic information is represented by modifications or markings of semantic structures. The success of retrieval then depends on semantic associations which existed prior to the experiment, rather than on episodic information which is specific to the experiment. Any item is represented by only one trace in semantic memory, or at the most a few traces, each corresponding to the different meanings possible to attribute to the item. What is stored is an item not an episode and if the item occurs in a recognition task access to the underlying trace is automatic. The subject simply has to decide whether or not this is appropriately marked. In contrast recall involves a search as well as a decision stage. Recall necessitates that a search along the semantic pathways recently activated is conducted, prior to the decision as to whether or not the relevant node is appropriately marked. Recall can thus never be superior to recognition.

Much evidence, however, has now accumulated which suggests that episodic information is best conceptualised as being represented in a system which is other than semantic. Weak semantic cues at output have been demonstrated to be more effective for recognition than strong associates, where weak cues were also present at input, (Tulving and Thomson, 1971; [illegible] Gobel, 1977 . 6). In addition recall performance has been shown to exceed recognition performance where recognition was tested in the presence of strong semantic associates that were not part of the encoding context, ([illegible], 1977; Tulving and Thomson, 1973).

A second line of evidence (and one which was pursued in the thesis), favouring episodic theory, comes from the work of Murdock and Anderson (1975). They found evidence to suggest that the retrieval of episodic information from memory involves a search within a system in which memory traces are organised according to a temporal dimension, rather in the fashion of a tape recorder. Reaction times in a recognition task were found to increase steadily as a linear function of the number of items intervening between presentation and test (lag) in the case of old target items, and of test position in the case of distractor items, (which have no lag). A serial self-terminating scan which proceeded backwards over both study and test items, seemed to be taking place within a memory system in which the temporal sequence of events was preserved in the format of the store. In the case of distractor items the search process was presumed to return to the beginning of the study list.

The model and paradigm of Murdock and Anderson were used in Chapter 4 to test a proposal derived from episodic theory. The idea that every event is unique, together with the assumption that any act of retrieval from either episodic or semantic memory constitutes an input into episodic memory (Tulving, 1972), led to the proposal that an act of recall results in the formation of a unique trace in episodic memory. This trace might differ from other representations of the item in terms of performance features derived from feedback of the response during recall, (Experiment 1) for instance. To test this proposal a list of items was presented for free recall in Experiment 5, and followed by a recognition task in which subjects were required to discriminate words they had recalled, words which had been presented but not recalled, and new distractor items. Assuming that recall is represented by a new trace in an episodic system in which the temporal sequence of events is preserved, and from which retrieval takes place by means of a backward, serial self-terminating scan, the following predictions were made: (i) reaction times where responses were correct and sure would be faster for items which had been recalled than those which had not been recalled, and (ii) that in turn reaction times to non recalled items would be faster than those to new items. In addition reaction times should bear a linear relationship to lag in the case of non recalled items, and recalled items, and to test position in the case of new items.

Reaction times were indeed found to be fastest to items which had been recalled. But while reaction times for non recalled and new items showed some evidence of a linear relationship to lag and test position respectively, reaction times in response to recalled items showed no such tendency. Whilst it was assumed that items which had been recalled would be represented by two episodic traces, one formed at presentation, and one at recall, it was thought that only the latter would need to be accessed in order to perform the recognition task. Lag was thus calculated as the number of items intervening between recall and test rather than presentation and test. It is possible however, that both traces influenced response latencies and so obscured the expected linear relationship between response latencies and lag. The evidence found in Experiment 5 was not especially supportive of Murdock and Anderson's model of memory, and further investigations of this were undertaken in Experiment 6.

Intuitively the model of memory which emerged from Murdock and Anderson's studies seemed to present an inefficient means of retrieving information from memory, particularly in the long term. Searching backwards over test as well as study items appeared to be a rather redundant procedure. To test whether the search process did proceed backwards, from the point of test and to examine whether the search was able to utilise semantic cues, as Tulving (1976) suggested, associated pairs of distractors were inserted in a recognition test sequence following the presentation of a list of targets. These distractor pairs were associated in one direction more strongly than in the reverse direction, so that if A is likely to evoke the implicit response B, B is much less likely to evoke A. These pairs were placed in the test sequence in the order B-A so that the association between them should only have become apparent if S searched backwards over the test items, and used semantic cues to aid the search. A control group received the same second members of each distractor pair as the experimental group, but the first member was replaced by another unassociated word. Lag was manipulated so that it varied at two different rates as a function of test position. Reaction times to targets were correspondingly predicted to vary at two different rates as a function of the lag manipulation.

Reaction times to distractors were expected to vary as a linear function of test position, as Murdock and Anderson found. Where distractors were associated pairs of words more false alarms to the second members of the pairs were expected. Presumably when the subject was tested with the second member of as associated pair of distractors, he would search backwards over the test items and find a close semantic match to the probe, which he might be inclined to mistake for a real match. Where this second member was correctly rejected however, reaction times were expected to be slower for the experimental group than the control group, since extra time was assumed to be needed for the decision that the close match was not in fact really a match.

Whilst false alarms were slightly more frequent in the experimental group than the control group,to second members of distractor pairs, this difference was not significant. Contrary to predictions reaction times were faster where associated pairs of distractors were in the test sequence, and this in response to both targets and distractors. This result was suggested to perhaps be due to a general effect the presence of associated pairs of words in the test sequence might have had. Their presence may have predisposed the subject to make a search based to a larger extent on semantic cues. Alternatively a more attractive possibility was that the subjects in the experimental group may have realised the trap posed by the distractor pairs and to ensure maximal recognition accuracy encoded the probe more deeply, (Craik and Lockhart, 1972), seeking a match in terms of the resultant semantic features. If either of these possibilities was justified the unexpected results did not necessarily contradict Murdock and Anderson's model. It was argued that the semantically associated pairs of words could only have influenced the speed and possibly the nature of the search if a backward search over test items had been made.

Whilst reaction times to targets for both groups of subjects did show some variation as a function of the lag manipulation, this was insignificant. This was attributed to a confounding of the lag manipulation with serial position effects which have been shown to influence response latencies, (Clifton and Birenbaum, 1970). Where lag was manipulated so that it increased at a greater rate across test position, recency items were tested first, and primacy items last. Conversely, where lag increased relatively steadily over test position, primacy items were tested first and recency items last.

Reaction times to both targets and distractors did show steady linear increases as a function of test position, in general reaction times to targets being faster than those to distractors. Although these findings were taken as consistent with Murdock and Anderson's model, the full pattern of results obtained in this experiment did not agree completely with the model.

The poor recognition of old but non recalled items in Experiment 5, was suggested to be due to the recall task interpolated between presentation and test, (Kay and Skemp, 1956; Packham, 1968; Brown and Packham, 1967). To test this, a series of lists was presented to a group of subjects in Experiment 7 and half of these followed by recall, and half by an anagram solving task. A final recognition test was presented in which old and new items were to be discriminated. Whilst there was no overall difference in recognition performance between the two conditions, a closer analysis of the data revealed that recognition performance, where preceded by recall, was more accurate for recalled items and not quite so accurate for old but non-recalled items. Where no recall trial had intervened recognition performance was uniform over comparable items. Recalling some items appears to influence the likelihood of other presented, but not recalled items being recognised.

Episodic theory is well able to accommodate the continual updating and transformation of memory traces which recall, as an important memorial event (Chapter 2) would necessitate. In contrast to the highly structured relative permanence of storage in semantic memory, information retained in episodic memory is highly likely to undergo transformation and loss, (Tulving, 1972). Since, according to episodic theory, every event is unique and is necessarily represented by the formation of a new trace in episodic memory, nominally identical events will be represented by a multiple number of traces. In the face of such multiplexing of traces summation, loss and transformation of traces must be essential. It then becomes necessary to understand how the traces stored in episodic memory are organised, if at all, and inter-related. For instance, conceivably, as Atkinson, Hermann and Wescourt (1974) suggest, a structure may evolve in semantic memory as a result of repeated representations in episodic memory. It remains indeterminate as to whether or not a new unique trace is formed in episodic memory as a result of recall. This was felt to be partly due to a lack of understanding of how the multiple traces representing nominally identical items would be related.

5.3 Summary of Conclusions drawn from the Thesis

The findings reported in the present thesis derive from three sets of studies. In the first, the role of response produced feedback during recall, in memory for past recall and in learning was considered. The second set of studies traced the development of the ability to assess past recall, whilst the third set attempted to understand the act of recall in terms of episodic theory, (Tulving, 1972, 1976). The main conclusions to emerge from these studies are summarised below.

1. An act of recall is an important memorial event which does not simply constitute a re-presentation of the stimulus or serve merely to increment the strength value of an already complete memory trace. Performance features become established during recall as part of the attribute ensemble defining an item. These performance features aid in the monitoring and assessment of past recall and influence learning in multi-trial recall. It was concluded that current models of memory need to make provision for the fact that (i) recall has an influence upon the underlying trace of an item, that is unique in some aspects, and (ii) that this influence will not be constant (Bjork, 1975) but will probably vary with serial position and across trials in multi-trial recall.

2. Children were found to be relatively inaccurate in assessing past recall despite the task being graded to suit the different age groups' estimated memory spans and vocabularies. In contrast to this, others (Moynahan, 1976; Kelly et al., 1976) have found children to be quite accurate in appraising their past recall performance. It was suggested that until the whole range of appraisal skills is adequately documented and inconsistencies in the literature understood, any possible relationships amongst appraisal skills and between these and the use of mnemonic strategies and memory performance cannot be properly investigated. A framework intended to clarify the theoretical issues and suggest questions for further research was outlined in Section 5.2 b(ii).

3. No conclusive evidence in favour of Murdock and Anderson's (1975) model of memory was found in Experiment 6. Although the functions describing the relationship between reaction times to new items and test position were linear in nature, and reaction times to targets were faster than those to distractors, the search process was speeded rather than slowed as expected, by the presence of associated pairs of distractors in the test sequence, and this for targets as well as distractors. It was suggested that the associated pairs of distractors could not have had any influence on the speed of the search, unless a backward search utilising semantic cues had taken place, but this remains to be substantiated.

Little evidence was found in Experiment 5 to support the notion derived from episodic theory, that an act of recall results in the formation of a unique trace in episodic memory. It was suggested that this trace might differ from other representations of nominally the same item in terms of performance features, as well as temporal and other attributes. Episodic theory is essentially a multiple trace theory, and one difficulty with investigating the above proposal lay in deciding what the relationship between various traces of nominally identical items might be - whether these would be dependent or independent of each other. If the traces are not independent the problem then becomes one of how the various traces interact and influence one another.

A final experiment demonstrated that although an interpolated recall task has no influence upon overall recognition scores, it does apparently influence the manner in which this score was arrived at. A recall trial appears to enhance the recognition of items which were successfully recalled, and to depress the recognition of other non recalled items.

REFERENCES

Anderson, J.R. and Bower, G.H. (1972) Recognition and retrieval processes in free recall. Psychological Review, 79, 97-123.

Anderson, J.R. and Bower, G.H. (1974) A propositional theory of recognition memory. Memory and Cognition, 2 (3), 406-412.

Appel, L.F., Cooper, R.G., McCarell, N., Sims-Knight, J., Yussen, S.R. and Flavell, J.H. (1972) The development of the distinction between perceiving and memorising. Child Development, 43, 1365-1381.

Atkinson, R.C., Hermann, D.J. and Wescourt, K.T. (1974) Search Processes in Recognition Memory. In: R.L. Solso (ed.) Theories in Cognitive Psychology. Winston, Washington, D.C.

Atkinson, R.C. and Shiffrin, R.M. (1968) Human Memory: A proposed system and its control processes. In: K.W. Spence and J.T. Spence (eds.) The Psychology of Learning and Motivation, Vol. 2. New York: Academic Press.

Bach, M.J. and Underwood, B.J. (1970) Developmental changes in memory attributes. Journal of Educational Psychology, 61, 292-296.

Bahrick, H.P. (1970) A two-phase model for prompted recall. Psychological Review, 77, 215-222.

Battig, W.F., Allen, M. and Jensen, A.R. (1965) Priority of free recall of newly learned items. Journal of Verbal Learning and Verbal Behavior, 4, 175-179.

Bilodeau, E.A. (1967) Experimental interference with primary associates and their subsequent recovery with rest. J.exp. Psychology, 73, 328-332.

Blake, M. (1973) Prediction of recognition when recall fails: Exploring the feeling-of-knowing phenomenon. Journal of Verbal Learning and Verbal Behavior, 12, 311-319.

Bjork, R.A. (1975) Retrieval as a Memory Modifier: An Interpretation of Negative Recency and Related Phenomena. In R.L. Solso (ed.) Information Processing and Cognition. The Loyola Symposium,

Winston, Washington, D.C.

Bousfield, W.A. and Rosner, S.R. (1970) Free versus uninhibited recall. Psychonomic Science, 20 (2), 75-76.

Bower, G. (1967) A Multicomponent Theory of the Memory Trace. In: K.W. Spence and J.T. Spence (eds.) The Psychology of Learning and Motivation. Vol. 1. New York: Academic Press.

Brown, A.L. (1975) The Development of Memory: Knowing, Knowing about Knowing, and Knowing How to Know. In H.W. Deese (ed.) Advances in Child Development and Behavior. Vol. 10. New York: Academic Press.

Brown, R. and McNeill, D. (1966) The 'Tip-of-the Tongue' Phenomenon. Journal of Verbal Learning and Verbal Behavior, 5, 325-337.

Brown, J. and Packham, D.W. (1967) Effect of prior recall on multiple response recognition. Quarterly Journal of Experimental Psychology, 19, 356-361.

Carroll, J.B. and White, M.N. (1973a) Age-of-acquisition norms for 220 pictureable nouns. Journal of Verbal Learning and Verbal Behavior, 12, 563-576.

Carroll, J.B. and White, M.N. (1973b) Word frequency and age of acquisition as determiners of picture-naming latency. Quarterly Journal of Experimental Psychology, 25, 85-95.

Clifton, C. Jr. and Birenbaum, S. (1970) Effects of serial position and delay of probe in a memory scan task. Journal of Experimental Psychology, 86 (1), 69-76.

Corballis, M.C. (1967) Serial order in recognition and recall. Journal of Experimental Psychology, 74, 99-105.

Craik, F.I.M. (1970) The fate of primary memory items in free recall. Journal of Verbal Learning and Verbal Behavior, 9, 143-148.

Craik, F.I.M. and Lockhart, R.S. (1972) Levels of processing: a framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

Darley, C.F. and Murdock, B.B. Jr. (1971) Effects of prior free recall testing on final recall and recognition. Journal of Experimental Psychology, 91 (2) 66-73.

Flavell, J.H. (1970) Developmental studies of mediated memory. In H. Reese and L.P. Lipsitt (eds.) Advances in Child Development and Behavior. New York: Academic Press.

Flavell, J.H. (1977) Cognitive Development. Prentice Hall, Englewood Cliffs: New York.

Flavell, J.H., Beach, D.H. and Chinsky, J.M. (1966) Spontaneous verbal rehearsal in a memory task as a function of age. Child Development, 37, 283-299.

Flavell, J.H., Friedrichs, A.G. and Hoyt, J.D. (1970) Developmental changes in memorization processes. Cognitive Psychology, 1, 324-340.

Gardiner, J.M., Craik, F.I.M. and Bleasdale, F.A. (1973) Retrieval difficulty and subsequent recall. Memory and Cognition, 1 (3),213-216.

Gardiner, J.M. and Klee, H. (1976) Memory for remembered events: An assessment of output monitoring in free recall. Journal of Verbal Learning and Verbal Behavior, 15, 227-233.

Gardiner, J.M., Klee, H., Redman, G. and Ball, M. (1976) The role of stimulus material in determining release from proactive inhibition. Quarterly Journal of Experimental Psychology, 28, 395-402.

Hagen, J.W. (1971) Some thoughts on how children learn to remember. Human Development, 14, 262-271.

Hagen, J.W., Jongeward, R.M. Jr. and Kail, R.V. Jr. (1975) Cognitive perspectives on the development of memory. In H.W. Deese (ed.) Advances in Child Development and Behavior, Vol. 10. New York: Academic Press.

Hart, J.T. (1967) Memory and the memory-monitoring process. Journal of Verbal Learning and Verbal Behavior, 6, 685-691.

Hudson, R.L. and Austin, J.B. (1970) Effect of context and category name

on the recall of categorised word lists. Journal of Experimental Psychology, 86, 43-47.

Huttenlocher, J. and Burke, D. (1976) Why does memory span increase with age? Cognitive Psychology, 8, 1-31.

Hyde, T.S. and Jenkins, J.J. (1973) Recall for words as a function of semantic, graphic and syntactic orienting tasks. Journal of Verbal Learning and Verbal Behavior, 12, 471-480.

Jarvella, R.J. (1971) Syntactic processing of connected speech. Journal of Verbal Learning and Verbal Behavior, 10, 409-416.

Kay, H. and Skemp, R. (1956) Different thresholds for recognition - further experiments on interpolated recall and recognition. Quarterly Journal of Experimental Psychology, 8, 153-162.

Keeney, T.J., Cannizzo, S.R. and Flavell, J.H. (1967) Spontaneous and induced verbal rehearsal in a recall task. Child Development, 38, 953-966.

Kelly, M., Scholnick, E.F., Travers, S.H. and Johnson, J.W. (1976) Relations among memory, memory appraisal and memory strategies. Child Development, 47, 648-659.

Keppel, G. and Strand, B.Z. (1970) Free association responses to the primary responses and other responses selected from the Palermo-Jenkins norms in L. Postman and G. Keppel (eds.) Norms of Word Association. Academic Press: New York.

Klee, H. and Gardiner, J.M. (1976) Memory for remembered events: contrasting recall and recognition. Journal of Verbal Learning and Verbal Behavior, 15, 471-478.

Klastrin Naron, N. (1978) Developmental changes in word attribute utilization for organization and retrieval in free recall. Journal of Experimental Child Psychology, 25, 279-297.

Kreutzer, M.A., Leonard, C. and Flavell, J.H. (1975) An interview study of children's knowledge about memory. Monographs of the Society for Research in Child Development, 40, 1 (159).

Levin, J.R., Yussen, S.R. and De Rose, T.M. and Pressley, M. (1977) Developmental changes in assessing recall and recognition memory capacity. Developmental Psychology, 13 (6) 608-615.

Light, L.L. and Carter-Sobell, L. (1970) Effects of changed semantic context on recognition memory. Journal of Verbal Learning and Verbal Behavior, 9, 1-11.

Loftus, E.F. and Suppes, P. (1972) Structural variables that determine the speed of retrieving words from long-term memory. Journal of Verbal Learning and Verbal Behavior, 11, 770-777.

Madigan, S.A., McCabe, L. and Itatani, E. (1972) Immediate and delayed recall of words and pictures. Canadian Journal of Psychology/Rev. Canad. Psychol., 26 (4) 407-414.

Martin, E. (1975) Generation-recognition theory and the encoding specificity principle. Psychol. Rev., 82, 150-153.

Masur, E.F., McIntyre, C.W. and Flavell, J.H. (1973) Developmental changes in apportionment of study time among items in a multi-trial free recall task. Journal of Experimental Child Psychology, 15, 237-246.

Means, B.M. and Rohwer, W.D. Jr. (1976) Attribute dominance. Developmental Psychology, 12 (5), 411-417.

Moely, B.E., Olson, F.A., Halwes, T.G. and Flavell, J.H. (1969) Production deficiency in young children's clustered recall. Developmental Psychology, 1 (1), 26-34.

Morton, J. (1970) A functional model for memory. In D.A. Norman (ed.) Models of Human Memory. New York: Academic Press.

Moynahan, E.D. (1973) The development of knowledge concerning the effect of categorisation upon free recall. Child Development, 44, 238-246.

Moynahan, E.D. (1976) The development of the ability to assess recall performance. Journal of Experimental Child Psychology, 21, 94-97.

Murdock, B.B. Jr. (1974) Human Memory: Theory and Data. Lawrence Erlbaum Associates, Potomac, Maryland.

Murdock, B.B. Jr. and Anderson, R.E. (1975) Encoding, storage and retrieval of item information. In R.L. Solso (ed.) Information Processing and Cognition. The Loyola Symposium, Lawrence Erlbaum Associates, Potomac, Maryland.

Murdock, B.B. Jr. and Dufty, P.O. (1972) Strength theory and recognition memory. Journal of Experimental Psychology, 94 (3), 284-290.

Murdock, B.B. Jr. and Okada, R. (1970) Interresponse times in single-trial free recall. Journal of Experimental Psychology, 86, 263-267.

Packham, D.W. (1968) Recognition of recalled and non-recalled items. Psychonomic Science, 11 (8), 291-292.

Postman, L. and Keppel, G. (eds.) (1970) Norms of word association. Academic Press: New York.

Reder, L.M., Anderson, J.R. and Bjork, R.A. (1974) A semantic interpretation of encoding specificity. Journal of Experimental Psychology, 102, 648-656.

Reese, H.W. (1962) Verbal mediation as a function of age level. Psychological Bulletin, 59, 6, 502-509.

Ritter, K., Kaprove, B.H., Fitch, J.P. and Flavell, J.H. (1973) The development of retrieval strategies in young children. Cognitive Psychology, 5, 310-321.

Robinson, J.A. and Kulp, R.A. (1970) Knowledge of prior recall. Journal of Verbal Learning and Verbal Behavior, 9, 84-86.

Rohwer, W.D. Jr. (1970) Images and pictures in children's learning: Research results and educational implications. Psychological Bulletin, 73 (6), 393-403.

Rumelhart, D.E., Lindsay, P.H. and Norman, D.A. (1972) A process model for long term memory. In E. Tulving and W. Donaldson (eds.) Organisation of Memory. London: Academic Press.

Rundus, D. (1973) Negative effects of using list items as recall cues. Journal of Verbal Learning and Verbal Behavior, 12, 43-50.

Rundus, D. (1974) Output order and rehearsal in multi-trial free recall. Journal of Verbal Learning and Verbal Behavior, 13, 656-663.

Ryan, S.M., Hegion, A.G.and Flavell, J.H. (1970) Non verbal mnemonic mediation in pre-school children. Child Development, 41 (3) 539-550.

Shiffrin, R.M. (1970) Memory search. In D.A. Norman (ed.) Models of human memory. Academic Press: New York.

Thomson D.M. and Tulving, E. (1970) Associative encoding and retrieval: weak and strong cues. Journal of Experimental Psychology, 86 (2), 255-262.

Tulving, E. (1967) The effects of presentation and recall of material in free-recall learning. Journal of Verbal Learning and Verbal Behavior, 6, 175-184.

Tulving, E. (1972) Episodic and semantic memory. In E. Tulving and W. Donaldson (eds.) Organisation of Memory. London: Academic Press.

Tulving, E. (1976) Ecphoric processes in recall and recognition. In J. Brown (ed.) Recall and Recognition. London: Wiley.

Tulving, E. and Colotla, V.A. (1970) Free recall of trilingual lists. Cognitive Psychology, 1, 86-98.

Tulving, E. and Madigan, S.A. (1970) Memory and verbal learning. in: P.H. Mussen and M.R. Rosenzweig (eds.) Annual Reviews of Psychology, Vol. 21, Annual Reviews Inc. U.S.A.

Tulving, E. and Pearlstone, A. (1966) Availability versus accessibility of information in memory for words. Journal of Verbal Learning and Verbal Behavior, 5, 381-391.

Tulving, E.and Thomson, D.M. (1971) Retrieval processes in recognition memory: Effects of associative context. Journal of Experimental Psychology, 87 (1) 116-124.

Tulving, E. and Thomson, D.M. (1973) Encoding specificity and retrieval processes in episodic memory. Psychological Review, 80 (5) 352-373.

Tulving, E. and Osler, S. (1968). Effectiveness of retrieval cues in memory for words. Journal of Experimental Psychology, 77, 593-601.

Tulving, E. and Watkins, M.J. (1975) Structure of memory traces. Psychological Review, 82 (4), 261-275.

Tulving, E. and Watkins, O.C. (1977) Recognition failure of words with a single meaning. Memory and Cognition, 5 (5), 513-522.

Tversky, B. and Teifler, E. (1976) Development of strategies for recall and recognition. Developmental Psychology, 12 (5) 406-410.

Underwood, B.J. (1965) False recognition produced by implicit verbal responses. Journal of Experimental Psychology, 70 (1) 122-129.

Underwood, B.J. (1969) Attributes of Memory. Psychological Review,76 (6),559-73.

Watkins, M.J. and Tulving, E. (1975) Episodic memory: when recognition fails. Journal of Experimental Psychology, 104, 5-29.

Watkins, M.J., Ho, E. and Tulving, E. (1976) Context effects on recognition memory for faces. Journal of Verbal Learning and Verbal Behavior,15, 505-157.

Wellman, H.W. (1977) Tip-of-the-tongue and feeling-of-knowing experiences: A developmental study of memory monitoring. Child Development,48,13-21.

Wellman, H.W. (1978) Knowledge of the interaction of memory variables: a developmental study of metamemory. Developmental Psychology, vol. 14, (1) 24-29.

Wells, J.E. (1974) Strength theory and judgements of recency and frequency. Journal of Verbal Learning and Verbal Behavior, 13, 378-392.

Wickens, D.D. (1970) Encoding categories of words: an empirical approach to meaning. Psychological Review, 77 (1), 1-15.

Wickens, D.D (1972) Characteristics of word encoding. In A.W. Melton and E. Martin (eds.) Coding Processes in Human Memory. London: Wiley.

Winograd, E. and Rivers-Bulkeley, N.T. (1977) Effects of changing context on remembering faces. Journal of Experimental Psychology: Human Learning and Memory, 3 (4), 397-405.

Yarmey, A.D. (1973) I recognise your face but I can't remember your name: further evidence on TOT phenomenon. Memory and Cognition, 1, 287-290.

Table A1 Experiment 1

Chi squared analysis performed on repetition errors

(observed frequencies are in top left of each cell, and expected frequencies in bottom right)

Response Mode Feedback Condition	Spoken	Written	Spoken & Written	Totals
Normal Feedback	52 28.12	5 21.59	1 8.29	58
Impaired Feedback	60 83.88	81 **64.40**	32 **24.71**	173
Totals	112	86	33	**231**

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(\sigma_{ij} - E_{ij})^2}{E_{ij}}$$

$$\chi^2 \quad (2) = 52.666 \qquad p < 0.001$$

Table A2 Experiment 1

The analysis of variance performed on the total number of words recalled as a function of response mode, feedback condition, and serial position

Source of variance	df	SS	MS	F	P
Between subjects	89	4568.93			
Response mode (R)	2	878.40	439.20	10.35	0.001
Subj.w.grps	87	3690.53	42.42		
Within subjects	450	18636.00			
Feedback condition (F)	1	1.25	1.25	0.23	NS
RF	2	3.21	1.61	0.29	NS
FxSubj.w.grp.	87	474.87	5.45		
Serial position (S)	2	13808.1	6904.07	436.49	0.001
SR	4	135.62	33.91	2.14	NS
SxSubj.w.grp	174	2752.23	15.82		
FS	2	14.96	7.48	0.91	NS
FSR	4	17.54	4.39	0.53	NS
FSxsubj.w.grp	174	1428.17	8.21		
Total	539	23204.93			

Table A3 Experiment 1

The analysis of variance performed on the d' scores, where these are calculated as a function of response mode, serial position and feedback condition

Source of variance	df	SS	MS	F	P
Between subjects	89	138.65			
Response mode (R)	2	17.37	8.69	6.23	0.01
Subj.w.grps	87	121.28	1.39		
Within subjects	450	255.79			
Feedback condition (F)	1	3.82	3.82	10.13	0.01
RF	2	0.11	0.05	0.14	NS
FxSubj.w.grp.	87	32.82	0.38		
Serial position (S)	2	71.03	35.52	79.26	0.001
SR	4	1.86	0.47	1.04	NS
SxSubj.w.grp.	174	77.97	0.45		
FS	2	0.49	0.24	0.63	NS
FSR	4	0.37	0.09	0.24	NS
FSxS.w.grp.	174	67.31	0.39		
Total	539	394.44			

Table A4 Experiment 1

Analysis of variance performed on d' scores for the primacy portion of the serial position curve only

Source of variance	df	SS	MS	F	P
Response mode (M)	2	3.49	1.75	1.77	NS
Subj.w.grp.	87	85.58	0.98		
Feedback condition (F)	1	2.60	2.60	7.58	0.01
FR	2	0.18	0.09	0.26	NS
FxS.w.grp	87	29.87	0.34		
Total	179	121.71			

Table A5 Experiment 1

Analysis of variance performed on d' scores for the asymptotic portion of the serial position curve only

Source of variance	df	SS	MS	F	P
Response mode (M)	2	3.49	1.75	2.32	NS
Subj.w.grps.	87	65.61	0.75		
Feedback condition (F)	1	1.32	1.32	3.06	NS
FR	2	0.18	0.09	0.21	NS
FxS.w.grps	87	37.63	0.43		
Total	179	108.24			

Table A6 Experiment 1

Analysis of variance performed on d' scores for the recency portion of the serial position curve only

Source of variance	df	SS	MS	F	P
Response mode (R)	2	12.24	6.12	11.08	0.001
Subj.w.grps.	87	48.06	0.55		
Feedback Condition (F)	1	0.39	0.39	1.04	NS
FR	2	0.12	0.06	0.17	NS
FxS.w.grp.	87	32.62	0.37		
Total	179	93.44			

Table A7 Experiment 1

Analysis of variance performed on false positive errors

Source of variance	df	SS	MS	F	P
Response mode (R)	2	0.72	0.36	7.17	0.01
S.w.grps.	87	4.37	0.05		
Feedback condition (F)	1	0.00	0.00	0.40	NS
RF	2	0.01	0.01	0.76	NS
FxS.w.grps.	87	0.85	0.01		
Serial position (S)	2	0.04	0.02	2.00	NS
SR	4	0.01	0.00	0.16	NS
SxSubj.w.grps	174	1.80	0.01		
FS	2	0.12	0.01	0.86	NS
FSR	4	0.02	0.00	0.44	NS
FSxS.w.groups	174	1.69	0.01		
Total	539	9.52			

Table A8 Experiment 2

Analysis of variance performed on number of words recalled as a function of trials and of response mode

Source of variance	df	SS	MS	F	P
Response mode (R)	1	176.33	176.33	1.95	NS
Subj.w.grp.	22	1990.92	90.50		
Trials (T)	7	9088.41	1298.34	182.61	0.001
RT	7	126.42	18.06	2.54	0.05
S.w.grp. x TR	154	1094.92	7.11		
Total	191	12477.0			

Table A9 Experiment 3

Analysis of variance performed on recall data

Source of variance	df	SS	MS	F	P
Age (A)	2	958.91	479.45	24.17	0.01
Subj.w.grps.	48	952.22	19.84		
Serial position (S)	8	314.57	39.32	14.66	0.01
SA	16	37.96	2.37	0.89	NS
S.Subj.w.grps.	384	1030.37	2.68		
Total	458	3294.01			

Table A10 Experiment 3

Analysis of variance performed on d' scores

Source of variance	df	SS	MS	F	P
Age	2	4.52	2.26	4.35	0.05
Within	48	24.97	0.52		
Total	50	29.49			

Table A11 Experiment 3

Analysis of false positive error rates

Source of variance	df	SS	MS	F	P
Age (A)	2	0.06	0.03	0.21	NS
S.w.grp.	48	6.46	0.13		
Serial position (S)	2	0.01	0.01	0.48	NS
AS	4	0.05	0.01	1.16	NS
SxS.w.grp.	96	1.01	0.01		
Total	152	7.58			

Table A12 Experiment 4

Analysis of variance performed on recall data

Source of variance	df	SS	MS	F	P
Age(A)	2	0.11	0.06	1.06	NS
S.w.grp.	39	2.05	0.05		
Serial position (S)	2	1.90	0.95	14.24	0.01
AS	4	0.27	0.07	1.02	NS
SxS.w.grp.	78	5.22	0.07		
Total	125	9.55			

Table A13 Experiment 4

Analysis of variance performed on the d' scores

Source of variance	df	SS	MS	F	P
Age (A)	2	9.64	4.82	3.40	0.05
S.w.grp.	39	55.27	1.42		
Serial position (S)	2	3.27	1.64	6.68	0.01
AS	4	0.57	0.14	0.58	NS
SxS.w.grp.	78	19.09	0.24		
Total	125	87.84			

Table A14 Experiment 4

Analysis of variance performed on false positive errors

Source of variance	df	SS	MS	F	P
Age (A)	2	0.26	0.13	1.17	NS
S.w.grps	39	4.33	0.11		
Serial position (S)	2	0.01	0.01	1.21	NS
AS	4	0.01	0.00	0.63	NS
SxS.w.grp.	78	0.44	0.01		
Total	125	5.05			

Table A15 Experiment 5

Analysis performed on reaction time data for vincentized thirds of lag and test position

Source of variance	df	SS	MS	F	P
Response type (R)	2	899846	449923	5.59	0.01
Vincentized thirds of lag or test position (V)	2	166991	83496	1.93	NS
S.w.grp.	14	12521995	894428.2		
VR	4	36839.2	9209.8	0.27	NS .
RxS.w.grp.	28	2254151	80505.4		
VxS.w.grp.	28	1214341	43369		
RxVxSw.grp.	56	1881560.7	33599.3		
Total	134	18975724.0			

Table A16 Experiment 6

Analysis of variance performed on reaction time data

Source of variance	df	SS	MS	F	P
Between subjects	31	46505655	1500182		
Associativeness of distractors (A)	1	6296148	6296148	4.70	0.05
S.w.grp.	30	40209507	1340317		
SS.w.subj.	480	13536647.1	28201.35		
Test position (P)	3	1431768	477256	24.08	0.01
Lag manipulation (L)	1	95293	95293	2.27	NS
Targets/ distractors (T)	1	2265123	2265123	24.91	0.01
AP	3	126937	42312	2.13	NS
AL	1	9670	9670	0.23	NS
AT	1	14.1	14.1	0.00	NS
PL	3	111709	37237	2.26	NS
PT	3	135639	45213	4.19	0.01
LT	1	1061	1061	0.11	NS
APL	3	41550.4	13850.1	0.84	NS
APT	3	14852.8	4950.9	0.46	NS
ALT	1	3703.4	3703.4	0.38	NS
PLT	3	14541.7	4847.2	0.57	NS
APLT	3	1124.3	374.8	0.04	NS
PxS.w.grp.	90	1783772	19819		
LxS.w.grp.	30	1259592	41986		
TxS.w.grp.	30	2727992	90933		
PxLxS.w.grp.	90	1483488	16483		
PxTxS.w.grp.	90	970747	10786		
LxTxS.w.grp.	30	289333	9644		
PxLxTxS.w.grp.	90	768736.4	8541.52		
Total	511	60042301.6			

Table A17 Experiment 6

Analysis of variance performed on d' scores

Source of variance	df	SS	MS	F	P
Associativeness of distractors (A)	1	0.30	0.30	0.24	NS
S.w.grp.	30	38.35	1.28		
Test position (P)	3	14.39	4.80	17.83	0.01
AP	3	2.84	0.95	3.52	0.05
PxS.w.grp.	90	24.22	0.27		
lag manipulation (L)	1	0.44	0.44	0.92	NS
LA	1	0.02	0.02	0.04	NS
LxS.w.grp.	30	14.40	0.48		
PL	3	3.41	1.14	3.60	0.05
APL	3	4.13	1.38	4.36	0.01
PLxs.w.grp.	90	28.41	0.32		
Total	255	130.91			

Table A18 Experiment 6

Analysis of variance on false alarm rate in response to second members of distractor pairs only

Source of variance	df	SS	MS	F	P
Associativeness of distractors (A)	1	0.06	0.06	1.42	NS
S.w.grp.	30	1.24	0.04		
Test position (P)	3	0.02	0.01	0.99	NS
AP	3	0.03	0.01	1.22	NS
PxS.w.grp.	90	0.63	0.01		
Total	127	1.98			

Table A19 Experiment 7

t test for dependent means performed on d' scores for tested and nontested trials

Tested	Nontested	d_i	d_i^2
1.28	1.15	0.13	0.0169
2.21	1.62	0.59	0.3481
1.66	1.43	0.23	0.0529
1.59	1.69	-0.10	0.01
1.69	1.83	-0.14	0.0196
0.94	1.15	-0.21	0.0441
1.82	1.80	0.02	0.0004
1.53	1.77	-0.24	0.0576
0.49	0.82	-0.33	0.1089
1.26	1.32	-0.06	0.0036
2.40	2.28	0.12	0.0144
1.39	1.75	-0.36	0.1296
1.72	1.31	0.41	0.1681
1.68	1.81	-0.13	0.0169
0.89	1.02	-0.13	0.0169
2.21	1.62	0.59	0.3481
1.98	1.56	0.42	0.1764
1.78	2.27	-0.49	0.2401
0.75	0.70	0.05	0.0025
2.05	2.17	-0.12	0.0144

$$\sum d_i = 0.25 \quad \sum d_i^2 = 1.7895$$

$$t_{obs.} = \frac{0.25}{\frac{20 \times 1.7895 - (0.25)^2}{19}}$$

$$= 0.182$$

df = 19 $p > 0.05$

Table A20 Experiment 7

Analysis of variance performed on hit rate for items with a high or low probability of recall under tested and nontested conditions

Source of variance	df	SS	MS	F	P
High or low probability of recall (P)	1	0.86	0.86	45.97	0.001
Tested/non tested (T) condition	1	0.06	0.06	4.71	0.05
S.w.grp.	19	0.88	0.05		
PT	1	0.60	0.60	87.23	0.001
PxS.w.grp.	19	0.36	0.02		
TxS.w.grp.	19	0.22	0.01		
PxTxs.w.grp.	19	0.13	0.01		
Total	79	3.12			

FIGURE 3.3. EXPERIMENT 3.
AGE PLOTTED AGAINST MEAN
d' VALUES.

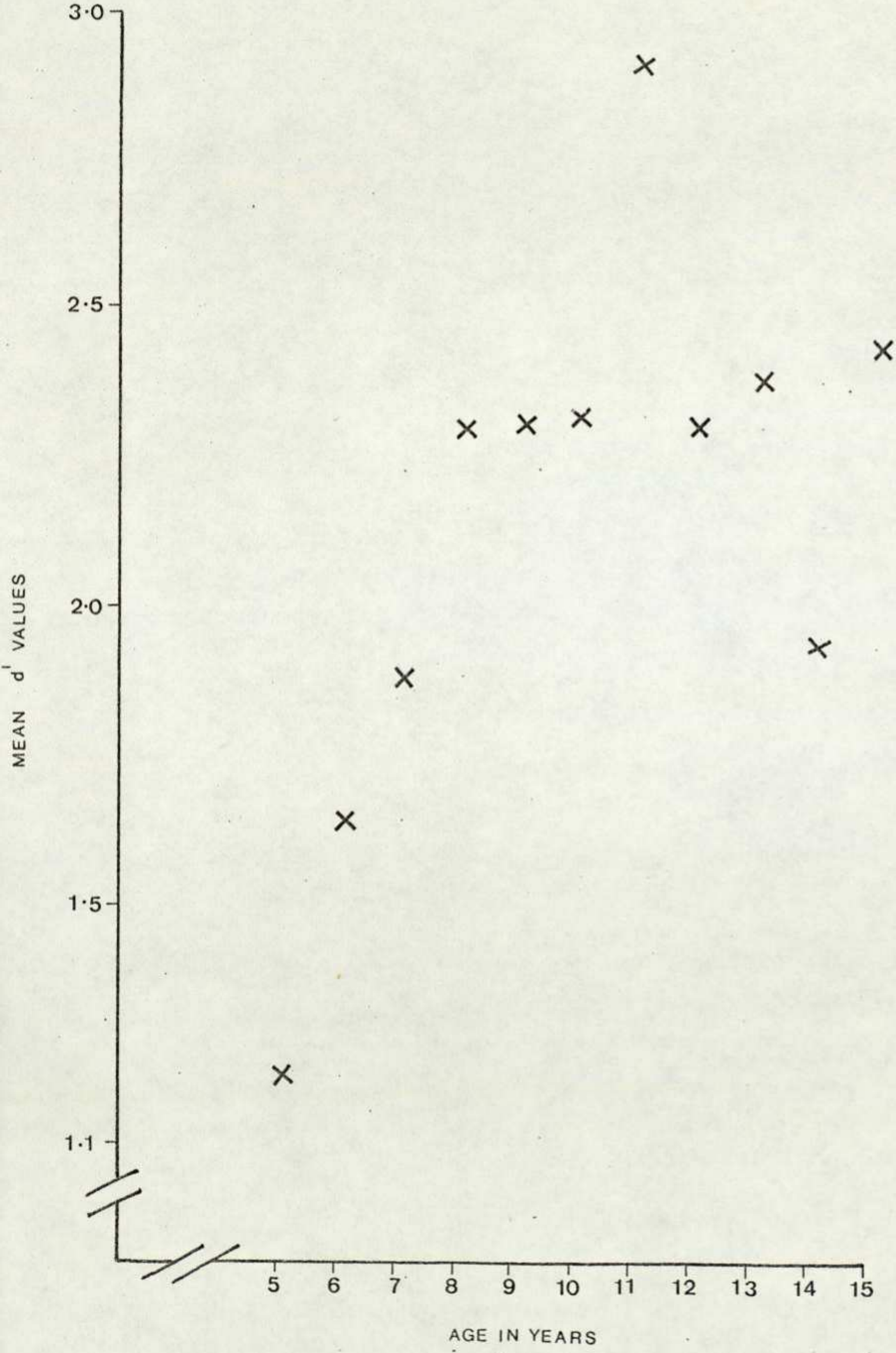


FIGURE 3.6. EXPERIMENT 4.

AGE PLOTTED AGAINST MEAN d' VALUES.