Comprehension and memory for everyday events by the elderly

M.Core, Jocelyn

M.Core, Jocelyn (1986) *Comprehension and memory for everyday events by the elderly*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/7106/

Comprehension and Memory for Everyday Events by the Elderly

Thesis submitted for the degree of
Doctor of Philosophy

Jocelyn M. Core B.Sc.
Trevelyan College

April 1986

Department of Psychology
The University Of Durham

For my family.

ACKNOWLEDGEMENTS

I should like to thank Professor P. M. A. Rabbitt who supervised this research and provided invaluable advice and encouragement.

There have been a number of other people who have also assisted me throughout this research project. Mrs C. Thompson of the Psychology Department, and Dr. W. R. Williams and Mr. J. Steele of the Computer Centre at Durham University have given considerable assistance and guidance with the use of computing facilities. Mr. N. Corr designed and constructed the electrical circuit for Experiments 4.1, 6.1, 6.2 and Mr. M. Rolling produced slides and photographs.

I am also grateful to Ms. V Abson, the entire North East Age Research team, and to the Department of Psychology at the University of Newcastle Upon Tyne for the frequent use of their rooms and equipment. I thank all the volunteers who participated in these studies for their time and effort. The maintenance grant from the Economic and Social Research Council for my research studentship is gratefully acknowledged.

Finally, I thank Dr. Stephen Bryan for his help in the production of this thesis and his support throughout.

ABSTRACT

A large literature has described the effects of advancing age on cognitive laboratory tasks, but there have been few attempts to investigate its effects on everyday cognitive performance. The experiments reported are an attempt to explore the effects of age on the everyday memory task of comprehending and remembering events as conveyed by television and in everyday perception.

The methodology used was cross-sectional with all subjects well-documented on a number of indices. These were assessed as predictors of performance on different cognitive tasks relating to the everyday memory task. Age per se was found to have a limited effect on performance, the best index of the cognitive effect of ageing being I.Q. test score. This index picked up most of the variance on the measurements taken.

Experiments were designed to examine the elderlys' recall of television news broadcasts. These demonstrated that elderly people with low I.Q. test scores have difficulty recalling facts and details from such an information source. Subsequent experiments attempted to identify the processes which explain groups differences on this task.

The employment of strategies as a source of group differences was considered. No evidence was found to show that the extent of strategy focusing was greater in the high I.Q. test score group than the low scoring group. However, differences in identifying thematically essential material and drawing plausible inferences to construct the appropriate theme were found. The low I.Q. test score group were less accurate in their decision as to whether a target slide was a plausible conclusion to a theme shown in previous slides. Further, they recalled fewer plausible inferences to photographs of everyday events.

Finally, the structure of long term memory and its effect on the speed of retrieval and access of information was examined. Search through both the verbal and semantic stores was faster for the high I.Q. test score group and the young. The contribution of all these individual processes to memory and comprehension of events is discussed.

CONTENTS

CHAPTER ONE

1 INTRODUCTION

Gerontology and the study of everyday memory have much in common. Both are areas of psychology which have experienced renewed interest and are expanding as well as undergoing change as a result. As with any science which is relatively new, the methodologies applied are still being questioned, and workable models are still being sought. This chapter explores the current state of these two areas in relation to the research reported in this thesis, which examines the effects of increasing age on performance in an everyday memory task. The task chosen is the comprehension and memory of everyday events. Models of the change in functioning associated with age are developed to explain group differences in the component processes of this task, and hence overall everyday performance.

There have been numerous publications on the effects of age on cognitive functioning, and there are ample reviews of recent work in this area (Birren and Shaie, 1977; Talland, 1968; Welford and Birren, 1965; Baltes, Reese and Lipsitt, 1980). However, one can not fail to notice, throughout this field of research, nagging difficulties with respect to the methodologies employed and the models applied.

The problems encountered range from methodological flaws to a lack of models or their misapplication. In the main, the models used to explain the effects have been borrowed from cognitive psychology, and have usually been based upon earlier studies of the performance of small groups of young, intelligent subjects performing an isolated task. It is therefore difficult to see how one can expect such models to explain individual differences, and changes in performance within an individual, which accompany an increase in age. The result of these misapplications is a failure to identify clearly the processes that change, for better or for worse, with age. As both Welford (1957) and Comfort (1957) have pointed out, the difficulty lies in identifying the mechanisms underlying the data. The present state is one in which the master processes of ageing are known, at least at a descriptive level from a functional and biological stance, but no explanations are offered as to the mechanisms causing the changes in performance levels.

1.0.1 Methodological Issues

Most commonly, gerontological research has compared single samples of individuals who fall into different age ranges. The data generated from these experiments have been used to conclude that any difference found on the dependent variable can be attributed to chronological age. This is based upon the assumption that all other aspects which may relate to the dependent variable under study are equivalent for the groups. This is clearly very seldom the case; subject samples are never homogeneous, which remains a persistent difficulty with such methodology. Further, the models of ageing from which these paradigms are derived are seldom referred to. This is a result of a failure to develop such models to any functional and useful extent.

Three points require consideration. First, what is the shape of the function of the change in cognitive performance with age. Second, which processes are affected by ageing, it may be some, it may be all; and how are these related to the neurobiological changes and lesions which occur with increasing chronological age. The third issue concerns the methodological approach to these questions, should we measure cohort differences or longitudinal change. The design of experiments that use age as a independent variable will inevitably be influenced by the basic assumptions accepted as characterising the nature of age-related changes. There are three models which are commonly referred to: irreversible decrement, stability and decrement with compensation.

Irreversible decrement is frequently applied to intelligence and creativity. It assumes that a maximum level of functioning is reached at a point in adulthood, and thereafter the age function proceeds along a linear, and irreversibly decrementing path. It implies that age changes will occur as a function of maturational events, regardless of environmental input, and is probably most appropriate for variables whose performance is dominated by psychomotor speed.

The second model, often used in personality studies, proposes that once adulthood is reached, adult behaviour remains stable. Such an assumption necessarily requires that the researcher is interested only in pathology-free populations. However, it does fit measures of crystallised intelligence and any biologically-founded behaviours which remain stable until the terminal drop (Riegel and Riegel, 1972) just before death. Once this model has been accepted, the age effect is no longer of interest and one has to look to generational shifts in asymptotic levels attained at maturity.

The third model, although it accepts an age decrement following maturity, permits the growing deficits of old age to be significantly compensated. The model is perhaps appropriate to fluid intelligence and other psychological variables involving speed of response. It fits cases where age changes can be expected due to biological events, but where environmental inputs can have moderating effects.

To enhance the validity of cross-sectional studies, and to take more account of the underlying assumptions derived from the chosen model, Baltes (1968), Baltes and Willis (1979) and Shaie (1977) have made some helpful suggestions with regard to the measurement of cohort differences. However, the degree of representation of the samples selected for the target populations remains problematic.

The models discussed indicate that there may be several patterns of ageing. Descriptive data are still needed on age changes in a population to define the underlying model more clearly. These can only be obtained properly by a longitudinal study of the same individuals. All performance measures support an increase in the variance of performance measures with age (Rabbitt,1981a). Longitudinal studies enable one to ascertain whether age changes are continuous, as suggested by the steadily falling averages of group decade samples, or whether the model of stability proposed by Riegel and Riegel (1972) is more appropriate. It is necessary to identify the different patterns of ageing and then to distinguish between the different factors which affect performance in different ways. Further, longitudinal studies have the added advantage of the availability of actual measures of intellectual functioning at different points in time, instead of being left with the problem of attempting to estimate past performance from current performance (Birren, 1968; Botwinick, 1967).

In addition to the adoption of models from cognitive psychology, gerontology necessarily involves the study of individual differences in performance. As Rabbitt (1981a) has strongly argued, these two fields do not complement each other as well as one would hope. Cognitive psychologists rarely seek to replicate their paradigms on large groups, or validate them against skilled performance in everyday life. In framing abstract models, they neglect the fact that performance is continually changing and that there is enormous variability between subjects. In contrast, psychometricians are eager to try to describe change and define individual differences. They are therefore conscientious in the validity and reliability of the tests they use, but fail to look for the causes of the relationships that their measures reveal. It is now imperative that work in ageing focuses on bringing these two fields closer together, so that casual relationships can be identified, relating individual differences to underlying processes. Rabbitt (1981b) has made attempts to develop models from assumptions underlying elementary control theory specifically developed to describe changing processes, and to suggest how self-controlling systems can adapt to alterations in task demands, as well as describing individual differences in strategies of control. The advantage of such a model is that, since it assumes an active and productive control by the subject over behaviour, it allows the extension of the interpretation of the data beyond a simple laboratory memory task to complex everyday situations. By assessing how the elderly access and use information stored in short and long-term memory to control their behaviour, and why the efficiency with which they do this declines with age, it provides a useful insight into their cognitive function in daily living. These ideas have been usefully developed in the study of the elderly's ability to keep track of laboratory analogues of everyday conversations, route finding and a computer analogue of supermarket shopping.

This thesis attempts to explain individual differences in memory and comprehension of everyday events by similarly identifying the underlying processes involved in these cognitive processes, and explaining age differences in performance on these indices. Thus, it seeks to assess how the elderly encode and comprehend events by accessing information in long-term memory. Further, it examines their ability to manipulate facts in short-term memory, and the efficiency with which the storage and retrieval systems operate.

1.0.2 Memory and Ageing

The majority of studies have used age comparisons on simple laboratory memory tasks, with the aim of locating the memory deficit in the system. Few studies have taken up the approach originated by Bartlett (1932), studying inferential memory and comprehension, and applying more global descriptions of memory concerned with use of memory in the comprehension of complex material akin to the information encountered in everyday life. Instead, researchers have confined themselves to studying basic verbal material ranging from a string of isolated words to text comprehension, and although concepts such as levels of processing (Craik and Lockhart, 1972) have been applied to the data, these have yet to be used for tackling the broad range of practical memory problems encountered by the elderly in everyday life.

Primary, or short-term memory functions show little, if any, decrement associated with ageing. This holds especially if either capacity or accuracy is measured. Age decrements in recency and the Brown-Peterson paradigm appear to be minimal (Craik, 1977, for a review). Forward memory span gives slight but reliable decrements (Parkinson, 1982) and retrieval times show little increase until the 70's are reached. It is only when the complexity of the task is increased that older people tend to perform less well (e.g. backward memory span gives a greater decrement with age than forward memory span: Bromley, 1958). However, these results must be treated with some caution, since, rather than showing a deficit in primary memory they illustrate that age decrements can be seen to increase with increasing manipulation and transformation of the material. This suggests reduction in the efficiency of a system in which recently acquired information can be re-ordered, selectively re-encoded, or rehearsed. These attributes fit the 'working memory' system postulated by Baddeley and Hitch (1974) rather than the contraction in the capacity of a static repository for gradually decaying information to which models of short term memory refer e.g. Waugh and Norman (1965).

The investigations of age differences in working memory are few (Baddeley and Hitch, 1974), although Rabbitt (1981a) has illustrated how a reduction in working memory capacity affects the use of control programmes in everyday cognitive function. In a more traditional form, Spilich (1983) has looked at age differences in working memory with respect to text processing. He found that whereas young subjects were able to carry four propositions forward in working memory, older subjects could only carry two. This implies a decline in the capacity of the system where division of attention, manipulation and novel integration of verbal materials are required.

Age differences have also been found in long term memory, for both the retrieval of specific events and knowledge of facts. There are large decrements found in episodic memory tasks (Burke and Light, 1981; Craik, 1977, for reviews). These have been explained by a reduction in encoding and retrieval efficiency (Rabinowitz, Ackerman and Craik, 1982). However, the evidence can be generalised to the extent that age losses are greatest where both encoding and retrieval operations require self-initiation, and there is consequently little help given from the stimulus or context. This means that free recall tasks show large age decrements, with the elderly using less effective retrieval processes. This age deficit can be reduced if encoding is aided by a semantic orienting task (Craik, 1977; Craik and Simon, 1980), by instructions to organise the list (Hultsch, 1969) or by using mediators in a paired associate task (Hulicka and Grossman, 1967).

In contrast to episodic memory, many studies have found that age differences are slight or non-existent when subjects are tested for factual knowledge (Botwinick and Storandt, 1980; Lachman and Lachman, 1980; Perlmutter, 1978). Likewise, Charness (1982) has found that performance on highly skilled games (e.g. chess, bridge) is well maintained to at least the age of sixty. There are, however, studies that contradict these; at the very least the elderly make frequent complaints about the retrieval of information. In addition, vocabulary tests, and tests of verbal fluency which both reflect semantic memory functioning, show age decrements. Alongside this, Cohen and Faulkner (1981) have found that the elderly show deficits in drawing inferences, and Taub (1979) has data indicating age-related decrements in comprehension. Both require the use of information in semantic memory.

To summarise, primary memory functions show only slight age differences, but working memory processes do seem to be impaired. This is highlighted when the task requires the manipulation of information, transformation information, or the division of attention. Episodic memory tasks show a large impairment, which is attributed to inefficient processing at both encoding and retrieval stages. It seems that the elderly fail to integrate the stimuli with the context, leaving the event to be less distinctive and resulting in contextual clues being less effective for retrieval. In contrast, many aspects of semantic memory seem to remain unimpaired, but higher order functions such as comprehension and inference construction do show some losses with age.

It appears from these studies that, in general, the decrements associated with ageing are most apparent when the subject is required to initiate processes, draw inferences, form associations, and manipulate the information presented. In contrast, the decrements are least when the processes are driven by the stimulus or supported by the context. If complex functions are required for the response, then only if they are habitual and overlearned is the deficit slight (Craik and Rabinowitz, 1984). This holds with Rabbitt's (1979,1981a) notion that data-driven processes show less decline with age, but 'memory' or 'conceptually' driven processes are impaired.

1.0.3 Everyday Memory

Much of the research already referred to can be considered rather sterile. In the main, it fails to address the question of how the elderly perform on everyday cognitive tasks. This is by no means a failing specific to ageing research, since one has only to glance over the last one hundred years of research in cognitive psychology to find plenty of evidence of the same failing. Yet, it would be harsh to conclude that a century has been spent discovering how undergraduates perform on cognitive games in the laboratory. The effects found from mainstream psychology, and the theoretical concepts which have been derived, are in many cases sufficiently robust to be applicable outside the laboratory. While results that have been obtained in the outside world cannot always be predicted from laboratory studies, the relationship is sufficiently close to be encouraging (Baddeley, 1981). Indeed, if this were not the case, traditional laboratory research would be seriously challenged, for as Neisser (1978) points out, the effects found are so robust that even ten year olds find them obvious.

Cognitive psychology has reached the point where the need for exploration beyond the laboratory is seen to be necessary. There is now a growing awareness of the need for 'ecological validity'. This new point is no longer controversial, and it is upheld by the fact that, ultimately, cognitive psychology is concerned with understanding human behaviour in general. By confining models to very simple and tightly controlled laboratory tasks, there is a danger of modelling trivial and incidental subroutines that are just part of a rich, complex and purposefully adaptive behavioural system (Rabbitt, 1981a). The 'reductionist' approach of dissecting complex tasks into apparently simpler tasks and more tractable component processes has been seen as one solution. The re-assembly of these 'building block' models of simple processes into models for complex processes is at least viable. However, it is surely better to have models for the entire process of memory, e.g.

it simply has not proved useful to have models for reading which are 'built up' from models for say line recognition, feature analysis etc. Why should models for memory be any different? There is a need for models of memory to be reconstructed so that they include plausible accounts of deductive inference, problem solving, and attentional selection, all of which are important processes in everyday memory tasks.

Another contributing factor to this move beyond the laboratory is derived from a deeper concern with applied problems. This has, in part, been the result of a changing political and economic climate. As research funds contract, the need to justify research has increased. From the perspective of ageing research, there is also the effect of the change in structure of the population, which has led to a pressing need for a greater knowledge of the patterns of ageing and the nature of the physical and psychological problems associated with it.

In order for any of these applied problems to be addressed successfully, there has to be an interaction between the laboratory and the studies carried out in the everyday world, in order that objective measures can still be obtained. In this way, models derived from laboratory research can be tried and tested outside, with problems and insights gained from this fed back to the laboratory. All these studies are faced with the difficulty of obtaining realism, but not at the complete expense of rigour which is necessary to make a useful interpretation from the data. A growing number of studies now strive to achieve the balance between realism and rigour in the examination of everyday memory. These extend from simulations of everyday tasks in the laboratory to natural experiments. Through these attempts, the techniques and methods used for working outside the laboratory will hopefully improve, so as to illuminate through objective studies more of the psychology of everyday life.

This thesis is concerned with how the elderly remember and comprehend everyday events. For this reason, the work reviewed will be confined to those aspects of everyday memory which are seen to be related to this, i.e. memory for television broadcasts, eyewitness testimony and the small amount of literature which covers attempts to study the everyday memory performance of the elderly. (Excellent reviews of other aspects of everyday memory can be found in Gruneberg, Morris, and Sykes, 1978; Neisser, 1980)

1.0.4 Memory for Television Broadcasts

The study of memory for television broadcasts is highly conducive to the integration of laboratory studies with natural studies, which is an important approach in the study of everyday memory. It is a task which a large proportion of the population engage in every day, but, from a natural study carried out by Neuman (1976), it appears that people are generally poor at recalling what they have seen and heard. The same is true of radio broadcasts; Stern (1961) found that half the listeners to a news programme could not recall a single item out of the nineteen presented.

The laboratory studies that have investigated recall performance give conflicting data, but this is put down to the fact that cued recall has been used instead of free recall. Even though individual differences are large, Findahl and Hoijer (1975) found that many subjects could retain a high percentage of story details when given multiple choice questions. Robinson, Dans, Salin and O'Toole (1980), using headlines to prompt recall, also found that subjects performed reasonably well, recalling about half the 'control themes' of the news broadcasts. However, there have been no attempts to identify the source of these individual differences in performance, or the cognitive processes which are part of the task. Berry, Gunter and Clifford (1981) complain that cognitive psychology offers few predictions about television's effectiveness as a medium of information and instruction, and blame this upon a lack of guidelines for applying psychological theory to real-life material.

Instead of taking up this challenge, research has considered a few social factors such as education level and motivational variables and the effect they have on recall, but mainly it has centred on programme design and organisational factors. This has allowed the application of paradigms from cognitive psychology to be transferred with relative ease and some success. Rates of delivery of spoken material have been examined, and also factors affecting optimal rates such as the topic of the material and the vocabulary used. Serial position effects have been found by Tannenbaum (1954) with radio newscasts, but Gunter (1979) found this to be weak in television broadcasts overall, but stronger in newscaster-only cases. This implies that picture items can easily offset this effect. However, although these and other effects such as the build up in proactive inference (Gunter, Clifford and Berry, 1980) have been found, there are still problems in dealing with such complex material and isolating the relevant factors. In some cases, but not all, the theme of the item in recall may be more important than the way in which it is presented.

The study of memory has paid little attention to situations where both visual and verbal information are simultaneously presented, as occurs in everyday life and in television broadcasts. Waagaaner et al. (1984) were prompted to investigate bisensory presentation of information following the finding that the presentation of weather forecasts on television did not lead to better recall than presentation on the radio. They examined 7 different conditions of presentation using words and pictures which were either matching or non-matching, from same or different categories, contiguous or noncontiguous, and found that in most conditions bisensory presentation led to poorer recall than the level expected in the case of independent combination. Facilitation was found only when pictures and words were from the same category and presented contiguously. Other researchers have also addressed the question of which presentation modes optimise recall of information. Katz, Adoni and Parness (1977) found that overall, listening to a television broadcast was as effective as listening and viewing. However, Findahl (1971) showed that recall was best for items when the pictures corresponded to the verbal information, and poorest when there was no illustration. It appears that cross-modal advantage only arises when elaboration of the material occurs (Waagaaner et al. 1984).

For the television producers, some of this research may be useful, but in the study of everyday memory it leaves one none the wiser as to why some people recall some broadcasts better than others, and what the main features of the broadcast, and the cognitive processes are.

1.0.5 Eyewitness Testimony

Eyewitness testimony necessarily involves the recall of events, and can be seen as an area of applied psychology which directly confronts theoretical issues in perception, learning and memory. Researchers have explored aspects of the event itself, the effect of intervening material between the event and the testimony, and individual differences between witnesses.

Many characteristics of the event have been investigated. First, the event exposure time. Loftus (1972) and Hintzman (1976) have found that picture recognition is a monotonically increasing function of inspection time. This is not linear since there are threshold type effects and ceiling effects. Second, the complexity of the event can lead to an increase in later recognition (Loftus, 1972; Wells, 1972). Third, the familiarity of the physical surroundings and the context of the event can affect recall, since lack of familiarity can produce large distortions in estimates of size and distance (Grether and Baker, 1972).

Loftus has examined the effect of the intervening material between the event and when the witness gives a testimony. Loftus and Palmer (1974) showed subjects a film of a traffic accident and varied the format of the intervening interrogation. When subjects were asked about the speed of the cars, they were either asked ' how fast were the cars going when they hit each other?' or the same question with 'smashed into' substituting the word 'hit'. A week later subjects who were questioned with the phrase 'smashed into' were more likely to indicate that they saw broken glass than those who were asked the question with the verb 'hit'. Similar studies illustrating effects of other intervening interrogation procedures and leading questions have also been carried out by Loftus (1975), Loftus, Altman and Geballe (1975) and Loftus, Miller and Burns (1978).

The work on individual differences has asked whether men are better witnesses than women, and how age affects the accuracy and completeness of accounts. There has been a series of conflicting results with respect to sex differences. Ellis, Shepherd and Bruce (1973), Lipton (1977) and Witryol and Kaess (1957) say that females are better than males, but, Clifford and Scott (1978) and Trankell (1972) maintain that males are better than females, Bird (1927), Cady (1924) and McKelvie (1976) on the other hand, could find no differences between the two.

Age differences have examined differences between groups of children as well as adult age differences. Older children are better at recognising events than younger children (Goldstein and Chance 1964, 1965; Kagan et al. 1973). In older adults the positon is more complicated and although Smith and Winograd (1977) have found that people over the age of 60 perform poorly compared with younger adults, more research is necessary with this complex material as well as a greater understanding of the processes involved in recalling events and which of these are affected by increases in chronological age.

Inspite of the successful research in this field, it receives similar critism to that of recall of television broadcasts. Clifford (1978) suggests that because eyewitness behaviour is at base a perceptual-memory phenomenon, the paradigms of visual memory have been over employed, which has restricted independent variable manipulation, dependent variable measurement and the ecological validity of experiments. There is a strong case for a re-examination of the methodology employed, since at present it tends to inflate the estimates of eyewitness ability. The use of static pictures is far removed from a dynamic event in which information is presented rapidly, and attention is diverted rather than focussed.

Both studies of television recall and eyewitness testimony may have begun to answer some of the questions asked by television producers and the judicial system respectively, however the need to identify the underlying cognitive subsystems employed in these tasks, and to explain the apparent individual differences on complex everyday memory tasks is yet to be achieved.

This thesis attempts to examine some of the cognitive processes involved in comprehending and remembering events. It investigates individual differences in performance on this task, by investigating differences in component processes.

1.0.6 Everyday Memory In The Elderly

The studies of everyday memory in the elderly fall into two categories. First, there are a few studies which have used questionnaires concerned with slips of absent-mindedness, of which the elderly frequently complain. These have sought correlations between self-assessment measures and performance on laboratory tasks, in the hope that the questionnaires may have some predictive value, and be of use in assessment of pre-senile people.

Second, a handful of laboratory studies have been carried out. These have been either traditional laboratory tasks which have been designed to explore a specific aspect of everyday cognitive performance, or laboratory tasks which simulate part of, or an entire task related to everyday life.

1.0.7 Questionnaire Studies

If one is interested in finding out if a person is functioning efficiently in their everyday environment, then perhaps the most simple and direct method of obtaining such information is to ask them. The use of everyday memory questionnaires is a cheap and simple means of surveying a large number of memory phenomena. They also have the advantage of frequently assessing covert memory lapses that everyday memory experiments fail to detect.

It is surprising that the history of everyday memory questionnaires is one of disappointment. In practice, people's ratings of their own cognitive performance have seldom, if ever been found to correlate either with their performance on laboratory memory tasks, or with their efficiency in everyday life. Both Broadbent's Cognitive Failure Questionnaire (Broadbent, Cooper, FitzGerald & Parkes 1982), and the Harris and Sunderland Memory Questionnaire (Sunderland, Harris & Baddeley 1983) have struggled to find significant correlations with laboratory measures. However, relatives' ratings were found to be more consistent and to yield higher correlations. A similar lack of success has also been encountered with Hermann and Neisser's (1980) inventory of everyday experiences.

There will always be important objections to the questionnaire method of assessment. A subject may not be aware of his own memory performance, may forget his failures of memory, or may not tell the truth. This can be applied to severly amnesic patients, but those with relatively normal memories are likely to have a clearer assessment of their cognitive functioning. Flavell and Wellman (1976) found that older children were able to predict fairly accurately their recall of two to ten items and Perlmutter (1978) found no age differences in the metamemory performance between old and young subjects on the numbers of words they thought they could recall and recognise.

Rabbitt (1982) has suggested that the lack of significant correlations between self-rating questionnaires and actual performance is related to the request for assessed absolute levels of competence. This poses a problem, since people can only assess their levels of competence in relative terms. To circumvent this, Rabbitt (1982) adapted the Harris and Sunderland Memory Questionnaire for elderly people with the requirement that subjects estimated the change in occurrence of absent-minded slips over the last thirty years. Using this form of the questionnaire, Rabbitt has obtained slightly stronger correlations between performance on laboratory memory tasks and perceived decrement, however there were no effects of increasing age or of I.Q. test score on the Cognitive Failure Questionnaire or the modified Memory Questionnaire.

Zelinski, Gilewski and Thompson (1980) developed a questionnaire to examine inter- and intra-age differences in self-assessment and performance on a variety of memory tasks. Their data are unusual since they obtained significant correlations between assessment and laboratory performance for the group of elderly subjects, but no such relationship was found for a sample of young college students. They suggest that younger subjects are just poorer at estimating their memory performance. A reason for this can be found by following Rabbitt's (1982) reasoning. The young may not have experienced sufficient change in memory performance to be able to form any relative judgement about how their memories perform.

Tenny (1984) has looked at age differences for a more specific lapse of memory, namely misplacing objects. The questionnaire was designed to examine incidents when objects were misplaced, and identify factors associated with this which might change with age. The findings showed that the young reported more incidents of misplacing objects than the old, although this difference was not significant. Search was the only problem identified by the elderly and not the number of incidents of misplacements. In a repetition of Tenny's study, Rabbitt and Abson (1984) used a questionnaire which was constructed to similate a recognition task in which a number of possible misplaced objects were listed to reduce the possibilty of subjects giving a low estimate of misplacements. They found that the youngest subjects made more complaints than the older subjects. However, there was no overall effect of I.Q. test score. This means that we can not assume that the elderly just 'forget that they forget' since the clever elderly complain no more than the less clever elderly. The young subjects complain more possibly because they still engage in a life style which exerts heavy cognitive demands on them, hence they are more likely to make errors and notice them more. Retired people are no longer stretched cognitively and their degree of cognitive exercise may be reducing faster than their capabilities..

These studies perhaps contribute more to the knowledge of questionnaires than to the everyday memory performance of the elderly.

However, Rabbitt (1984a) has found consistency in the responses given to questionnaires by the elderly, which Harris, Sunderland and Baddeley (1983) found to be low in severe head injury patients. Further, Rabbitt offers a plausible explanation as to why correlations with performance are difficult to obtain, and why the younger subjects report a similar decrement to the older subjects. Bennett, Levy and Powell (1980) have suggested that further improvement in correlations could be obtained by using tests which are closer to the real-life skills being assessed in the questionnaire. Nonetheless the difficulty remains with obtaining significant correlations and establishing questionnaires as methods of assessment. Much replication of results and validation is still required.

1.0.8 Everyday Memory In The Elderly: Laboratory Studies

Simon (1979) notes that research on ageing has been restrictive in its investigation of 'semantically rich' domains, which has left much of the work lacking external validity. Of the studies which have considered everyday cognitive performance or applications of their findings, two approaches are evident. Some studies have merely considered the declines in information processing ability and how they relate to performance in everyday situations, and others have considered the development of compensatory mechanisms to reduce the effects of the decrement in processing skills. These experiments have left behind a doubt associated with the word list learning tasks and paired associate learning tasks, namely whether they are appropriate tasks for investigating ageing effects and evidence of developmental plasticity (Scheidt and Schaie, 1978).

There are additional advantages to be gained from the ecological validity associated with tasks, as Rabbitt (1982) has indicated. Tasks with high ecological validity are more easily understood and less likely to confuse the elderly. They are usually familiar and hence are likely to yield more realistic performance levels than laboratory tasks, which invariably appear boring and pointless.

Laboratory experiments on visual search and monitoring have been used, notably by Rabbitt and Vyas (1980), to contribute to our knowledge of how the elderly adopt search strategies in everyday life. Sanford and Maule (1971,1973a, 1973b) compared young and elderly people on a simulated industrial monitoring task in which they were required to search up to three 'locations' to detect 'signals'. The absolute and conditional probabilities of the signals differed between the locations, which the young quickly learned and adopted for their search strategies. Although the elderly could describe the variations in signal probabilities, they did not seem to use this information in their search strategy. Rabbitt (1979) reports a similar finding where targets occurred more often at some display locations than others. The elderly were able to produce the three probable target locations in the correct rank order, but again seemed unable to use this information to guide search. This has implications for the strategies the elderly adopt in everyday living; even though an elderly person may find a misplaced item in the same location numerous times, that is not necessarily where he or she starts their search.

The distinction between the possession and the use of necessary information in performing tasks was apparent in another everyday memory task reported by Rabbitt (1981b). Subjects aged over fifty years were asked 'how would you explain to a tourist how to get from X to Y ?', and as a control, 'how would you point out the correct route to the tourist using this city plan?' (all subjects were very familiar with the city in question). The subjects were mostly very efficient at both tasks. However, subjects older than seventy years often gave inadequate descriptions and frequently forgot which section of the route they had just described, resulting in either back-tracking or skipping forward and forgetting essential information. These subjects were, nonetheless , able to specify useful landmarks and give a lot of detail of the shops etc. which occurred on this mental walk, suggesting that all the information was available to them but they were not always able to use it efficiently in tasks.

The problem of following group conversations is another aspect of everyday life which causes difficulties for the elderly, although they generally seem able to talk freely and sensibly to a single person. Rabbitt (1982) has investigated this using a task in which subjects were required to recall sentences spoken over a tape recorder and also who spoke them. As the number of speakers increased, the elderly suffered far more than the young, finding it difficult to recall either context or source. In group conversations it appears that the elderly have difficulty following what is said and who said it, whereas in one-to-one situations this is markedly reduced.

These studies all relate to the elderly's ability to comprehend everyday events. An important part of this is to use information available in long-term memory to understand and process efficiently the event encountered. The first two studies referred to above suggest that the elderly are not always able to use the information they have available to perform a task. Secondly, the 'conversation experiment' indicates that the elderly find it difficult to follow and keep track of more complex situations, which is the very nature of everyday events and television broadcasts.

Both Charness (1981) and Salthouse (1985) provide interesting data from a different approach to everyday performance in the elderly. Their concern is with compensatory processes which have been suggested as part of the developmental plasticity associated with ageing. If the development of these processes exists, then it follows that there are different mechanisms underlying equivalent performance levels in the old and the young. This approach has been termed the Molar Equivalence-Molecular Decomposition Procedure (Salthouse, 1984), as it involves equating individuals of different ages on the proficiency of a molar task, and then measuring age trends on the effectiveness of molecular components. The drawback with these studies is that the subject sample is selected so as to eliminate any relationship between age and molar performance, and therefore cannot really be used to make inferences about the presence or absence of age relationships in the wider population. However, age differences in molecular components are of more relevance to normative ageing research.

Charness (1981,1983) originated this approach in studies of chess playing and bridge. In the chess study there was only a correlation of 0.09 between age and skill. The task was to select a move from a mid-game configuration and then later to recall the positions of the pieces in that configuration. Older players recalled fewer pieces in correct locations, but they also spent less time evaluating the configurations when they chose their moves. Thus the age differences can be assigned to differences in study time. Charness suggests that older players compensate for their poor memory by more efficient search of problem space.

In the study of bridge playing, older players were slower at registering and encoding bridge information, taking longer to make bidding decisions. However, Charness was unable to identify any compensatory processes on which the elderly performed better than the young. The reason for their equal overall performance in this task was left unexplained.

Salthouse (1985) has performed a similar study on transcription typing. He found older typists were able to achieve high levels of typing speed despite the handicap of slower perceptual-motor speed, through the compensatory mechanism of anticipatory perception, which minimizes the limitations of the slower perceptual-motor processes. However there is a problem with these data. Salthouse is unable to identify whether the advantage of anticipatory perception, which is characteristic of the older subjects, is a compensatory mechanism which has been developed as such, or if it is a process which has maintained its efficiency. Without knowing the levels the older subjects achieved on the measures taken at a younger age, these conclusions remain in doubt.

These experiments and those on everyday memory illustrate the problems in methodology which both applied areas of study are attempting to resolve. The link with experimental psychology is a vital one, if objective measures are to be made, and the necessary control of variables achieved. However, there is a danger it may become restricting. There is a need in ageing research for models of change in performance (Rabbitt 1981a). Cognitive psychology per se does not offer such a model and so the time has come for gerontology to develop its own models, instead of borrowing inappropriate models and just accumulating effects.

1.0.9 The Approach and Aims of This Thesis

Against the background of research in ageing and everyday memory, which has been reviewed previously, the studies reported here will now be briefly introduced. There follows an attempt to explore the effects of increasing age on the everyday memory task of comprehending and remembering events as conveyed by television or everyday perception. The methodology used is necessarily cross-sectional, due to the time limit imposed on this work. The benefits and limitations of this method have already been covered.

The critical aspect of cross-sectional studies is subject selection. Studies have generally compared groups of young college students with groups of elderly people. The groups are frequently matched for health, education level, and I.Q. (usually by using age-adjusted scores), as an attempt to eliminate confounding factors. However, Rabbitt (1984b) has questioned the practice of using age-adjusted I.Q. test scores. Performance on a number of cognitive tasks has been found to correlate highly with raw I.Q. test scores, therefore matching subjects on age-adjusted I.Q. test scores means that one is deliberately selecting old people with lower raw scores than their young controls. If one knows that age adjustments correctly predict former young adults' I.Q. scores then this may be permissible, however this is not the case because age adjustments have always been made on the basis of cross sectional data. This means that effects found which are attributed to age differences are likely to be a consequence of the difference in I.Q. test score. Indeed, Rabbitt (1982) has found I.Q. test score to be a better predictor of performance on a number of indices than age per se.

In the studies reported here, subjects were selected from volunteers taking part in a longitudinal research project organised by North East Age Research. This provided a pool of over two thousand well documented subjects, on verbal and performance I.Q. tests. For all studies, large groups of elderly subjects were used within the age range 50-85 years. Their selection was based upon age and AH4 part one test score. In this way both age and I.Q. test score effects could be assessed on the indices measured. Three bands of AH4 test score were filled with reasonably even numbers of subjects in each, and an even distribution of ages within each.

A frequent problem with the elderly is that they are easily confused by what is happening around them or by what has happened. They find it difficult to comprehend exactly what is 'going on'. Television offers an ideal medium to begin investigating this problem. The goal of television is to transmit events and information to mass audiences: thus it provides suitable testing material which can be used in the laboratory.

The first studies examine the elderly's performance on recalling television news broadcasts. The effects of recall condition and the types of detail recalled with respect to age and I.Q. test score groups are investigated. The experiments which follow this seek to explore possible differences in processing and the structure and organisation of memory which may account for the group differences on this everyday memory task.

An important component process in the comprehension of events is the access and retrieval of information stored in long-term memory. This knowledge of factual information, previous experiences, and personal information is used to identify present situations, draw inferences in order to comprehend them,and also to predict what event is likely to occur next. These uses of long-term memory and the nature of the systems associated with its role are explored in Chapters 4 to 7. The structure and organization of long-term memory, and how this affects the performance of the elderly, is the focus of this thesis.

The laboratory studies offer explanations of group differences in comprehension and memory which are related to a decline in inferential reasoning ability and the richness of the association network and its organisation in long-term memory. Models of ageing and individual differences in intelligence are applied to the data, and an explanation is offered of the processes underlying the elderly's competence levels in the comprehension and memory of everyday events.

CHAPTER TWO

2 MEMORY FOR TELEVISION NEWS BROADCASTS

2.1 INTRODUCTION

The task of watching and recalling television broadcasts is particularly relevent to the elderly, since watching television is one of the dominant leisure activities of older people (Cowgill and Baulch 1962 ; Schramm 1969). However, little is known about the ability of the elderly to comprehend and recall the information conveyed to them via this medium. Indeed, it is only recently that the characteristics of television as an information medium have been investigated.

The overwhelming feature of television is that it allows a lot of complex information to be conveyed in a relatively short time using simultaneously the visual and the auditory modalities. In the main, it depends upon immediate comprehension and effective organization for later retrieval. One immediately obvious problem for the elderly is the speed of presentation. Findahl and Hoijer (1977) found that while American students could cope with delivery rates of 250-300 words per minute older volunteers had difficulty in coping with half this transmission rate (Fairbanks, Guttman and Miron, 1957; Foulke, 1968). Another related problem for the elderly, particularly in news broadcasts, is the inadequacy of parsing in the material, such that they are faced with an 'unrelenting flow of information'. Many of the production variables and artistically pleasing visual effects may only serve to confuse the elderly, with the use of flashbacks and flashforwards leaving the elderly hopelessly floundering in their effort to construct the story line. Nevertheless, the elderly derive pleasure from television, and, according to Rubin and Rubin (1974), the most frequently expressed motive for watching television in the elderly was information gain.

There are a number of advantages to be gained from using material which is familiar to the elderly. One of the problems with most laboratory memory tasks is that they are tedious, especially to the elderly who fail to see the connection between recalling a list of unassociated words and the problems they encounter due to their slips of memory in everyday life. For this reason, the elderly may frequently perform poorly on simple laboratory tasks, not because of any cognitive deficit, but because they often formulate bizarre experimental hypotheses and respond accordingly, instead of following the experimenter's instructions. Once corrected they may begin to perform well, since the task is frequently easier than they first thought. Welford (1958) has suggested that much of the apparent difficulty in learning by older people is not due to an incapacity to learn or recall as such, but rather to an inability to comprehend the material or deal with the conditions under which it is presented. A task such as recall of television news broadcasts has more ecological validity than many laboratory memory tests, is more easily understood through its familiarity, and therefore runs less risk of these confusions. Relative to the young, the elderly improve more with practice on choice reaction time tasks (Rabbitt, 1980), however one cannot generalise from this simple task to more complex tasks. It appears that the initial performance of the elderly on unfamiliar tasks may seem exceptionally poor, so that valid comparisons between old and young subjects require task familiarity. Watching television is a familiar everyday cognitive task for the elderly and therefore should more accurately reflect actual performance levels.

The complexity of the medium of television has already been referred to, but although this task has an apparently complex structure, it can be dissected into various cognitive skills, such as speed of theme construction, identifying thematic and irrelevant details, and effective organisation through association. Two studies on the recall of television news broadcasts by the elderly are reported here. The first was a pilot study which involved 65 elderly volunteers, and the second a larger study designed to examine not only the recall of such material, but also the effects of serial position and interest on recall performance. Both also investigate the intercorrelations between performance on this everyday memory task and a number of indices.

Rabbitt (1982, 1984a) found that correlations between laboratory memory tests and chronological age per se were poor, and that IQ test scores yield a far stronger correlation with performance measures. Hence the best index of cognitive change with age appears to be I.Q. test score. All the subjects selected for these studies had previously been screened on the AH4 test of general intelligence (part I and II), the Mill Hill vocabulary scale and two self rating memory questionnaires; hence the value of these indices as predictors of performance on this task could be assessed.

2.2 EXPERIMENT 2.1- PILOT STUDY

2.2.1 Introduction

As mentioned in the general introduction (1.1), research into ageing and performance on everyday memory tasks, such as watching and remembering television programmes, has been extremely limited. Cavanaugh (1983) considered comprehension and retention of television broadcasts in a group of sixty year olds; he found that only the low verbal adults were consistently poorer than their younger counterparts on tests of free recall, probe recall and recognition. The order of task presentation was such that free recall was always administered first. However, as there was no interaction between verbal ability and test type, Cavanaugh suggests that the poor performance of the low verbal subjects may not be due to retrieval since retention differences were not related to delay, rather there may be a deficit in initial programme comprehension.

This pilot study set up two different recall conditions, free recall and cued recall with questions referring to news story details from the news broadcast shown. Unlike Cavanaugh (1983) who used short clips from different broadcasts, one entire evening news broadcast was used as recall material. There was a further manipulation of the recall condition in that questions were either random or corresponded in sequence to the temporal order in which the facts had been given. Age and a number of other indices obtained on the subjects were explored as predictors of performance following the finding that it was the low verbal subjects who performed poorly and not the old. Of particular interest were the two self-rating memory questionnaires because although memory questionnaires have failed to predict performance at simple, abstract laboratory memory tasks, this study with its use of 'real life' material might show the questionnaires to be more predictive. Watching television is an everyday life task and the questionnaires require subjects to rate their competence on such tasks, hence their impressions of everyday life performance may correlate with their performance on the everyday memory task of comprehending and recalling news broadcasts to a higher level of significance than found with simple laboratory tasks.

The study also looked for serial position effects and duration of episode effects. Serial position curves for real life events and material have been found by Baddeley and Hitch (1975) for rugby matches and by Gunter (1979) with television news items with some presentation formats, although radio news broadcasts show the strongest effects (Tannenbaun 1954). One factor which may interfere with this effect is the duration of the episodes within the broadcast, hence both duration and serial position effects were examined.

In each experiment except Experiment 2.1 the subjects were not randomly selected, rather they were selected according to their age and AH4 score on part one of the test. A 3X3 matrix was constructed of three age bands and three AH4 test score bands. The selection sought to fill each of the cells in this matrix with an equal number of subjects. By using this selection both age and I.Q. test score group differences could be examined and their relative contributions to the variance in the data.

2.2.2 Subjects

Sixty five volunteers were selected from those taking part in a longitudinal study organised by North East Age Research. The criterion for selection was that the subjects had already received the two preliminary tests administered for the longitudinal research, which provided useful documentation on the subjects. All had received the AH4 test of general intelligence, the Mill Hill vocabulary scale and two self-rating memory questionnaires, Harris/Sunderland Memory Questionnaire and the Broadbent Cognitive Failure Questionnaire. The mean age of the group was 64.72 years, age range 51-84 years, std. dev. 9.505,and mean AH4 part one test score 30.3 out of 65, std. dev. 9.505, range 5-47 (see Table 2.1). All subjects were paid travelling expenses at the flat rate of 2 pounds per session.

2.2.3 Materials

A B.B.C. news broadcast was recorded in December 1982, approximately three months prior to testing subjects. The news broadcast contained eight items and filled the ten minute slot allocated to it.

2.2.4 Apparatus

A National Panasonic Cassette Video Recorder (NV366,VHS) was used to record the broadcast and it was shown to subjects through a Sony Trinitron Monitor (CV,2000 PSB) with a 21 inch screen.

Table 2.1

3x3 Matrix Of Subjects For Age and AH4 Test Score Groups

AH4 Test Score Group Age Group	Low Group 1	Middle Group 2	High Group 3	Total Mean (x) Std. Dev. Range (r)
Young Group 1	1	8	2	N=11 x=53.27 std.dev.=2.15 r=51-58 years
Middle Group 2	3	27	11	N=41 x=64.32 std.dev.=2.92 r=60-69 years
Old Group 3	3	10	0	N=13 x=75.69 std.dev=5.406 r=70-84 years
Total Mean (x) Std. Dev. Range	N=7 x=13.4 std.dev.=4.99 r=5-19/65	N=45 x=29.56 std.dev.=5.24 r=20-38/65	N=13 x=44.08 std.dev.=3.95 r=40-55/65	N=65 x=30.7 64.7years std.dev.=9.70 7.65 r=5-55/65 51-84yrs

2.2.5 Procedure

Subjects were tested in groups of approximately twelve. Each group was shown the broadcast, and immediately afterwards instructed to write down in as much detail as possible all they recalled of the programme. There was no time limit imposed on this task, but typically it took ten to fifteen minutes. Following a teabreak, subjects were asked to answer a set of questions referring to the broadcast. Half of the group were given booklets with the questions in random order and half with serial ordered questions. Again no time limit was imposed for completion of the task. The two conditions were always administered in the same order. The alternative balanced design was not used as it would have led to assymetrical practice effects. Cued recall in trial one would have benefitted uncued recall in trial two more than the reverse. So one knew that [illegible] cued recall was going to be better than uncued recall anyway, and one did not mind the facilitation by the previous free recall since the main comparison of interest in this study was between groups. Examples of cued recall questions can be found in Appendix 1.

2.2.6 Results

Overall percentage correct free recall of details, and percentage correct answers to the questions relating to details from the broadcast, were computed for each subject. Mean overall scores for the group were 19.5%, and 35.9% respectively, the difference between conditions being highly significant (t= 16.05 $p < 0.01$). This was expected since the improvement in recall score in a cued condition is well documented (Craik 1971, Smith 1980), and the cued recall may in any case have benefitted from the previous free recall.

The effect of question order was also highly significant. Mean recall for random questions was 31.54% n=33, and 40.48% for serial questions n=32, (t= 2.81, f=1.9177 $p<0.01$: p 0.03 respectively).The probability of the first mean being smaller than the second is 0.9964 as computed by a t-test.

In addition to these overall effects of recall condition, the relationship between indices previously obtained on the subjects and recall on the task was examined. A correlation matrix was computed (Table 2.2). It is clear that recall for both conditions is highly correlated (r=0.799). The other indices show that both parts of the AH4 test of intelligence correlate highly with recall, and the Harris/Sunderland Memory Questionnaire also gives a significant correlation with recall performance. Interestingly age comes out very poorly as a predictor of cognitive performance. The Mill Hill vocabulary scale was also a weaker predictor than might have been expected from Cavanaugh's (1983) results with correlations of 0.58, 0.54, 0.38 for young subjects and 0.65, 0.60, 0.35 for old subjects between verbal ability and recall performance on free recall, probe recall and recognition tests respectively.

Table 2.2

Correlation Matrix

	F.R	C.R	M.Q	C.F.Q.	AH41	AH42	M.H.1	M.H.2	AGE
Free Recall (F.R.)	1.00	0.80	-0.30	-0.13	0.59	0.50	0.34	0.15	-0.05
Cued Recall (C.R.)		1.00	-0.40	-0.23	0.50	0.51	0.28	0.18	-0.10
Memory Questionnaire (M.Q.)			1.00	0.23	-0.22	-0.22	0.13	0.19	0.11
Cognitive Failure Questionnaire (C.F.Q.)				1.00	-0.10	-0.02	-0.20	0.17	-0.10
AH4 (part I)					1.00	0.69	0.45	0.38	-0.18
AH4 (part II)						1.00	0.36	0.26	-0.18
Mill Hill (part I) (M.H. 1)							1.00	0.57	0.25
Mill Hill (part II) (M.H. 2)								1.00	0.11
Age									1.00

This was examined further with multi-linear regression analysis using the MIDAS statistical package (see Table 2.3, 2.4). AH4 parts one and two, age, Memory Questionnaire, and Cognitive Failure Questionnaire were included in the analysis as predictors. In both cases AH4 part one was the best predictor, with partialled correlations of 0.54, and 0.407; the predictors accounted for 29.88% and 30.7% of the variance for the free recall and cued recall conditions respectively. The strength of Memory Questionnaire in the cued recall condition is particularly interesting, and the weakness of age in both conditions indicates that it is a particularly poor predictor of cognitive performance.

From Table 2.1 of the subjects' details, it is clear that the distribution of age and AH4 part one scores within this sample is uneven, with the majority of subjects falling within the middle bands. However, as a form of exploratory analysis group differences were examined. Firstly, a 1X2 analysis of variance was computed on age group, I.Q. test score group and recall condition. Separate analyses for age and I.Q. test score group were computed since not all subject cells were complete. There was an overall effect of I.Q. test score group(F=9.104, df=2,61 $P<0.01$), and of recall condition (F=124.083, df=1,61 $P<0.01$). However, no differences between age groups were found (see Figures 2.1, 2.2 and Tables 2.5, 2.6).

The two orders of questions were also tested with respect to age and I.Q. test score group. A one-way analysis of variance was performed on each between subjects variable and the condition. There was a

Table 2.3

Multiple Regression Analysis For Percentage Correct Free Recall.

Analysis of Variance N=51 out of 65

SOURCE	DF	SUM SQRS	MEAN SQR	F-STAT	SIGNIF
REGRESSION	1	1168.1	1168.1	20.885	0.0000
ERROR	49	2740.6	55.9		
TOTAL	50	3908.7			

MULT R = 0.54667 R-SQR = 0.29885 SE = 7.4787

VARIABLE	PARTIAL	COEFFICNT	STD ERROR	T-STAT	SIGNIF
CONSTANT		2.6266	3.8526	0.6817	0.498
AH4 1	0.54667	0.5528	0.1209	4.5700	0.0000

REMAINING	PARTIAL	SIGNIF
AH4 2	0.1177	0.4155
MQ	-0.2145	0.1346
CF	0.0717	0.6207
AGE	-0.0301	0.8356

Table 2.4

Multiple Regression Analysis For Percentage Correct Cued Recall.

```
Analysis of Variance       N=51 out of 65

 SOURCE            DF        SUM SQRS    MEAN SQR    F-STAT    SIGNIF

 REGRESSION         2        2382.8      1191.4      10.645    0.0001
 ERROR             48        5374.7       111.9
 TOTAL             50        7757.5

 MULT R = 0.55422            R-SQR = 0.30716         SE =10.582


 VARIABLE          PARTIAL     COEFFICNT STD ERROR   T-STAT    SIGNIF

 CONSTANT                        44.495    12.734    3.4941    0.0010
 AH4 1             0.40683        0.548     0.1778   3.0855    0.0034
 MQ               -0.33535       -0.274     0.1112  -2.4661    0.0173


 REMAINING         PARTIAL      SIGNIF
 AH4 2             0.2262       0.1181
 CF                0.0305       0.8350
 AGE              -0.1677       0.2494
```

Table 2.5

1X1 Analysis Of Variance Age Group X Recall Condition

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Age	290.561	2	145.280	0.660	0.520
S-Within	13418.063	61	219.968		
Cond.	5299.945	1	5299.945	164.674	0.001
Age X Cond.	83.717	2	41.858	1.301	0.280
Cond. X S-Within	1963.250	61	32.184		

Table 2.6

1X1 Analysis Of Variance I.Q. Test Score Group X Recall Condition

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	3271.492	2	1635.746	9.104	0.001
S-Within	10960.625	61	179.682		
Cond.	4187.035	1	4187.035	124.083	0.001
I.Q. X Cond.	35.1811	2	17.591	0.521	0.596
Cond. X S-Within	2058.375	61	33.744		

Fig. 2.1

Effect Of I.Q. Test Score Group & Age Group On Recall.

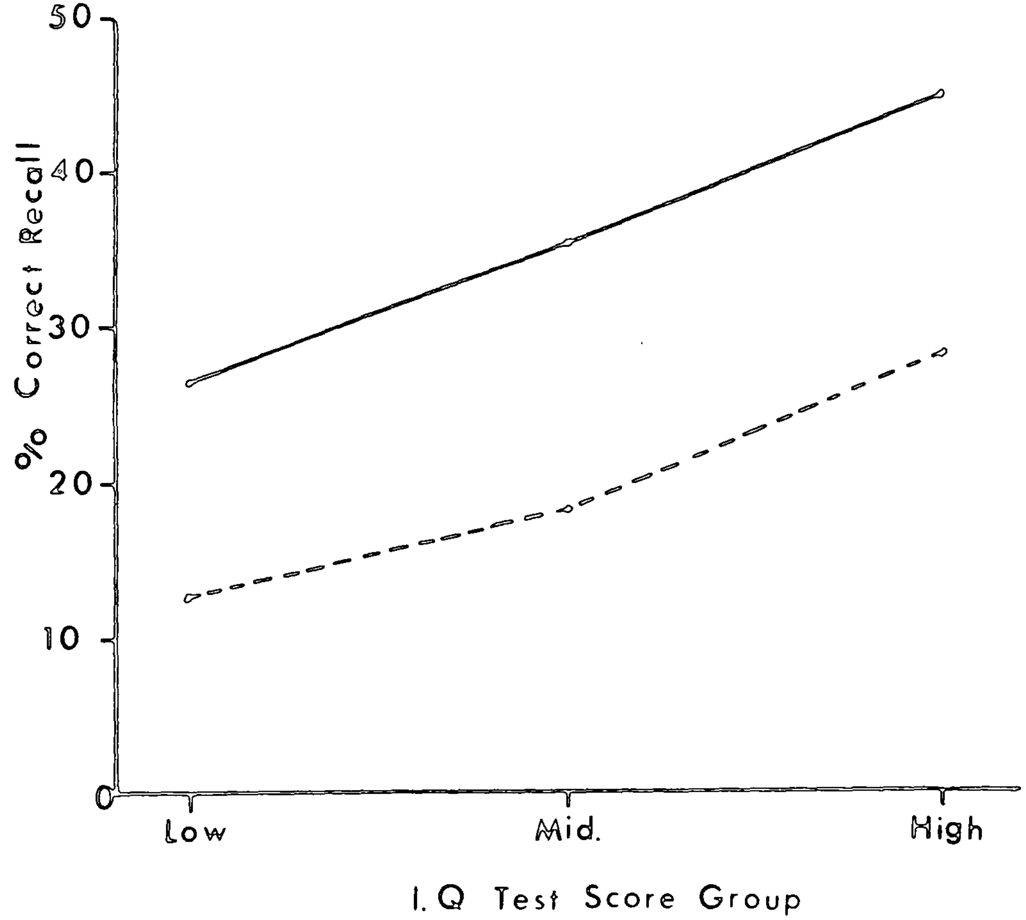
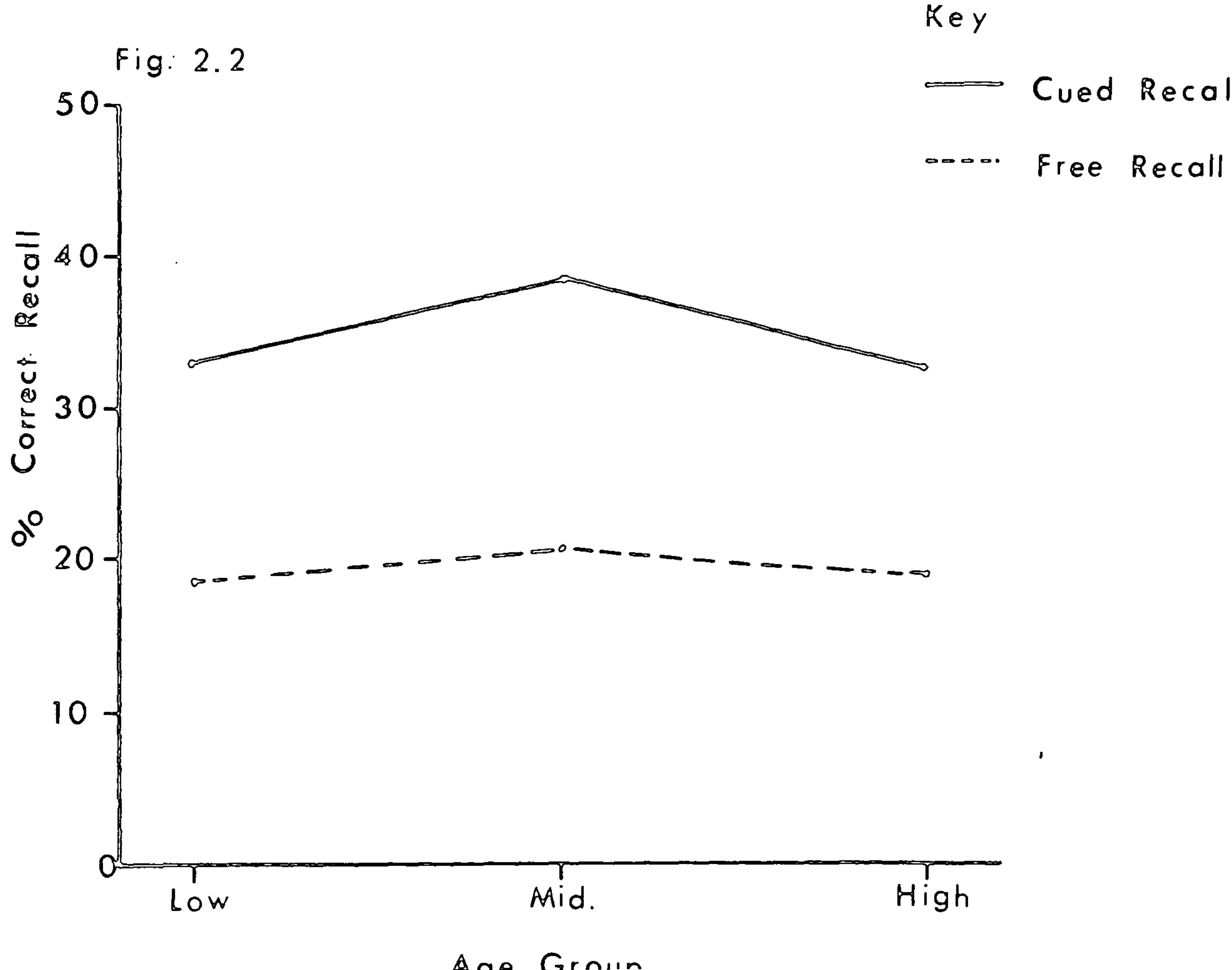


Fig: 2.2



significant difference between I.Q. test score groups (F=6.875 df=2,61 P<0.0035) in the random question condition but not between age bands. No group differences were found in the serial questioning condition.

A serial position curve was plotted for both the free recall and cued recall data (Figure 2.3), and for both question sequences in the cued recall condition (Figure 2.4). Both recency and primacy effects can be seen. No effect of episode duration was found which was disappointing, the correlation between episode time and percentage correct recall was insignificant. For longer episodes in which more information is given there could be two possible effects: first, these episodes may contain more irrelevant details which only confuse the elderly and disguise the relevant information: second, the longer episodes may contain more contextual information which aids comprehension and reinforces the important points in the story. The lack of duration effects indicates that other factors such as visual presentation modes and the subjects interest in the news items may be over-riding this effect.

2.2.7 Discussion

Both the correlation matrix and the multiple regression analysis strongly indicate that I.Q. test score is the best predictor of cognitive performance on this everyday memory task: this agrees with Rabbitt's (1982,1984a) findings on laboratory memory tasks and reaction time data. The correlations obtained between performance and I.Q. test score are exceptionally high compared with those obtained by Rabbitt (1984b) using

Fig. 2.3

Serial Position Curves For Free & Cued Recall.

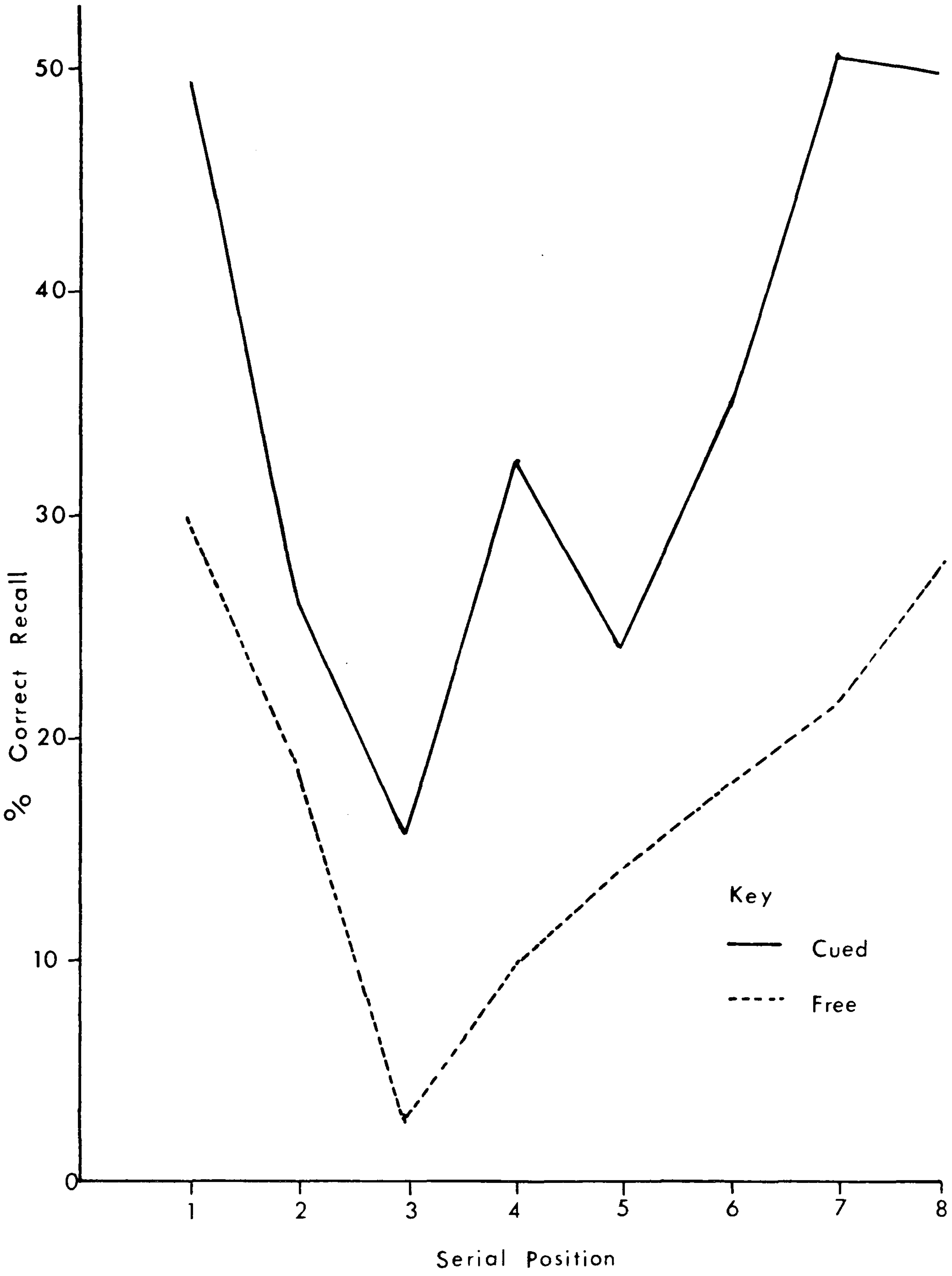
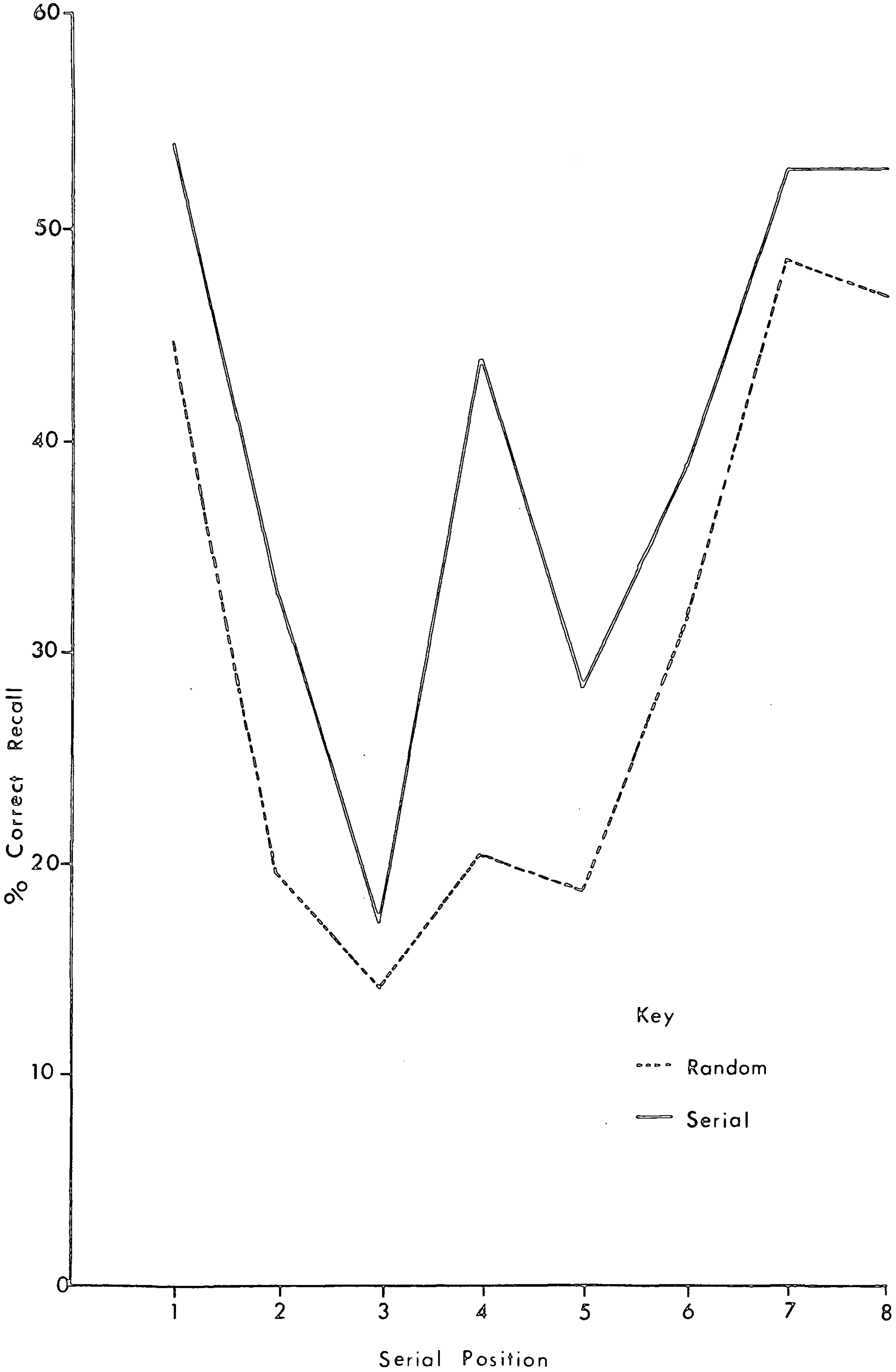


Fig. 2.4

Serial Position Curves For Random & Serial Questioning.



standard laboratory memory tests. This is very encouraging, since it indicates that this particular test is likely to be very sensitive to changes in cognitive performance with increasing age. Further, this indicates that differences due to I.Q. test score are not flattened by overall low motivation as perhaps they are in laboratory tests, or obscured by uncharacteristic poor performance due to test novelty as may occur in unfamilar laboratory tasks. Task familarity seems to be an important contributing factor to the high correlation indices between recall performance and I.Q. test score. The task is complex and requires the elderly to use many different cognitive processes; likewise the AH4 test of intelligence taps many different cognitive skills such as manipulation in working memory, identifying similarities between figures and numbers, and selecting the correct response from distractors.

The emergence of Memory Questionnaire scores as predictors of performance is encouraging but needs to be tested further given the history of failure to obtain significant correlations between the ratings from the questionnaire and actual performance. The questionnaire was in a modified form, asking volunteers to rate on each question the extent of change they perceived in their efficiency at everyday cognitive tasks such as remembering names, faces, telephone numbers etc. The reason for this estimate of perceived change followed the suggestion by Rabbitt (1982) that volunteers probably can not make absolute judgements. The basis for relative judgements are often outside the experimenters control e.g. colleagues at work, friends, spouse etc., by asking each volunteer to estimate their perceived change since the age of thirty, each will be using the same basis for their judgement. This adaptation has been used by Rabbitt (1982), and in some samples it improved correlations between self-ratings and objective scores on laboratory memory tasks. The correlations were most likely increased in this study by the nature of the recall material, and the ecological validity of the test itself. The processes involved in performing this task are used frequently by the elderly, and similar processes are tapped by the questionnaire with its referral to absent-minded slips and recall of information necessary in everyday settings. It is also interesting that Harris, Sunderland and Baddeley (1983) obtained their best predictions of memory performance with self reports when subjects were asked to recall a prose passage. The parallels between recall of complex text and news stories are evident in that both require subjects to follow and construct a story line, using inference construction and comprehension.

Only tentative conclusions can be drawn from the group differences analysis, since the numbers in the groups were so uneven. As expected all the groups improved in the cued recall condition. The difference between groups was not lower on the cued recall than on the uncued recall, as one would expect if the notion of a retrieval deficit were correct (Schonfield 1967). On Schonfield's hypothesis groups who were relatively poor at free recall should have been helped more by the cues provided, than groups who were relatively competent. In this experiment no such age effects were found, and it is perhaps more useful to interpret such findings with reference to models of intelligence.

The group differences in recall with different question orders were again limited to I.Q. test score group and also to the random question condition. The high I.Q. test scorers answered more questions correctly. This points to their superior organization and flexibility of memory; they are able to quickly and successfully retrieve successions of unrelated facts. It also indicates that the volunteers with a high I.Q. test score have a wider range of alternative context from which they can retrieve information or they are able to generate effective contextual cues for themselves. In contrast, the low I.Q. test scorers seem to be seriously handicapped by random questionning. They have a much narrower range of context, and must be helped by linear questionning, indicating that they rely on more primitive and simple means of cueing. This is emphasised by the fact that the differences between groups seen on random questionning disappears when serial questionning is used.

The overall effects were as predicted, with the difference in recall conditions shown clearly in the data. The serial position curve shows both recency and primacy in both recall conditions. These were also present for both sequential and random questioning in cued recall. However, these effects need be tested further using random sequences for presentation, since the structure of the broadcast imposed by producers may well be responsible for the effect For example, high impact stories are always shown first, and this will undoubtedly exaggerate any possible primacy effect, one would expect the recency effect to come through in an immediate recall paradigm, such as that used here. There was no effect of episode duration, but again this may have been lost by the imposed structure of the broadcast.

The most striking feature of this pilot study is, without doubt, the absences of any effects of chronological age per se. In the study of ageing, it appears that it is difficult to separate the effects of increasing chronological age from those of reductions in I.Q. test scores which may accompany the ageing process. It means, rather, that as far as these data show, cognitive changes accompanying old age are better captured by a simple rather crude I.Q. test score than by any other performance index. In order to understand 'age' changes it is useful to try to unpack the nature of the I.Q. tests and to look to models for 'intelligence' and for the articulation of individual differences in 'intelligence' in terms of cognitive function than to seek for models for of cognitive 'ageing'. Accordingly in all subsequent studies subjects were selected on the basis of their age and AH4 test score to ensure that within each I.Q. test score band there was an even distribution of age. This allowed group differences in I.Q. test score and age to be jointly examined in the same design.

2.3 EXPERIMENT 2.2

2.3.1 Introduction

In the second study, subjects were selected according to their age and score on AH4 part 1. As far as possible, the selection sought to produce three I.Q. test score bands of high, middle and low rating and, a rectangular age distribution within each of these groups. Thus age could also be banded into high, middle and low groupings. These groupings allowed age and I.Q. differences to be examined and the information processing skills which distinguished the high I.Q. scores from the lower ones to be investigated.

In accounting for group differences, various possibilities require consideration. First, there may be differences in the general processing efficiency of the organism. Some individuals may simply process information more rapidly and efficiently than others. In the case of the memory system, this would mean a more rigid and efficient means of filtering, encoding, storing, and retrieving information i.e. both control and structural features might be superior in the high I.Q. test score groups. Second, the difference may lie in the relative efficiency of higher order strategies which high and low I.Q. individuals acquire and use to encode, store and retrieve information, solve problems and manipulate information held in memory. Third, there may be a difference in superordinate control. This can be interpreted as control being exercised with the advantage of a more sophisticated database, in which associations are more elaborate and a much richer knowledge base has been built up. The models can not be considered in isolation, but by looking to an interactive and interdependent model the task of tracing the processes which contribute to individual or group differences becomes increasingly difficult. Allowing for this, it is still important to recognise that the efficiency of lower order processes could well determine the efficiency of higher order processes. The converse might also be true e.g. the efficiency ofthe elementary processes of the system could well be affected by the efficiency of the control processor, since possession of a well organised database can easily be seen to enhance the speed and accuracy of elementary access and retrieval processes.

Three aspects of the second study were considered in the light of these models. First, following the finding in the pilot study that the group differences in recall performance were not reduced in the cued recall condition, if the general efficiency of the system is better for high I.Q. test scorers, they should perform well on both a recognition/cued recall test and a free recall test. However, if differences between high and low test scorers lie in the relative complexity and efficiency of the higher order strategies they have developed and use, then the performance of the high test scorers will be enhanced in free recall but not significantly in recognition/cued recall, where cues are provided and successful performance may depend less on the availability of acquired retrieval strategies.

The second point to consider was whether there was any qualitative difference in the nature of the information recalled by high and low test scorers. If one of the factors responsible for differences between high and low test score groups is the relative degree of the organisation of the databases they possess, we might expect that the high test scorers would recognise and use the themes of the news stories to organise their recall. In contrast, the low I.Q. test scorers might be incapable of such organisation. Thus high I.Q. test scorers would preferentially recall thematically relevant events, while low test scorers would be more random in this respect, recalling thematically related and non-thematically related events with equal probability.

A third point was to develop the analysis of some internal features of the task. A serial position curve had been obtained in the pilot study but this needed more rigorous testing, since the structural order imposed by the producer and the overall structure of the broadcast with respect to the position of salient items might have artifically produced this effect. Tannenbaum (1954) found the usual serial learning curve in the recall of material in a radio newscast with both primacy and recency effects. But, Gunter (1979) compared serial position effects for recall of material from a television broadcast when both sound and vision were available, and when sound alone was presented. Serial position effects were much weaker in the former than in the latter case. This may imply that enhanced recall of pictorial material can substantially affect the normally powerful serial position effect for auditorially presented material. Accordingly we generated four different random sequences to investigate the effects of serial position over a ten minute broadcast. In addition we attempted to measure motivational variables asking each subject to rate his or her interest in each news episode shown. Neuman (1976) found that viewers who said that they watched television to be entertained recalled significantly less than those who watched to get information. One would intuitively predict that viewers learn and retain most information about themes in which they are interested as found by Funkhoser and McCombs (1971). However, Genova and

Greenberg (1979) have found that the educational levels and the relative interest expressed on topics presented, only gave a small and inconsistent prediction of the individuals' tested learning of news information in general. We hoped that by asking subjects to assess their interest before and after viewing the items broadcast, we might find a clearer relationship between expressed interest and recall.

2.3.2 Subjects

Ninety four volunteers aged between 51 and 84 years were selected to fill a 3X3 matrix providing approximately equal groups in each of three age and three I.Q. test score bands. Details of this matrix are shown in Table 2.7. All of the subjects had previously received the Broadbent Cognitive Failure Questionnaire, and the amended form of the Harris/Sunderland Memory Questionnaire, referred to in Experiment 2.1, the Mill Hill Vocabulary Scale parts one and two, and the AH4 test of intelligence parts one and two. All subjects were paid travel expenses of 2 pounds for their participation in this study.

2.3.3 Apparatus

The news broadcasts were recorded on a National Panasonic Cassette Video Recorder (NV366,VHS), and they were shown through a Sony Trinitron Monitor (CV,2000 PSB) with a 21 inch screen.

Table 2.7

3x3 Matrix Of Subjects For Age and AH4 Test Score Groups

AH4 Test Score Group Age Group	Low Group 1	Middle Group 2	High Group 3	Total Mean (x) Std. Dev. Range (r)
Young Group 1	9	10	10	N=29 x=54.04 std.dev.=2.44 r=51-58 years
Middle Group 2	7	13	12	N=32 x=63.02 std.dev.=2.66 r=60-69 years
Old Group 3	11	11	11	N=33 x=71.08 std.dev=7.66 r=70-84 years
Total Mean (x) Std. Dev. Range	N=27 x=15.59 std.dev.=4.76 r=5-19/65	N=34 x=30.85 std.dev.=6.15 r=20-38/65	N=33 x=46.06 std.dev.=4.83 r=40-61/65	N=94 x=31.81 63.505 std.dev.=13.2 7.6 r=2-61/65 51-84yrs

2.3.4 Materials

Two early evening news broadcasts were recorded with two weeks intervening,and approximately five months prior to the first subjects being tested. The first broadcast contained nine news items and lasted for 8 minutes 45 seconds, the second broadcast contained ten items and lasted for 9 minutes 12 seconds. Both were edited to provide three random sequences of each, and these sequences with the original order of presentation gave four presentation sequences for each broadcast.

2.3.5 Procedure

Subjects were tested in groups of approximately twelve. Allocation to a particular group merely depended on when it was convenient for a subject to attend an experimental session. Before viewing the first broadcast, subjects were required to rate their interest in forty news items on a 1-5 scale (1= no interest, 5= very interested). Within this list the nineteen items the subjects were about to be shown were embedded. Items were listed in a 'headline' form, giving no detail but indicating the theme of the news item in an obvious manner e.g. i) the open golf championship, ii) M.P.s receive a pay rise, iii) the inflation rate. Once interest ratings were completed, the subjects were shown the first broadcast: the presentation order of the broadcasts was counter-balanced across testing groups so that each broadcast was shown in each temporal position an equal number of times.

Immediately after the first broadcast had been shown, subjects were given one of two recall tasks. This was either i) a recognition test in which subjects were required to answer a question of fact or detail by underlining the correct answer from a set of five alternatives, or ii) a task in which subjects were required to answer a question of detail referring to each news item and then recall the item it referred to in as much detail as possible. Since there were three random sequences for each broadcast for the majority of subjects the questions were not asked in the order of presentation. No time limit was imposed for the completion of recall but the majority of subjects appear to exhaust their search within ten minutes. After the first recall task subjects were given a break of fifteen minutes in which tea and biscuits were served. The presentation of the second broadcast followed the break, and recall for this was tested by the remaining recall task, which was not used following the first broadcast. The position of the two news broadcasts in the testing session was counter-balanced across groups as were the recall tasks concerning each braodcast.

Finally, subjects rated their interest in the news items they had actually been shown on the same 1-5 scale as they had used for the previous ratings. The entire experimental session lasted for approximately 75 minutes. The duration of the episodes and overall length of the broadcast were measured using a stop watch. Examples of probe questions and multiple choice questions can be found in Appendix 1.

2.3.6 Scoring

2.3.7 Free Recall

Free recall data following the probe question were scored into three categories of detail similar to those used by Collins (1979). Nine undergraduates viewed each broadcast, and afterwards they were given a transcript. They were instructed to place each phrase/sentence into one of the three categories listed below.

A: A statement central to the understanding of the news item it refers to, that must be included in any retelling of the news item, (e.g. the main points of the story or the setting of the story).

B: A statement of detail relevant to the news item, which enhances the understanding of the news item but may be omitted without distracting from the basic understanding (e.g. subplots).

C: A statement of detail irrelevant to the news item; this neither enhances understanding nor detracts from understanding, regardless of whether it is included.

Agreement at the 70% level was obtained for each statement.

Each detail recalled was scored as 1 if it was correct, and coded into the relevant category. Percentage correct recall for each category was calculated, for each story, and for the entire news broadcast. An overall percentage correct recall of details regardless of detail category was also obtained.

Probe question answers were scored as 1 for a correct response, and when a response could be considered as only partly correct as ocurred in a few cases a score of 1/2 was given. Incorrect details were scored 1 for each to give an error score.

2.3.8 Multiple Choice

Scoring was simply 1 for a correct response, and 0 for an incorrect response. In all cases, percentage scores were computed because news items varied with respect to number of details and duration so that absolute scores would not have been informative.

2.3.9 Results

Overall percentages of content details given in free recall after the probe question, and percentages of correct response to the multiple choice questions, were computed for each subject. Mean overall scores for the 95 subjects were: 15.76 std. dev. 10.19, and 52.95 std. dev. 15.70 for free recall and multiple choice respectively, $P < 0.01$. This highly significant difference in percentage correct recall under the two conditions was expected, given the numerous studies reporting the difference between cued and free recall (Drachmann and Leavitt, 1972).

A correlation matrix was computed to assess the relationship between the indices measured on the subjects and their recall of the news broadcast in both conditions (see Table 2.8). From the matrix, it is

Table 2.8

Correlation Matrix

	F.R.	C.R.	M.Q.	C.F.Q.	AH41	AH42	M.H.1	M.H.2	AGE
Free Recall (F.R.)	1.00	0.60	0.17	0.04	0.64	0.56	0.48	0.66	-0.02
Cued Recall (C.R.)		1.00	0.14	-0.01	0.62	0.55	0.43	0.55	-0.17
Memory Questionnaire (M.Q.)			1.00	0.56	0.26	0.13	0.34	0.45	0.07
Cognitive Failure Questionnaire (C.F.Q.)				1.00	-0.06	-0.07	-0.06	0.00	-0.09
AH4 (part I)					1.00	0.82	0.67	0.67	-0.12
AH4 (part II)						1.00	0.46	0.47	-0.26
Mill Hill (part I) (M.H. 1)							1.00	0.77	0.32
Mill Hill (part II) (M.H. 2)								1.00	0.28
Age									1.00

clear that both the AH4 part 1 test score and Mill Hill part 2 test score correlate highly with recall under both conditions (r=0.64, 0.62, 0.66, 0.55 for free recall and multiple choice, AH4 and Mill Hill respectively). AH4 part one test score and Mill Hill parts one and two also correlate highly with each other (r=0.67, 0.67), as does recall performance on the two recall tasks (r=0.60). It is important to note the low correlations of age with recall, but the correlations are negative as expected.

A multi-linear regression analysis was carried out omitting Mill Hill scores and AH4 part 2, due to their high correlations with AH4 part 1. AH4 part 1, age, Memory Questionnaire and Cognitive Failure Questionnaire were included as the predictors. These accounted for 45% and 43% of the variance for multiple choice and free recall conditions respectively (R SQ=0.453, 0.430). AH4 part 1 was the best predictor of performance with partialled correlations 0.57, 0.58 : see Tables 2.9, 2.10.

Both self-rating memory questionnaires correlated poorly with performance in either condition (r=0.017, 0.14, 0.04, -0.01) for free recall and cued recall and Memory Questionnaire and Cognitive Failure Questionnaire respectively but they correlate highly with each other (r=0.56), and significantly with AH4 test score and Mill Hill test score (r=0.26, 0.34).

Recall condition was analysed further with respect to I.Q. test score (AH4 part 1) and age. Subjects were grouped into three I.Q. test score

Table 2.9

Multiple Regression Analysis For Percentage Cued Recall.

```
Analysis of Variance        N=67 out of 95

  SOURCE            DF        SUM SQRS    MEAN SQR    F-STAT    SIGNIF

  REGRESSION         4        0.45301     0.11325     12.837    0.0000
  ERROR             62        0.54699     0.00882
  TOTAL             66        1.00000

  MULT R = 0.67306            R-SQR = 0.45301         SE = 0.0939


  VARIABLE            PARTIAL    BETA WT     STD ERROR    T-STAT     SIGNIF

  AH4 1               0.57694     0.60643     0.10785     5.6230    0.0000
  C.F.Q.             -0.11314    -0.10831     0.12080   -0.89664    0.3734
  M.Q.                0.07228     0.07276     0.12751    0.57064    0.5703
  AGE                -0.23441    -0.19000     0.10007   -1.8986     0.0623
```

Table 2.10

Multiple Regression Analysis For Percentage Correct Free Recall.

Analysis of Variance N=67 out of 95

SOURCE	DF	SUM SQRS	MEAN SQR	F-STAT	SIGNIF
REGRESSION	4	0.43096	0.10774	11.739	0.0000
ERROR	62	0.56904	0.00918		
TOTAL	66	1.00000			

MULT R = 0.65648 R-SQR = 0.43096 SE = 0.0958

VARIABLE	PARTIAL	BETA WT	STD ERROR	T-STAT	SIGNIF
AH4 1	0.58115	0.60643	0.10785	5.6230	0.0000
C.F.Q.	0.01716	0.01665	0.12321	0.13516	0.8929
M.Q.	0.01738	0.01780	0.13005	0.13690	0.8916
AGE	-0.17961	-0.14673	0.10207	1.4376	0.1556

bands and three age bands (see Table 2.7). An analysis of variance was performed on percentage recall scores for each condition. There was a main effect of I.Q. test score group (F=40.59, df=2, 85 P<0.01), and also a significant interaction between I.Q. test score group and recall condition (F=4.077 df=4, 85 P<0.01). The high and middle I.Q. test score groups recall correctly significantly more facts than the low group in the multiple choice condition, although they do not differ significantly from each other. In the free recall condition there was a significant difference between all groups. As expected the effect of recall condition was also significant (F=882.065 df=1, 85 P<0.01). There were no significant differences between age groups (see Table 2.11 for complete summary table and Figure 2.5). This is consistent with the pattern of correlations described above, since it will be recalled that the three age groups were matched for raw AH4 I.Q. test scores.

The free recall data were analysed further with respect to percentage recall of the categorised details; I.Q. test score group and age were the between subject variables. An analysis of variance was computed. Once again there was a main effect of I.Q. test score group (F=29.797, df=2, 85 P<0.01), and detail type (F=5.844, df=2, 170 P<0.01). There was also a significant interaction between recall of detail type and I.Q. test score (F=3.126, df=8, 170 P<0.02). Fig 2.6 shows that the high and middle groups distinguish between detail types. The high I.Q. test score group give a significant difference between type A and type B details (t= 4.067 P<0.0003), but not between types A and C (t= 1.897 P< 0.067), nor types B and C (t=1.367, P<0.181). The middle group show a significant difference

Table 2.11

Analysis Of Variance I.Q. Test Score Group / Age Group X Recall Condition

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	1.179	2	0.590	40.595	0.001
Age	0.075	2	0.037	2.571	0.082
I.Q. X Age	0.056	4	0.014	0.956	0.436
S-Within	1.234	85	0.015		
Cond.	6.243	1	6.243	882.065	0.001
I.Q. X Cond.	0.058	2	0.029	4.077	0.020
Age X Cond.	0.036	2	0.018	2.520	0.086
I.Q. X Age X Cond.	0.035	4	0.009	1.221	0.308
Cond. X S-Within	0.602	85	0.007		

Fig. 2.5

Interaction Between Recall Condition & I.Q. Test Score Group.

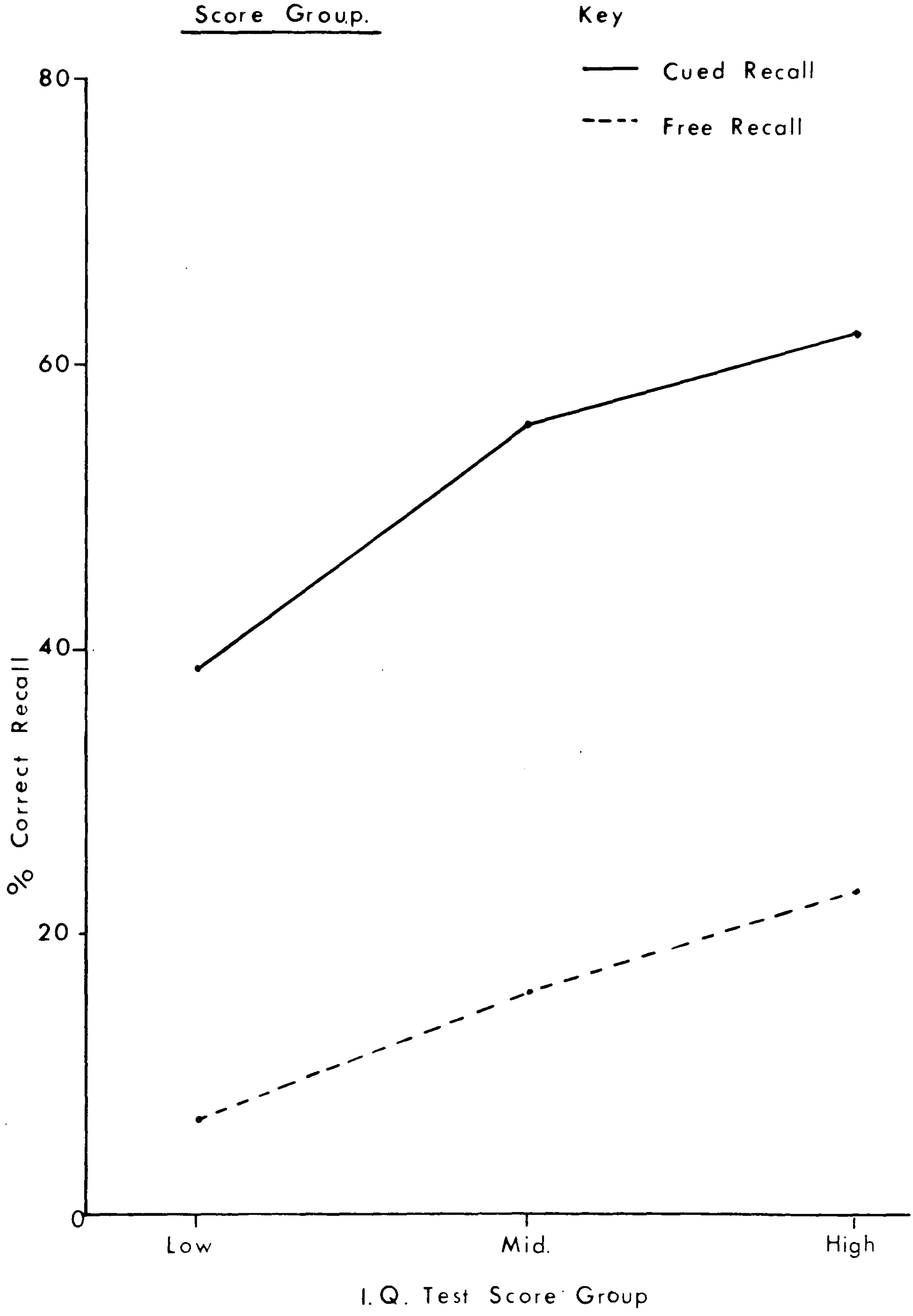
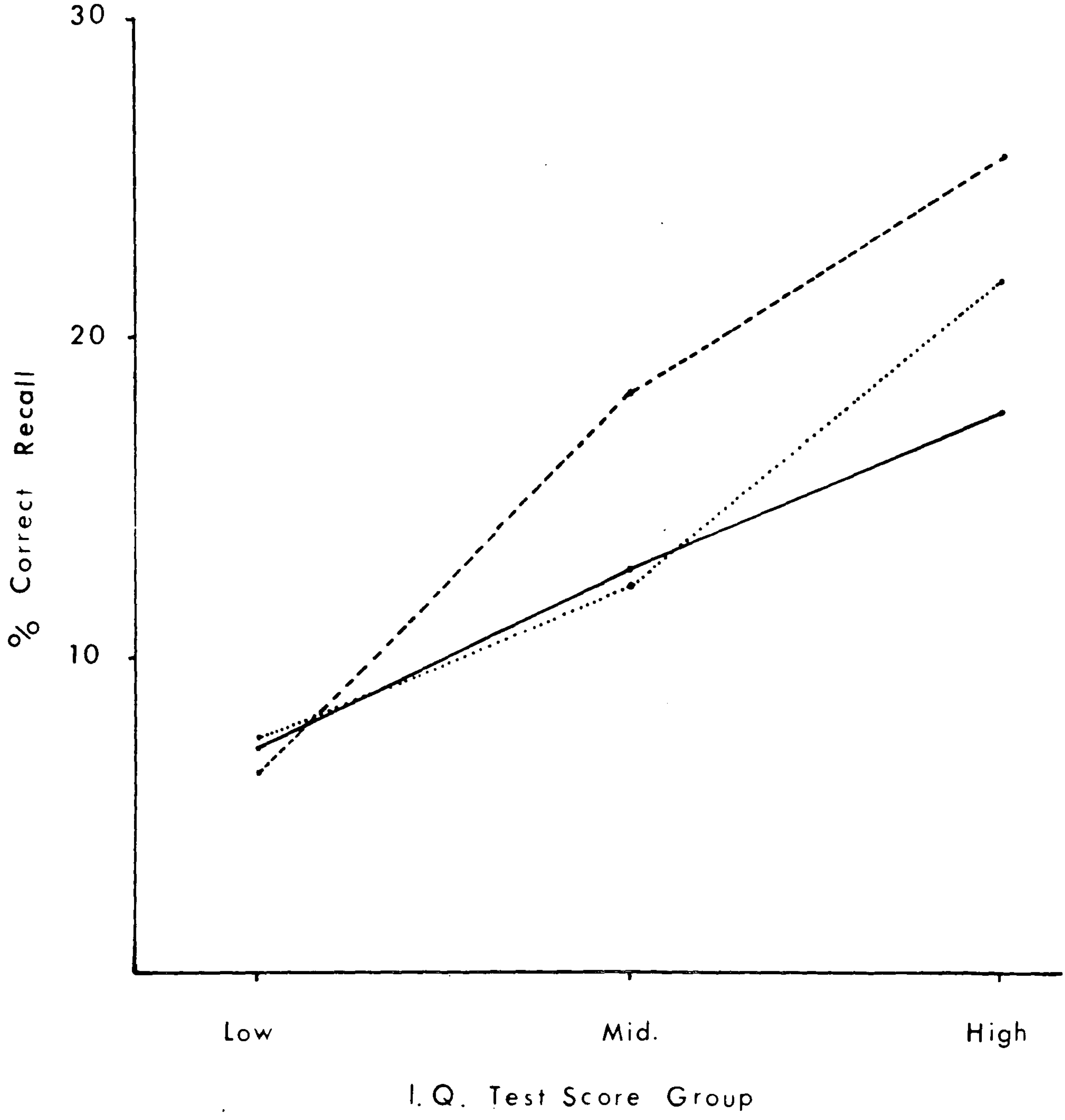


Fig. 2.6

Interaction Between I.Q. Test Score Group & Detail Types Recalled.



between type A and B and type B and C, ($t= 4.00$ $P< 0.01$, $t=2.747$ $P<0.01$). There is a significant difference in percentage recall for all detail types except the middle and high groups on the irrelevant details (type C) (see Table 2.12 and Figure 2.6). From these data it can be seen that the high and middle I.Q. test score groups make better use of the thematic structure of the material in their recall performance.

Figures 2.7 and 2.8 show the serial position curves for the free recall and cued recall of both broadcasts.

The ratings made by subjects of their interest in each news item before and after viewing were correlated with recall for each subject. The correlation coefficients obtained for each subject were then used in an analysis of variance with age and I.Q. test score group as between subject variables; recall and the order in which the interest rating was made i.e. before or after viewing, were within subject variables. The only significant main effect was the position of interest rating ($F=10.722$, $df=1,85$ $P<0.002$). All interactions were insignificant. Interest after viewing the broadcast correlates more highly with recall than interest rated before viewing the news items. However mean correlation coefficients were low ($r=0.124$ after, $r=0.043$ before). The variance of the correlation coefficients was large, and they were not significant. We therefore find no evidence that rated interest affects probability of recall in this experiment.

Table 2.12

Analysis Of Variance I.Q. Test Score Group / Age Group X Detail Type

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Age	0.072	2	0.036	2.178	0.120
I.Q.	0.988	2	0.494	29.797	0.001
Age X I.Q.	0.106	4	0.026	1.596	0.183
S-Within	1.409	85	0.017		
Cond.	0.075	2	0.037	5.844	0.004
Age X Cond.	0.035	4	0.009	1.358	0.251
I.Q. X Cond.	0.080	4	0.020	3.126	0.016
Age X I.Q. X Cond.	0.044	8	0.006	0.860	0.552
Cond. X S-Within	1.088	170	0.006		

Fig. 2.7 Serial Position Curves For Free Recall Task.

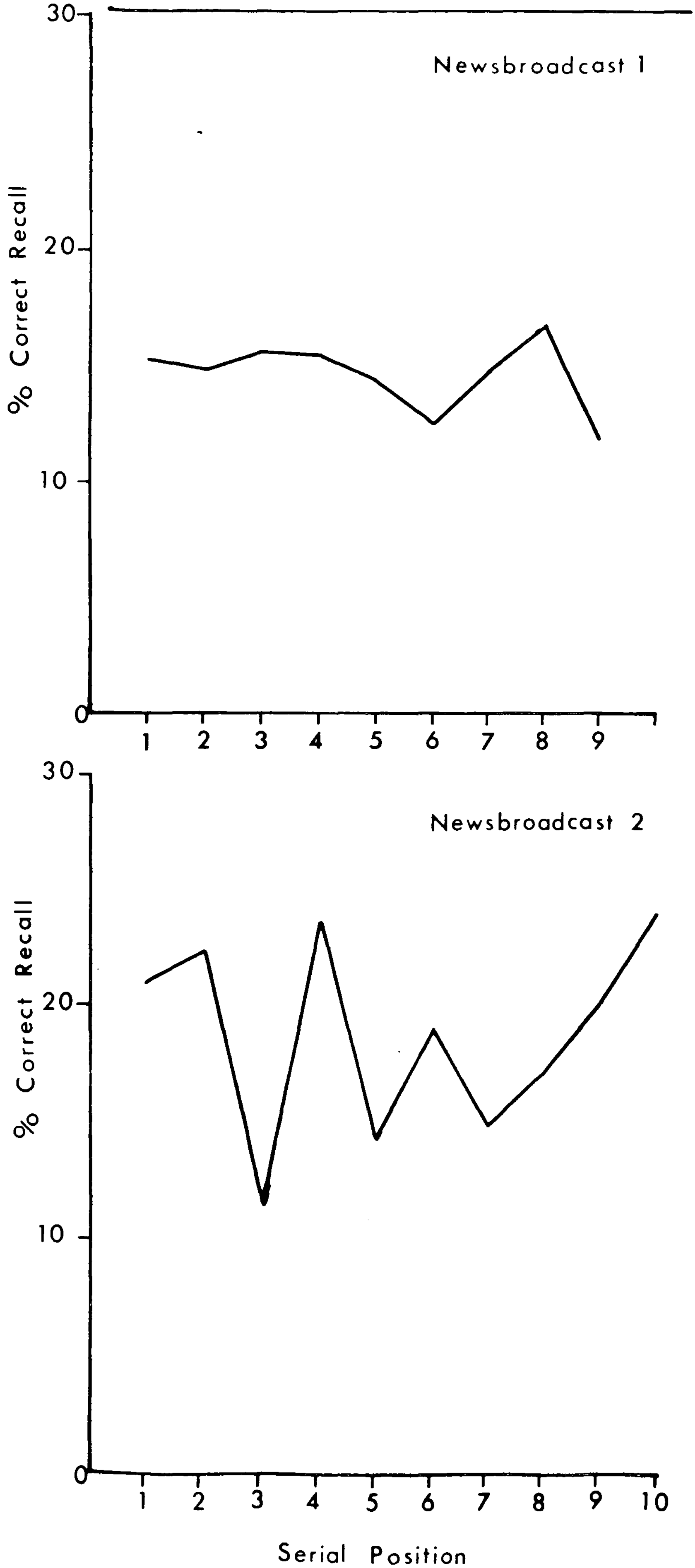
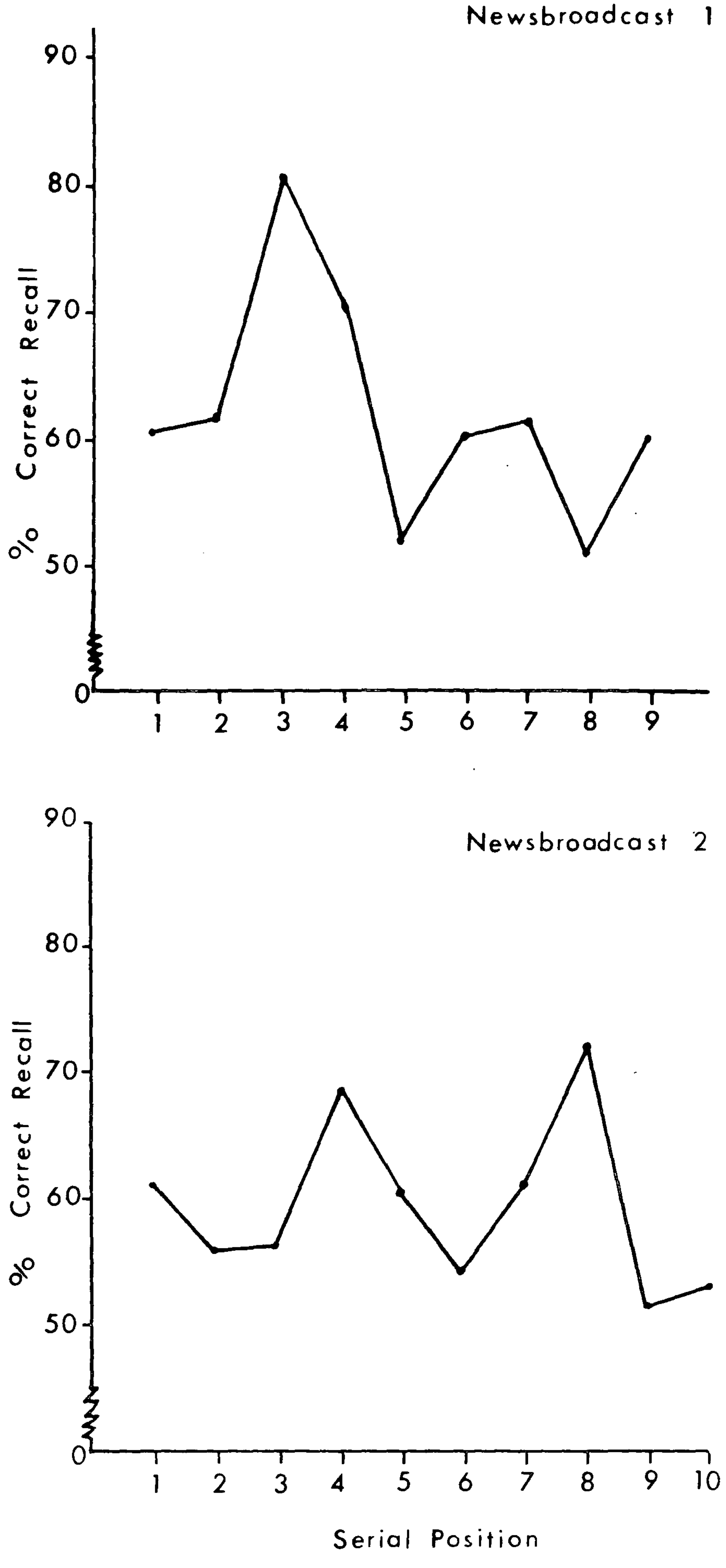


Fig. 2.8

Serial Position Curves For Cued Recall.



Duration effects were sought by taking four pairs of stories from each presentation with a difference in duration of less than 8 seconds. Free recall scores were averaged for the pair, and an analysis of variance computed separately for each broadcast. In both news broadcasts there was a main effect of I.Q. test score group and duration, but these effects were different for each broadcast, see Fig.2.9 and Tables 2.13, 2.14.

2.3.10 Discussion

As in the pilot study the correlation matrix and the regression analysis clearly refute the predicted age effect on recall of this fast, information-rich material. Rather than chronological age, I.Q. test scores and Mill Hill Vocabulary scores are the best predictors of performance on this everyday memory task. They account for as much as 45% of the variance. This finding is again consistent with Rabbitt(1982,1984a).

The implications of the contribution of I.Q. test score to the variance are substantial since once again, we find that variance in cognitive efficiency which may be due to age differences is far better captured by differences in I.Q. test scores than by differences in years lived through since birth. Crudely, if groups of volunteers in their 50s and their 70s are matched for I.Q. test scores, no differences in memory efficiency appear. Thus, instead of studying 'age' effects per se, we are forced to study the effects of I.Q. test score which may diminish

Fig. 2.9

Effect Of Episode Duration On Recall.

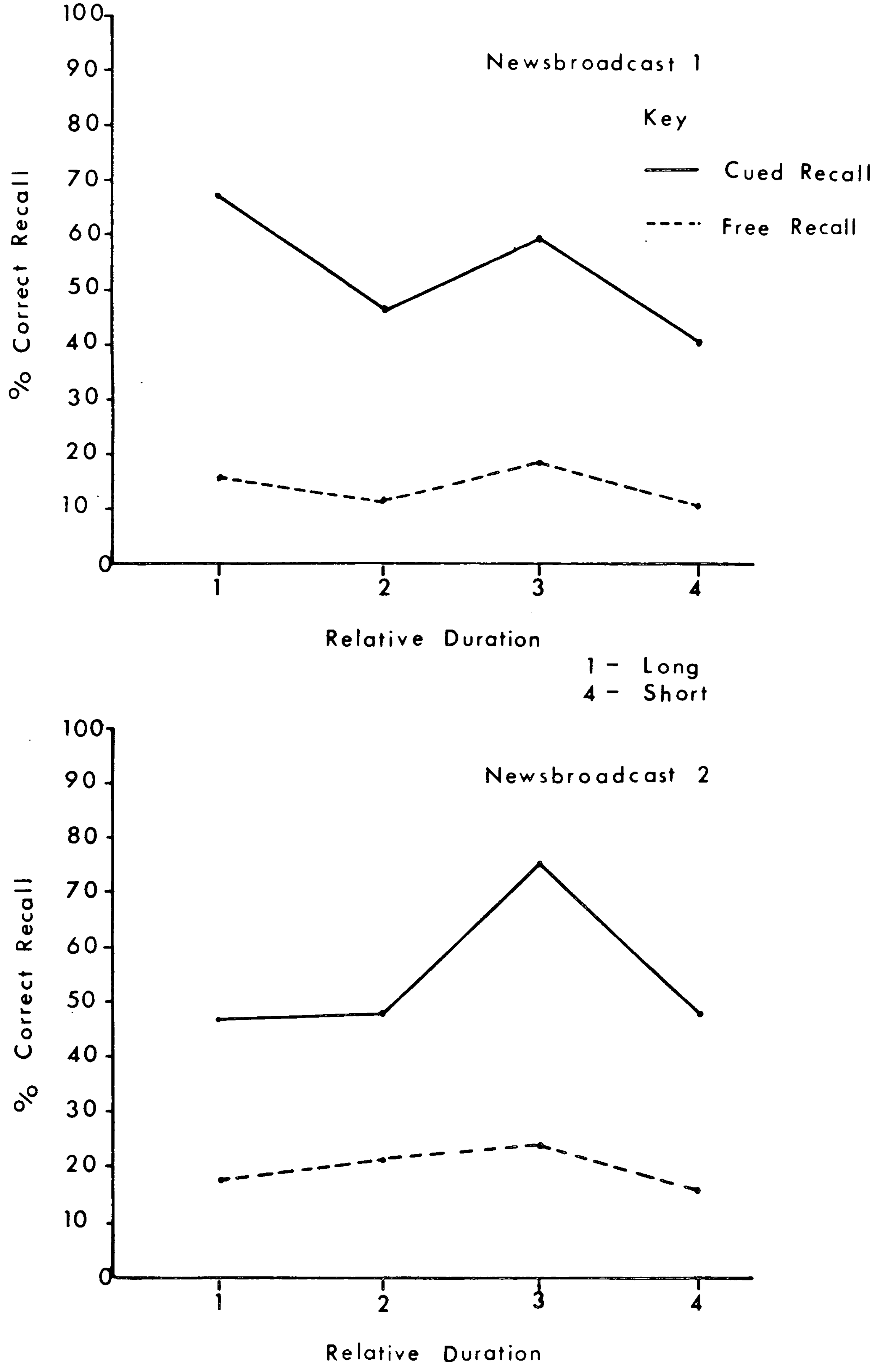


Table 2.13

Experiment 2.2

. . . .

Percentage Correct Free Recall (Newsbroadcast 1)

Analysis Of Variance I.Q. Test Score Group / Age Group X Episode Duration

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Age	0.021	2	0.010	00.418	0.661
I.Q.	0.817	2	0.408	16.360	0.001
Age X I.Q.	0.218	4	0.055	02.186	0.087
S-Within	1.074	43	0.025		
Duration	0.159	3	0.053	09.425	0.001
Age X Duration	0.036	6	0.006	01.055	0.393
I.Q. X Duration	0.047	6	0.008	01.403	0.218
Age X I.Q. X Duration	0.056	12	0.005	00.832	0.617
Duration X S-Within	0.724	129	0.006		

Table 2.14

Experiment 2.2

Percentage Correct Free Recall (Newsbroadcast 2)

Analysis Of Variance I.Q. Test Score Group / Age Group X Episode Duration

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Age	0.070	2	0.035	00.948	0.398
I.Q.	1.017	2	0.508	13.686	0.001
Age X I.Q.	0.061	4	0.015	00.410	0.800
S-Within	1.226	33	0.037		
Duration	0.146	3	0.050	05.314	0.002
Age X Duration	0.054	6	0.009	00.959	0.457
I.Q. X Duration	0.098	6	0.016	01.749	0.117
Age X I.Q. X Duration	0.153	12	0.013	01.364	0.196
Duration X S-Within	0.926	99	0.009		

with age. The two are difficult to separate. Negative correlations between age and I.Q. are well documented: Botwinick (1977) provides a thorough review. In this study, no age effect independent of an I.Q. test score effect has been found. Rather, the data show that the best index of the cognitive effects of ageing appears to be I.Q. test score. The high correlations between I.Q. test score and recall performance were replicated and emphasise that this task is sensitive to changes or, at least, to measure individual differences in cognitive performance and efficiency.

The significant correlations between Memory Questionnaire and recall performance were not replicated with this larger and more carefully selected sample: in fact both questionnaires failed as significant predictors of recall performance. This problem is widely acknowledged in the literature (Broadbent, Cooper, FitzGerald & Parkes,1982; Harris, Sunderland & Baddeley,1983; Herrman & Neisser,1978). Although all these authors emphasise failures to predict subjective estimates of memory performance from laboratory memory studies and vice versa, the results from our pilot study encouraged us to pursue the matter further in this experiment. However, the correlations between such indices are obviously elusive and if found, difficult to replicate with new samples. For the moment, we must conclude from this analysis that the most reliable predictor of cognitive performance is I.Q. test score. There are many possible reasons why I.Q. test scores should be good predictors of recall of organised material. Perhaps the most obvious of these is that I.Q. test scores may correlate with capacity and efficiency of sub-processes necessary in comprehension and organised retrieval of complex material, such as the rapidity and efficiency with which inferences can be updated and manipulated in working memory.

For this reason our analysis of group differences focused on the use of cues and the organisation of the material. The evidence from both studies clearly does not support the notion of a retrieval deficit supported by Schonfield and Robertson (1966), who showed that age differences were not found in recognition tests but appeared in free recall tests. Although these interactions between age and the method of test have been replicated by other investigators (Craik 1971; Smith 1975). While age differences have sometimes been found in recognition tests they are typically smaller than those found in recall tasks (Erber 1974).

However, several points concerning the Schonfield and Robertson (1966) paper need to be raised. First, they matched their age groups on age-adjusted I.Q. test scores. This means their effects can be interpretted as I.Q. test score effects rather than age effects, since age-adjustments mean that their older subjects have lower raw I.Q. test scores. Second, in their experiments, recognition memory scores appear to be near ceiling for all age groups.

The data shown here conflict markedly with the position of Schonfield and Robertson. There is an overall I.Q. test score effect, and also an interaction between I.Q. test score and recall condition. All three I.Q.

test score groups show better performance in the recognition condition, but this improvement is greatest for the middle and high scoring groups. If Schonfield and Robertson's hypothesis of a retrieval deficit in age disadvantaged groups were correct we should have found maximum differences between conditions for low I.Q. test scorers: we found the reverse.

These results also provide some suggestion of qualitative differences in the structure of memory in high and low I.Q. test scorers. Free recall can be considered as a synthetic process, by which the subject constructs the to-be-remembered material by the use of internally generated cues and associations. The data imply that having a high I.Q. test score means one is more capable at this reconstruction process. Furthermore, a high I.Q. test score is also associated with better recognition performance. This shows superior use of the cues provided in the more elementary process of matching the presented material with the material held in memory. In terms of the models outlined in the introduction for this thesis, these results can be seen as supporting the notion that group differences occur because of differences in the knowledge structures of the groups. High I.Q. test scores are associated with a very rich knowledge structure or data base, which allows the formation of a larger number of optional associations, and of more elaborate associations, thus allowing the absorbtion of more information which can be more 'deeply' processed in the terminology used by Craik and Lockart (1972). Alternatively, differences in recall associated with I.Q. test score differences could simply be explained by the entire system being more efficient. This would result in more information being encoded, in storage and retrieval operating more efficiently so increasing overall recall performance under both recall conditions. Both of these explanations can be seen as passive in nature, and they invoke characteristics of the system, rather than features of the individuals' method of approach to the task. The analysis of the type of details recalled allows the examination of a more active model.

One possible reason why the high I.Q. test score group perform better on the free recall task is that they may use more effective strategies to encode the material. In this context, this would mean selecting the thematically relevant information for the news stories and structuring additional information around this. Conversely, the low I.Q. test score group were predicted to be more random in their extraction of information, not necessarily identifying it as relevant or irrelevant. The overall I.Q. test score effect for percentage recall of each detail type was as expected, and the interaction between detail type and I.Q. test score group showed that only the low scorers did not show a significant difference between their recall of the three detail types. The middle and high scorers recalled details in the same proportions as the composition of the news broadcast. Both these groups recalled more type B than type A details and the middle, but not the high scoring group recalled more type B than type C details. Note that there was a higher proportion of type B than type A and C details in the compositon of the broadcast. No obvious strategy for selecting any particular detail type is obvious in either group. Nevertheless this suggests that the low I.Q.

scorers were in fact more random in their encoding and recall of details, since they show less apparent differentiation between detail types in their recall than do the other two groups. In contrast the middle and high group were able to recall the information more fully reflecting the composition of the stories in the mean percentage recall of the detail types. However, this in no way helps to eliminate or support any one of the models. The high scorers could have achieved the same result because their memory systems are generally more efficient or because their data bases were richer. However, these results allow us to keep in play an 'active' model for individual differences in these tasks, according to which active selection of thematic material provides sets of cues for later retrieval. A less interesting possibility is that the high and low test scorers were responding in terms of different perceptions of task demands. But since all subjects were well aware that a test of recall would follow presentation of material, this statement reduces to a paraphrase of the 'active' model, i.e. the high test scorers respond to the task demands by recognising and using thematic cues to guide their encoding and retrieval of the broadcasts.

The serial position effect was not replicated in this more rigorous examination, but the relative salience of news stories might well have distorted any peak effects which might have been present. Item length, different visual effects and the dramatic impact of the story are other contributing factors which need to be controlled in future experiments before serial position effects are understood.

It was disappointing that the effects of relative interest were so weak. The interest ratings made after viewing the broadcast were better predictors of performance than those made before. This is understandable since interest is often aroused by the reporter's style of presentation or the accompanying pictures to a seemingly dull story. The effect was most probably lost by the fact that subjects were expecting to be tested on recall of the whole broadcast and hence any effect of interest was lost because they responded loyally to task demands by concentrating equally on all the material presented to them.

Duration effects although present have to be treated with caution since the method of analysis was at best crude with mean recalls for four episodes within a time range being used. Unless we can find more or less homogeneous material with similar content and visual effects, effects of duration are confounded with the effects of other sorts of salience and with differences in styles of presentation of long amd brief episodes.

2.4 GENERAL DISCUSSION

The most interesting point which has been supported by both the pilot study and the more extensive study is the lack of age effects. To some extent this was expected, because although our knowledge of the neurobiology of ageing is at best patchy, paradigms from experimental psychology and psychometric tests alike, show a picture of increasing variability with age. Rabbitt (1981a) reports a test on 1800 people on reaction time and memory tests for cross-sectional decade samples from 50 to 80+ years; he found mean scores declined by 20% and intersubject variance increased by 160% over the same range. The implication is that the range of individual differences in performance measures increases sharply with age and so chronological age alone is likely to be a poor predictor of performance.

In contrast to the lack of age effects, both sets of data are dominated by I.Q. test score effects. This has important implications for future research, since age and I.Q. test scores are correlated and our studies of recall for news broadcasts have failed to separate their effects on performance. But, in practice this means that models of cognitive changes in performance with age become isomorphic with models for individual differences, which explain how differences in I.Q. test scores are articulated in terms of differences in performance at simple tasks.

The main study considered three models of intelligence, by posing the question 'what processes or features of the cognitive system can account for the the high I.Q. test scorers' high performance levels?'. The increase in the number of cues available at recall did not eliminate the group differences found, which suggests that the difference is not only due to encoding and retrieval strategies. The interaction obtained between percentage recall of detail type indicates that the low I.Q. test scorers were more random in their selection and recall, and there were differences in the percentage recall of detail types for the other two groups, but it is difficult to base a theory of strategic differences in performance on this finding alone. Overall the analysis of group differences gave few indications as to which of the three models is most appropiate.

The internal analysis proved disappointing, as in all cases the salience of the material and the use of structural variables introduced by television producers and editors together with the differences in editing and presentation of visual background material employed in different episodes, overwhelmed any possible effects of temporal order of occurrence (serial position) and of any vaild conclusion being draw from the relative temporal duration of individual episodes.

This task was a very complex one. It had the advantages of complexity in that it threw up a number of interesting effects which require further analysis. However, these are probably better studied in isolation, even if this isolation tends to require somewhat artificial laboratory experiments, - since they are very difficult to investigate among the rich interactions present in these experiments.

The experiments which follow attempt, in a variety of ways, to find out how individual differences in intelligence are articulated in terms of cognitive subsystems which must be deployed for everyday memory to be possible. In particular the studies focus on individual differences in the processes employed in the comprehension and memory of everyday events. The recall of television news broadcasts has already shown individual differences in everyday memory, the following experiments attempt to locate the source of these individual differences.

CHAPTER THREE

3 TASK DEMANDS, FREE RECALL AND RECOGNITION

3.1 INTRODUCTION

The television recall studies in Chapter 2 clearly show that it is more useful to consider the effects of ageing in terms of a decline in I.Q. test score than an increase in chronological age. Variance between individuals of different ages is better captured by I.Q. test score differences than by differences in chronological age, whether cognitive changes with age are abrupt or continuous. This has directed the focus of the interpretation of the data towards models of intelligence instead of models of ageing.

The previous chapter raised several issues concerning the structure of memory and the strategies that subjects used to encode and retrieve material. Three models of differences in I.Q. test score were proposed, two passive and one active. However, our analysis of data from this deliberately unstructured and naturalistic experiment could not allow us to decisively eliminate any of these. It was hoped that since recognition depends less on encoding and retrieval strategies, the group differences in this condition would be reduced, supporting the active model. However group differences were still significant in this condition, with the high I.Q. test score group showing greatest improvement. In essence, this would support either of the two passive models, since either increased efficiency in the system, (a 'rich resource' model for intelligence) or a more absorbent, rich data base, (

a 'structural' model for intelligence) could claim to achieve higher levels of performance in both conditions. Nor did analyses of the recall of different types of detail clarify the position. Although there was an interaction between the type of detail and I.Q. test score group, this occurred because both the high and middle test score groups recalled significantly more of all types of detail than the low test score group, and both recalled more of the relevant but not essential details than the other detail types. However, since recall of the essential details did not differ significantly from recall of irrelevant details, the proposed strategy of selecting the essential details from the mass of information does not appear to operate in any of the groups. In fact the recall loosely corresponded to the proportions of the details in the broadcast in the high and middle I.Q. test score groups.

There are difficulties in distinguishing between the three models in terms of cognitive performance, since all three give plausible accounts of individual differences in cognitive competence. One possibilty is to consider the ways in which features of the system such as processing efficiency, capacity, and the use of strategies, act as limiting factors, and explore how as such they can be overcome. Another option is to find ways in which the effect of these features can be prevented. The efficiency of the system can be tested by varying the presentation rates: a fast presentation rate will eliminate the opportunity for the use of strategies such as rote rehearsal, thus only efficiency will be measured. The use of strategies can be tested by comparing performance on tasks with fast presentation rates with performance on a task where there is sufficient time to employ strategies. Alternatively, subjects can be trained to use a particular strategy and the associated increase in performance with the training used as an index of the effect of the implementation of the strategy. More simply, a strategy can be suggested through task instructions and its effect measured by comparison with a control condition.

An association-rich data-base is obviously a feature which will determine the capacity of the system. It may reflect the existence of richer information processing resources (i.e. people with rich resources will tend to develop a more complex data-base) or have developed independently of 'basic' resources as a function of individual interest or opportunities offered by idiosyncratic environments. Relative richness or sparcity of association networks can be objectively and directly determined. However, differences in capacity (or resources) which may provide pre-conditions for the assimulation of information are more difficult to ascertain. Any residual group differences observed even when asymptotic performance is reached must be ascribed to differences in capacity. However, capacity limits can not be overcome by psychological manipulation, and will affect the acquisition of strategies for learning and retrieval.

Craik (1977) provides some evidence for the decline in learning strategies with age. Young adults and people over the age of 65 years were presented with lists of words under four conditions, one involved intentional learning, and the others incidental learning under three different orienting tasks. Age differences were greatest with intentional learning, and free recall as opposed to recognition. The former suggests differences in learning strategies, and the latter, although this result was not replicated in Experiments 2.1 and 2.2, suggests differences in free recall strategies. Retrieval is primarily a reconstruction of encoding (Tulving and Thompson 1975) and therefore deficits attributed to ineffective retrieval may be due to differences in encoding strategies. Recall will obviously be more sensitive to these strategies than recogition. No differences were found in the incidental and recognition tests, adding support to the difference in strategies. Nevertheless, the effect of capacity must also be considered, as there is every reason to expect that it declines with age, and little reason to suppose that forgetting from disuse of general strategies for learning and retrieval can account for differences.

One of the main strategical problems encountered by the elderly discussed in the literature, is the organisation of the material. Smith (1980) argued that the effect of increasing list length on recall could be explained by organisation strategies. Using the 'delayed recall' method developed by Shiffrin (1970), list length was varied by asking subjects to recall not only the list just presented but also the list that occurred immediately prior to this. Hence there was always another to-be-remembered list interpolated between the presentation of any list and its subsequent recall. In this way the affect of a) the length of the list being recalled (retrieval), and b) the length of the list interpolated between presentation and recall (storage) could be investigated. The list lengths were either 10, 20, or 40 items in length. The findings were that effects of list length during the retention were significant, but did not interact with age effects. However, there was a significant interaction between age and the length of the list being recalled, with a greater age effect on shorter lists. Smith interpreted this to show that the elderly are less spontaneous in employing organisational strategies so they were less affected by list length, i.e. increasing list length caused organisational difficulties, only affecting the young who could use organisation. However, Craik (1968) had found the opposite effect, with larger age effects for longer lists: his explanation is that the number of items interpolated between presentation and retrieval causes special difficulties for the old. However, an alternative explanation would be that longer lists are more difficult to organise, and because the elderly do not operate organisational strategies as effectively as the young, age effects are greater with longer lists which depend more on such strategies for effective recall.

Hultsch (1971) used a different method to investigate organisational strategies. Subjects sorted words into two to seven categories, until they had achieved two identical categorisations, in the control condition they just inspected the words. There were no age effects in the free classification of words for the sorting condition, all subjects sorted the words into the same number of categories. However, age effects were found for recall. The largest age differences occurred in the control condition when the words had only been inspected and no organisational strategy such as categorisation had been imposed. This clearly provides evidence that the differences in recall are due to organisational strategies, and that older people benefit relatively more than the young when organisational strategies are imposed upon them, and they do not have to devise them for themselves. However, the data can also be explained in terms of the depth of processing model (Craik and Lockhart, 1972), which was developed after Hultsch's experiment. Craik (1977) cites some unpublished data from White which shows age effects in recall associated with the depth to which words are processed. In the sorting condition, all subjects would have processed the words to a similar level in order to categorise them; but in the control condition the depth to which the words were processed would possibly vary according to the age of the subjects tested. Hultsch (1969) also found a similar effect when specific instructions which helped the material to be better organised were given: these appeared to compensate for the organisation deficit.

Essentially, age differences have been found in the manner in which the information is organised which transforms it in some meaningful way and enters it into the long-term store. The difficulties highlighted are that older adults do not spontaneously use organisational strategies as extensively as younger adults, or that they do not use them as effectively (Canestrari 1968, Eysenck 1974, Hulicka & Grossman 1967, Hultsch 1969,1971,1974). When various organisational strategies are built into the task the performance of the old significantly improves (Canestrari 1968, Hulicka & Grossman 1967, Hultsch 1969,1971).

The studies stress how the decrements can be modified and reduced by the use of instructions and the organization of the material presented for recall, which implies that the difficulties in encoding and retrieval encountered by old people can, to some extent, be compensated for by improved organisation of material presented. If the passive models of intelligence are upheld, a compensation view of ageing can not operate since in these models age decrements are due to absolute limitations imposed by the efficiency of the cognitive system or the richness of the residual knowledge structure which the elderly can use to encode material. Organisational strategies alone, would not be expected to eliminate or greatly reduce these differences.

Two experiments are reported which seek to consider the active model of differences in intelligence by examining the use of strategies. Both are based on a simple memory task with the implementation of strategies being manipulated by instructions. In both experiments it was expected that the effect of instructions would be greatest in the high I.Q. test score group, and if an age effect was found, that it would be greatest in the young group.

3.2 EXPERIMENT 3.1

3.2.1 Introduction

Rabbitt (1965a), investigating age differences in the ability to ignore irrelevant information, gave old and young subjects a task which involved sorting two packs of cards labelled with the letters 'A' or 'Z' into separate piles. Besides the letter 'A' or 'Z', each card had an additional eight irrelevant letters printed on it, the set of additional letters being different for each pack. Sorting times were measured for each pack and it was found that the older group showed a markedly greater increase in sorting time on transfer to a new set of irrelevant letters. This implies that the elderly improve their performance as a result of learning specific features of the set of irrelevant symbols. Since the old take longer than the young to ignore irrelevant symbols in a display, their dependence on learning irrelevant symbols can not represent an improvement in strategy. Rather, the elderly appear to use information which the young find unnecessary.

The notion of ignoring what is apparently irrelevant is an important strategy in watching and comprehending television, where each scene potentially overflows with information and selection becomes essential. One of many ways of examining the selection strategy of elderly people is to instruct them to select particular stimuli from an array, and then test these in a recall test, as well as testing their recall for the irrelevant stimuli.

Experiment 3.1 was designed using the paradigm outlined above. Subjects were instructed to categorise each of a sequence of words into one of a choice of three categories. This ensured that all words were processed to a similar depth. Craik and Tulving (1975) have shown that a category placement task leads to a better memory trace than a task requiring a decision about some physical feature of the word. The semantic task is assumed to give a more elaborate encoding of the item, which leads to a better and more stable memory trace. Eysenck (1974) found support for this elaboration-deficit hypothesis, in finding that the old do not encode material to the same level of elaboration as the young. Eysenck tested incidental learning using different types of orienting task. Two of these tasks required subjects to process the meaning of words and two did not involve semantic processing. There was a significant Age x Orienting Task interaction. Incidental learning by old subjects relative to younger subjects depends on the depth of processing of stimulus materials required by the orienting task, where this processing is non semantic, old and young subjects show equivalent recall performance. However, when this processing is semantic the old are disadvantaged.

In addition to processing the words to make a semantic categorisation decision, the subjects were instructed to pay particular attention to one of the three categories because they would have to recall it later. Hence the task was structured so as to require the subjects to categorise, but also to selectively attend to one of the three categories more than the other two. Subsequent recall tests examined the extent to which effective strategies to achieve this are related to chronological age and to I.Q. test score.

3.2.2 Method

3.2.3 Subjects

96 subjects were selected according to their age and score on the AH4 test of intelligence part one. The mean age of the subjects was 64.4 years (std. dev. 7.7), and the mean AH4 part one score 30.4 out of 65 (std. dev. 11.9). The subjects were divided into three age groups and three AH4 test score groups, the composition of which is shown in Table 3.1.

3.2.4 Materials

Nine categories were selected from Battig and Montague (1969), all within the potency range 6.82-7.02. The nine categories were grouped to form three groups of three categories: fifteen words were selected from each category to be balanced for phonemic length across their group, and the experiment used the most frequently generated words as far as was possible within this constraint (Appendix 2).

Table 3.1.

3x3 Matrix Of Subjects For Age and AH4 Test Score Groups

AH4 Test Score Group Age Group	Low Group 1	Middle Group 2	High Group 3	Total Mean (x) Std. Dev. Range (r)
Young Group 1	11	9	12	N=32 x=55.97 std.dev.=2.40 r=52-59 years
Middle Group 2	8	17	12	N=37 x=64.62 std.dev.=2.72 r=60-69 years
Old Group 3	9	12	6	N=27 x=74.15 std.dev=4.19 r=70-85 years
Total Mean (x) Std. Dev. Range	N=28 x=15.59 std.dev.=3.76 r=6-20/65	N=38 x=30.85 std.dev.=5.10 r=21-39/65	N=30 x=46.06 std.dev.=4.54 r=40-61/65	N=96 x=30.41 64.61 std.dev.=11.91 7.77 r=6-61/65 52-85 years

3.2.5 Apparatus

Slides were made of the fifteen words selected from each category using Letraset (Futura Medium 1874), and these were projected onto a large white screen using a Kodak Carousel 'S' Projector fitted with a 100 mm lens.

3.2.6 Procedure

The categories were always shown in the same groups to maintain the phonemic balance of the words in each set. The words were randomly sequenced for each group of categories. Subjects were tested under three conditions and performed an intervening task of naming and matching pictures;(this was part of another study between conditions). The intervening task lasted for approximately 15 minutes. The category group occupying the various conditions was randomly allocated for each subject. Most subjects were tested individually, although a few were tested in pairs.

The first trial was the control condition in which the subject was simply instructed to categorise the words shown on the screen, into one or other of the three categories which they had been told the words would belong to. Each category was given an appropriate symbol, in most cases this was the first letter of the category, (e.g. animal 'A' , vehicle 'V' , kitchen utensil 'K') and the subject was instructed to write the category symbol of the word in the space provided on the response sheet. Each slide appeared for approximately three seconds which allowed the subject time to categorise the word and write the category symbol in the space provided on the response sheet. Once the subject had completed the categorization task, he/she was immediately instructed to write down as many of the words as they could recall. They were instructed to refrain from guessing and just generating words associated with the three categories.

Following the intervening task, Condition I was carried out. The subject was once more instructed to categorise the words shown into three new categories. On this occasion they were given the addition instruction to give the words belonging to one of the categories more attention because they would have to recall them later. This emphasised category was randomly allocated to each subject from the three possible categories. After completing the categorization task, the subject was asked to recall as many of the words belonging to the emphasised category as he/she could. Once this recall was exhausted, on a separate recall sheet they were asked to recall as many words from the two remaining categories as possible.

The final trial (Condition II) once again followed an intervening task, and was identical to the second condition except for the recall condition. The subject was instructed to recall the emphasised category after previously recalling the two remaining categories.

3.2.7 Results

First, the categorisation task was analysed. Mean number of words correctly categorised over the three trials were calculated for each subject. These were placed in an [illegible] analysis of variance with age and I.Q. test score group as between factor variables. There was a significant main effect of I.Q. test score group (F=4.201 df=2,87 P<0.02), but no effect of age group, and the interaction term was also insignificant (F=0.934, P<0.729). Table 3.2 shows that overall categorisation performance for all groups was high, however the subjects with high I.Q. test scores made fewer errors than the other subjects.

Overall percentage correct recall was computed for the control condition, the emphasised categories, and the remaining categories in conditions I and II. Table 3.3 gives the overall means for the recall of each, and also measures of skewness and kurtosis for each distribution. Neither differed sufficiently from zero to cast doubt on the normality of the distributions.

An [illegible] analysis of variance was computed on the total percentage recall of all three experimental conditions with I.Q. test score group and age group as between subject factors. Table 3.4 shows a complete summary table. The main effects of I.Q. test score group and recall condition are shown in Fig. 3.1, and there was no significant interaction. There were no age group effects or significant age group interactions. As expected, recall in the control condition was significantly poorer than

Table 3.2.

Descriptive Statistics Of Categorisation Performance For I.Q. Test Score Groups & Age Groups

Percentage correct scores

I.Q. Test Score Group	N	MIN.	MAX.	MEAN	STD. DEV.	SKEW.	KURT.
Low	28	94.81	100.00	99.31	01.56	-2.273	3.622
Middle	38	98.52	100.00	99.79	00.45	-1.930	2.409
High	30	99.26	100.00	99.97	00.13	-5.199	25.034
Age Group							
Low	32	95.55	100.00	99.70	00.85	-3.821	15.258
Middle	37	98.52	100.00	99.86	00.38	-2.679	6.028
High	27	94.81	100.00	99.51	01.39	-2.968	7.289

Table 3.3

Descriptive Statistics For Percentage Correct Recall In Each Condition

Condition	N	MIN.	MAX.	MEAN	STD. DEV.	SKEW.	KURT.
Control	96	2.22	62.22	31.32	11.66	0.286	-0.193
Stated Category Position 1	96	26.66	93.33	58.61	16.53	-0.206	-0.956
Stated Category Position 2	96	13.33	93.33	53.40	16.88	-0.094	-0.452
Remaining Categories Position 1	96	3.33	76.66	30.52	12.47	0.507	0.865
Remaining Categories Position 2	96	6.66	73.33	38.71	15.32	0.117	-0.643

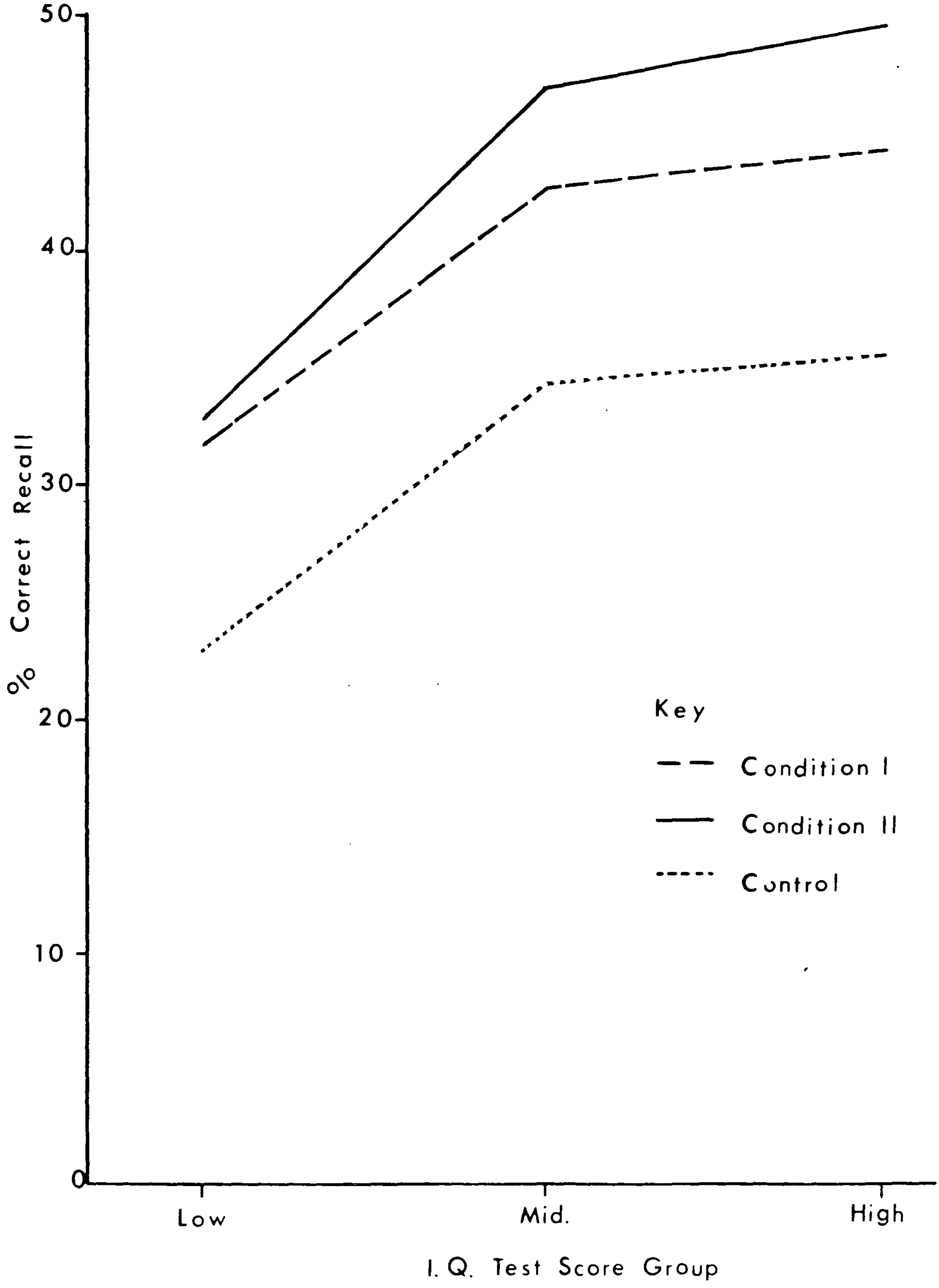
Table 3.4

Analysis Of Variance I.Q. Test Score Group / Age Group X Condition

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	10040.766	2	5020.383	18.903	0.001
Age	90.005	2	45.002	0.169	0.844
I.Q. X Age	571.583	4	142.896	0.538	0.708
S-Within	23105.813	87	265.584		
Cond.	6943.578	2	3471.789	58.468	0.001
I.Q. X Cond.	167.119	4	41.780	0.704	0.590
Age X Cond.	183.973	4	45.993	0.775	0.543
I.Q. X Age X Cond.	132.295	8	16.537	0.278	0.973
Cond. X S-Within	10332.063	174	59.380		

Fig. 3.1

Effect Of Experimental Condition & I.Q. Test Score Group On Recall.



either of the other conditions ($t = -10.312, -8.315$ $P<0.01$ $P<0.01$) and condition II had significantly higher recall than both the control condition and condition I ($t = 3.315, 3.549$ $P<0.01$ $P<0.01$).

The two experimental conditions were analysed separately from the control condition. This was done to examine the effect of the instruction to attempt to remember the emphasised category. In addition, the effect of position of recall was investigated, (i.e. whether the emphasised category was recalled before or after the remaining category was recalled). An analysis of variance was computed using total percentage recall, which included both the number of remaining category words and emphasised category words (see Table 3.5 for summary table). Once again there was a main effect of I.Q. test score group, and both of the within factors show significant effects but do not interact significantly (Fig. 3.2). There is however a significant interaction between order of recall and I.Q. test score group, with the middle and high groups recalling significantly more than the low scorers in both conditions, as well as showing a greater effect of order of recall (Fig. 3.3).

It is interesting that age group interacted significantly with emphasis (Fig. 3.4.), the young age group clearly showing the effect of the instruction to remember the emphasised category: the difference between emphasised and remaining category recall was greatest in their case and least in the oldest group. The three way interaction between age, I.Q. test group and emphasis is more difficult to disentangle

Table 3.5

Analysis Of Variance I.Q. Test Score Group / Age Group X Position Of Recall & Emphasis

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	15287.992	2	7643.996	16.824	0.001
Age	471.612	2	235.806	0.519	0.597
I.Q. X Age	504.243	4	126.061	0.277	0.892
S-Within	39528.000	87	454.345		
Posn.	3671.926	1	3671.926	41.565	0.001
I.Q. X Posn.	836.711	2	418.356	4.736	0.011
Age X Posn.	54.180	2	27.090	0.307	0.737
I.Q. X Age X Posn.	273.978	4	68.495	0.775	0.544
Posn. X S-Within	7685.688	87	88.341		
Emph.	38735.613	1	38735.613	287.837	0.001
I.Q. X Emph.	34.478	2	17.239	0.128	0.880
Age X Emph.	1274.461	2	267.230	4.735	0.011
I.Q. X Age X Emph.	2042.832	4	510.708	3.795	0.007
Emph. X S-Within	11700.000	87	134.575		
Posn. X Emph.	173.622	1	173.622	1.51	0.222
I.Q. X Posn. X Emph.	37.557	2	18.778	0.16	0.849
Age X Posn. X Emph.	220.414	2	110.207	0.96	0.387
I.Q. X Age X Posn. X Emph.	210.563	4	52.641	0.46	0.766
Posn. X Emph. X S-Within	9988.063	87	114.805		

Fig. 3.2

Effects Of i) Position Of Recall Of Emphasised Category & ii) Emphasis On Overall Recall.

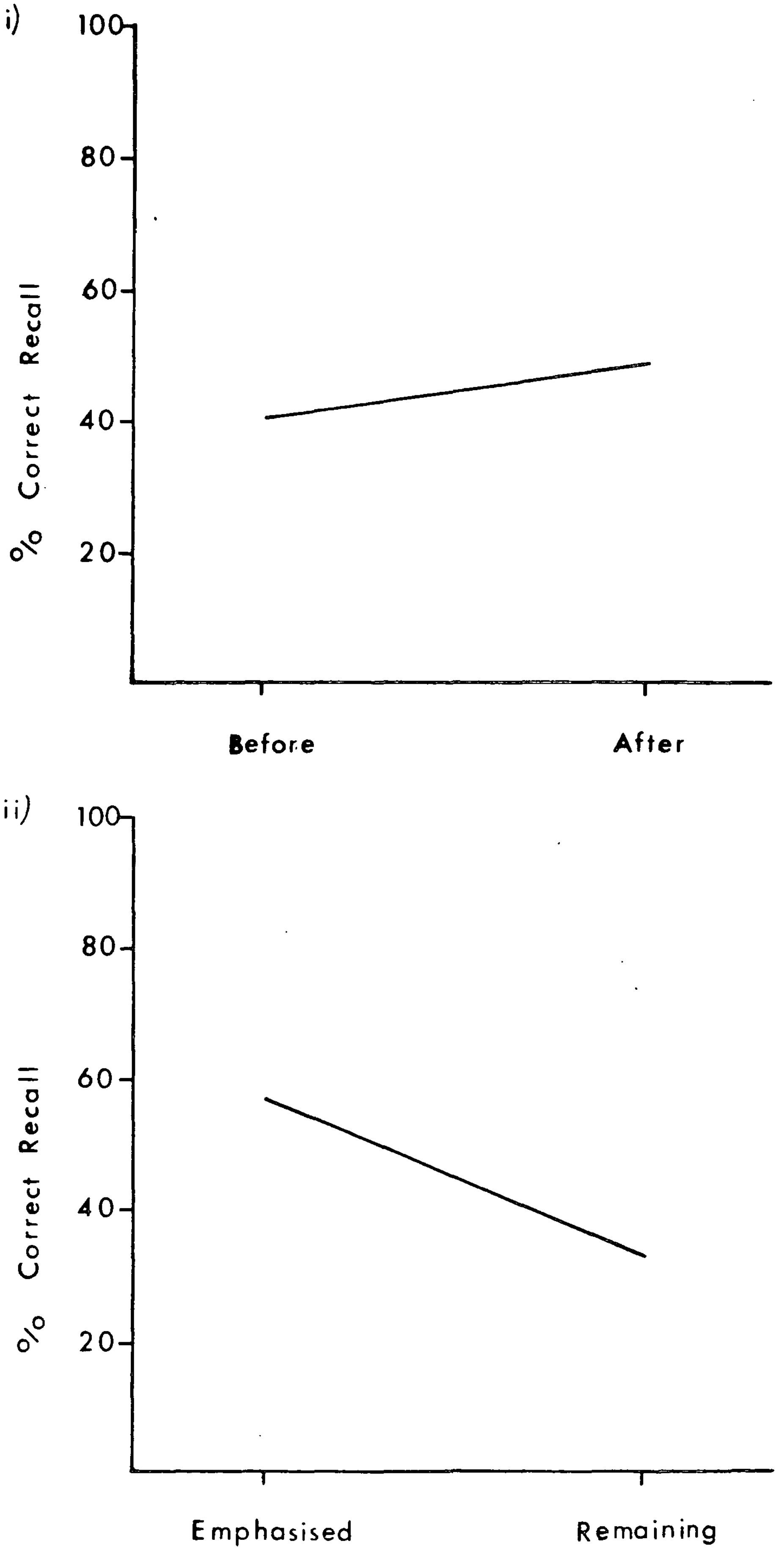


Fig. 3.3

Interaction Between I.Q. Test Score Group & Recall Position Of Emphasised Category.

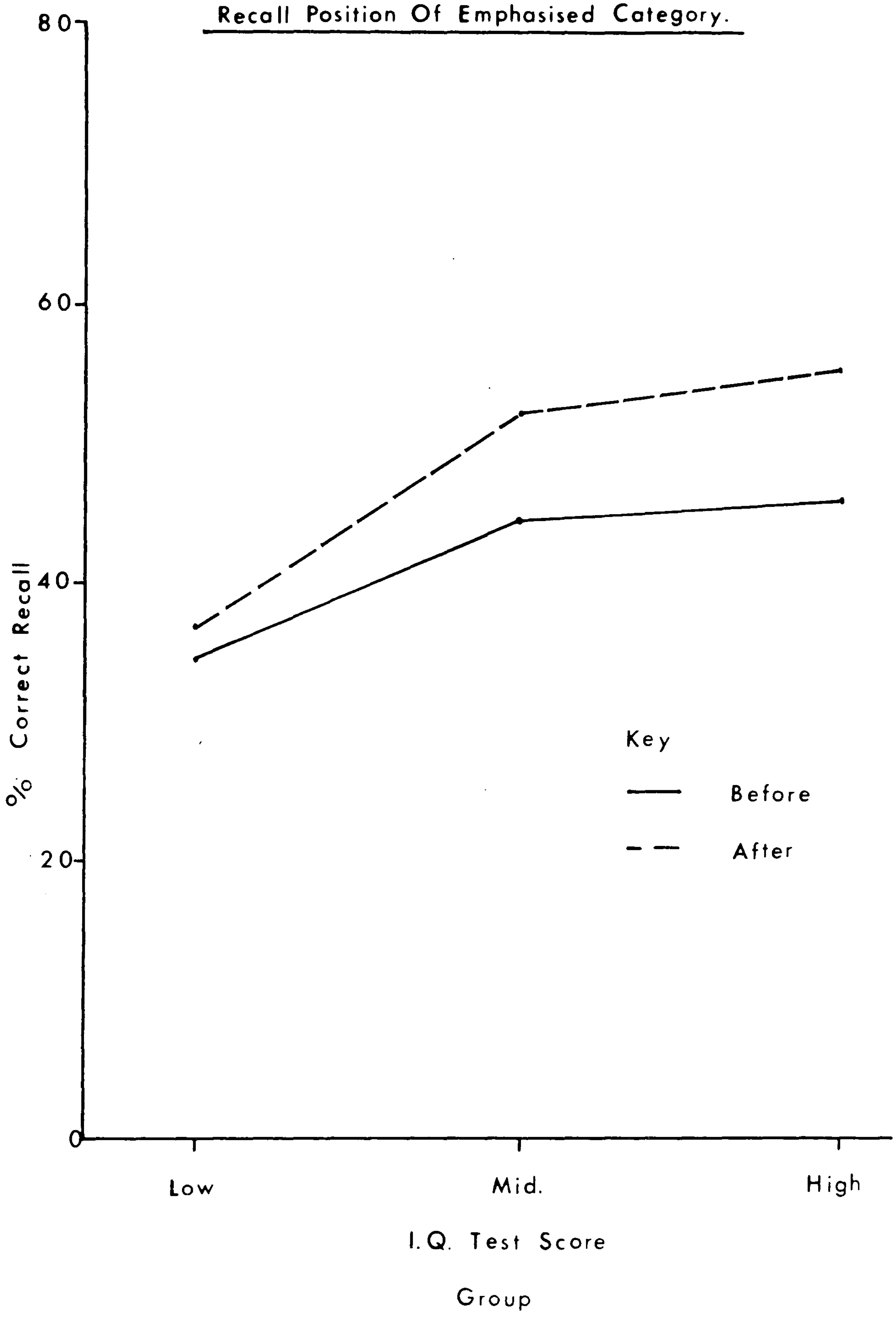
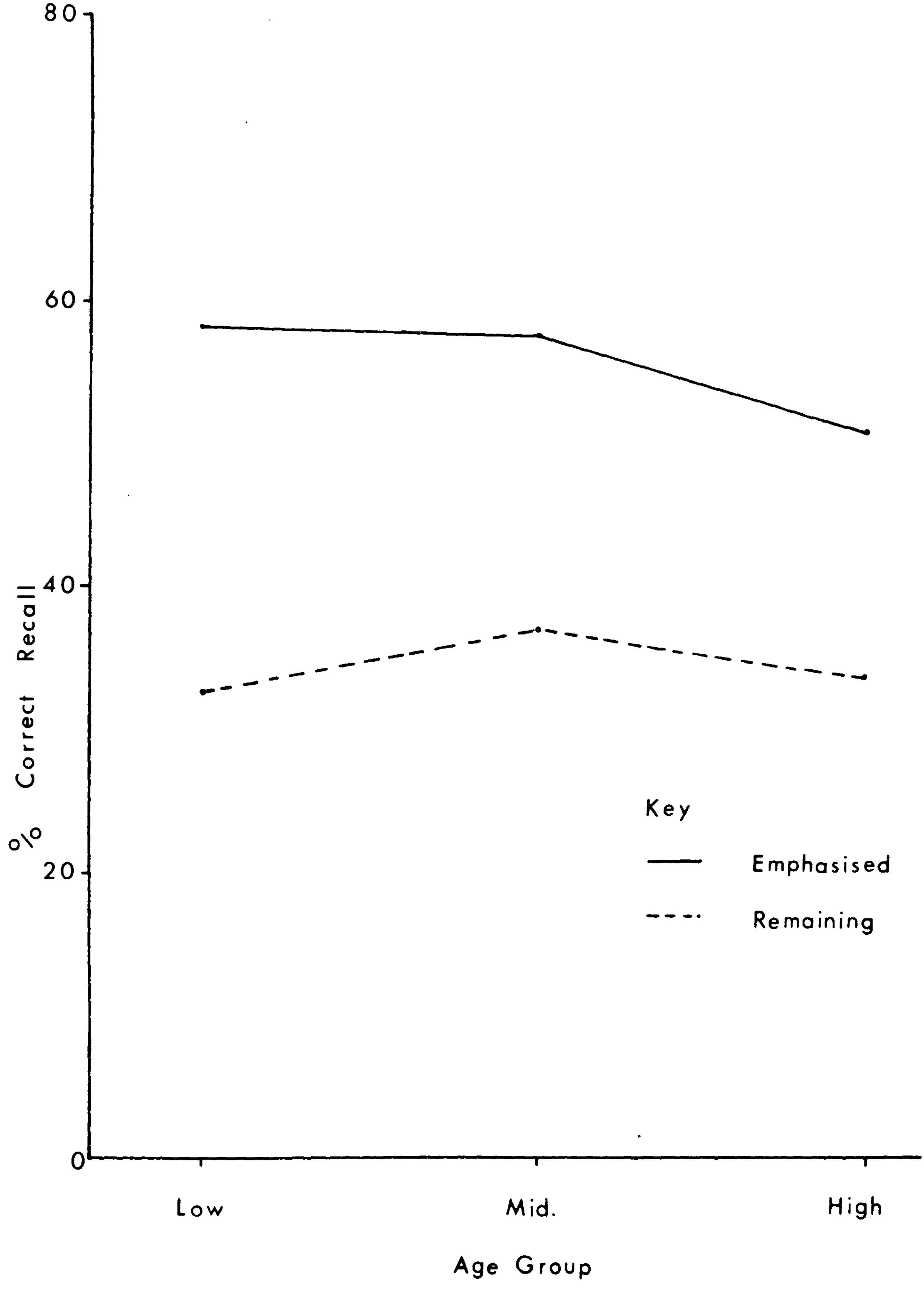


Fig. 3.4

Interaction Between Age Group & Emphasis.



(Fig.3.5). This seems to reflect the peculiar behaviour of the middle age subgroup of the low I.Q. test score group. The performance of the low I.Q. test score group is clearly significantly poorer in both category sets than the other two groups, but the effect of the instruction to selectively attend to one category over the others is just as apparent in this group.

The difference between the emphasised and remaining categories was examined in a different way by computing the difference in percentage recall between the emphasised and the remaining categories in conditions I and II. An analysis of variance was then computed using these (Table 3.6). There was no main effect of I.Q. test score but there was an effect of age. This was to be expected given the interaction between age and emphasis in the previous analysis. There was also a significant interaction between age and I.Q. test score, again as might be expected from the previous analysis (Fig. 3.6). However, it is interesting that I.Q. test score group interacted significantly with recall order. The difference between I.Q. test score groups was greatest when the emphasised category was recalled first However, when the emphasised category was recalled after the remaining categories the difference is markedly reduced in the high and middle I.Q. groups, but only marginnally in the low I.Q. test groups.

Fig. 3.5

3-Way Interaction Between Emphasis, Age & I.Q. Test Score Group.

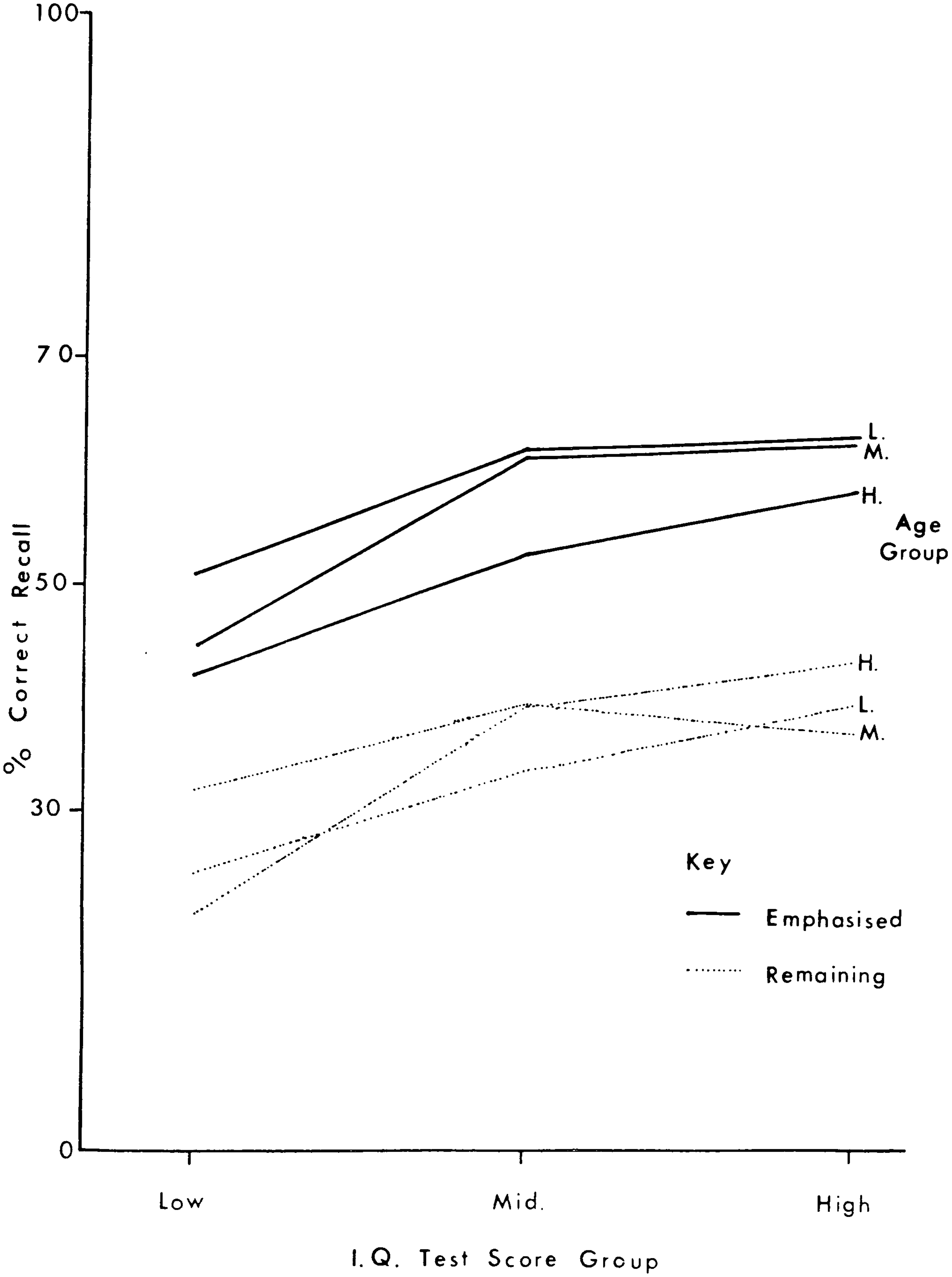


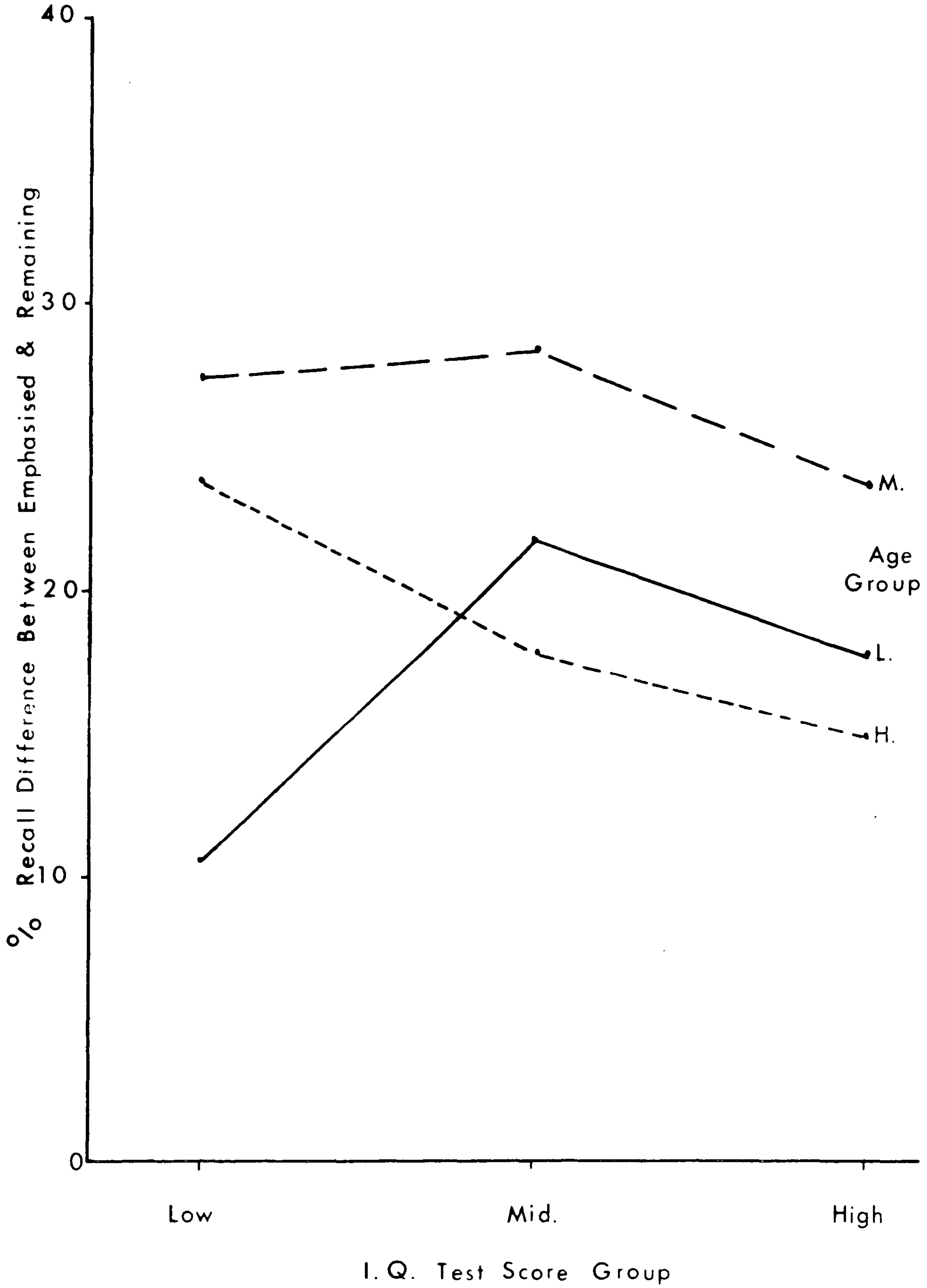
Table 3.6

Analysis Of Variance I.Q. Test Score Group / Age Group X Position For Percentage Difference In Recall Between Remaining and Emphasised Categories

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	114.156	2	57.078	0.215	0.807
Age	1866.544	2	933.272	03.517	0.034
I.Q. X Age	3022.849	4	755.712	2.848	0.029
S-Within	20699.938	78	265.384		
Posn.	5551.762	1	5551.762	32.233	0.001
I.Q. X Posn.	1665.589	2	832.795	4.835	0.010
Age X Posn.	200.954	2	100.477	0.583	0.560
I.Q. X Age X Posn.	576.381	4	144.095	0.278	0.506
Posn. X S-Within	13434.500	78	172.237		

Fig. 3.6

Interation Between Age & I.Q. Test Score Group.



3.2.8 Discussion

The main feature of all three conditions in this experiment is that all the words were categorised by the subjects and thereby elaborately encoded so that they were likely to have a better memory trace than words which were merely inspected, or required a decision on the basis of some physical feature of each word (Craik and Tulving, 1975). In addition to this imposed depth of processing, in two conditions the subjects were asked to attempt to remember the words of one of the presented categories as they were being shown for categorisation. This instruction was given in order to assess how well the subjects could ignore the other categories presented and adopt an effective strategy of encoding the 'emphasised' words over and above the others. The first important point to make is that subjects were indeed capable of this. There was an overall effect of emphasis, with percentage recall for the emphasised categories being significantly higher than for the remaining categories.

The second feature, after the emphasised category instruction, was the position of recall of the emphasised categories, and reciprocally of the remaining categories. This had a significant effect in both cases, but whether this is due to the interference of the previously recalled items, or to the delay caused by recalling these leading to trace decay, remains unresolved. Taub and his colleagues (Taub, 1968; Taub & Grief, 1967; Taub & Walker, 1970) have suggested that older subjects might be unduly affected by response interference, with larger age effects found in the recall of the second half of the sequence. The data here can be interpreted as showing response interference, but the effect of delay requires greater control.

After the consistent I.Q. test score group effects in the television studies, the main I.Q. test score effects found here were as expected; once again the recall performance of the high I.Q. test scorers was significantly higher than the low scorers. As in the television study this held across all conditions in the experiment. However, there was no significant interaction between I.Q. test score group and facilitation due to category emphasis, such as might be predicted if I.Q. differences were interpreted in terms of a model presupposing more efficient active selection of the information for later recall. Nevertheless, the significant interaction occurred with age group. This suggests that the older subjects were least effective in operating a strategy which encoded the emphasised category to a deeper or more elaborate extent than the other categories, or they were more prone to distraction from the other words. Both are plausible, since Rabbitt (1965a) has shown that the elderly are less capable of ignoring irrelevant information, and Smith (1980) concludes his review of encoding, storage, and retrieval with the remark that age differences in recall are due to the spontaneous use of different strategies by young and old subjects, with the older subjects engaging in less spontaneous organization.

The difference derived by subtracting the recall of the remaining from the emphasised categories not only highlights the age effect, with the difference being significantly less in the oldest group, but it also draws attention to another factor requiring consideration. Besides the overall age effect, I.Q. test band interacts with the order of recall, with the computed difference in recall being greatest in condition I, dramatically so in the case of the high and middle I.Q. test score groups and marginally in the low I.Q. test score group. This poses problems, since two possible factors combine to generate this effect. The first is the position of recall of the emphasised category. When it is recalled after the remaining categories the subjects' recall is as shown higher, and the emphasised category lower, even though overall the emphasised categories' recall is significantly higher than that of the remaining ones. This obviously reduces the difference between emphasised and remaining categories in this condition. But there is another necessary consideration: in conditions I and II, the instructions are identical, as all subjects have after condition I have been asked to attempt eventually to recall all of the previously presented words. It is highly likely that following the same instructions for condition II subjects might have become suspicious, and anticipate the subsequent recall of all the words. Hence, they might have adopted a different encoding strategy to enable more effective retrieval in response to these 'new' task demands. Obviously, if the subjects are now attempting to encode with the prospect of retrieving all the words the difference between emphasised and remaining categories will be reduced. The interesting point is that this effect is very weak in the low I.Q. test score group, who are more likely to respond more rigidly to the experimenter's task demands, or are unable to anticipate the last condition as the other two groups appear to have done.

There are two important points to be made from this study. First as a technique of implementing a more advanced strategy from subjects, categorisation of presented word lists and selective emphasis of one out of three categories of words presented appears to have been successful. Second, the interaction between age group and emphasis shows that the older group are clearly less able to make use of the organisational strategy forced upon them. However, the effect of I.Q. test score is still puzzling since its overall effect indicates that the benefits lie both in enhanced retrieval of the emphasised category words, and also in enhanced retrieval of remaining category words. We can not say whether this is due to more effective encoding strategies or to availability of a richer data structure in memory with more elaborate associations. The third possibilty of processing resources simply being better in high I.Q. volunteers also remains open.

There are, additionally, obvious problems with the experimental design. It fails to isolate the effect it was designed to test, and fails to disentangle the effects of response interference from those of delay. More importantly, in a study which is attempting to make subjects respond strictly to the experimenter's task demands, it allowed subjects to anticipate the likely recall demands when ignorance was necessary, and some of the data suggest that they responded according to this. Experiment two was designed to rectify these faults.

3.3 EXPERIMENT 3.2

3.3.1 Introduction

The main effect of emphasis in Experiment 3.1 illustrated that the instruction to remember the words of one of the categories to be categorised was an effective means of using task demands to cause the implementation of strategies. However, there were undoubtedly problems with the design of the study, since subjects by condition II were able to anticipate the recall demands, and this probably contributed to the high recall scores in this condition, especially in the remaining categories. The experiment was complicated further by the effects of recall position, i.e. whether the emphasised category was recalled before or after the remaining categories. This was seen as having two possible effects, i) it increased the amount of response interference, and ii) it increased the delay between presentation and recall. This effect was strongest in the high and middle I.Q. test score bands who recalled more words, thus increasing both factors.

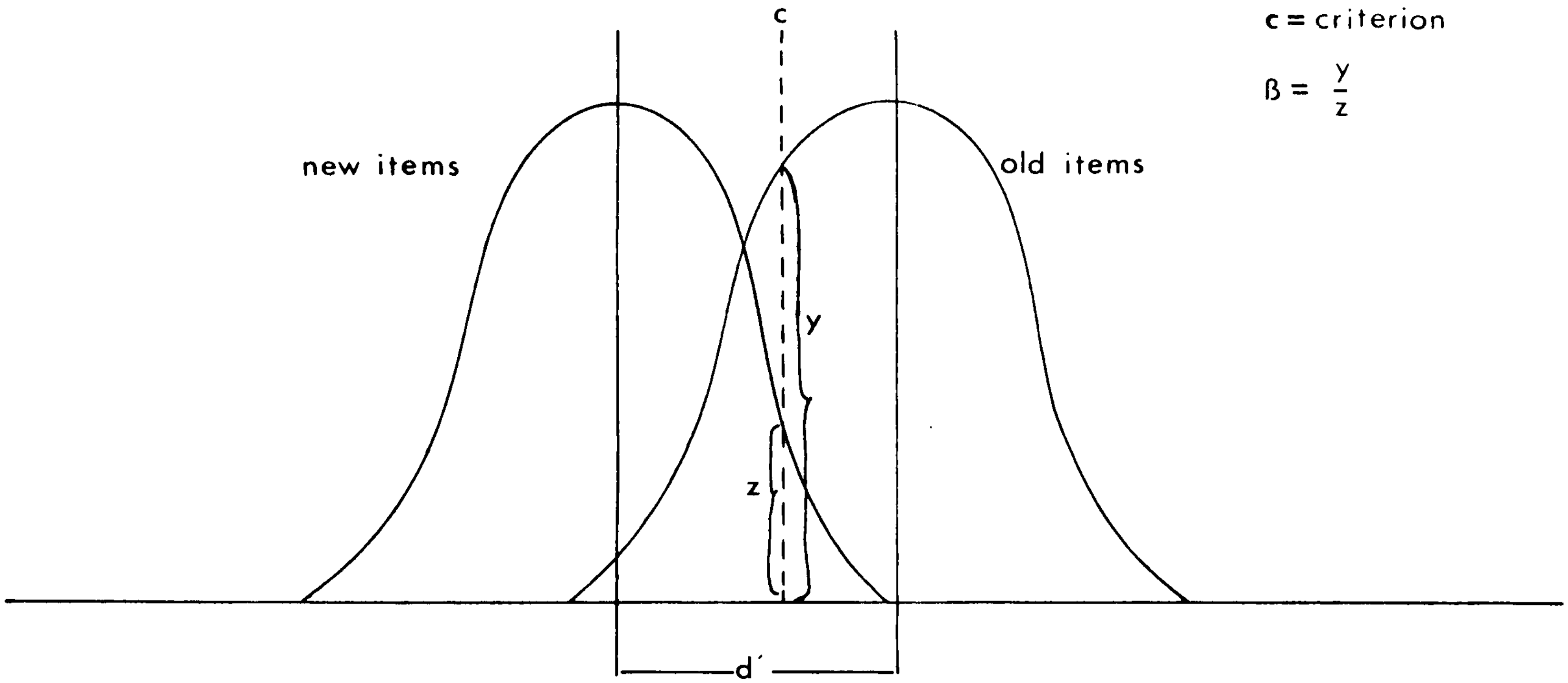
To compensate for this a recognition task was used, with the occurrence of emphasised words, remaining words and distractors randomly distributed, equating the effects of delay and interference across word groups. This method was also chosen because it allowed the application of signal detection theory in the analysis of the data, and hence d' and beta scores as measures of sensitivity and the criterion respectively could be examined.

In the mid-1960s signal detection theory was taken from psychophysics and applied to the study of memory in the analysis of recognition. Items were presented for learning and again among distractors. On the second presentation subjects were required to indicate which of the items they had seen previously and which were 'new' items. In this application of the theory there is the assumption that the tendency for any item to be indicated as an 'old' item varies from moment to moment due to random central activity or neural noise, meaning that over a period of time the strength of the tendency forms a roughly normal distribution. The presentation of any item for learning is assumed to add a constant to this tendency. The distributions for the items in the recognition tests can therefore be represented as two overlapping normal distributions as in Fig 3.7. It follows from this, that the subject is assumed to adopt a criterion such that any item with a momentary tendency above this, or to the right of this, is treated as having been presented before, and anything with a momentary tendency to the left, or below the criterion, is treated as 'new'. If normal distributions with equal variances are assumed, then d' can be calculated as a measure of recognition. The z scores are summed which correspond to two probabilities, these are i) the probability that items presented before will be recognised as such or will be regarded as 'new', and ii) the probability that a 'new' item will be recognised as such or be regarded as 'old'. For the criterion (beta), this is the ordinate of the corresponding z of i)/ii) (McNichol 1972)

Signal detection theory has been used in the literature on ageing Gregory (1974) and Vickers, Nettelback and Wilson (1972) found that the

Fig. 3.7

Normal Distributions For Old & New Items When Noise Is Added



signal to noise ratio in discrimination tends to be lower in older people. One explanation for this is the effect of sensory deficits, which reduce signal levels and possibly increase neural noise. However, as Welford (1980) points out, the application of these ideas to memory does not imply that memory deficits are attributable to poorer perception of the to-be-learned material or in relation to which recall is required, but that memory itself depends on the signal to noise ratios of memory traces.

Once the measures of d' and beta are obtained, group differences can be examined with respect to age and I.Q. test score bands. Of particular interest is the change in sensitivity and criterion with emphasis. If group differences in recall are due to differences in encoding strategy then one would predict an interaction between emphasis and beta levels for the different I.Q. test score or age groups. However if the passive models provide the correct explanation then one would expect group differences in sensitivity alone. This experiment tests these two proposals.

3.3.2 Subjects

150 subjects were selected on the basis of their age and AH4 part one test score. The subjects were then divided into three age groups, and three AH4 test score groups, the composition of which is shown in Table 3.7. The subjects used in this experiment were different from those used

Table 3.7

3x3 Matrix Of Subjects For Age and AH4 Test Score Groups

AH4 Test Score Group Age Group	Low Group 1	Middle Group 2	High Group 3	Total Mean (x) Std. Dev. Range (r)
Young Group 1	13	17	16	N=46 x=56.02 std.dev.=2.02 r=52-59 years
Middle Group 2	16	19	18	N=53 x=64.07 std.dev.=2.58 r=60-69 years
Old Group 3	17	18	16	N=51 x=72.74 std.dev=3.47 r=70-85 years
Total Mean (x) Std. Dev. Range	N=46 x=16.65 std.dev.=3.89 r=2-20/65	N=54 x=30.15 std.dev.=5.88 r=21-39/65	N=50 x=46.68 std.dev.=5.58 r=40-61/65	N=150 x=31.53 64.55 std.dev.=13.19 7.28 r=2-61/65 53-83 years

in Experiment 3.1.

3.3.3 Materials

Three categories of words were selected from Battig and Montague (1969), differing from those used in Experiment 3.1, but within the same potency range. From each category 30 words were selected and these were arbitrarily designated targets or distractors under the constraints of the targets and distractors being balanced for phonemic length. All sets of targets and distractors were balanced in this way both within and across categories. The categories and words selected are shown in Appendix 3.

3.3.4 Apparatus

A B.B.C. microcomputer was programmed to give a non-replacing random sequence of the 45 target words, followed by a random presentation of the target words embedded in 45 distractors. Four of the these sequences were recorded on colour video using a National Panasonic WVP-100 E colour camera with the targets and distractors exchanged on half of them. The sequences were shown using a Panasonic VHS video recorder through a Sony CVM.2000 PSB 19 inch colour monitor.

3.3.5 Procedure

Subjects were tested in small groups of approximately twelve. The procedure was similar to that of Experiment 3.1, in that initially subjects were told to categorise the 45 words presented into one of the three categories given by writing the corresponding symbol in the space provided on the response sheet. They were also instructed to give more attention to one of the categories because they would have recall words from this category later on. The words were shown for three seconds each with two seconds I.S.I.s between successive words to allow for the response. A loud bleep was dubbed onto the tape just before each word appeared to warn subjects that the next word was due. After the categorization task, the subjects were given a recognition test of all the words they had previously categorised, amongst distractors. The emphasised category was randomly allocated to each group, as were the categorization sequences.

The analysis applied signal detection theory to calculate d' and beta levels for each subject. Group differences were examined on these indices, and on the effect of the task demands on these two measures of memory and the strategies employed by the subjects.

3.3.6 Results

The word categorisation data was analysed first. Mean number of correct word categorisations were calculated for each subject, and then placed in an analysis of variance with age and I.Q. test score as between factor variables. Both age group and I.Q. test score had significant main effects (F=3.457 df=2,93 P<0.03 , F=9.172 df=2,93 P<0.01). However, the interaction was insignificant (F=1.273 df=4,93 P<0.28). As in Experiment 3.1 the performance level was generally high (Table 3.8), with mean correct response for the low I.Q. score group being 41.976/45 and for the elderly subjects 42.894/45. All subjects' responses were recorded as hits (correct responses), misses, false positives, and false negatives (type I and type II errors). From these d' and beta were computed after the computation of z scores for hits and false positives. Decriptive statistics of these are provided in Table 3.9

An analysis of variance was used for the d' values with emphasis as a within subject factors and age and I.Q. test score group as between subject factors. Table 3.10 gives the full summary table and Table 3.11 decriptive statistics. The only significant effect was I.Q. test score group, and no age effects or effects of emphasis were found. All interactions were also insignificant. The I.Q. effect shown in Fig. 3.8 shows the high and middle groups were significantly more sensitive than the low I.Q. test score group

Table 3.8

Descriptive Statistics Of Categorisation Performance For I.Q. Test Score Groups & Age Groups

Actual Scores (Maximum Score = 45)

I.Q. Test Score Group	N	MIN.	MAX.	MEAN	STD. DEV.
Low	41	18.00	45.00	41.976	05.90
Middle	52	39.00	45.00	44.423	01.45
High	50	41.00	45.00	44.860	00.61
Age Group					
Low	46	35.00	45.00	44.239	01.94
Middle	51	38.00	45.00	44.392	01.67
High	47	18.00	45.00	42.894	05.44

Table 3.9

Descriptive Statistics For Age Groups

Percentage Scores

Age Group 1

	N	MIN.	MAX.	MEAN	STD. DEV.
POSITIVES	44	44.444	97.778	73.737	12.032
FALSE POSITIVES	44	22.222	66.666	18.939	14.009
NEGATIVES	44	33.333	97.778	80.152	14.875
FALSE NEGATIVES	44	22.222	55.555	25.758	11.545

Age Group 2

	N	MIN.	MAX.	MEAN	STD. DEV.
POSITIVES	53	35.556	93.333	70.776	14.769
FALSE POSITIVES	53	0.000	51.111	20.008	12.601
NEGATIVES	53	48.889	100.000	79.455	12.802
FALSE NEGATIVES	53	06.667	64.444	29.266	14.810

Age Group 3

	N	MIN.	MAX.	MEAN	STD. DEV.
POSITIVES	51	37.778	100.000	71.373	14.115
FALSE POSITIVES	51	0.000	100.000	24.440	18.311
NEGATIVES	51	0.000	100.000	76.166	17.939
FALSE NEGATIVES	51	0.000	55.555	27.320	13.431

Table 3.9 (Continued)

Descriptive Statistics For I.Q. Test Score Groups

Percentage Scores

I.Q. Group 1

	N	MIN.	MAX.	MEAN	STD. DEV.
POSITIVES	46	37.778	97.778	62.850	14.369
FALSE POSITIVES	46	00.000	66.666	24.783	15.935
NEGATIVES	46	33.333	100.000	75.604	16.014
FALSE NEGATIVES	46	02.222	57.778	35.556	14.241

I.Q. Group 2

	N	MIN.	MAX.	MEAN	STD. DEV.
POSITIVES	52	35.556	100.000	75.256	11.898
FALSE POSITIVES	52	02.222	100.000	21.538	15.704
NEGATIVES	52	00.000	97.778	77.650	15.875
FALSE NEGATIVES	52	00.000	64.444	24.573	11.820

I.Q. Group 3

	N	MIN.	MAX.	MEAN	STD. DEV.
POSITIVES	49	51.111	93.333	77.007	10.624
FALSE POSITIVES	49	0.000	51.111	17.778	13.691
NEGATIVES	49	48.889	100.000	81.995	13.731
FALSE NEGATIVES	49	06.667	48.889	22.902	10.727

Table 3.10

Analysis Of Variance I.Q. Test Score Group / Age Group X Emphasis for D prime Scores

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	7.204	2	3.602	7.329	0.001
Age	2.562	2	1.281	2.606	0.079
I.Q. X Age	3.584	4	0.896	1.823	0.131
S-Within	45.708	93	0.491		
Emph.	0.235	1	0.235	0.884	0.350
I.Q. X Emph.	0.138	2	0.069	0.259	0.772
Age X Emph.	0.083	2	0.042	0.157	0.855
I.Q. X Age X Emph.	1.093	4	0.273	1.027	0.397
Emph. X S-Within	24.733	93	0.266		

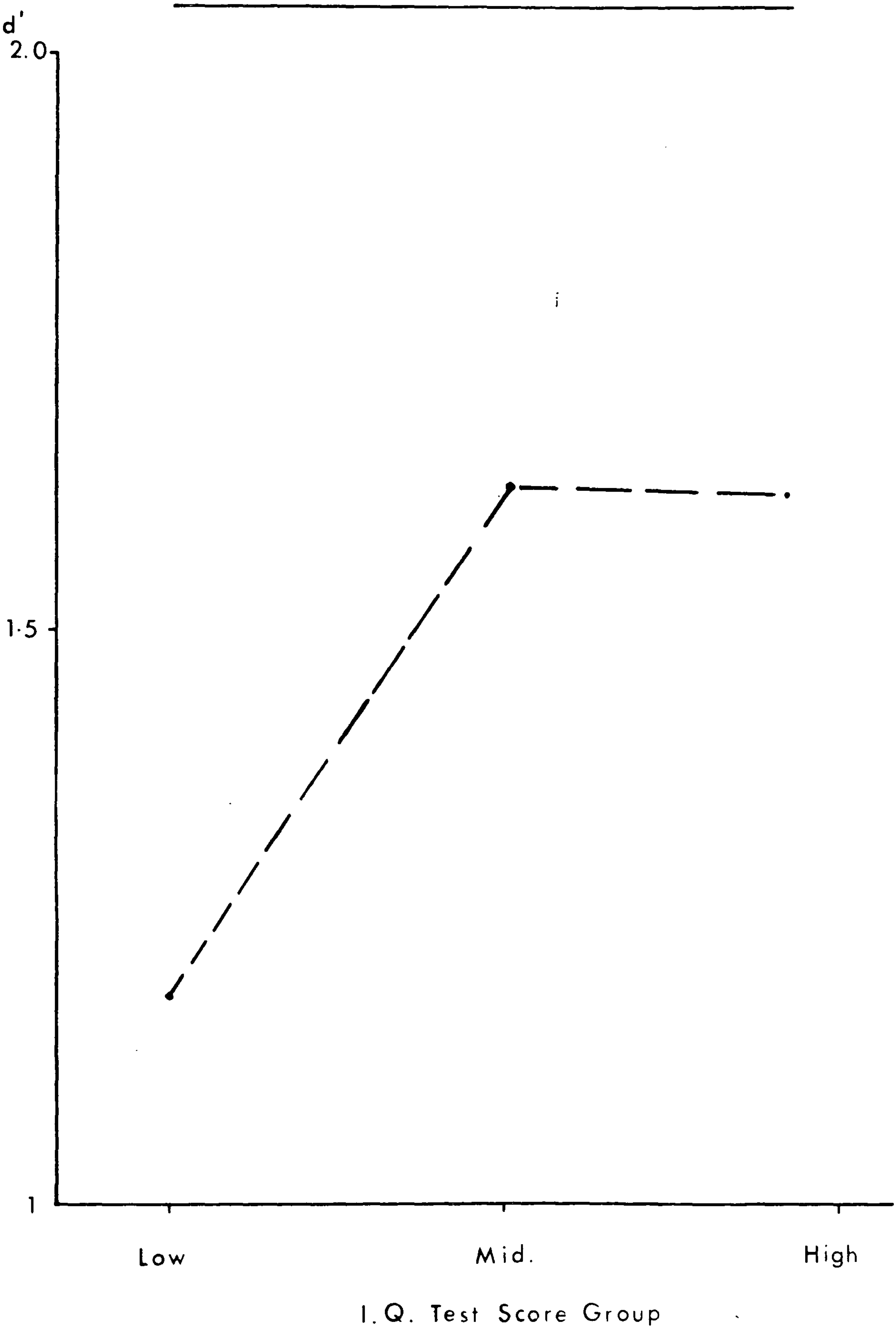
Table 3.11

Descriptive Statistics Of D Prime And Beta For Stated And Remaining Categories

Condition	N	MIN.	MAX.	MEAN	STD. DEV.	SKEW.	KURT.
Control	96	2.22	62.22	31.32	11.66	0.286	-0.193
Stated Category Position 1	96	26.66	93.33	58.61	16.53	-0.206	-0.956
Stated Category Position 2	96	13.33	93.33	53.40	16.88	-0.094	-0.452
Remaining Categories Position 1	96	3.33	76.66	30.52	12.47	0.507	0.865
Remaining Categories Position 2	96	6.66	73.33	38.71	15.32	0.117	-0.643

Fig. 3.8

Effect Of I.Q. Test Score Group On d'



A similar analysis was computed for the beta values (see Table 3.12). In this case there was once again a significant effect of I.Q. test score group, but also a significant effect of emphasis, with the beta value being significantly closer to one in the case of the remaining categories, and thus the critierion is more lax for the emphasised categories. The I.Q. effect showed the high and middle I.Q. test score group as adopting the most lax criteria (see Figs. 3.9-3.10).

3.3.7 Discussion

Once again, the effect of I.Q. test score was significant for both the measures made in this study. In the case of d', the only significant effect was of I.Q. test score group with the high and middle group being significantly more sensitive than the low scorers. Hence these groups are better at the recognition task, and therefore show a higher signal to noise discrimination. This provides more support for the passive models, with the sensitivity of the system being more sophisticated in the case of the high and middle I.Q. test scorers. A data base which is more absorbent would process material to a deeper level and form richer associations. Its elaborate network of associations enables the subject to be more aware of the items stored within the data-base.

It was surprising that there was no interaction between sensitivity and category type, since one would have expected subjects to have increased sensitivity to the emphasised category in comparison with the remaining categories.

Table 3.12

Analysis Of Variance I.Q. Test Score Group / Age Group X Emphasis for Beta Scores

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	0.139	2	0.069	10.080	0.001
Age	0.011	2	0.006	0.832	0.438
I.Q. X Age	0.020	4	0.005	0.740	0.566
S-Within	0.786	114	0.007		
Emph.	0.146	1	0.146	25.621	0.001
I.Q. X Emph.	0.001	2	0.001	0.115	0.891
Age X Emph.	0.006	2	0.003	0.530	0.590
I.Q. X Age X Emph.	0.005	4	0.001	0.224	0.925
Emph. X S-Within	0.648	114	0.006		

Fig. 3.9

Effect Of Emphasis On ß Values.

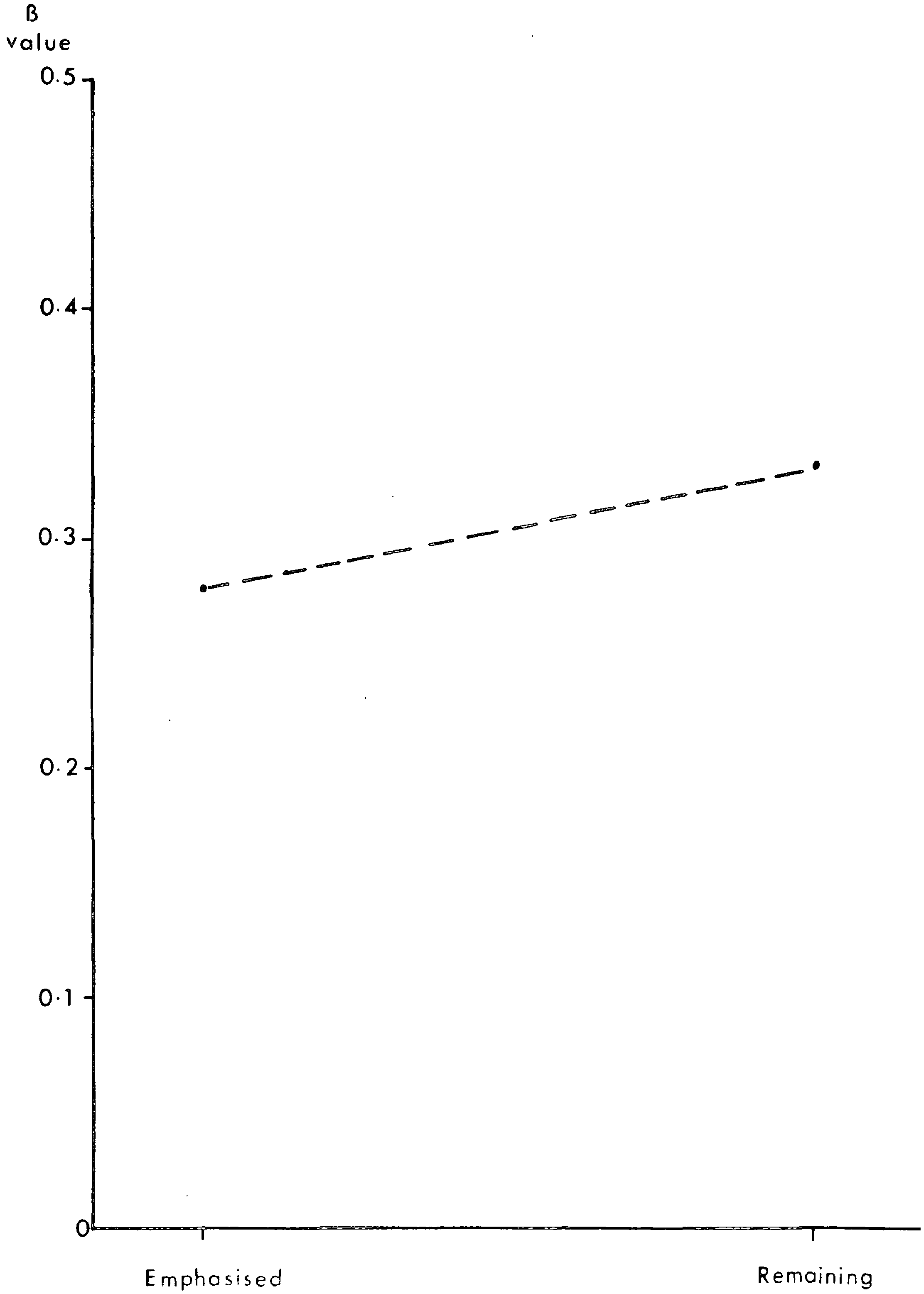
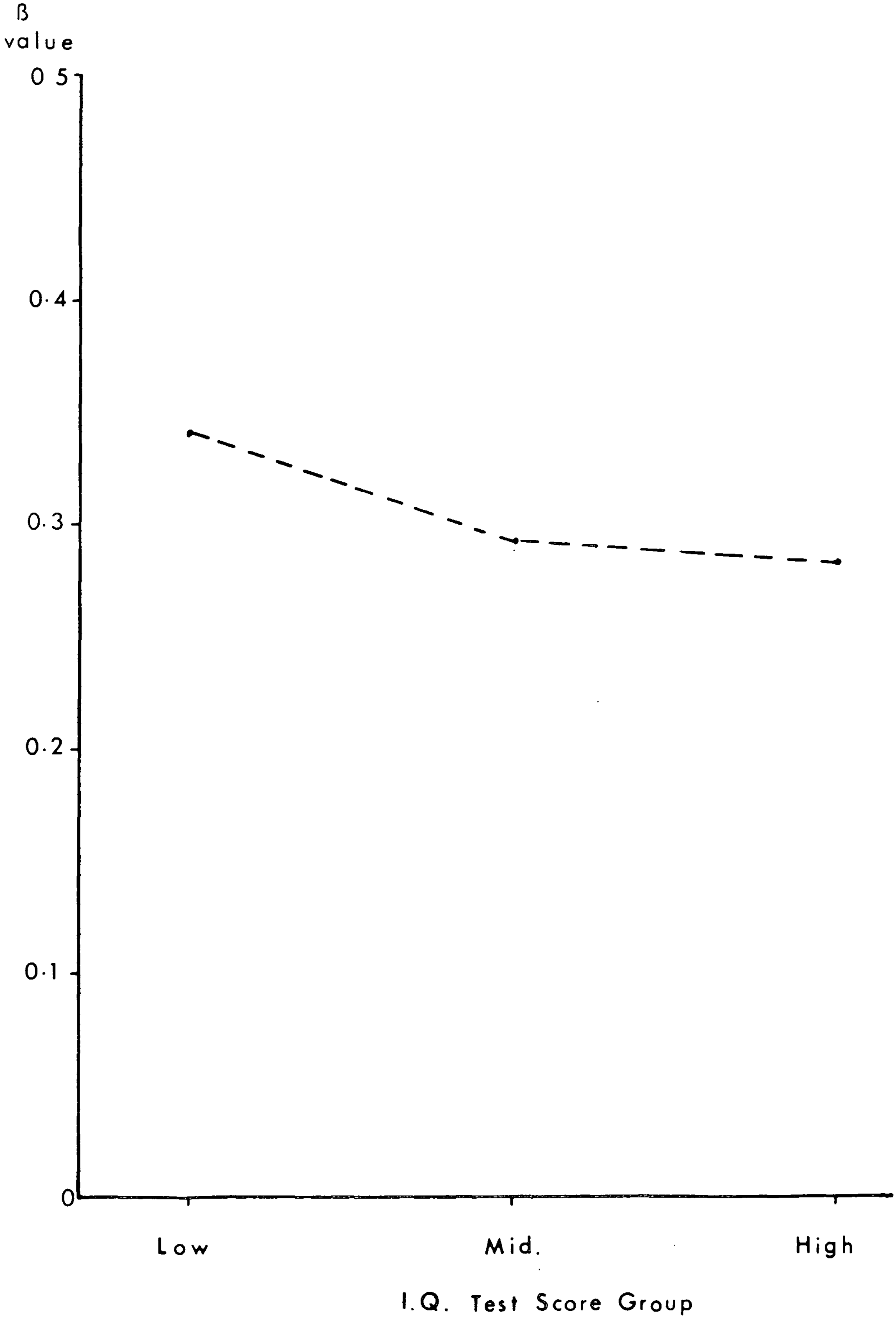


Fig. 3.10

Effect Of I.Q. Test Score Group On ß Values.



The beta values also showed a main effect of I.Q. test score:the high I.Q. group and middle I.Q. test score group adopt a lax criterion in comparison with the low I.Q. test score group, which they can afford to do this given their higher sensitivity. In addition to the effect of I.Q. test score group, there was an effect of category type, with the emphasised category having a more lax criterion than the remaining categories. This was adopted by all I.Q. test score bands and all age bands. This indicates that for each group the strategy adopted was to be more relaxed with the emphasised category in the hope of scoring more highly, but at the risk of making more false positives. Conversely, a stricter criterion could be adopted for the remaining categories, which subjects were not instructed to remember: hence the motivation to recognise these as being present was lower. Alternatively subjects were less certain about items in these categories and so were more stringent in their guessing.

The lack of interaction with age or I.Q. test score leaves open the hypothesis motivating these experiments. It may be that in some tasks group differences in strategic encoding occur, but there is no evidence in the present data to support this, while there is definitely a shift in criteria associated with instruction to remember. The effects of this shift are not specific to or greater for any age or I.Q. group. Therefore, all groups are adopting the same strategy. It may be that the high and middle groups are more successful at operating this strategy - but it is equally possible that because of their higher sensitivity, their levels are greater.

The data show the already established trend of the low I.Q. group being less sensitive (Gregory, 1974) and, consequently, since they are more cautious (Botwinick 1972), they are likely to have set a stricter criterion as a result of their awareness of their lower sensitivity. The converse may be true for the high I.Q. test scorers.

Overall, the data leave the question of the value of strategies in explaining group differences unresolved. Indeed, it appears that the results may give more support to the passive models reviewed earlier.

3.4 GENERAL DISCUSSION

The difference between emphasised and remaining categories in both experiments confirms that the method used was effective as a technique for investigating the ability to attend selectively to a set of words.

The overall I.Q. test score effect is as prevalent in these studies as it was in the television recall experiments. Once again, performance on the AH4 part one test of intelligence can be seen to be accounting for a high percentage of the variance in this task, which requires subjects to attend selectively to one of three categories in a word categorisation task. However, the AH4 test has not accounted for all of the variance, since there was a significant age effect interacting with emphasis. This clearly indicates that the more elderly subjects are less able to attend selectively to the emphasised category, and thus the difference between the remaining category recall and emphasised category recall is less for this group. This confirms Rabbitt's (1965a) finding of the elderly being less able to ignore irrelevant information.

There are two models of intelligence which are appropriate to these data. One is an active model, the other passive. In the former, group differences can be explained as follows: the high I.Q. test scorers are capable of a more effective deployment of their resources using strategies of encoding and retrieval. These are part of the control processes of the system, and they require an input of processing capacity or attention. Both experiments tested the subject's ability to attend selectively to words from a large set. The passive model is based on the overall efficiency of the system and explains group differences, not in terms of the way features of the system are allocated, and strategies implemented, but it assumes that all these systems are essentially more efficient in the high I.Q. test scorers. The overall I.Q. effect, without an interaction with emphasis, again corresponds to the effect found in the television experiments, where not only was the high I.Q. group better at recalling thematically relevant information, but also irrelevant details. In this case we find a similar effect; the high I.Q. test score group recall the emphasised categories better than the low scorers, but the same is also found for the other as well. Hence, we cannot interpret the difference as being due to better selective attention. Rather, the passive model is found to be more appropriate, since if the entire system is more efficient than [illegible], both the emphasised and remaining categories would be recalled better. The effect of strategic use is undoubtedly evident, but the data here cannot be used for supporting the active model.

Experiment 3.2 provided supporting evidence for the above analysis, in showing that the high I.Q. group were significantly more sensitive than the low group, since they had higher d' values. However, d' did not interact with emphasis. An I.Q. effect with beta was seen, the high test score group adopting a more risky criterion. This is what one would expect since Korchin and Basowitz (1957) found that during a series of rote learning tasks, older subjects tended to respond accurately or not at all, which implies a high cut-off. However, interpretation of these results must also consider sensitivity. It appears from the data that there is an effect of metamemory; since the high I.Q. group are more sensitive, and hence better at the recognition task, they can afford to adjust their criterion to a lower cut off point. Further, this notion is supported by the effect of emphasis on beta with the beta levels for the emphasised category being significantly lower than for the remaining categories. This pattern was found for all I.Q. groups. Again, what appears to be happening is that all the groups are probably more confident of their recall for the emphasised category, and hence adopt a relaxed criterion, whereas the remaining categories are treated with more caution. From this we can see that d' and beta are not always independent; one's beta level is fixed by taking one's sensitivity and confidence into account.

To summarise, these studies set out to test the active model of intelligence more rigorously, following the surprising lack of support found in the television studies, where it was explored in two different analyses. There is still no evidence to isolate it as responsible for group differences. Rather, all groups appear to be capable of effectively allocating attention, although this ability declines with age. Further, the high I.Q. test scorers were also better at overall recall for the non-emphasised categories, which provides support for a more passive model. The study also provides evidence for the use of metamemory in recognition, with a more risky criterion being adopted by subjects in the conditions where they are more confident of correct recognition.

CHAPTER FOUR

4 DRAWING INFERENCES FROM PICTORIAL SEQUENCES

4.0.1 Introduction

In order to recall not only television broadcasts but also everyday events, episodes must be identified and the nature of the thematic and logical links between episodes must be comprehended. The structure of events is based upon an underlying theme, which requires identification from the superfluous details which inevitably go with it. From the television news broadcasts recall studies, it is clear that the elderly encounter problems with recalling events, and this particularly applies to those of low I.Q. test score. Volunteers with low I.Q. test scores may have had difficultly with these events for two separate, though related, reasons. First, they might have had difficulty in detecting logical connections (i.e. 'theme' or a 'plot') linking segment scenes portrayed in the news broadcast. This would be in effect, a decrement in selective attention for critical, thematically relevant, cues. Second, they might have found it difficult to pick out cues or details which were important to the theme of the news stories because they were poor at using inferences (Cohen, 1979, 1981) to reconstruct complex themes from fragmentary details presented.

Rapid and accurate inferential reasoning is an important cognitive skill primarily related to comprehension. In the case of comprehending television broadcasts its use is notable in that scenes change rapidly, with plots frequently moving haphazardly through time, so that complex inferences may be necessary to get the drift of a story and these may have to be made very quickly indeed. The relative complexity of inferences that elderly people can draw, and the relative speed with which they can make inferences are, therefore separate factors which may limit their efficiency in understanding and remembering broadcasts.

Information received either linguistically or visually is usually impoverished, and therefore requires considerable supplementation from knowledge held in long-term memory in order for comprehension to occur. Cohen (1979) and Light, Zelinski, and Moore (1982) have shown how vulnerable inferential reasoning is to the effects of ageing when comprehension of discourse was tested; although the elderly subjects performed as well as the young subjects on verbatim questions, they were significantly poorer on questions which could only be answered correctly by making inferences. However, the majority of studies have focused on the comprehension of prose and spoken discourse, which have been followed by recall and recognition tests. Pictorial material on the other hand, has been rarely tested and yet inference construction from scenes is an important component of comprehension in everyday life.

The experiment reported in this chapter, investigates the speed and accuracy with which elderly subjects can make judgements based upon the theme they have constructed from a sequence of picture slides. Other studies have used a comic strip or sequence of pictures to study. In the main, these have sought to study the integration of presented material with stored knowledge (Jenkins, Wald, and Pittenger, 1978). Such experiments provide ample evidence to show that any set of slides implies some set of plausible options concerning the outcome or continuation of the event depicted. Hence, a set of slides which are coherent, and relate to each other in some systematic fashion, is seen to specify other stimuli and actions by inference. We tested this by using coherent sequences of pictures from a French Essay writing exercise book (Underwood, 1979) which followed a storyline, and asking subjects to decide whether the last slide in each sequence was consistent with the theme illustrated in the previous slides, i.e. did it follow the theme of the previous slides?

The difficulty experienced by the old and low I.Q. test score subjects in inferential reasoning may, plausibly increase as the load of information on which the inference has to be based increases. This was tested by presenting subjects with slide sequences of differing lengths, i.e. sequences were composed of four, three or two slides. In this way subjects were given slide sequences with different information loads and different amounts of contextual information. Cohen and Faulkener (1984) have tested the use of contextual cues on different age groups. They found no age and context interaction, so it seems that both old and young subjects use the contextual cues to facilitate encoding as efficiently as one another. However, the old were found to be disadvantaged when they were required to integrate information from different places in the text. This reinforces the position of a decrement in inferential processing, and suggests that the elderly fail to integrate material effectively into a global structure, since they fail to form necessary links between various elements in the material presented. Here, we test whether the elderly find it more difficult to integrate information when information load is increased from two to four pictures in a thematic sequence.

The effect of information load upon inferential reasoning can be interpreted as a feature of the working memory system, as this system is used to maintain information in consciousness so that computations can be performed, hence if the working memory capacity is reduced the effect of information load will increase. Arguing along these lines Light, Zelinski and Moore (1982) suggest that observed decrements in inferential reasoning associated with increasing age are explained by a reduction in working memory capacity. They find support for this hypothesis in the effect of presentation rate upon inference construction. Cohen (1979) has found that the elderly are disadvantaged at faster presentation rates, and also when spoken text as opposed to written text is used (Cohen 1981). There are two possible reasons for these findings; either the capacity of the working memory system is reduced so that it can manipulate and maintain less information, or there is an overall reduction in the cognitive processing efficiency of the system. Even if the working memory capacity of the elderly and the low I.Q. test scorers is adequate to allow them to make correct inferences, they may still have difficulty because they are slow at drawing inferences and can not keep pace with the presentation of material. In this experiment, viewing times for each picture in each sequence were recorded as well as the decision latency for a consistency/inconsistency judgement concerning the final slide in the sequence and its relationship with the previously shown slides in the sequence, which reflected the speed of inference construction. This allowed the possibiltiy of an interaction between information load and relative slowing of decision latency to be examined, since it is plausible that as information load increases the decision latency may increase in the low I.Q. test score group and the elderly subjects.

A final factor which may affect inference construction needs to be referred to here, although it is examined in detail in chapters 5, 6, and 7, this is the access and retrieval of information in long-term memory. When a person is required to 'fill in the gaps' in the information presented, two types of information contribute to their comprehension: first there is the information from the stimulus material and second the information activated internally by the stimulus, which is the associated knowledge held in long-term memory. The combination of these two information sources form the basis of inference construction. Therefore, both the speed of access to and retrieval of information stored in long-term memory and the quality of this information will affect the speed and accuracy of inference construction.

If, as expected, the old and the low I.Q. test score group are slow to comprehend these pictorial sequences, then this holds important implications for the difficulties they experience in everyday life. It is likely that they can not keep pace with the fast moving world, and as events become more and more complex, their difficulties mount until eventually, when the information load becomes too great their inference construction becomes too slow to allow adequate comprehension. By examining the effects of age and I.Q. test score on viewing times and decision latencies in relation to information load, this experiment attempts to disentangle the points outlined above, and to highlight some of their implications for everyday cognitive functioning.

4.0.2 Method

4.0.3 Subjects

Seventy five subjects were selected according to their age and AH4 part one test score. Three AH4 test score bands of high, middle and low were set on the same basis as in the previous experiments, and an even distribution of age was achieved in each. The mean age of the group was 63.75 years, std. dev. 7.713, with a mean AH4 part one score of 30.19/65, std. dev. 12.98, Table 4.1 shows the composition of the age and I.Q. test score bands.

4.0.4 Materials

36 sequences of black and white pen-and-ink drawings were composed from picture sequences used for French essay writing (Underwood, 1979) by the author for further rating by a panel of fifteen undergraduates. For half of the sequences, the final picture was inconsistent with the sequence theme and for the other half it was consistent. Six sequences

Table 4.1

3x3 Matrix Of Subjects For Age and AH4 Test Score Groups

AH4 Test Score Group Age Group	Low Group 1	Middle Group 2	High Group 3	Total Mean (x) Std. Dev. Range (r)
Young Group 1	9	7	7	N=23 x=55.13 std.dev.=1.89 r=52-58 years
Middle Group 2	8	8	8	N=24 x=62.12 std.dev.=2.49 r=59-67 years
Old Group 3	8	10	10	N=28 x=72.21 std.dev=3.86 r=68-81 years
Total Mean (x) Std. Dev. Range	N=25 x=16.76 std.dev.=3.91 r=6-20/65	N=25 x=28.04 std.dev.=5.59 r=21-39/65	N=25 x=45.76 std.dev.=5.25 r=40-61/65	N=75 x=30.19 63.75 std.dev.=12.98 7.71 r=6-61/65 52-81 years

were eliminated as being too ambiguous or difficult for theme construction by the panel. Booklets were made containing the 30 chosen sequences and 50 of these were sent out to volunteers who were randomly selected. Only 42 returned the completed booklets, mean age= 66.90 years, std. dev.= 6.47 and mean AH4 part one score= 32.38/65 , std. dev.=15.85. They were asked to decide whether the last picture in each sequence fitted in with the story theme of the previous pictures i.e. whether it was consistent with the previous pictorial material. As soon as he had decided this, the subject was simply required to write 'yes' or 'no' in a space provided. In addition, the subjects were asked to rate on a scale of 1 to 7 (1=very easy, 7=very difficult), how easy it was to make a decision as to the consistency of the last picture in relation to the previous pictures. Four examples which fully explained the procedure were presented at the beginning of the booklet.

The raters were grouped according to I.Q. test score and age in the same way as subjects were in the previously reported experiments (Experiments 2.1, 2.2, 3.1, 3.2) see Table 4.2. Mean ease of decision ratings were computed for each sequence, and these were ranked, the top 15 sequences were then classified as 'easy' sequences and the remaining 15 as 'difficult'. One way analyses of variance were computed for I.Q. test score group and the difficulty ratings referring to the decision on the last picture in each sequence for each sequence length separately. Only difficulty ratings for the four slide sequences showed an effect of I.Q. test score group, with the low I.Q. test score group finding the decision more difficult than the other groups (Table 4.3). The same

Table 4.2

3x3 Matrix Of Raters For Age and AH4 Test Score Groups

AH4 Test Score Group Age Group	Low Group 1	Middle Group 2	High Group 3	Total Mean (x) Std. Dev. Range (r)
Young Group 1	0	1	5	N=6 x=55.68 std.dev.=1.97 r=52-57 years
Middle Group 2	6	6	5	N=17 x=64.82 std.dev.=2.48 r=61-68 years
Old Group 3	7	6	6	N=19 x=72.32 std.dev=3.42 r=69-79 years
Total Mean (x) Std. Dev. Range	N=13 x=13.61 std.dev.=5.35 r=1-20/65	N=13 x=30.46 std.dev.=4.72 r=23-39/65	N=16 x=49.19 std.dev.=6.06 r=42-60/65	N=42 x=32.38 66.90 std.dev.=15.85 6.47 r=1-60/65 52-79 years

Table 4.3

1 Way Analysis Of Variance I.Q. Test Score Group X Difficulty Rating 4 Slide Sequences

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Between	07.8807	2	3.9404	7.0377	0.0025
Within	21.836	39	0.5599		
Total	29.717	41	(Random Effects Statistic)		

I.Q. Group	N	Mean	Variance	Std. Dev
Low	13	2.7060	1.0398	1.0200
Middle	14	1.8571	0.3602	0.6002
High	15	1.7060	0.3339	0.5778
Grand	42	2.0659	0.7248	0.8513

analyses were computed for the age groups of the raters, but there was no significant effect of age group on the ratings for the three different slide length sequences including the four slide sequences (Table 4.4). In addition the consistency of each final 'target' frame of the sequence was established from the raters responses, for all sequences there was atleast 70% aggreement, which was validated by the high agreement found later with subjects responses. Appendix 5 gives some examples of sequences used in this experiment.

4.0.5 Apparatus

A Kodak 'S' carousel projector with 100mm lens was connected to an Apple Euro II microcomputer which had an Apple Clock inserted in slot 05. Three switches were wired to the Apple, a 'yes' and 'no' switch and a slide 'change' switch. In order to ensure that viewing and decision times were recorded from the time the slide appeared on the screen, a small hole was drilled in the surround of each slide. Projection light passing through this hole activated a diode photoreceptor in the projector which in turn started the microcomputer clock. The clock continued counting in milliseconds, until either the 'change' switch was pressed, or a decision switch closed. The switches were monitored, and all responses recorded through the software. When a target slide appeared on the screen a bleep sounded from the microcomputer, the software being arranged to count slides and sound a tone on each target slide onset. (For circuit diagram see Appendix 4.)

Table 4.4

1 Way Analysis Of Variance Age Group X Difficulty Rating 4 Slide Sequences

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Between	02.3988	2	1.994	1.7123	0.1937
Within	27.318	39	0.700		
Total	29.717	41	(Random Effects Statistic)		

Age Group	N	Mean	Variance	Std. Dev
Low	6	1.5357	0.2389	0.4887
Middle	17	2.2700	1.0179	1.0089
High	19	2.0508	0.5465	0.7392
Grand	42	2.0659	0.7248	0.8513

4.0.6 Procedure

Each subject was tested individually, the nature of the task being explained at the beginning of the session. The subjects were told to view each slide individually from the thirty sequences, spending as brief a time viewing each slide as they considered necessary for comprehension. The slide changes were performed by the operator, the subjects simply saying 'next' when they wanted the next slide in the sequence. At the occurrence of a target slide, the subject was required to press either the 'yes' or the 'no' button as soon as he or she had decided whether the slide presented was consistent or inconsistent with the theme illustrated by the earlier members of the sequence. A 'yes' response indicated that the subject judged the final slide in the sequence to be consistent with the previously shown slides, and a 'no' response corresponded to the target slide being judged as inconsistent with the previous slides in its sequence. The sequences were randomised, and varied in length with 2, 3 or 4 slides. A further four practice sequences were used to familiarise the subject with the process before testing began. Both the viewing times for each slide and the response times for the final slides were recorded by the computer on disk files.

4.0.7 Results

There were two possible approaches to the analysis of these data. If one takes an extreme point of view, and allocates a 'liberal criterion', the material can be considered as highly subjective, and all responses can be considered 'correct', since all terminal frames of sequences might represent logically possible (though very highly implausible) sequences. If we take this point of view, the data can be analysed in terms of decision latencies alone, since all responses are considered equally valid. The factors of sequence length, ratings made by the panel of judges of 'difficulty', and the consistent/inconsistent interpretation of the final slide can still be applied to the data. The other approach, is simply to apply the raters responses to the 'target' frame, in deciding the 'correct' response to each sequence. The responses made by the subjects in the experimental task can then be considered as 'correct' or 'errors' and latency data can be analysed accordingly.

The number of sequences in the final analysis was reduced to 24. These were made up of eight sequences of each length, each length grouping with equal frequencies of consistent/inconsistent sequences and easy and difficult. This reduction in the number of sequences was necessary since slides from two sequences had proved to be faulty, because mis-aligned holes in the slide surround led to unreliable timing of the decisions by the experimental equipment.

Mean response times were used, since the number of stories in each condition was so few that a median score could not be used. In the first analysis the 'liberal criterion' was used and all decision latencies to the target slides were pooled, regardless of whether the response made was 'correct' or 'wrong' in terms of the raters judgements. An analysis of variance was computed on this repeated measures design, with age and I.Q. test score group as within factor variables and sequence length, difficulty and consistent/inconsistent as within factors. The main finding was an overall effect of age group, with no effect of I.Q. test score group. There were no significant interactions between the within and between factors. All three within subject factors had main effects; the inconsistent and difficult sequences gave rise to longer response times than the easy and consistent ones. The effect of sequence length proved interesting, since the three length slide sequences generated longest latencies and the four slide sequences the shortest latencies. (See Table 4.5 for complete summary table of the analysis of variance).

In addition to main effects there were second order interactions between the variables of sequence length and the other two within factor variables. Sequence length interacted significantly with the consistency of the target slide. This apparently occurs because the inconsistent response sequences with fewest frames led to shorter response times than their consistent counterparts (Fig. 4.1) An interaction between sequence length and difficulty apparently occurred because the difference between easy and difficult sequences was largest for three frame sequences and least for two slide sequences (Fig. 4.2). There was also a significant third order interaction (Fig. 4.3). This seems to occur for a number of reasons, a) the four frame sequences show a larger difference in decision latencies between easy and difficult sequences in consistent sequences than inconsistent ones, as well as difficult consistent sequences leading to longer latencies than difficult inconsistent sequences, b) in three

Table 4.5 (All Responses Times Included)

Analysis Of Variance I.Q. Test Score Group / Age Group X Consistency Of Target Sequence Length & Difficulty

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Age	548.356	2	274.178	4.591	0.014
I.Q.	31.889	2	15.945	0.267	0.767
Age X I.Q.	158.058	4	39.514	0.662	0.621
S-Within	3941.695	66	59.723		
Len	153.703	2	76.852	8.770	0.001
Age X Len	7.431	4	1.858	0.212	0.931
I.Q. X Len	43.538	4	10.884	1.242	0.296
Age X I.Q. X Len	35.163	8	4.395	0.502	0.853
Len X S-Within	1156.766	132	8.763		
Con	73.029	1	73.029	6.425	0.014
Age X Con	2.444	2	1.222	0.109	0.898
I.Q. X Con	11.572	2	5.786	0.509	0.603
Age X I.Q. X Con	24.305	4	6.076	0.535	0.711
Con X S-Within	750.180	66	11.366		
Len X Con	216.512	2	108.256	16.246	0.001
Age X Len X Con	41.497	4	10.374	1.557	0.190
I.Q. X Len X Con	2.460	4	0.615	0.092	0.985
Age X I.Q. X Len X Con	33.269	8	4.159	0.624	0.756
Len X Con X S-Within	879.586	132	6.664		
Diff	310.265	1	310.265	31.956	0.001
Age X Diff	3.326	2	1.663	0.171	0.843
I.Q. X Diff	6.559	2	3.280	0.338	0.715
Age X I.Q. X Diff	20.429	4	5.107	0.526	0.717
Diff X S-Within	640.801	66	9.709		

Table 4.5 (Continued)

Analysis Of Variance I.Q. Test Score Group / Age Group X Consistency Of Target Sequence Length & Difficulty

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Len X Diff	88.343	2	44.171	4.009	0.020
Age X Len X Diff	21.275	4	5.319	0.483	0.748
I.Q. X Len X Diff	21.976	4	5.494	0.499	0.737
Age X I.Q. X Len X Diff	41.134	8	5.142	0.467	0.878
Len X Diff X S-Within	1454.227	132	11.017		
Con X Diff	71.884	1	71.884	11.955	0.001
Age X Con X Diff	22.976	2	11.488	1.911	0.156
I.Q. X Con X Diff	17.945	2	8.972	1.492	0.232
Age X I.Q. X Con X Diff	24.785	4	6.196	1.030	0.398
Con X Diff X S-Within	396.855	66	6.013		
Len X Con X Diff	264.215	2	132.107	19.189	0.001
Age X Len X Con X Diff	30.815	4	7.704	1.119	0.350
I.Q. X Len X Con X Diff	29.696	4	7.424	1.078	0.370
Age X I.Q. X Len X Con X Diff	39.983	8	4.998	0.726	0.668
Len X Con X Diff X S-Within	908.773	132	6.885		

Fig. 4.1

Interaction Between Sequence Length & Consistency Of Target Slide.

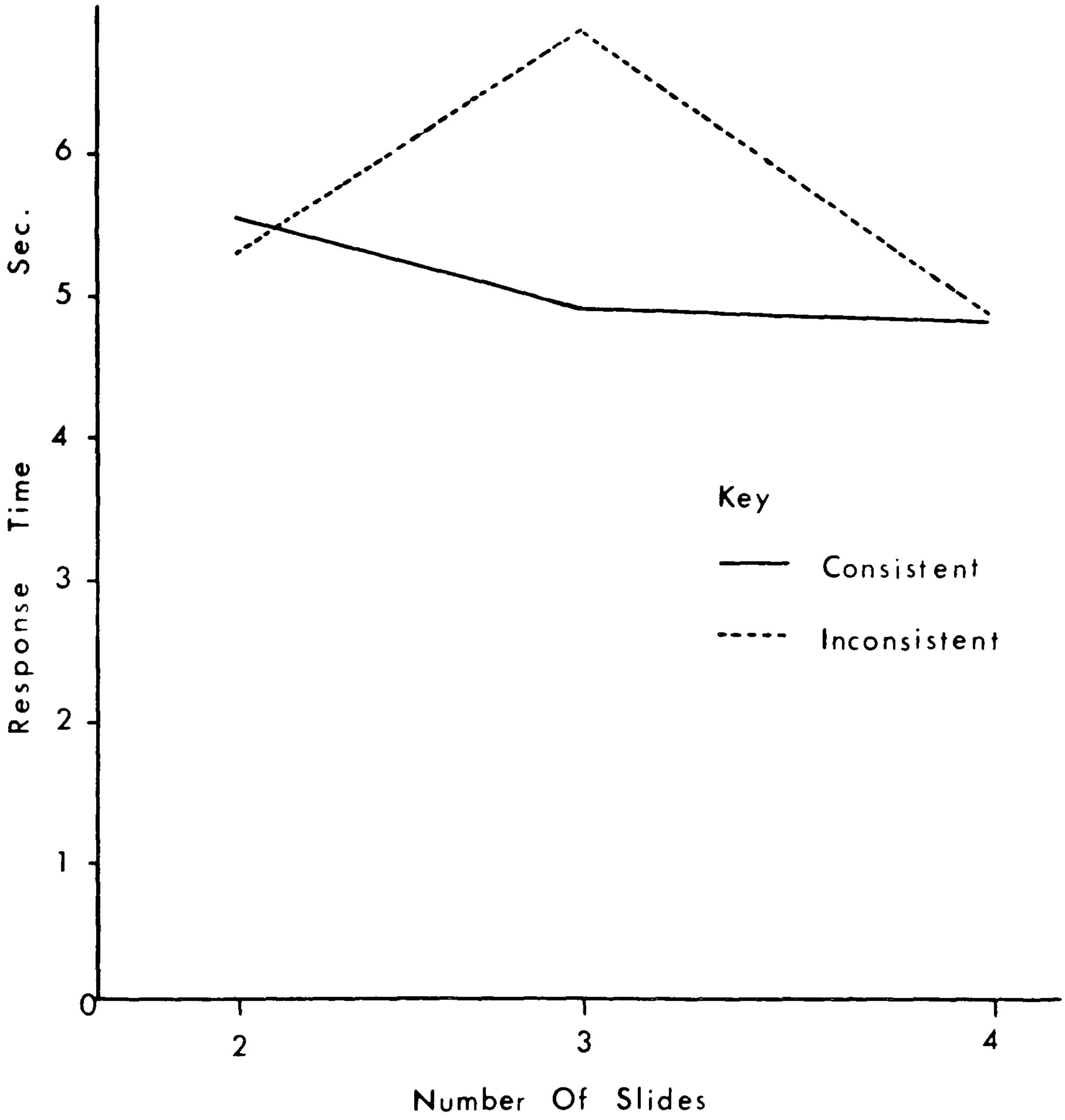


Fig. 4.2

Interaction Between Task Difficulty & Sequence Length.

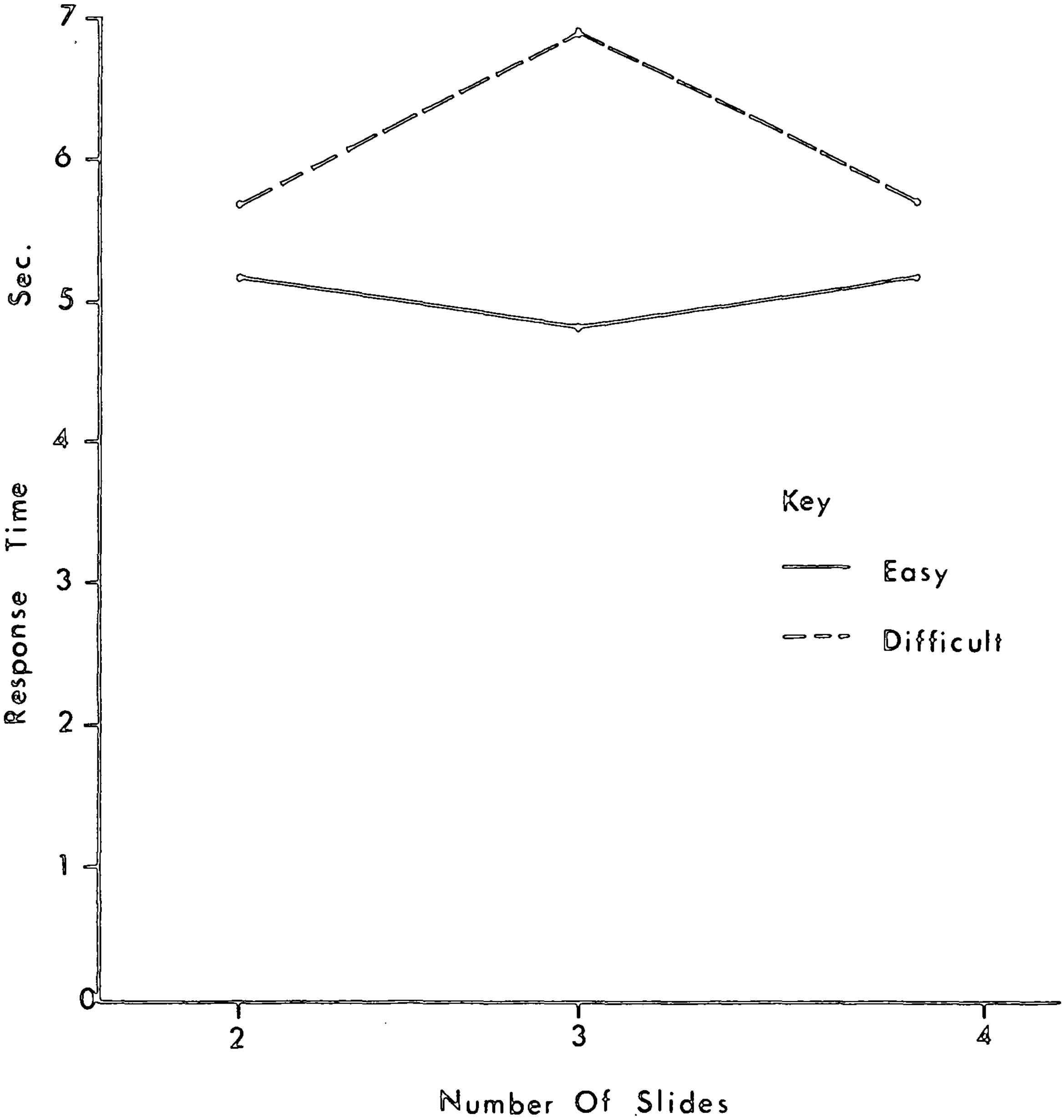
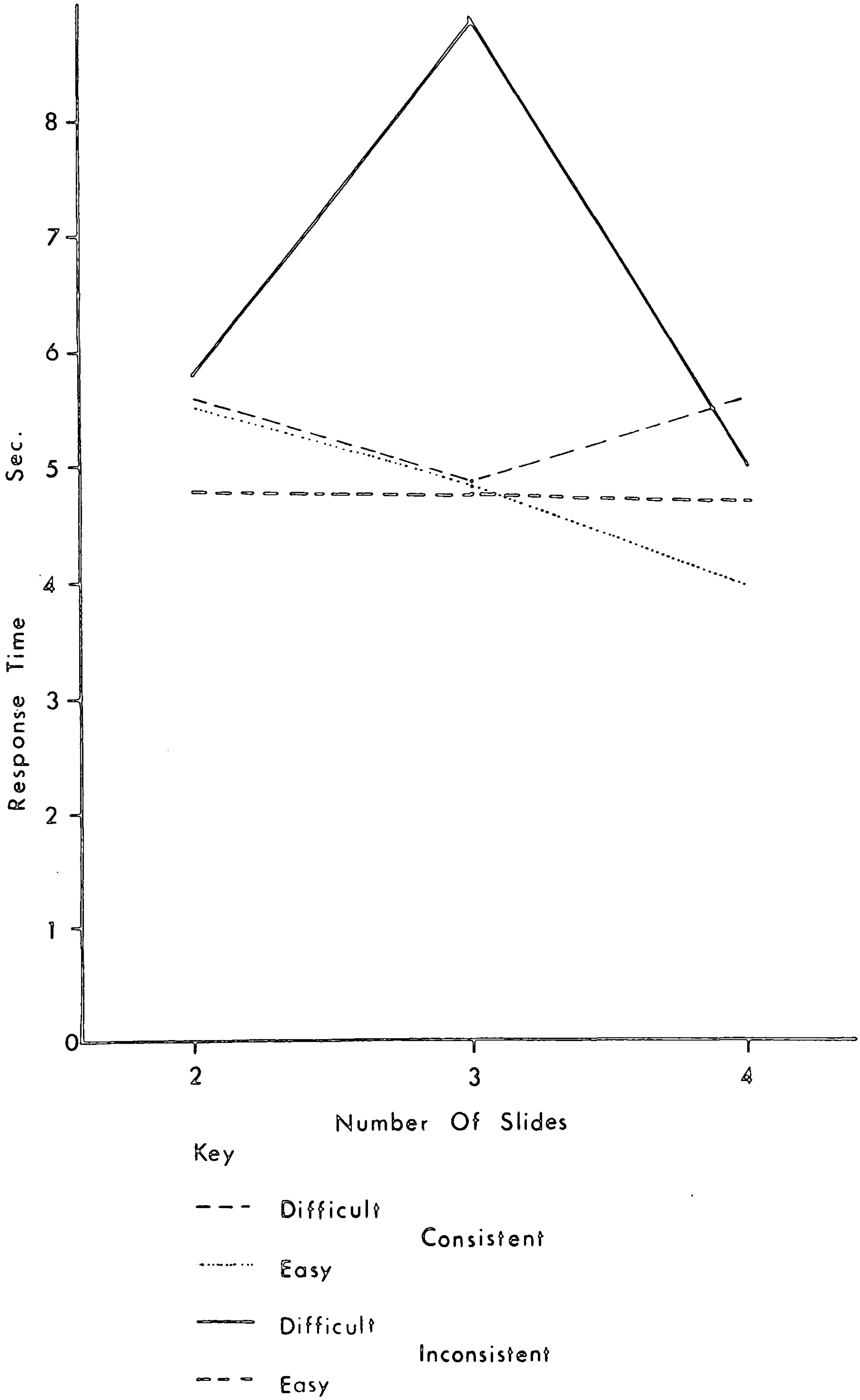


Fig. 4.3

3-Way Interaction Between Sequence Length, Consistency & Difficulty.

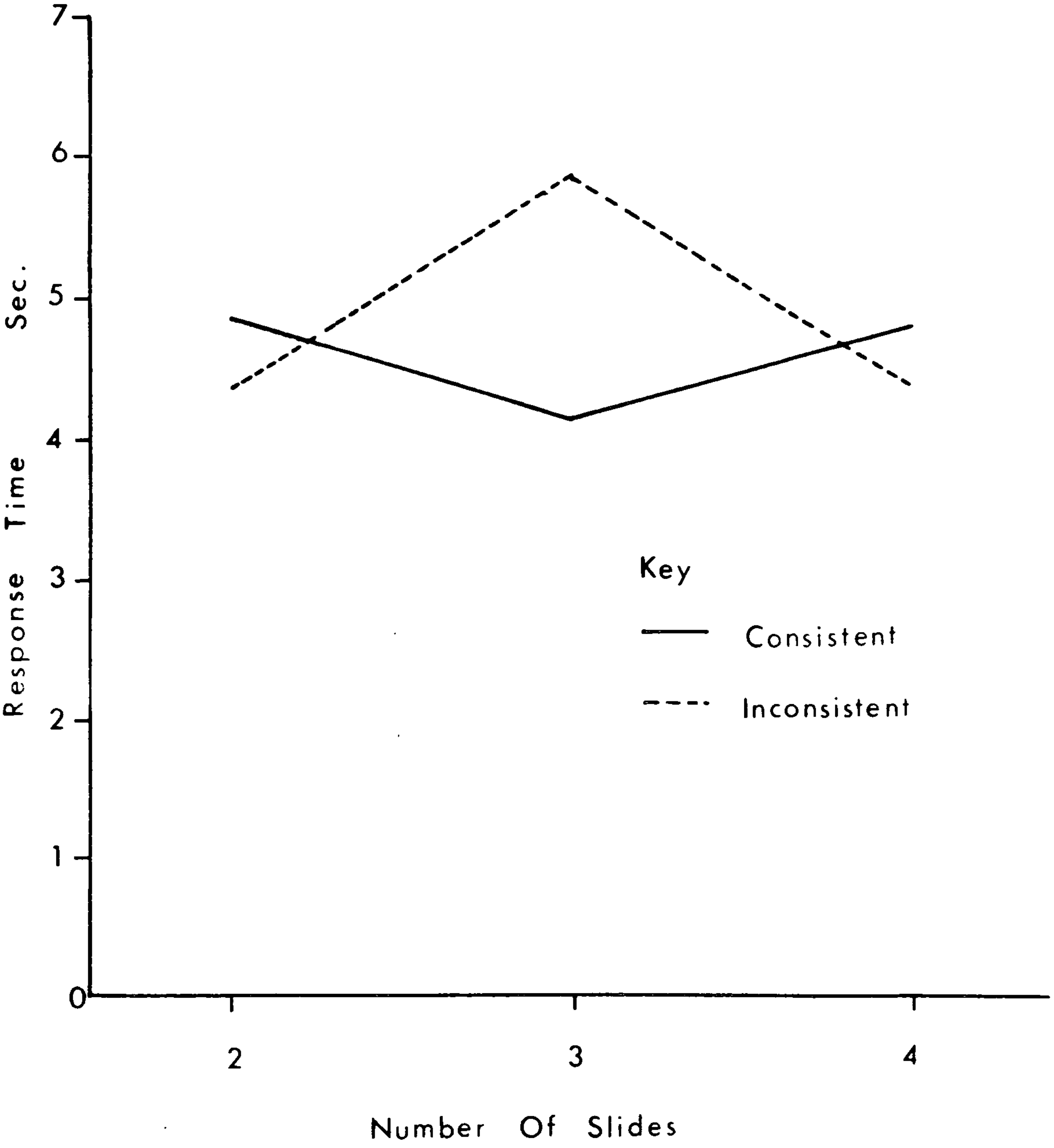


frame sequences, the differences between easy and difficult sequences for consistent target slide sequences is marginal i.e mean difference of 0.003 seconds, however there is a dramatic difference in the case of inconsistent sequence with a mean difference of 4.1 seconds, c) the two frame sequences, the small difference between easy and difficult sequences is repeated for consistent response sequences, but is more pronounced for inconsistent target slide sequences. Easy consistent sequences appear to generate longer response latencies than easy inconsistent target slide sequences. This third order interaction is undoubtedly difficult to disentangle, and illustrates how different amounts of context affect the response latency especially in relation to the judged difficulty of the sequences.

The second analysis was based on a 'strict' criterion, and we only considered decision latencies associated with 'correct' responses (in the terms defined by the panel who had rated the sequences). The same analysis of variance as just described was carried out on these data. Once again there was a significant effect of age group, but no interactions with this variable were significant. With respect to the within factors, there were significant interactions between sequence length and the consistent/inconsistent factor (Fig. 4.4). The reason for this is that the inconsistent responses are marginally shorter than the consistent responses for both the two and four slide sequences, but the three slide sequences show a significant difference in the opposite direction. An effect of difficulty was found in the expected direction, with sequences rated as easy by the panel giving rise to significantly

Fig. 4.4

Interaction Between Sequence Length & Consistency Of Target Slide.



shorter response times than those rated as difficult. The third order interaction was significant, as in the first analysis, again it is difficult to disentangle, but it seems to be significant because a) in four frame sequences the difference between easy and difficult sequences is most pronounced in the sequences requiring a consistent response, b) for three slide sequences, as in the previous analysis, the sequences with inconsistent target slides show a dramatic effect of difficulty, but the consistent target slide sequences show the unpredicted trend of sequences judged to be easy generating longer mean response times than those judged to be difficult, c) two frame sequences, show the effect of difficulty in the expected direction, with consistent and inconsistent target slide sequences leading to longer response times for those sequences judged to be difficult by the raters, however, the consistent sequences gave rise to longer response latencies than the inconsistent ones which was not expected (Table 4.6 and Fig. 4.5).

The number of correct responses for each subject was converted to a percentage score. This gave an index of the accuracy of the subjects' responses, descriptive statistics are given in Table 4.7. An analysis of variance was computed with age and I.Q. test score group as between subject factors, and accuracy as the within factor. Both age and I.Q. test score had a significant effect on accuracy (Table 4.8), and there were no significant interactions. A multiple regression analysis was also carried out which showed AH4 to be a better predictor of performance than age (Table 4.9). In addition chi-square distributions were computed on each sequence to investigate the hypothesis that the elderly and the

Fig. 4.5

3-Way Interaction Between Sequence Length, Consistency & Difficulty.

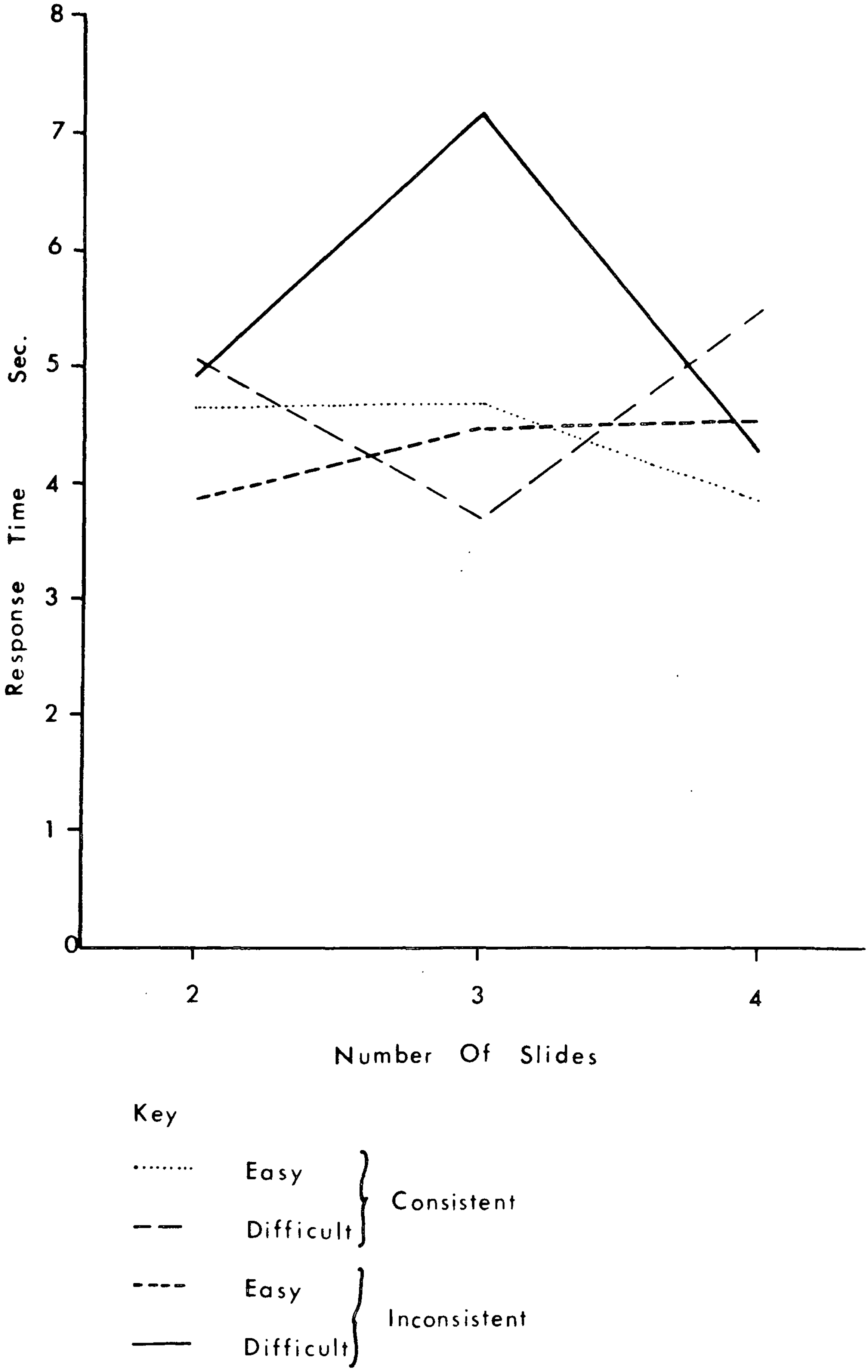


Table 4.6 (Correct Responses Times Only)

Analysis Of Variance I.Q. Test Score Group / Age Group X Consistency Of Target Sequence Length & Difficulty

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Age	525.739	2	262.869	4.946	0.010
I.Q.	22.349	2	11.174	0.210	0.811
Age X I.Q.	184.387	4	46.097	0.867	0.488
S-Within	3507.840	66	53.149		
Len	36.422	2	18.211	1.558	0.214
Age X Len	59.496	4	14.874	1.272	0.284
I.Q. X Len	90.349	4	22.587	1.932	0.109
Age X I.Q. X Len	113.012	8	14.126	1.209	0.299
Len X S-Within	1542.969	132	11.689		
Con	17.631	1	17.631	1.444	0.234
Age X Con	0.795	2	0.398	0.033	0.968
I.Q. X Con	9.450	2	4.725	0.387	0.681
Age X I.Q. X Con	4.770	4	1.192	0.098	0.983
Con X S-Within	806.043	66	12.213		
Len X Con	212.618	2	106.309	10.996	0.001
Age X Len X Con	31.011	4	7.753	0.802	0.526
I.Q. X Len X Con	20.047	4	5.012	0.518	0.722
Age X I.Q. X Len X Con	109.775	8	13.722	1.419	0.194
Len X Con X S-Within	1276.133	132	9.668		
Diff	152.345	1	152.345	9.636	0.003
Age X Diff	0.958	2	0.479	0.030	0.970
I.Q. X Diff	1.885	2	0.943	0.060	0.942
Age X I.Q. X Diff	157.145	4	39.286	2.485	0.052
Diff X S-Within	1043.406	66	15.809		

Table 4.6 (Continued)

Analysis Of Variance I.Q. Test Score Group / Age Group X Consistency Of Target Sequence Length & Difficulty

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Len X Diff	0.268	2	0.134	0.011	0.989
Age X Len X Diff	19.178	4	4.794	0.409	0.802
I.Q. X Len X Diff	62.986	4	15.746	1.343	0.257
Age X I.Q. X Len X Diff	128.225	8	16.028	1.367	0.217
Len X Diff X S-Within	1547.770	132	11.726		
Con X Diff	39.763	1	39.763	3.508	0.066
Age X Con X Diff	17.659	2	8.830	0.779	0.463
I.Q. X Con X Diff	9.936	2	4.968	0.438	0.647
Age X I.Q. X Con X Diff	26.345	4	6.586	0.581	0.677
Con X Diff X S-Within	748.176	66	11.336		
Len X Con X Diff	283.358	2	141.679	12.348	0.001
Age X Len X Con X Diff	28.353	4	7.088	0.618	0.651
I.Q. X Len X Con X Diff	50.818	4	12.704	1.107	0.356
Age X I.Q. X Len X Con X Diff	40.969	8	5.121	0.446	0.891
Len X Con X Diff X S-Within	1514.520	132	11.474		

Table 4.7

Descriptive Statistics Of Number Of Correct Responses For Age and I.Q. Test Score Groups

Age Group	N	MIN.	MAX.	MEAN	STD. DEV.	SKEW.	KURT.
1 Young	23	50.00	95.83	80.25	11.53	-0.961	0.683
2 Middle	24	50.00	91.67	75.17	11.88	-0.503	-0.371
3 Old	28	45.83	91.67	72.92	11.76	-0.584	-0.227

I. Q. Test Score Group	N	MIN.	MAX.	MEAN	STD. DEV.	SKEW.	KURT.
1 Low	25	45.83	87.50	70.83	11.28	-0.552	-0.485
2 Middle	25	50.00	95.83	75.17	13.90	-0.448	-0.865
3 High	25	66.67	95.83	81.67	7.70	0.040	-0.530

Table 4.8

Analysis Of Variance I.Q. Test Score Group X Age Group For Number Of Correct Responses

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	1573.694	2	786.847	6.617	0.002
Age	907.125	2	453.562	3.814	0.027
I.Q. X Age	396.364	4	99.091	0.833	0.509
S-Within	7848.250	66	118.913		

Table 4.9

Multiple Regression Analysis Of Number Of Correct Responses.

```
Analysis of Variance       N=75 out of 75

SOURCE             DF        SUM SQRS    MEAN SQR    F-STAT    SIGNIF

REGRESSION          2        0.23229     0.11614     10.893    0.0001
ERROR              72        0.76771     0.01066
TOTAL              74        1.00000

MULT R = 0.48196             R-SQR = 0.23229         SE = 0.10326


VARIABLE           PARTIAL   BETA WT     STD ERROR   T-STAT    SIGNIF

AH4 1              0.39666    0.38260     0.10435     3.6666   0.0005
AGE               -0.26475   -0.24309     0.10435    -2.3296   0.0226
```

low I.Q. test score group would make more errors on the difficult sequences, however, no evidence was found to support this.

Finally, since the experiment was self-paced, viewing times were recorded for each slide shown leading up to the final, critical slide which required the consistency/inconsistency decision. Total viewing times were compared within each sequence length, in an analysis of variance in the repeated measures design. Age and I.Q. test score group were between factor variables, and difficulty and consistency as the within factors. There was a significant overall effect of age group for each sequence length, with the oldest subjects taking longest to view the slides in each case. No significant effects of I.Q. test score group were found. There was a main effect of difficulty for the four frame sequences with 'easy' sequences requiring less viewing time than the sequences rated as difficult. Both the three frame and four frame sequences had main effects of consistency, in the case of the three frame sequences, this was due to sequences with consistent target slides requiring shorter viewing time than inconsistent sequences, however in the four frame sequences the opposite effect was found. There was no main effect of consistency for two slide sequences. Three and four frame sequences also showed a significant interaction between difficulty and consistency, again this was insignificant for two frame sequences. In the instance of four frame sequences, the difference in viewing times between easy and difficult sequences is more pronounced for consistent sequences than inconsistent sequences, with consistent sequences generally leading to longer viewing times for both easy and difficult sequences than their inconsistent counterparts. However the interaction for three frame sequences seems to occur because the difference between viewing times for consistent sequences is less with respect to easy and difficult sequences than for inconsistent sequences (Tables 4.10, 4.11, 4.12).

4.0.8 Discussion

The findings of this experiment provide evidence not only on the accuracy with which the elderly can draw inferences from picture sequences, but also on the speed with which they can deduce and use thematic information. The old group and the low I.Q. test score group make, as predicted, more errors in the inferences they draw than do the young and the high I.Q. test score group. This effect of I.Q. was expected, as a similar task forms part of the WAIS I.Q. test battery. Furthermore, the errors made by the subjects were not due to fast rates of presentation of the material to be understood, since the task was self-paced. From the distributions of the consistency judgements, it is clear that disagreements with the raters judgements were unaffected by the difficulty ratings assigned to the sequences. The chi-square distributions illustrated that the frequency of disagreements or 'errors' were similar for both easy and difficult rated sequences. These data imply that both the old and low I.Q. test score volunteers find it more difficult to construct themes from pictorial sequences using inferential reasoning. This fully supports the findings of Cohen (1979, 1981) and Cohen and Faulkener (1984) who found similar effects with both written

Table 4.10 (Viewing Times For Two Slide Sequences)

n=75

Analysis Of Variance   I.Q. Test Score Group / Age Group X Consistency Of Target & Difficulty

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	22.647	2	11.324	0.235	0.791
Age	634.322	2	317.161	6.594	0.002
I.Q. X Age	201.942	4	50.486	1.050	0.389
S-Within	3174.516	66	48.099		
Con	35.216	1	35.216	1.852	0.178
I.Q. X Con	11.044	2	5.522	0.290	0.749
Age X Con	60.524	2	30.262	1.592	0.211
I.Q. X Age X Con	16.203	4	4.051	0.213	0.930
Con X S-Within	1254.934	66	19.014		
Diff	121.095	1	121.095	14.581	0.001
I.Q. X Diff	22.800	2	11.400	1.373	0.261
Age X Diff	11.629	2	5.815	0.700	0.500
I.Q. X Age X Diff	11.791	4	2.948	0.355	0.840
Diff X S-Within	548.145	66	8.305		
Con X Diff	10.569	1	10.569	1.583	0.213
I.Q. X Con X Diff	5.709	2	2.855	0.427	0.654
Age X Con X Diff	39.661	2	19.831	2.969	0.058
I.Q. X Age X Con X Diff	22.653	4	5.663	0.848	0.500
Con X DiffX S-Within	440.758	66	6.678		

Table 4.11 (Viewing Times For Three Slide Sequences)

n=75

Analysis Of Variance   I.Q. Test Score Group / Age Group X Consistency Of Target & Difficulty

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	114.685	2	57.342	0.607	0.548
Age	962.896	2	481.448	5.093	0.009
I.Q. X Age	112.730	4	28.182	0.298	0.878
S-Within	6239.125	66	94.532		
Con	142.314	1	142.314	9.019	0.004
I.Q. X Con	19.007	2	9.504	0.602	0.551
Age X Con	25.386	2	12.693	0.804	0.452
I.Q. X Age X Con	28.431	4	7.108	0.450	0.772
Con X S-Within	1041.438	66	15.779		
Diff	525.473	1	525.473	24.792	0.001
I.Q. X Diff	6.731	2	3.366	0.159	0.854
Age X Diff	15.353	2	7.677	0.362	0.698
I.Q. X Age X Diff	11.379	4	2.845	0.134	0.969
Diff X S-Within	1398.875	66	21.195		
Con X Diff	323.861	1	323.861	23.658	0.001
I.Q. X Con X Diff	14.520	2	7.260	0.530	0.591
Age X Con X Diff	25.482	2	12.741	0.931	0.399
I.Q. X Age X Con X Diff	28.463	4	7.116	0.520	0.721
Con X DiffX S-Within	903.500	66	13.689		

Table 4.12 (Viewing Times For Four Slide Sequences)

n=75

Analysis Of Variance   I.Q. Test Score Group / Age Group X Consistency Of Target & Difficulty

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	38.624	2	19.312	0.258	0.773
Age	890.009	2	445.004	5.945	0.004
I.Q. X Age	212.926	4	53.232	0.711	0.587
S-Within	4940.688	66	74.859		
Con	608.778	1	608.778	31.802	0.001
I.Q. X Con	67.311	2	33.655	1.758	0.180
Age X Con	26.604	2	13.302	0.695	0.503
I.Q. X Age X Con	29.360	4	7.340	0.383	0.820
Con X S-Within	1263.438	66	19.143		
Diff	916.805	1	916.805	83.339	0.001
I.Q. X Diff	41.508	2	20.754	1.887	0.160
Age X Diff	32.213	2	16.107	1.464	0.232
I.Q. X Age X Diff	7.532	4	1.883	0.171	0.952
Diff X S-Within	726.063	66	11.001		
Con X Diff	287.449	1	287.449	23.576	0.001
I.Q. X Con X Diff	2.308	2	1.154	0.095	0.910
Age X Con X Diff	16.123	2	8.061	0.661	0.520
I.Q. X Age X Con X Diff	51.445	4	12.861	1.055	0.386
Con X Diff X S-Within	804.688	66	12.192		

and auditorally presented prose.

The difficulty may be due to working memory limitations. Older subjects and those of low I.Q. test score may i) forget the provisional continuation inferences concerning themes which they derive from earlier frames in a series as later frames come up for inspection, or ii) if the elderly delay making inferences until later in the series they may forget earlier frames and so base their conclusion on inadequate evidence. However, if a source of difficulty for the elderly and low I.Q. test score subjects was storing information, in working memory, then we would expect fewer 'errors' to occur in the shorter two slide sequences. This was not the case. In order to perform the task, the subject identifies the scene presented in the first slide and then, using knowledge stored in long term memory, derives a series of possible alternative inferences. These are held in working memory and, when the next slide is presented, some or none of these inferences are confirmed. If none are confirmed, the subject then has to refer back to the previous slide and switch between the two scenes presented in order to realise the correct thematical line that associates the two. Hence, information must be constantly updated and retained in working memory along with the knowledge of the processes which have been in operation on the material. Failures in control processes, and in the manipulation and retention of information, can occur because of loss of working memory capacity as suggested by Cohen (1979), Eysenck and Eysenck (1979) and Rabbitt (1981a).

Both response latencies and viewing times show age effects and indicate that the old take longer to make a decision on the consistency of the final target slide, and take longer to visually encode each of the slides in the preceeding sequence which led up to it. Pezdek and Miceli (1982) found a similar effect with young children (8 years) and older adults (65-78 years). These groups did not integrate pictures and sentences when they were presented individually at a rate of 8 seconds, but if the rate was increased to 15 seconds/ item cross-modality integration resulted. A slower rate of visual search may be responsible for this effect; both Rabbitt (1964) and Anders, Fozard and Lillyquist (1972) have shown that visual search becomes less efficient in the elderly especially when targets must be held in memory, or the number of targets in the stimulus array are varied (Rabbitt 1965b). In addition, the elderly are more likely to be distracted by irrelevant information (Rabbitt 1965a, 1965b, 1968), and this might also contribute to the longer viewing times necessary before adequate identifications can be made.

For all groups, the more thematical and contextual information was given, the faster became the decisions about the consistency of the final target slide. A possible explanation for this is that in the longer slide sequences, the theme is more likely to be more established and less ambiguous because of the increased context provided. Thus there is more chance for subjects to eliminate 'alternative' themes before the final slide appears. However sequences of three frames generated longer latencies for decisions about the final frame than both 2 and 4 frame sequences. Possibly this was because with these sequences subjects only had (the first) two frames to help them to establish a 'storyline', so that the 'theme' of the sequence was still ambiguous and unresolved at the moment when the third, 'target', frame appeared. In two frame sequences, 'themes' were necessarily more limited. The subjects' task was simpler, since they only had to either accept or reject the second frame as being a plausible sequel to the first. Moreover, in two frame sequences the load in working memory was very slight. Indeed it is arguable that subjects did not have to hold in memory a variety of possible themes, since a single frame usually could not on its own generate any 'theme' which might be followed through in a series. The subjects task for these two frame sequences thus amounted to briefly holding a single picture in memory; a very simple task indeed (Standing, Conezio, and Haber, 1970).

The within factor effects of the design gave the results predicted. Final 'target' slides which were consistent with the previously presented slides in their sequence gave faster decision times, as did the sequences in which this decision was rated as 'easy' to make by the panel of raters. Both results validate the ratings provided by the elderly raters. There were no significant interactions between these factors and age and I.Q. test score group.

The primary interest of this study was the effect of age and I.Q. test score on speed of inferential reasoning and overall performance in the task. Both age and I.Q. test score affect the ability to draw inferences, but I.Q. test score accounts for more of the variance in this task than does age. This is to be expected, given that similar tasks are used in I.Q. test score batteries. However, if the speed at which inferential reasoning occurs is considered, age is the main determinant. Age and I.Q. test score are thus shown to have separate effects in this task. I.Q. test score accounts for more of the variance than does age on the accuracy of inference, as seen in the linear regression analysis (Table 4.9), with the two factors accounting for 23.23% of the variance and AH4 part one having a partial correlation of 0.3966 compared with the partial correlation for age of -0.2647. However both age and I.Q. test score group have significant main effects on the number of 'correct' responses made by subjects (F=6.617 df=2,66 P<0.01, F=3.814 df=2,66 P<0.03) for I.Q. test score group and age group respectively. However age seems to account for more of the variance than I.Q. test score with regard to the speed of inference construction. In both analyses of variance, in which all response times were analysed (Table 4.5) or just 'correct' response times, where there was aggreement between the subject's consistency judgement of the final frame and that of the raters (Table 4.6), there were no significant effects of I.Q. test score group, but a significant main effect of age group in each case.

These data have obvious implications for everyday comprehension of television broadcasts and events. In paced tasks such as television news broadcast monitoring, it may be difficult for the elderly to make inferences because events happen so fast that they do not have time to do so. This is suggested by the fact that when, as in this task, the elderly are allowed to take as long as they like to inspect material and to draw inferences from it, they take much longer than the young. However, we see that when subjects are allowed all the time they need, age effects in accuracy of inference are weaker than those of I.Q. test score.

There are three separate limitations on performance on this type of task. First, there is the efficiency and capacity of the working memory system: if adequate working memory is not available available, then good inferences can not be made since both mental computations and the maintenance of the necessary material will be adversely affected. Second, as Cohen (1979, 1981) has shown, even if memory is apparently intact, there may be computational difficulties such that correct inferences can not be made. Finally, there is the additional factor of speed of computation and speed of decision, which were measured separately in this experiment by viewing times and decision latencies respectively. It is possible that older people can make correct decisions if they are allowed enough time. The data presented in this chapter can be interpreted as showing that while I.Q. test score primarily affects the first and second of these factors, age primarily affects the third. Hence, there is a useful distinction to be made between age and I.Q. test score effects in this sort of task.

The next chapter gives an account of accuracy of memory for factual and inferential information from photographs of both unfamiliar and very familiar events. This experiment was designed to assess whether the elderly and low I.Q. test scorers retain less factual information from scenes, and whether they are also less able to use information in long term memory to draw plausible inferences from pictorial information. In using more complex pictorial information, and a richer amount of detail in the material, the study seeks to assess how well the elderly comprehend and retain both implicit and explicit information contained in scenes.

CHAPTER FIVE

5 COMPREHENSION AND MEMORY FOR PICTURES

5.0.1 Introduction

In Chapter 4 we saw that as people grow older they may suffer from two separate kinds of comprehension problem, namely recognising the scenario in each picture, and then putting together changes in successive scenarios to deduce a theme.

Each scene or frame provides the perceiver with a number of different cues, which can be used to infer the location, time of year, country and general context of the scene. Television producers frequently use the minimum number of such cues. Comprehending television broadcasts requires the rapid formation of complex inferences from minimal cues, as does rapidly decoding and understanding everyday life, where prior knowledge is often assumed and little elaboration of contextual cues is offered.

The data from Chapter 4 suggested that the elderly are slower at performing inferential reasoning, and less able to draw correct inferences. It was suggested that this may be due to a reduced working memory capacity. Another possible contributing factor is examined here. The elderly may have impaired inferential reasoning because the information they hold in long-term memory is less well structured and has fewer associations. When a scene is presented the search through long-term memory is less efficient since there are fewer associations available to direct the search and also less information from which inferences can be drawn.

To comprehend and recall events effectively, rapid identification and inferential reasoning are necessary. Both of these processes are dependent on accessing information stored in long-term memory. When this information is used efficiently, a mere glance at a scene which has never been seen before generates a clear subjective impression, and an almost instantaneous perception and comprehension of what is being looked at (Beiderman, 1972). This automatic perception is hypothesised as being mediated by structures called frames (Minsky, 1975), which store stereotypical scenes and events. These lead to the relatively automatic comprehension of events and scenes. However, there are occasions when unique events and scenes are perceived. In these cases processing is assumed to be data-driven, with the requirement of greater cognitive resources (Friedman, 1979). If the elderly retain frames which are less rich with respect to their information content and their association network, then more data-driven processing will be required in order for comprehension to occur. This will necessarily put more demand on the cognitive resources available. Another plausible explanation is that the elderly will require more attentional capacity to maintain the appropriate script in the state of activation which is required for comprehension. This effort would impair their memory for such information (Light & Anderson,1975).

The experiment reported in this chapter considers the immediate recall of pictorial information as presented in photographs of everyday events. It is one of the few studies to use cued recall instead of recognition memory for an examination of the retrieval of two types of information from pictures i.e. information explicitly shown in the photographs, and information which could be plausibly inferred from the photographs. Group differences are examined for the two proposed information types, which are termed 'specific feature' and 'holistic' information. 'Specific feature' information refers to the details shown in the picture e.g. the red rose in the vase on the table. 'Holistic' information can be classified as the type of scene of which the picture is a particular representative case e.g. a dining room. It is the 'gist' or 'schema' information from which our subjective impressions of the scene are derived, and from which we comprehend what we are looking at.

Memory for pictures as investigated by recognition studies, overwhelmingly suggests that visual memory is both virtually unlimited, and also remarkably precise with regard to details (Shepard, 1967; Standing, Canezio, and Haber, 1970). Standing (1973) has estimated a total retention of 6600 items. However, it is likely that under free recall conditions or cued recall a much lower level of performance would be obtained. Friedman (1979) suggests that the high level of performance obtained in picture recognition studies is due to subjects not so much remembering actual details, but rather remembering an abstract theme. There are data to support this view that people generally remember just the themes of pictures from false recognitions occurring to distractors that are conceptually consistent with the themes of the target pictures (Bower and Glass, 1976; Jenkins, Wald and Pittenger, 1979; Mandler and Johnson, 1976).

It appears that the two types of information derived from pictures are extracted at different rates. People are able to identify the 'gist' or the 'schema' of a picture probably during the first eye fixation after the picture has appeared for as little as 100 msec. Loftus, Nelson, and Kallman (1983), used a two-alternative forced-choice paradigm, in which the target/distractor relationship was manipulated so that either 'holistic' or 'specific feature' was required for the correct response. They found that following one fixation on a picture, performance based on 'holistic' was superior to that based on 'specific feature' information. The reverse was true following sufficient time for studying the pictures. Hence, the two types of information seem to be acquired at different rates, most 'holistic' information is extracted at a glance (Bierderman, Rabinowitz, Stacy and Glass, 1974; Palmer, 1975), whereas subsequent fixations seek out specific informative features.

Friedman (1979) and Goodman (1980) have developed the 'script theory' from Schank and Abelson (1975) to provide a model for comprehension and extraction of different detail types from pictures. The general idea behind the theory is that when a picture is presented, a preformed schema (or frame) is quickly activated in long-term memory, after identification has occurred, specific details are extracted from the scene and stored in memory. The schema guides the search for these features, since the probability of a particular feature being selected for extraction and storage is a decreasing function of the 'a priori' probability that the feature already exists in the preformed schema. In this way, previously stored frames in long-term memory guide the search for unusual features in a scene which are unlikely to be present in the relevant frame. Specific details are extracted if they are considered to be unusual and these are most likely to be recalled. The remaining information i.e. that which includes the 'holistic' aspects of the scene, what the scene represents, the knowledge of the particular scenario, and the physical information such as spatial relationships is extracted from the picture initially and requires less processing.

In summary, this experiment looks for group differences in immediate memory for holistic and specific information extracted from photographs of everyday events. This is considered in relation to the elderly's ability to draw correct inferences and comprehend everyday scenes. It is predicted that the elderly and the low I.Q. test scorers will be poor at recalling both the explicit and implicit information, showing that not only are their factual memories poor at recalling visual information, but also that they are less able to draw plausible inferences from the scenes they encounter. In this way they are easily confused by idiosyncratic events. An explanation is sought in terms of a frame theory of perception.

5.0.2 Method

5.0.3 Subjects

150 subjects were selected from the same pool of volunteers as used in previous studies (Experiments 2.1, 2.2, 3.1, 3.2, 4.1). They were selected according to their age and AH4 part one test score, to fill three AH4 test score groups of 'high', 'middle' and 'low' and three age bands, in the same way as for the previously reported experiments. The composition and group criterion of the 3X3 matrix are shown in Table 5.1.

5.0.4 Materials

Fifteen black and white pictures were selected from a photography book together with a collection of photographs used as conversation pieces in schools, 5 of which are printed in Appendix 6. The selection sought to achieve a diversity of subjects with regard to events, countries and dates. Slides were made of the chosen photographs.

5.0.5 Apparatus

A Kodak Carousel 'S' projector fitted with a 100mm lens was used to display the slides, and a Smith's stop clock timed events in seconds.

Table 5.1

3x3 Matrix Of Subjects For Age and AH4 Test Score Groups

AH4 Test Score Group Age Group	Low Group 1	Middle Group 2	High Group 3	Total Mean (x) Std. Dev. Range (r)
Young Group 1	13	17	16	N=46 x=56.02 std.dev.=2.02 r=52-59 years
Middle Group 2	16	19	18	N=53 x=64.07 std.dev.=2.58 r=60-69 years
Old Group 3	17	18	16	N=51 x=72.74 std.dev=3.47 r=70-85 years
Total Mean (x) Std. Dev. Range	N=46 x=16.65 std.dev.=3.89 r=2-20/65	N=54 x=30.15 std.dev.=5.88 r=21-39/65	N=50 x=46.68 std.dev.=5.58 r=40-61/65	N=150 x=31.53 64.55 std.dev.=13.19 7.28 r=2-61/65 53-83 years

5.0.6 Procedure

Subjects were tested in groups of approximately twelve. Four random sequences were generated for the presentation of the 15 photographs. The sequences were designated to groups so as to achieve similar numbers of subjects for each sequence. The photographs were projected onto a large white screen at a angle of approximately 45 degrees above eye level. Each photograph was shown for 15 seconds. Subjects were required to answer a series of questions concerned with each picture immediately after the slide was removed from the screen. The questions were printed in a booklet. Subjects turned the pages after each picture was presented. They did not see the questions about each picture until after it was presented.

There were six questions for each photograph. Three required the retrieval of information presented explicitly in the photograph (e.g. What was the boy carrying?, How many chairs were there in the picture?) and three required the subjects to draw inferences (e.g. Which country was the photograph taken in?, What event were the people attending?):see Appendix 6. All of the questions required a few words of response, and the subjects were given 90 seconds to answer the questions, which seemed long enough for pacing not to be a relevant factor since subjects easily completed answering the questions if this was not the case it was due to the subjects inability to answer the questions. The booklets were arranged in the sequence of presentation.

5.0.7 Scoring

Three undergraduate raters were given black and white photographs of the fifteen pictures used in the study. Each marked all the responses of the 150 subjects. A mark of one was given for a correct response, and zero for an incorrect response. If the subjects had made no attempt to answer the question, an 'X' was allocated.

The agreement between the raters was measured using Cohen's (1960) Kappa statistic: the smaller the value of Kappa obtained by this method, the greater the disagreement between raters. Kappa values were calculated for each individual question using the ratings assigned for all the subjects. In addition, mean values were calculated for each story, and for each set of implicit and explicit fact questions across all stories. The data are shown in Table 5.2. It can be seen that for the majority of questions the Kappa values were very high, suggesting that the raters' closely agreed with each other on their categorisations of the subject's responses.

5.0.8 Results

The numbers of correct answers, incorrect answers and missing responses, on the basis of the raters' judgements, were computed for overall performance. These were then broken down to give the same indices for factual questions and questions of inference. The means of all these are shown in Table 5.3. A paired t-test was computed to

Table 5.2

Mean Kappa Values For Each Picture From Three Raters

Picture	Mean Kappa
1	0.733
2	0.765
3	0.707
4	0.694
5	0.589
6	0.474
7	0.701
8	0.734
9	0.516
10	0.731
11	0.703
12	0.507
13	0.699
14	0.672
15	0.639

Table 5.3

	N	MINIMUM	MAXIMUM	MEAN	STD DEV	SKEWNESS	KURTOSIS
ALL DATA							
NO. RESPONSES	150	26.000	90.000	79.867	9.9843	-1.839	5.118
% CORRECT	150	14.444	96.667	70.467	14.958	-1.208	1.638
NO. MISSING	136	1.0000	64.000	11.176	9.9136	1.868	5.255
% INCORRECT	133	1.1111	52.222	11.988	9.1298	1.788	4.161
FACTUAL QUESTIONS							
NO. RESPONSES	150	19.000	45.000	41.180	3.8722	-1.996	6.727
% CORRECT	150	17.778	93.333	67.793	13.163	-1.102	1.711
NO. MISSING	124	1.0000	26.000	4.6210	3.7990	2.134	7.441
% INCORRECT	124	6.6667	62.222	24.158	9.2907	.930	1.935
INFERENTIAL QUESTIONS							
NO. RESPONSES	150	7.0000	45.000	38.687	6.7716	-1.671	3.323
% CORRECT	150	11.111	100.00	73.141	18.415	-1.159	1.057
NO. MISSING	124	1.0000	38.000	7.6371	6.7346	1.634	3.150
% INCORRECT	124	2.2222	51.111	13.728	9.1692	1.732	3.626

compare performance on the two question types. There was a significant difference between the means ($t=5.75$, $p < 0.01$), the mean number of correct answers to inference questions being significantly higher than for factual questions.

Group differences were considered in terms of age and I.Q. test score band. An analysis of variance was computed and main effects of age, I.Q. test score group and question type were found. The elderly subjects and the low I.Q. test score group recalled fewer correct details than the other groups, and overall more correct answers were given to inference questions than to factual questions. There was also a significant interaction between I.Q. test score group and question type (Fig 5.1; full summary shown in Table 5.4). The interaction apparently occurs because the difference in percentage recall for the two question types is greatest for the high I.Q. test score group and least for the low I.Q. test scorers.

The percentage 'error' score and percentage of missing answers were also analysed with respect to question type (i.e. factual or inferential), and age and I.Q. test score group. An analysis of variance was performed with age and I.Q. test score as between factor variables and question type as the within factor variable. For the percentage 'error' score, there were significant main effects of all three factors, the elderly and low I.Q. test scorers had the highest percentage of errors, and there were significantly more errors made on factual questions than inference questions. None of the interaction

Fig. 5.1

Interaction Between Question Type & I. Q. Test Score Group.

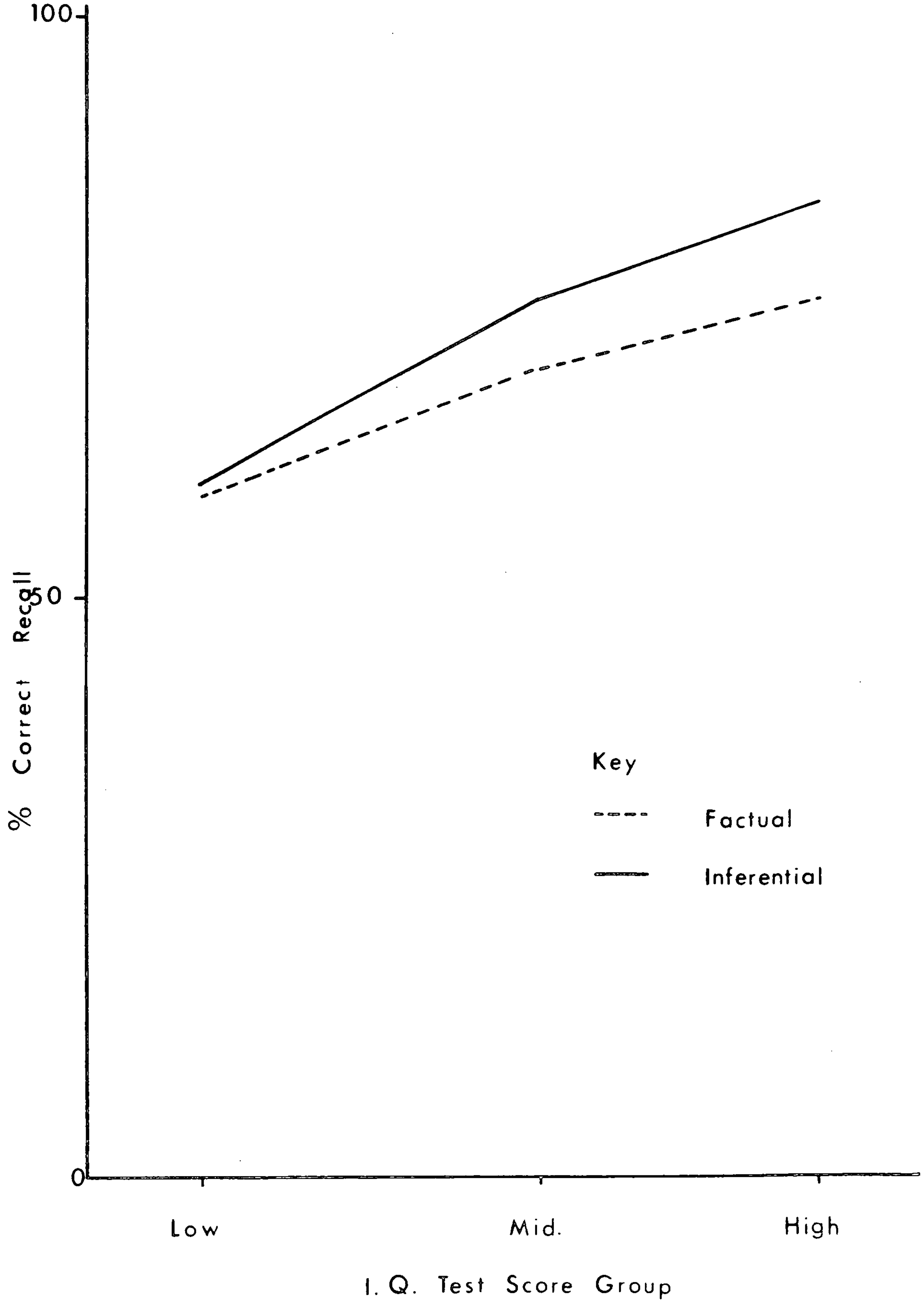


Table 5.4

Analysis Of Variance   I.Q. Test Score Group / Age Group X Question Type

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	19770.766	2	9885.383	39.055	0.001
Age	4408.957	2	2204.479	8.709	0.001
I.Q. X Age	398.004	4	99.501	0.813	0.813
S-Within	33411.000	132	253.114		
Question	1691.518	1	1691.789	26.801	0.001
I.Q. X Question	780.552	2	390.276	6.184	0.003
Age X Question	92.739	2	46.369	0.735	0.482
I.Q. X Age X Question	88.875	4	22.219	0.352	0.842
Question X S-Within	8331.000	132	63.114		

terms were significant. The missing data analysis also had main effects of age, I.Q. test score group and question type, again the old and the low I.Q. test scorers left more missing answers than the other groups. The significant main effect of question type shows that there were more missing responses to inference questions than factual questions, even though inference questions had a higher percentage correct response score than factual answers which generated more incorrect responses. In addition the interaction between I.Q. test score group and question type was significant, this is probably due to the difference between missing responses being greatest for the low I.Q. test score group and least for the high I.Q. test scorers (see Tables 5.5, 5.6, and Figures 5.2, 5.3).

A final analysis of errors and missing respones, combined the error score and missing score for each question type and considered this to represent an overall percentage error score. Again, an analysis of variance was computed with age and I.Q. test score group as between subject factor variables and question type as the within factor variable. There were significant main effects of all three factors. As before percentage error score was highest in the low I.Q. test score group and for the elderly subjects. The effect of question type shows that more errors occur for the factual questions than the inferential ones. None of the interaction terms were significant (Table 5.7).

A multiple regression analysis was performed, with overall number of correct responses as the dependent variable and age and AH4 score on both parts one and two as the predictors. These indices accounted for 47.0%

Table 5.5

Analysis Of Variance For Percentage Errors I.Q. Test Score Group / Age Group X Question Type

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Age	1571.627	2	785.813	5.424	0.005
I.Q.	6575.992	2	3287.996	22.697	0.001
Age X I.Q.	643.596	4	160.899	1.111	0.354
S-Within	19846.500	137	144.865		
Question	5371.387	1	5371.387	99.736	0.001
Age X Question	17.528	2	8.764	0.163	0.850
I.Q. X Question	293.604	2	146.802	2.726	0.069
Age X I.Q. X Question	103.794	4	25.948	0.482	0.749
Question X S-Within	7378.250	137	53.856		

Table 5.6

Analysis Of Variance For Missing Responses I.Q. Test Score Group / Age Group X Question Type

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Age	398.392	2	199.196	6.472	0.001
I.Q.	1181.044	2	590.522	19.186	0.001
Age X I.Q.	39.663	4	9.916	0.322	0.813
S-Within	4216.582	137	30.778		
Question	510.879	1	510.879	55.702	0.001
Age X Question	4.733	2	2.366	0.258	0.003
I.Q. X Question	270.965	2	135.483	14.772	0.482
Age X I.Q. X Question	20.015	4	5.004	0.546	0.842
Question X S-Within	1256.508	137	9.172		

Fig. 5.2

Mean Number Of Missing Responses By Question Type.

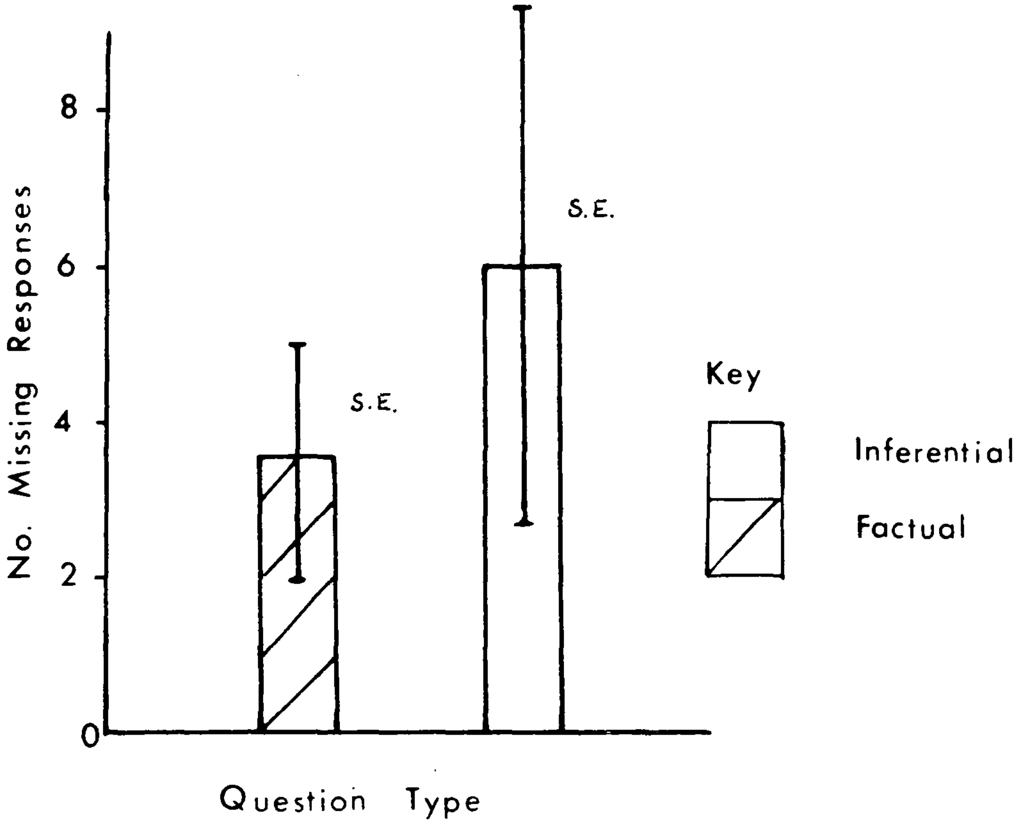


Fig. 5.3

Interaction Between I.Q. Test Score Group & Question Type.

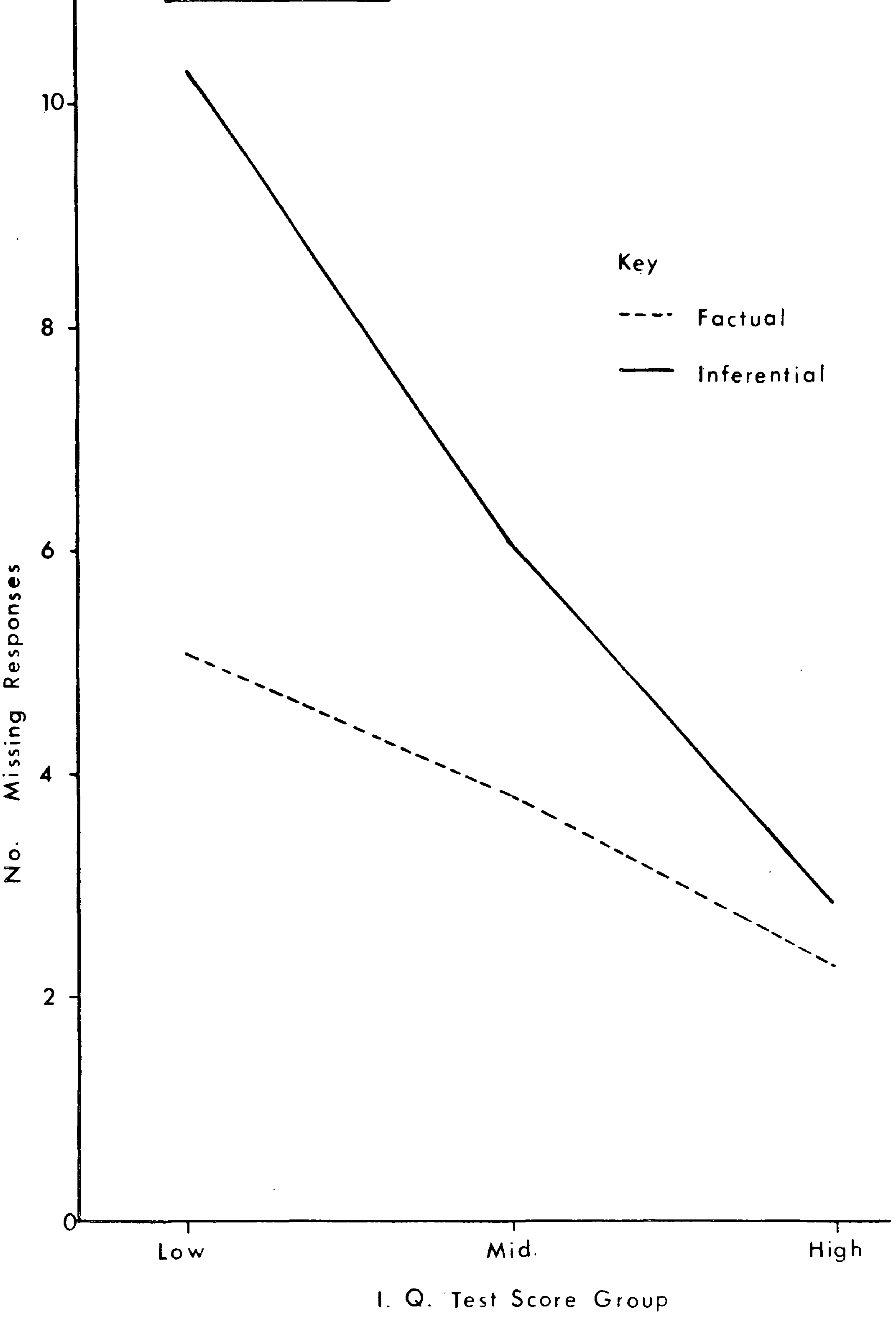


Table 5.7

Errors = Missing & Incorrect Responses

Analysis Of Variance For Percentage Errors I.Q. Test Score Group / Age Group X Question Type

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Age	7039.199	2	3519.600	9.552	0.001
I.Q.	24755.934	2	12377.965	33.592	0.001
Age X I.Q.	163.800	4	40.950	0.111	0.978
S-Within	50481.625	137	368.479		
Question	531.818	1	531.818	5.832	0.017
Age X Question	18.089	2	9.045	0.099	0.906
I.Q. X Question	382.927	2	191.463	2.099	0.126
Age X I.Q. X Question	101.735	4	25.434	0.279	0.891
Question X S-Within	12493.875	137	91.196		

of the variance, of which the AH4 part two accounted for 41.0%, with the inclusion of age increasing this to 44.7%. These two indices likewise account for 39.9% of the variance for fact retrieval and 39.7% for inference retrieval, with AH4 part two score accounting for 37.9% and 36.6% of the variance respectively (Table 5.8, 5.9, 5.10).

5.0.9 Discussion

The study tests directly memory for pictorial details, as well as recall of implicit features which go beyond what is shown explicitly. In this way, the elderly were assessed on their comprehension of the pictures, their understanding of the context of the picture, and the event it illustrated.

Studies using linguistic material clearly show that memory for gist is likely to be more complete and longer lasting than recall for specific details (Bartlett,, 1932; Cofer,1977; Dooling and Christaansen,1977; Frederiksen,1975; Kintsch,1976). Friedman (1979) has suggested that the same is probably true for pictures. However, the available data in support of this notion is based upon false positive responses to distractors which are conceptually consistent with the theme of the target pictures in a recognition task (Bower and Glass,1976; Jenkins, Wald and Pittenger,1977; Mandler and Johnson,1976; Mitterer and Rowland,1975). These data have to be considered as circumstantial only in their support of this notion. However, this study is perhaps the first to provide convincing support for the superior recall of 'gist' and

Table 5.8

Stepwise Regression Analysis Of Number Of Correct Responses.

```
Analysis of Variance       N=141 out of 150

STEP 1
......

SOURCE            DF         SUM SQRS    MEAN SQR    F-STAT    SIGNIF

REGRESSION         1         10030.0     10030.00    99.677    0.0000
ERROR            139         13987.0       100.36
TOTAL            140         24017.0

MULT R = 0.64624             R-SQR = 0.41762         SE = 10.031


VARIABLE          PARTIAL   COEFFICIENT    STD ERROR    T-STAT    SIGNIF

CONSTANT                       44.172       2.1164      20.871    0.0000
AH4 2             0.64624       0.649       0.0650       9.984    0.0000


REMAINING         PARTIAL    SIGNIF

AH41              0.16413    0.0526
AGE              -0.22562    0.0074

STEP2
.....

SOURCE             DF        SUM SQRS    MEAN SQR    F-STAT    SIGNIF

REGRESSION          2        10742.0     5371.00     55.834    0.0000
ERROR             138        13275.0       96.19
TOTAL             140        24017.0

MULT R = 0.66878             R-SQR = 0.44727         SE = 9.8079

VARIABLE          PARTIAL   COEFFICIENT    STD ERROR    T-STAT   SIGNIF

CONSTANT                     67.032        8.6536        7.7461  0.0000
AH4 2              0.5989     0.590        0.0671        8.7853  0.0000
AGE               -0.2256    -0.327        0.1203       -2.7205  0.0074

REMAINING         PARTIAL   SIGNIF

AH4 1              0.2082   0.0139
```

Table 5.9

<u>Stepwise Regression Analysis Of Number Of Correct Responses To Factual Questions</u>

<u>Analysis of Variance</u>   N=141 out of 150

```
STEP 1
......

SOURCE           DF       SUM SQRS   MEAN SQR   F-STAT   SIGNIF

REGRESSION        1        1765.2     1765.20   84.977   0.0000
ERROR           139        2887.4       20.772
TOTAL           140        4652.6

MULT R = 0.61595          R-SQR = 0.37940       SE =  4.5577


VARIABLE         PARTIAL   COEFFICIENT   STD ERROR   T-STAT   SIGNIF

CONSTANT                       22.511      0.9616    23.410   0.0000
AH4 2            0.61595        0.272      0.0295     9.218   0.0000


REMAINING        PARTIAL    SIGNIF

AH41             0.09222    0.2785
AGE             -0.17824    0.0351

STEP2
.....

SOURCE             DF     SUM SQRS   MEAN SQR   F-STAT   SIGNIF

REGRESSION          2      1856.9     928.45    45.831   0.0000
ERROR             138      2795.6      20.258
TOTAL             140      4652.6

MULT R = 0.63176          R-SQR = 0.39912       SE = 4.5009

VARIABLE         PARTIAL   COEFFICIENT   STD ERROR   T-STAT  SIGNIF

CONSTANT                      30.716      3.9712      7.7346  0.0000
AH4 2             0.56989      0.251      0.0308      8.1471  0.0000
AGE              -0.17824     -0.117      0.0552     -2.1279  0.0351

REMAINING        PARTIAL   SIGNIF

AH4 1             0.1243   0.1450
```

Table 5.10

Stepwise Regression Analysis Of Number Of Correct Responses To Inferential Questions

```
Analysis of Variance        N=141 out of 150

STEP 1
......

SOURCE             DF         SUM SQRS    MEAN SQR    F-STAT    SIGNIF

REGRESSION          1          3379.8      3379.8     80.169    0.0000
ERROR             139          5860.0        42.158
TOTAL             140          9239.2

MULT R = 0.60480              R-SQR = 0.36579         SE =  6.4929


VARIABLE           PARTIAL    COEFFICIENT    STD ERROR    T-STAT    SIGNIF

CONSTANT                          21.662       1.3699      15.812    0.0000
AH4 2              0.60480         0.376       0.0421       8.954    0.0000


REMAINING          PARTIAL     SIGNIF

AH41               0.18884     0.0254
AGE               -0.22345     0.0080

STEP2
.....

SOURCE              DF        SUM SQRS    MEAN SQR    F-STAT    SIGNIF

REGRESSION           2         3672.4     1836.2      45.514    0.0000
ERROR              138         5567.4       40.344
TOTAL              140         9239.8

MULT R = 0.63044              R-SQR = 0.39745         SE = 6.3517

VARIABLE           PARTIAL    COEFFICIENT    STD ERROR    T-STAT   SIGNIF

CONSTANT                         36.316       5.6041       6.4802  0.0000
AH4 2               0.55279       0.339       0.0435       7.7926  0.0000
AGE                -0.22345      -0.210       0.0779      -2.6930  0.0080

REMAINING          PARTIAL    SIGNIF

AH4 1               0.2334    0.0057
```

'inferential' information from pictures, in comparison with the recall of factual or specific detail information. The experiment is also important since it uses a cued recall test of memory, instead of the commonly used recognition paradigm for picture memory. Although this task was more difficult, with fewer cues provided than in a recognition test, the difference between recall of the two information types extracted from pictures comes through clearly.

The group differences were as expected for recall of both explicit and implicit information. The young and the high I.Q. test score groups obtained significantly higher percentage recall than the old and low I.Q. test score groups. The main point of interest in these data is the significant interaction between I.Q. test score group and the type of information recalled. The difference in recall performance for 'gist' and factual information is greatest in the high I.Q. test score group, and is insignificant in the low test score group. This means that not only are the low I.Q. test score group poor at recalling the explicit facts, which are notably the most difficult to remember, but they are equally poor at recalling the implicit information derived from the picture:- a simpler task. This suggests that this group have difficulty comprehending the picture presented, since they were unable to identify the scene and retrieve the relevant information from long-term memory in order to draw plausible inferences. The trend is similar in the old subjects, although the interaction was not significant.

If we consider the stepwise regression analysis in which AH4 test scores and age are assessed as predictors of overall performance on this task, we find AH4 part II accounts for 41.7% of the variance, and this is then increased to 44.7% when age is included. It is clear from this that the AH4 part II test of spatial intelligence accounts for more of the variance in this task than any other of the measures of individual differences taken. Further, it is age and not AH4 part I test score which picks up most of the remaining variance. The strength of the spatial intelligence task of AH4 part II as a predictor of performance is unusual given the high correlations obtained with AH4 part I and the large amount of variance it accounts for on some of the previous measures taken. However, one of the dominant cognitive processes in picture comprehension is the interpretation of spatial relationships, and the manipulation of these relationships in order to mentally test possible inferences. This may well explain the strength of this index as a predictor of memory for pictorial information.

This finding has clear implications for the explanation of group differences in the cognitive processing of everyday events. The younger subjects, and those of high I.Q. test score, are better able to encode and retrieve both specific and inferential information from pictorial representations of events. The group differences are, however, most pronounced in the simpler task of recalling inferential information. The difference can be explained in two ways. Firstly, the low I.Q. test score group and the elderly may just be poor at inferential reasoning, and indeed Cohen (1981) has provided evidence to support this. Secondly, the group differences may be due to a difference in the ability to access relevant information in long-term memory, and the quality of the information available in memory.

The analysis of the number of error responses and missing data responses indicated that the low I.Q. test score subjects and the older subjects are likely to make more of both. It appears that they have difficulty accessing the appropiate material in long-term memory, or that the material which is activated is impoverished, in the sense that it lacks sufficient associations and data for these groups to attempt to form plausible inferences. The interaction between number of missing responses and I.Q. test score group, is probably caused by the low I.Q. group either having so little confidence in the inferences they are able to construct that they do not write them down, or that they are simply unable to draw inferences and comprehend the photographs, on the basis of the information stored in their long-term store. When they attempt to construct inferences, they are frequently incorrect, either because of their impoverished information store or because of poor inferential reasoning ability, as shown by Cohen (1979, 1981) and Light and Anderson (1975).

The proposed group differences in the access and retrieval of the relevant information from long-term memory will have two possible effects on subject's comprehension. First, the search for the relevant schema to enable identification to occur, will be prolonged. If the elderly subjects and those with low I.Q. test scores have fewer 'slots' (Minsky,

1975) filled with 'specific' instances of actions and schemes, as well as other data, then the search for the relevant schema will be poorly directed, since there are fewer identifying features available in long-term memory to aid the matching process between presented and stored information. Second, after identification has taken place, the number of 'slots' indicating associations within and between schemata will be less than for the young and high I.Q. test scorers. This means that there will be fewer expected items generated from the activated schema, and more data-driven, specific feature analysis will be necessary. Friedman (1979) suggests that this is costly in terms of cognitive resource, and therefore fewer resources will be left available for remembering inferential and factual information.

The emphasis on the comprehension and memory of scenes in this study, implies that it has direct relevance to cognitive experiences in everyday life. The identification and interpretation of visually presented scenes is a frequent cognitive demand for the comprehension of what is going on. Hence, we spend much of our time identifying scenes and drawing inferences from the component items. A typical example is exploring an unfamiliar town, inference leads us to expect that the banks, and the post office are near the town centre or market square, identification confirms or rejects this inference for the particular town in question. By accessing information in long-term memory concerning the general structure of towns, a set of expectations is available for the locations of the main buildings, these guide our exploration of the new town, with the process of identification confirming or rejecting the inferences we have formed.

Further experiments are needed to test the suggestion that the elderly possess schemata which are impoverished, and incorporate fewer embedded associations, and to probe the effect of this on search and access in long-term memory. Nevertheless, this explains the longer processing times required for comprehension of picture story sequences seen in Chapter 4, and the effect found by Cohen (1979) that written text produces better performance than spoken text. The inferential reasoning of the elderly is further impaired by a decrement in the working memory capacity (Light and Anderson, 1975: see Chapter 4), which is particularly important when information requires manipulation in memory.

The following chapters investigate the access of information in long-term memory, and the nature of the schemata stored there.

CHAPTER SIX

6 NAMING LATENCIES AND MATCHING LATENCIES WITH PICTURES

6.1 INTRODUCTION

A characteristic of television broadcasts is the speed with which information is conveyed to its audience. In order to comprehend the programme, the viewer is required very rapidly to identify the scenes presented and then draw the correct inferences as one scene leads to the next. Identifying people and scenes requires very fast access and retrieval of information from long term memory, since the viewer must frequently draw on the 'knowledge about the world' which he has built up over many years and, presumably, stores in his 'long-term memory' in order to comprehend the information which is being presented to him from moment to moment. Lachman and Lachman (1980) identify at least three components of 'knowledge actualization'. Retrieval is the principal component: knowledge is obviously useless in this context if it can not be brought to consciousness. Inference is another important component, involving the ability to synthesise new information from existing information. Finally, metamemorial control processes have been identified as mechanisms which direct efforts to retrieve and draw inferences by assessing whether or not a particular piece of knowledge is available in memory, how likely one is to access it, and how reliable the particular piece of information is as a basis for any inference that the situation may demand.

This chapter examines the efficiency of retrieval of information from long-term memory in the elderly, as a process which is necessary before effective comprehension of television broadcasts can occur. However, this can also be considered in much broader terms, since retrieval from long term or permanent memory is essential for effective functioning in the daily contingencies of everyday life. It is necessary to retrieve one's knowledge of the world both distant and immediate in order to cope with the current situation and to predict what is going to happen next. Retrieval of the appropriate knowledge from long term memory is usually automatic, unconscious, and does not appear to depend on any strategies, or require effort. It also seems to proceed without interference from secondary tasks. However, there are some occasions when retrieval appears to become unusally effortful and demanding. These occasions, as when one gropes for a proper name or a word, or vainly searches for a piece of information which one knows is perfectly familiar, become more frequent as old age advances. The elderly commonly report that they are plagued by 'tip of the tongue' experiences, and the feeling of knowing, with the inadequacy of retrieval.

The effect of this reduced retrievability may require the elderly to develop compensatory strategies. Hence, they come to rely more and more on drawing inferences in order to construct the facts which they often know to be in their store but which are, for the moment, inaccessible. This undoubtedly takes more time than direct retrieval (Haviland and Clark, 1974), so their rate of information processing is reduced. They may take longer to retrieve information and to make useful inferences, and they become more error prone.

Another characteristic of the elderly which affects the efficiency of their access to stored information is that they tend to be cautious (Botwinick, 1973): this leads to them wasting time construing facts by inference, when often they could have successfully accessed them directly had they been more confident of the contents of their memory store and the efficiency with which they could access information within it. It appears that the elderly are much less efficient than the young at using the useful 'data base' which they have built up over many years. This occurs because of a reduced ability to retrieve information, greater dependency on inference drawing, and poorer metamemorial control processes.

To investigate possible age changes in the efficiency of the retrieval process from long term memory, we carried out two experiments which compared groups of people aged 50 to 85 on their ability to retrieve names of objects, and to match a given name to one of four pictures. Picture naming and matching seemed an appropriate task since Oldfield and Wingfield (1965) have shown it to be a useful paradigm for investigating retrieval systems and the organization of memory (Oldfield, 1966). The task is additionally attractive because, although it is apparently simple, it permits generalization to everyday memory tasks with retrieval from long term memory as a major component.

It has already been established that as people grow old, they find it harder to retrieve words from their vocabulary, (Thomas et al., 1977) but their vocabulary score appears to remain reasonably stable (Botwinick, 1967; Doppelt and Wallace 1955). This suggests that while the size of their memory data base (i.e. vocabulary) may be unimpaired, they begin to find difficulty accessing the material within it. This has been supported by recent results obtained by Rabbitt (1984a), who found that vocabulary test scores (i.e. Mill Hill parts I and II) apparently remain unchanged until very late in life. This allows vocabulary test scores to be used to estimate 'youthful' scores on age sensitive tasks. These findings make name retrieval an interesting area for investigation with an elderly population. There is undoubtedly a relationship between verbal intelligence and vocabulary score. Perfetti (1983) compared adults of different verbal intelligence on various tasks, and found that name retrieval was a rate-limiting process, suggesting that verbal intelligence is associated with an efficient word access and retrieval system.

All the subjects used in these experiments had been screened on the AH4 test of general intelligence, the Mill Hill vocabulary scale and their ages were known. This allowed us to examine efficiency of retrieval of words from memory (names) as a function of these individual differences.

Both of the tests used were graded clinical tests which were adapted to form a naming latency and a matching latency task. The McKenna and

Warrington (1983) Graded Naming Test allowed us to compare naming latencies for the recall of names cued by pictures. This test is designed to detect vocabulary loss, or aphasia, following neurological damage. It uses 30 pictures of relatively unfamiliar objects, the (rarely used) names of which may be vulnerable to word retrieval difficulties (Rochford and Williams, 1965 ; Newcombe, Oldfield and Wingfield, 1965), sampling items on the boundary of the individual's naming vocabulary. Since the test is graded it permits the comparison of different age and I.Q. test score groups for correct response time with respect to item difficulty. The slope and the intercept of a response time against difficulty plot will give estimates of the memory search time, and all other processes, such as encoding and response execution, respectively.

The British Picture Vocabulary Scale (B.P.V.S.) (Dunn, Dunn, Whetton and Pintille 1982) allowed us to examine the times our volunteers took to select which of 4 pictures simultaneously presented on a single slide matched a word (name) read out by the experimenter. It was therefore a forced-choice matching, or recognition task. The B.P.V.S., which is sometimes used as a simple test of intelligence, is also a graded test, in which there is a gradual progression from 'easy' items which can be matched by nearly all humans taking it, to 'difficult' items which can be matched only by a few people with unusually large vocabularies. Response time plotted against item difficulty will in this instance, give a slope reflecting speed of memory search and match for names, with an intercept which may reflect time taken for encoding the pictorial information and for response production. The effects on these indices of vocabulary size (from Mill Hill scores), I.Q. (from AH4 scores) and age could be investigated independently in these studies.

Norman and Rumelhart (1970) have proposed a model for perception and memory which includes a naming system. The model supposes that information is initially represented as multicomponent vectors by the perceptual system, and is then translated by the naming process into different multicomponent vectors of attributes in memory. Into this model they incorporate a similar relationship to that postulated by Sternberg for search through short-term memory i.e. that names are retrieved by sequential search through a mental lexicon or dictionary.

A different model has been proposed by Shiffrin (1970), who suggests that latency should be a linear function of the number of draws made into a search set prior to response output. A draw is a search into the set by which information is extracted and used as the basis for subsequent decisions. According to Shiffrin memory search follows the following procedures: first, a search set is selected; this may be based upon such factors as the stimuli provided, cues available, and the nature of the task. A draw is then made from this set. The information recovered by the draw can then be used to guide the subsequent decisions. These include whether to give a response, which response to give, or whether to continue the search. If a decision to continue the search is made, then the recursor loops back to the executive decision maker for a possible selection of a new search set, before the process continues and the next draw is started.

Oldfield (1966) derived a theoretical relationship between naming response latency and memory ensemble size which was subsequently tested by Briggs and Swanson (1969). Oldfield suggested that Zipf's law (Zipf, 1953) provides a critical insight into the nature of the organisation of human memory. This states that in a large sample of words there is an inverse linear relationship between the logarithm of the numbers of words with the frequency of occurrence 'f' and log 'f'. Oldfield made the simple observation that if we possess vocabularies of 75,000 words, which is a plausible estimate, then a serial search procedure would take an implausibly long time. Search must proceed by successive choices which progessing restricts the search size set, i.e. in terms of a search 'tree'. Oldfield suggested that search through this tree was carried out by successive dichotomisations or 'choices' between 'ensembles' of words. He felt that organisation in terms of ensembles of different relative frequencies was a plausible model, so that, memory for object names can be seen as organised into ensembles of names with the largest ensemble holding the least commonly used names and smallest ensembles for the most commonly used names. Memory can be conceived as involving first, a selection between ensembles and then a dichotomous search within the selected ensemble. This means that object naming proceeds at a rate well within the empirical estimates of skill limits found with linguistic material.

In our studies naming and matching latencies were analysed to provide evidence for choices between these models. Of particular interest were the instances where the same words were correctly retrieved by volunteers in all groups. In this case, differences in response time should show differences in search strategy since different groups are searching for words present in all sets, but in different set sizes.

Lorge (1956) has stressed the influences of changes in the speed of functioning with age, and differentiated between 'speed' or 'performance' tests and 'power' tests. Lorge's main conclusion was that the former showed a decrement with age, whilst on 'power' tests subjects up to the age of 60 years did not demonstrate any decline in intellectual ability. Waugh and Barr (1980) also emphasise that most observed age-related decrements in cognitive function, and specifically in memory, can be accounted for by age-related decrements in speed of performance. This has been found in a paired associate task. When given sufficient time the elderly have performed as well as their younger counterparts (Monge and Hultsch, 1971 ; Treat and Reese, 1976). However the instances of speed only effects are limited and there is plenty of evidence to suggest that speed is not the only factor accounting for the decrements observed. The design of these experiments allows both speed and power factors to be assessed, speed by latency measurement amd power by the number of correct responses.

As mentioned previously, the elderly show a tendency to be cautious, although Lachman, Lachman and Thronesbery (1979) found them to be reasonably accurate at assessing their performance levels in comparison with younger subjects. The naming latency task looked at this from a slightly different view point. Subjects were given three attempts to name the picture correctly, but were also permitted to abort a trial if they considered that they did not have the name in their store, or they were unable to retrieve it. Hence data were obtained on how long it took subjects to abort trials, and how many attempts they needed to retrieve the correct response, or before aborting the trial.

To summarise, two studies investigate the effect of age and I.Q. test score on the ability to retrieve and match names accessed in long term memory. Models of intelligence are considered to explain group differences in performance on these tasks. The power and speed of the tests are considered as factors influencing age and I.Q. test score effects, and the elderly's metamemorial ability is assessed by their use of the 'abort' option, and the number of attempts subjects make to reach the correct response.

6.2 EXPERIMENT 6.1

6.2.1 Subjects

96 volunteers, aged 52 - 85 years, mean age 64.42 years, were selected on the basis of their age and AH4 part one test score. Both of these factors were divided into three bands, and subjects were fitted into a 3X3 matrix as shown in Table 6.1. All the subjects had previously

Table 6.1

3x3 Matrix Of Subjects For Age and AH4 Test Score Groups

AH4 Test Score Group Age Group	Low Group 1	Middle Group 2	High Group 3	Total Mean (x) Std. Dev. Range (r)
Young Group 1	9	10	10	N=29 x=54.04 std.dev.=2.44 r=52-59 years
Middle Group 2	7	13	12	N=32 x=63.02 std.dev.=2.66 r=60-69 years
Old Group 3	11	11	11	N=33 x=71.08 std.dev=3.75 r=70-85 years
Total Mean (x) Std. Dev. Range	N=27 x=15.59 std.dev.=4.76 r=2-20/65	N=34 x=30.85 std.dev.=6.15 r=21-39/65	N=33 x=46.06 std.dev.=4.83 r=40-61/65	N=94 x=31.81 63.05 std.dev.=13.29 7.66 r=2-61/65 52-85 years

received the AH4 test of intelligence parts 1 and 2, and also the Mill Hill vocabulary scale parts 1 and 2.

6.2.2 Apparatus

An Apple II microcomputer with an Apple clock card inserted in slot 5 was connected to a Kodak carousel slide projector. Appendix 4 illustrates the circuit used, in which a slide progressor switch and three response press switches were connected to the Apple microcomputer, and the responses recorded on disk. The response latency was measured from the onset of the picture on the screen. This was achieved by means of a hole in the plastic surround of each slide, through which light passed when the slide was in the projector, thereby activating a photoreceptor which started the Apple clock. The clock was stopped when either the 'yes' switch, which indicated a correct response, or the 'abort' switch which indicated that the subject said that he or she did not know or could not retrieve the name of the displayed item, was pressed. The 'no' switch indicated an incorrect response. Each time it was pressed, the elapsed time from the slide onset (response time) was recorded, but the clock continued until it had been pressed three times, giving further records if necessary, for each of three attempts that subjects were allowed to name each item. If none of these was correct the trial was automatically labelled an 'abort' trial. All switches were operated by the experimenter.

6.2.3 Materials

Slides were produced from the McKenna and Warrington graded naming test. This consists of thirty pictures of items to be named, which increase in difficulty. McKenna and Warrington had selected these pictures on the following criteria: a) the pictorial representation of the name must be unambiguous, in that only a single response should be elicited. b) The pictorial representation should be distinctive and not perceptually or verbally confusable with other items. c) No specialist knowledge should be required. d) Items in very common usage should be excluded. The items selected are listed below, in increasing order of difficulty.

1. Kangeroo
2. Scarecrow
3. Buoy
4. Thimble
5. Handcuffs
6. Tweezers
7. Corkscrew
8. Sporan
9. Tassel
10. Sundial
11. Chopsticks
12. Periscope
13. Boar
16. Turtle
17. Trampoline
18. Bellows
19. Shuttlecock
20. Ant-eater
21. Pagoda
22. Radius
23. Leotard
24. Mitre
25. Yashmak
26. Sextant
27. Centaur
28. Cowl

14. Blinkers
15. Monocle
29. Tutu
30. Retort

6.2.4 Procedure

Subjects were all tested individually by the author. The experimenter operated all the response press switches, with the pictures being projected onto a large screen at approximately 35 degrees above subject eye level. Subjects were told that they would be shown 30 pictures, and that their task was simply to speak aloud the name of the object portrayed as soon as they could manage to do so. A few illustrative examples were given, and the subject was given an opportunity to question the experimenter about the procedure at his or her leisure. Subjects were allowed up to three errors which were timed. If they did not know the name of any item, they were permitted to say so and move on to the next slide. The clock started when the picture appeared on the screen, and the experimenter pressed the response switch appropriate to the subject's response as soon as he uttered the name or requested to 'abort' the trial. Subjects were only told whether their response was correct or incorrect; no other information was given. The picture remained on the screen until after all the responses had been made, and was subsequently replaced by the next picture.

6.2.5 Results

Response times were recorded for every response given hence, times for correct, incorrect and abort responses were all available. The number of times each type of response occurred was counted, so that the number of attempts at each picture was known. Mean response times for each type of response were computed for each subject and then placed in a correlation matrix with age, AH4 test score and Mill Hill vocabulary scale (see Table 6.2). The number of correct responses correlates highly with AH4 parts I and II (R=0.605, R=0.541 respectively), and also with Mill Hill vocabulary scale parts I and II (R=0.609, R=0.605 respectively). Negative correlation coefficients from AH4 and Mill Hill vocabulary scale with correct response time were as expected, showing that high AH4 test scorers and those with high vocabulary scores, took less time to produce the correct response (R=-0.460, -0.535, -0.374, -0.349 for parts I and II of each test respectively). Chronological age correlates poorly with the number of correct responses (R=-0.167), as we would expect from the evidence that vocabulary remains stable into old age. However, age correlates significantly with correct response time (R=0.416 P<0.001) and also with the number of incorrect responses made (R=0.294 P<0.001), (note that these may be errors made before a correct response was finally given.) Thus it seems that older people take longer to retrieve words and make more errors in retrieval before the correct word is eventually retrieved.

Table 6.2

Correlation Matrix

	AH41	AH42	M.H.1	M.H.2	AGE	CRT	C	IRT	I	ART	A
AH4 (Part 1)	1.00	0.81	0.65	0.68	-0.17	-0.46	0.60	-0.26	-0.49	0.31	-0.59
AH4 (Part 2)		1.00	0.54	0.53	-0.30	-0.53	0.54	-0.27	-0.56	0.20	-0.52
Mill Hill (Part 1)			1.00	0.77	0.11	-0.37	0.61	-0.03	-0.44	0.26	-0.60
Mill Hill (Part 2)				1.00	0.15	-0.35	0.60	-0.04	-0.41	0.26	-0.59
Age					1.00	0.42	-0.17	0.31	0.29	0.15	0.18
Correct Response Time (CRT)						1.00	-0.51	0.44	0.59	0.17	0.50
Number Of Correct Responses (C)							1.00	-0.19	-0.59	0.17	-0.99
Incorrect Response Time (IRT)								1.00	0.27	0.46	0.15
Number Of Incorrect Responses (I)									1.00	-0.03	0.61
Abort Response Time (ART)										1.00	-0.17
Number Of Abort Responses (A)											1.00

The data were then examined with respect to the subjects' allocated I.Q. test score group and age group. Table 6.3 shows the mean response times for the three I.Q. test score groups and the three age groups. The group differences were firstly examined using one-way analyses of variance, in which each of the measures taken were analysed. For the I.Q. test score groups, there was a significant main effect of I.Q. test score group on correct response time (F=12.712 df=2,92 $P<0.01$), number of correct responses (F=27.067 df=2,92 $P<0.01$), incorrect response time (F=3.2076 df=2,92 $P<0.04$), number of incorrect responses (F=14.266, df=2,92 $P<0.01$), abort response time (F=3.0944 df=2,92 $P<0.05$) and number of abort respones (F=25.084 df=2,92 $P<0.01$). Planned comparisons were used to establish which groups differed significantly. Correct response time, and incorrect response time were significantly faster for the middle than for the low I.Q. test score groups. By inference the high I.Q. test score group is also faster than the low test score group (F=14.385 df=(1,92) $P<0.001$, F=4.17 df(1,92) $P<0.05$), however the difference between the high and middle I.Q. test scorers did not differ significantly on these indices (F=1.863 df(1,92) $P \gg 0.05$ F=0.0215 df(1,92) $P \gg 0.05$). Differences in the number of correct responses (F=5.8 df(1,92) $P<0.05$, F=27.41 df(1,92) $P<0.001$), incorrect responses (F=4.061 df(1,92) $P<0.05$, F=12.867 df(1,92) $P<0.01$) and the number of abort responses (F=4.465 df(1,91) $P<0.05$, F=28.6488 df(1,91) $P<0.001$) are significant between all groups. The figures in parthenses are for high and middle group, and middle and low group comparisons respectively.

Table 6.3

Descriptive Statistics For Age Groups

Age Group 1

	N	MIN.	MAX.	MEAN	STD. DEV.
CORRECT RESPONSE TIME (SEC)	31	2.747	8.514	4.412	1.521
NO. CORRECT RESPONSES	31	4.000	30.000	23.323	4.996
INCORRECT RESPONSE TIME (SEC)	29	3.159	13.346	7.325	2.786
NO. INCORRECT RESPONSES	29	1.000	12.000	4.862	2.735
ABORT TIME (SEC)	30	7.750	26.907	17.936	5.415
NO. ABORT TRIALS	30	1.000	26.000	6.900	4.922

Age Group 2

	N	MIN.	MAX.	MEAN	STD. DEV.
CORRECT RESPONSE TIME (SEC)	36	2.623	7.539	4.544	1.369
NO. CORRECT RESPONSES	36	12.000	29.000	23.972	4.123
INCORRECT RESPONSE TIME (SEC)	36	3.443	15.863	8.099	3.378
NO. INCORRECT RESPONSES	36	2.000	10.000	4.667	2.305
ABORT TIME (SEC)	36	9.796	63.758	21.148	11.366
NO. ABORT TRIALS	36	1.000	18.000	6.028	4.123

Age Group 3

	N	MIN.	MAX.	MEAN	STD. DEV.
CORRECT RESPONSE TIME (SEC)	26	2.807	15.195	6.635	3.055
NO. CORRECT RESPONSES	26	10.000	30.000	21.731	6.271
INCORRECT RESPONSE TIME (SEC)	26	3.427	18.695	10.173	3.880
NO. INCORRECT RESPONSES	26	1.000	15.000	6.461	3.614
ABORT TIME (SEC)	25	4.008	61.000	22.497	11.366
NO. ABORT TRIALS	25	1.000	20.000	8.480	6.152

Table 6.3 (Continued)

Descriptive Statistics For I.Q. Test Score Groups

I.Q. Group 1

	N	MIN.	MAX.	MEAN	STD. DEV.
CORRECT RESPONSE TIME (SEC)	28	3.443	15.195	6.608	2.683
NO. CORRECT RESPONSES	28	4.000	28.000	18.643	5.826
INCORRECT RESPONSE TIME (SEC)	28	4.718	15.863	9.788	2.786
NO. INCORRECT RESPONSES	28	2.000	15.000	7.250	3.284
ABORT TIME (SEC)	28	9.796	27.565	17.196	5.324
NO. ABORT TRIALS	28	2.000	26.000	11.321	5.844

I.Q. Group 2

	N	MIN.	MAX.	MEAN	STD. DEV.
CORRECT RESPONSE TIME (SEC)	37	2.684	11.218	4.721	1.860
NO. CORRECT RESPONSES	37	17.000	29.000	24.000	3.091
INCORRECT RESPONSE TIME (SEC)	36	3.159	18.695	8.022	3.743
NO. INCORRECT RESPONSES	36	1.000	9.000	4.917	2.260
ABORT TIME (SEC)	37	4.008	63.758	20.633	10.376
NO. ABORT TRIALS	37	1.000	13.000	5.973	3.059

I.Q. Group 3

	N	MIN.	MAX.	MEAN	STD. DEV.
CORRECT RESPONSE TIME (SEC)	28	2.684	7.420	4.042	1.157
NO. CORRECT RESPONSES	28	16.000	30.000	26.464	2.987
INCORRECT RESPONSE TIME (SEC)	27	3.395	15.588	7.616	3.593
NO. INCORRECT RESPONSES	27	1.000	9.000	3.593	2.117
ABORT TIME (SEC)	26	7.750	61.000	23.728	11.943
NO. ABORT TRIALS	26	1.000	14.000	3.769	2.847

The one way analyses of variance were repeated for the three age groups. In this case, significant group differences were present for correct response time (F=10.630 df=2,92 $P<0.01$ incorrect response time (F=5.2464 df=2,92 $P<0.01$) and also the number of incorrect responses (F=3.338 df=2,92 $P<0.04$). There were no significant group differences in the number of correct responses (F=1.500 df=2,92 $P<0.22$), abort response time (F=1.6195 df=2 $P<0.2038$) or number of abort responses (F=1.7768 df=2,92 $P<0.175$). Planned comparisons were performed where a significant group difference had been found. The old and middle range subjects differ significantly on correct response time (F=16.127 df(1,92) $P<0.001$), incorrect response times (F=5.756 df(1,92) $P<0.05$), and number of incorrect responses (F=5.944 df(1,92) $P<0.05$), the middle and young range subjects did not differ significantly on any of these (F=0.007, 0.0872, 0.077). It is interesting to note that the subjects in the middle age group have a higher mean number of correct responses than the young age group in the range, but their response times are marginally slower.

Correct response times were considered in relation to increasing item difficulty. Since the task was graded, median correct response times were computed over each successive six pictures in the graded sequence. This provided five median scores for each subject corresponding to five successive levels of increasing difficulty. Using a multiple regression analysis with difficulty as a predictor of performance for each age and I.Q. test score group, the gradients of the linear functions of the relationship were calculated, together with the intercepts on the X and Y axes. The regression lines for each age and I.Q. test score group are plotted in Figures 6.1 and 6.2.

An analysis of variance was computed on the five medium scores with age, I.Q. test score as the between subject variables and difficulty with five levels as the within factor variable. There were main effects of all three factors, the overall effect of age group ($F=11.26$ $df=2,73$ $P<0.01$) and I.Q. test score group ($F=15.341$ $df=2,73$ $P<0.01$) can be explained by the expected result that response time increases as age increases and I.Q test score decreases. However, there was a significant interaction between these two variables ($F=2.807$ $df=4,73$ $P<0.03$) which was probably because of subjects belonging to the middle age group in the range with low I.Q. test scores responded faster than the younger subjects with similar I.Q. test scores (Fig 6.3). The significant effect of difficulty can be seen in Figure 6.4, it is clear that the 'easy' pictures give rise to the fastest response times and the 'difficult' pictures to the longest response times. All between, and within interactions were insignificant (Table 6.4).

Finally, to discover whether age or I.Q. test score was the better predictor of performance with respect to the increasing levels of difficulty associated with the task, linear multiple regression analyses were carried out using forward selection for each level of difficulty with age and AH4 part one test score as predictors of response time. A summary of each separate regression can be found in Table 6.5. It is clear from Table 6.5 that, at each level of difficulty, except level 3, AH4 part one test score accounts for more of the variance in response

Fig. 6.1

Regression Lines For Age Groups With Mean Scores.

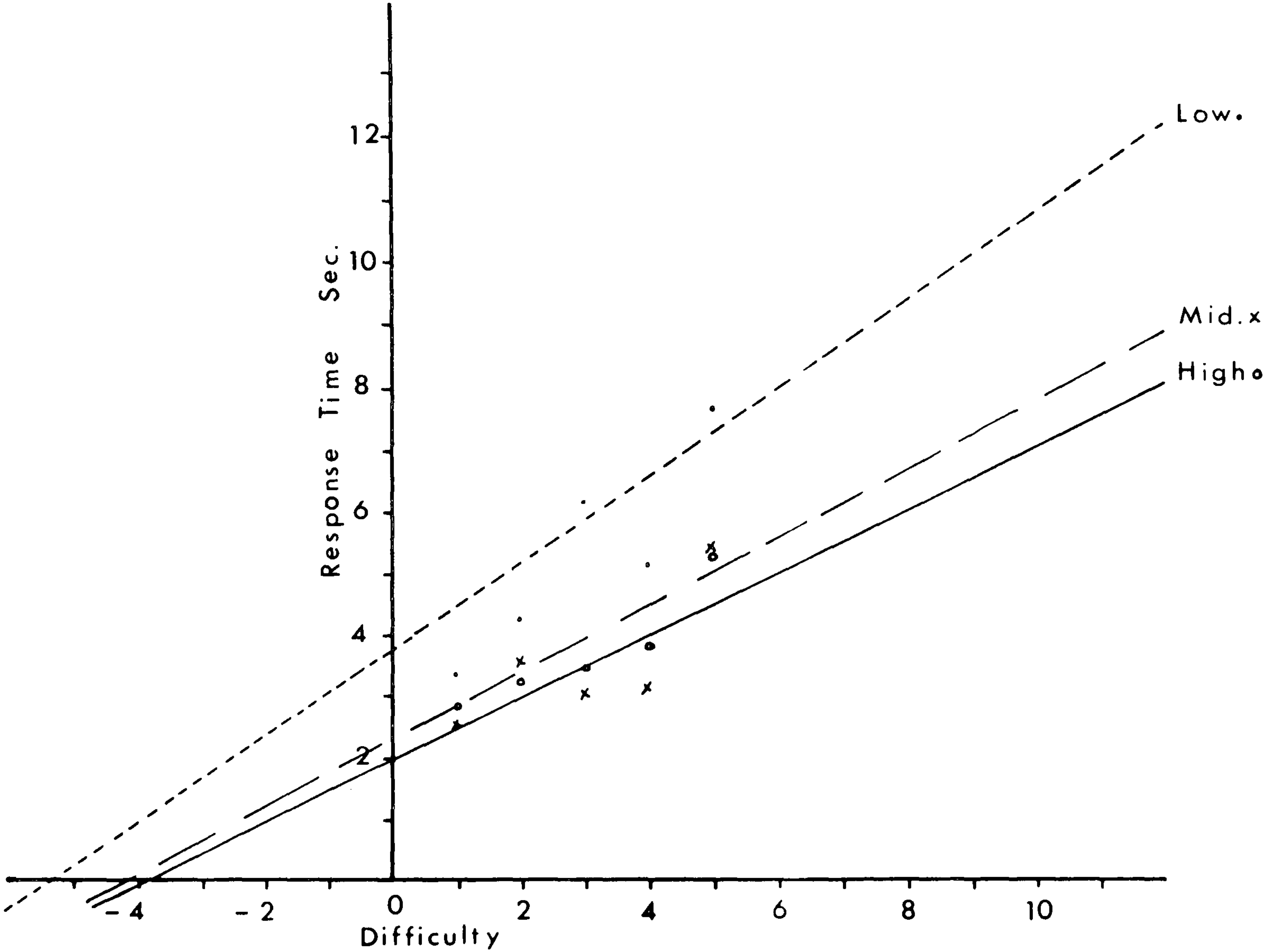


Fig. 6.2

Regression Lines For I.Q. Test Score Groups With Mean Scores.

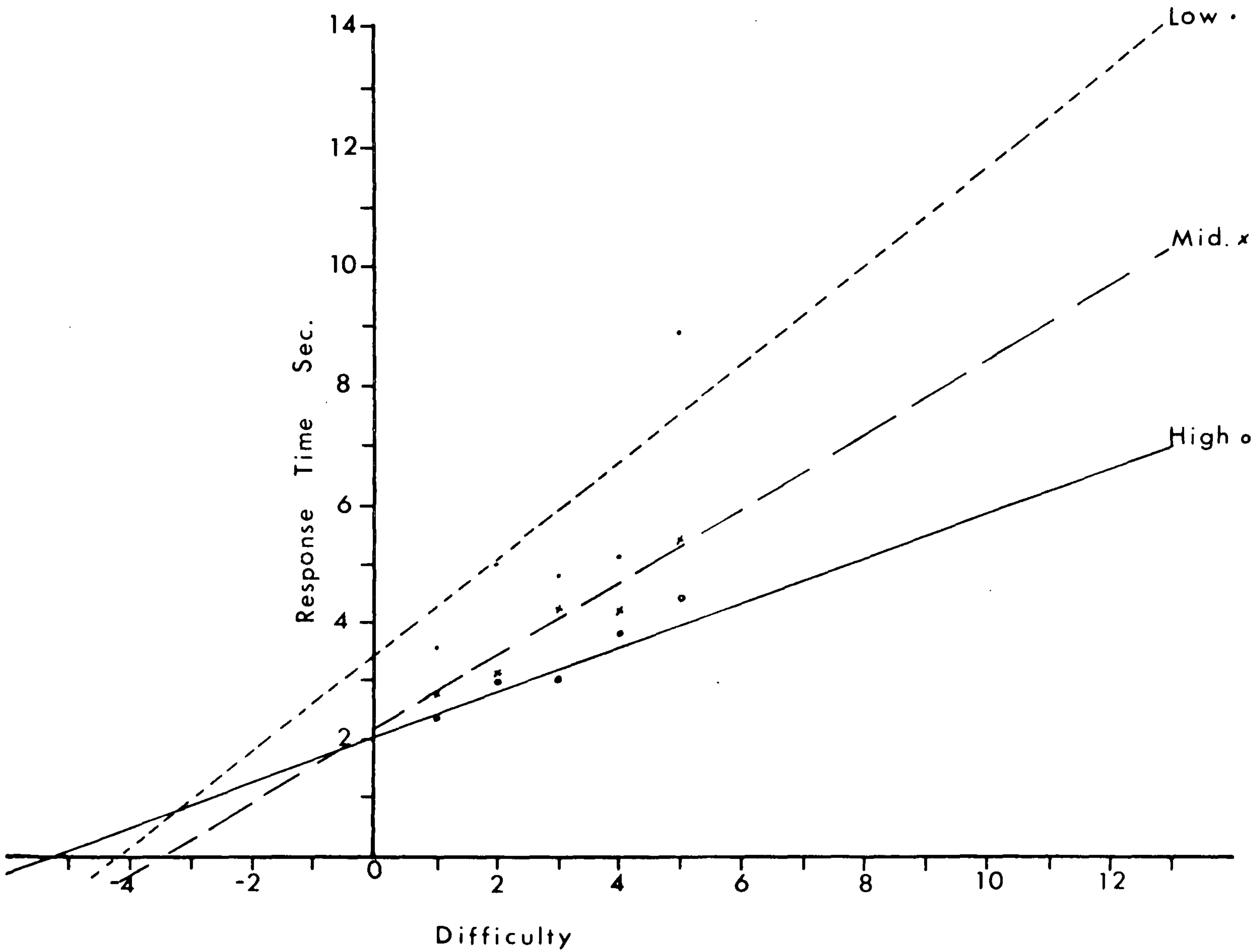


Fig. 6.3

Interaction Between Age & I.Q. Test Score Group.

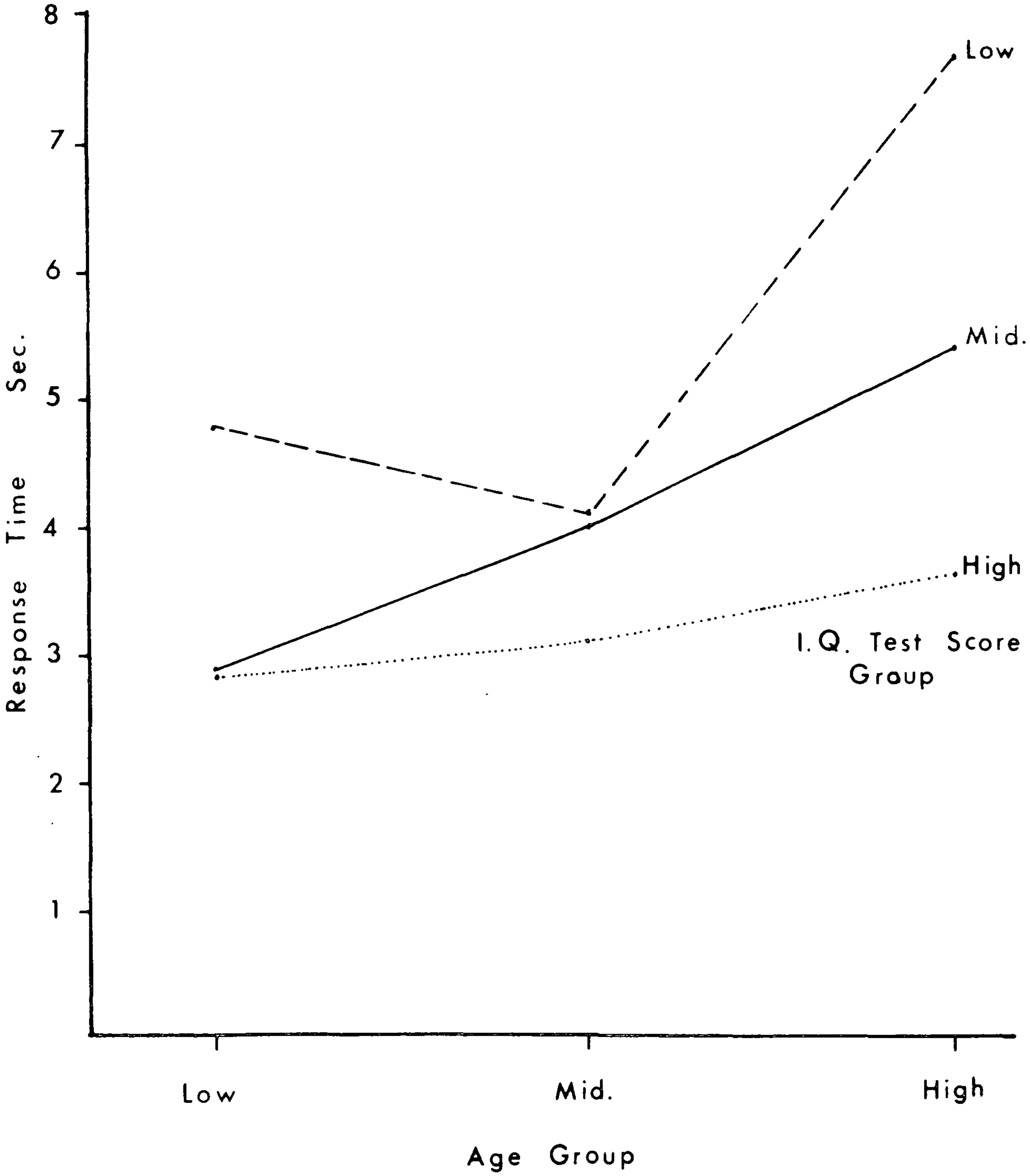


Fig. 6.4

Effect Of Difficulty.

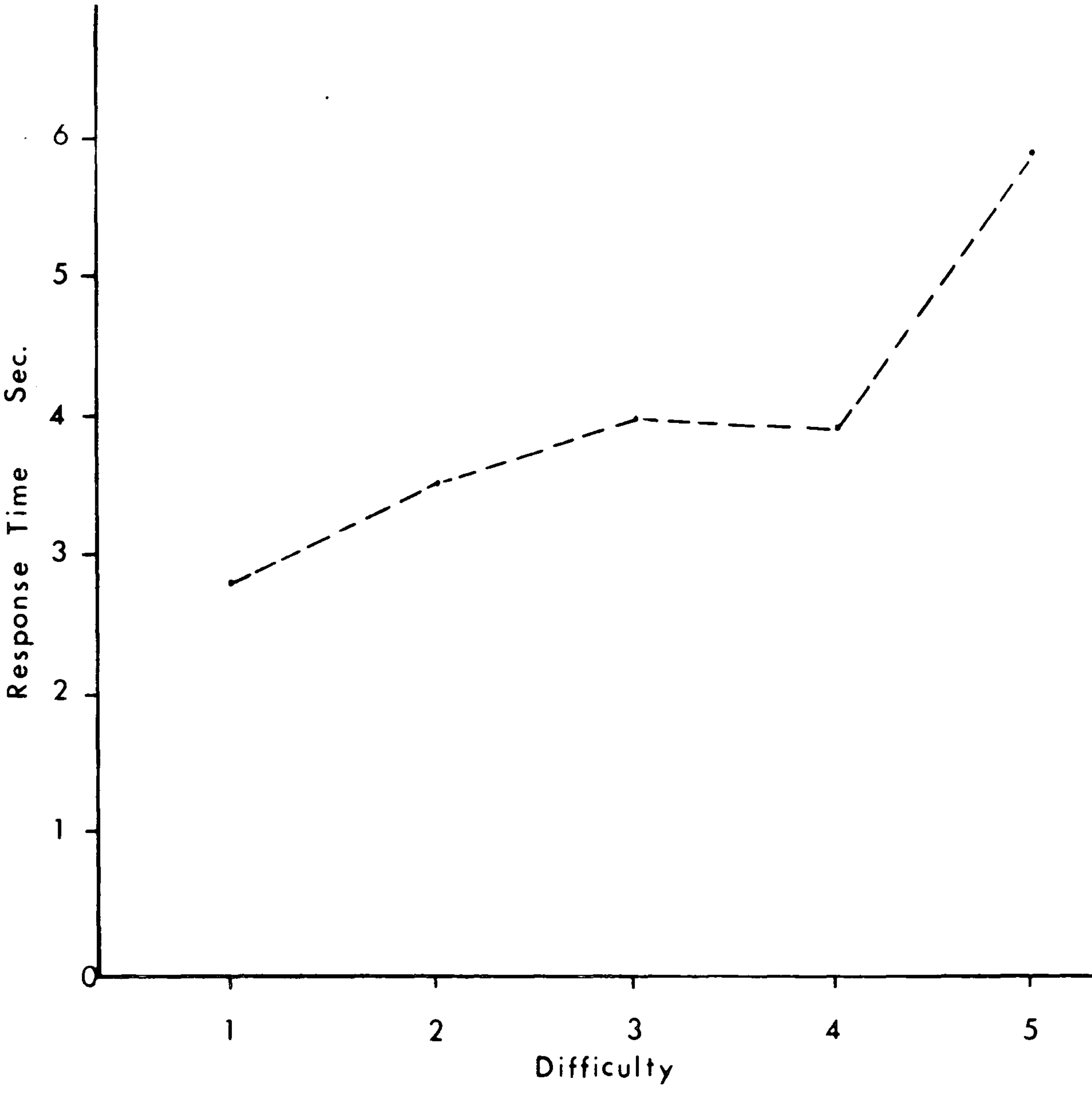


Table 6.4

Experiment 6.1

Analysis Of Variance   I.Q. Test Score Group / Age Group X Difficulty

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	337.181	2	168.591	15.341	0.001
Age	247.488	2	123.744	11.260	0.001
I.Q. X Age	123.400	4	30.850	2.807	0.032
S-Within	805.250	73	10.990		
Difficulty	438.188	4	109.547	12.148	0.001
I.Q. X Difficulty	95.397	8	11.925	1.322	0.232
Age X Dificulty	49.954	8	6.244	0.692	0.698
I.Q. X Age X Difficulty	83.635	16	5.227	1.322	0.898
Difficulty X S-Within	2633.242	292	9.018		

Table 6.5.

Experiment 6.1

Regression Analysis For Each Level Of Difficulty Using Forward Selection

Level 1

....

Step	R-SQR	STD ERROR	VARIABLE		PARTIAL	SIGNIF
1	0.2490	1.6722	AH4	IN	-0.4990	0.0000
2	0.3440	1.5714	AGE	IN	0.3556	0.0003

Level 2

....

Step	R-SQR	STD ERROR	VARIABLE		PARTIAL	SIGNIF
1	0.1426	2.3546	AH4	IN	-0.3776	0.0002
2	0.1875	2.3049	AGE	IN	0.2289	0.0291

Level 3

....

Step	R-SQR	STD ERROR	VARIABLE		PARTIAL	SIGNIF
1	0.0753	3.7759	AGE	IN	0.2745	0.0078
2	0.1040	3.7375	AH4	IN	-0.17604	0.0932

Level 4

....

Step	R-SQR	STD ERROR	VARIABLE		PARTIAL	SIGNIF
1	0.1292	3.1232	AH4	IN	-0.3594	0.0004
2	0.1866	3.0353	AGE	IN	0.2567	0.0135

Level 5

....

Step	R-SQR	STD ERROR	VARIABLE		PARTIAL	SIGNIF
1	0.1178	4.6311	AH4	IN	-0.3431	0.0015
2	0.1179	4.6599	AGE	IN	-0.0028	0.9798

time than does age. There is no obvious reason why age should account for more of the variance on level 3, but even here age accounts for only 7.534% of the variance, so that chronological age is, overall, a weak predictor. In levels 1, 2, and 4 age can be seen to pick up some of the remaining variance, but in the case of the most difficult level (5), age does not account for any residual variance already accounted for by differences in AH4 test scores.

6.3 EXPERIMENT 6.2

6.3.1 Subjects

The subjects used in Experiment 6.1 were also run in this experiment during the same testing session. Between this experiment and Experiment 6.1 subjects experienced one of the trials used in Experiment 3.1 which involved categorising and recalling lists of words. This provided an interval of ten minutes between the two picture naming experiments.

6.3.2 Apparatus

The apparatus and circuit design was the same as in Experiment 6.1, however in this case the abort switch was redundant as a response recorder, since subjects were allowed one response and were told to guess if they did not know the correct answer, thus all reponses were scored as either correct or incorrect, by the experiment pressing the appropriate switch according to the subjects response.

6.3.3 Materials

38 slides were made of items from the short form of the British Picture Vocabulary Scale, these included six training plates. Each plate consisted of four bold line drawings of objects or actions. The drawings were numbered from one to four, and the numbers were printed beneath each. The numbers were associated with the same position for every frame.

1. Bucket
2. Ball
3. Car
4. Wooden
5. Camera
6. Envelope
7. Circle
8. Furniture
9. Nostril
10. Dangerous
11. Furious
12. Athelete
13. Artist
14. Weary
15. Socket
16. Antler
17. Pulley
18. Inflated
19. Assisting
20. Collision
21. Floral
22. Goblet
23. Utensil
24. Talon
25. Confiding
26. Innoculation
27. Consuming
28. Gable
29. Apparition
30. Emission
31. Ambulation
32. Sa$\overset{u}{l}$tation

6.3.4 Procedure

Subjects were tested individually by the author. The basic format of the slides was decribed to each subject, and they were instructed to speak aloud the number of the picture that corresponded to the target word they had been given just before each slide appeared. They were told that each of their responses would be timed and that they should try to select a response as fast as possible. The six training slides allowed the subjects to become familiar with the format and to grasp the simple task. It also allowed the experimenter to modify the loudness and pitch of her voice to ensure that the subject could hear the target word clearly. As in Experiment 6.1, the experimenter controlled the response switches. Subjects were only allowed one response to each trial and no feedback was given by the experimenter.

6.3.5 Results

Mean correct and incorrect response times were computed, (see Table 6.6 for group descriptive statistics) and these together with the total number of correct and incorrect responses for each subject placed in a correlation matrix. Age, Mill Hill vocabulary scale parts I and II, and AH4 part I and II test scores were the measures of individual differences also placed in the matrix (Table 6.7). Both parts of the AH4 test correlate highly with the number of correct responses (R=0.565, 0.5774)

Table 6.6

Descriptive Statistics For Age Groups For Experiment 6.2

Age Group 1

	N	MIN.	MAX.	MEAN	STD. DEV.
CORRECT RESPONSE TIME (SEC)	31	2.708	5.112	3.712	0.619
NO. CORRECT RESPONSES	31	21.000	32.000	29.290	3.002
INCORRECT RESPONSE TIME (SEC)	29	2.356	49.270	11.634	11.495
NO. INCORRECT RESPONSES	29	1.000	12.000	3.345	2.931

Age Group 2

	N	MIN.	MAX.	MEAN	STD. DEV.
CORRECT RESPONSE TIME (SEC)	36	2.467	6.029	4.066	0.960
NO. CORRECT RESPONSES	36	25.000	32.000	29.444	1.812
INCORRECT RESPONSE TIME (SEC)	35	3.475	21.090	9.828	4.350
NO. INCORRECT RESPONSES	35	1.000	7.000	2.857	1.817

Age Group 3

	N	MIN.	MAX.	MEAN	STD. DEV.
CORRECT RESPONSE TIME (SEC)	25	2.810	7.704	4.893	1.307
NO. CORRECT RESPONSES	25	20.000	32.000	28.120	3.153
INCORRECT RESPONSE TIME (SEC)	23	3.039	26.760	13.803	7.271
NO. INCORRECT RESPONSES	23	1.000	12.000	4.478	2.967

Table 6.6 (Continued)

Descriptive Statistics For I.Q. Test Score Groups For Experiment 8

I.Q. Group 1

....

	N	MIN.	MAX.	MEAN	STD. DEV.
CORRECT RESPONSE TIME (SEC)	28	2.935	7.376	4.425	1.103
NO. CORRECT RESPONSES	28	20.000	32.000	27.000	3.255
INCORRECT RESPONSE TIME (SEC)	27	3.039	26.137	8.512	4.662
NO. INCORRECT RESPONSES	27	1.000	12.000	5.592	2.938

I.Q. Group 2

....

	N	MIN.	MAX.	MEAN	STD. DEV.
CORRECT RESPONSE TIME (SEC)	35	2.467	7.704	4.240	1.095
NO. CORRECT RESPONSES	35	22.000	32.000	29.514	2.020
INCORRECT RESPONSE TIME (SEC)	33	3.622	26.760	11.474	5.824
NO. INCORRECT RESPONSES	33	1.000	10.000	2.939	1.919

I.Q. Group 3

....

	N	MIN.	MAX.	MEAN	STD. DEV.
CORRECT RESPONSE TIME (SEC)	29	2.587	6.472	3.844	0.965
NO. CORRECT RESPONSES	29	27.000	32.000	30.414	1.323
INCORRECT RESPONSE TIME (SEC)	27	2.356	49.270	14.457	11.765
NO. INCORRECT RESPONSES	27	1.000	6.000	1.925	1.328

Table 6.7

Correlation Matrix

	AH41	AH42	M.H.1	M.H.2	AGE	CRT	C	IRT	I
AH4 (Part 1)	1.00	0.81	0.65	0.68	-0.17	-0.29	0.56	0.27	-0.63
AH4 (Part 2)		1.00	0.54	0.53	-0.30	-0.45	0.57	0.25	-0.61
Mill Hill (Part 1)			1.00	0.77	0.11	-0.16	0.54	0.31	-0.58
Mill Hill (Part 2)				1.00	0.15	-0.19	0.51	0.24	-0.56
Age					1.00	0.46	-0.22	0.10	0.22
Correct Response Time (CRT)						1.00	-0.31	0.19	0.32
Number Of Correct Responses (C)							1.00	0.17	-0.98
Incorrect Response Time (IRT)								1.00	-0.21
Number Of Incorrect Responses (I)									1.00

for parts I and II respectively, as do parts I and II of the Mill Hill vocabulary scale (R=0.543, 0.516). Hence the high I.Q. test scorers and subjects with a large vocabulary score are able to make more correct matchings of the presented name with the presented pictures. It follows that these indices have strong negative correlations with the number of incorrect responses (R=-0.6308, -0.6061, -0.5843, -0.567) for AH4 test score and Mill Hill vocabulary scale respectively. Mean correct response time has negative correlations with these indices but the relationship is not as strong as for the number of correct responses (R=-0.2927, -0.4504, -0.1631, -0.1902). Note that throughout, the correlations between age are weak (R=-0.221, 0.0991, 0.2162) for number of correct responses, incorrect response time, and number of incorrect responses respectively. However, there is one major exception to the lack of age effects: age has a strong relationship with correct response time (R=0.4576).

Group differences for age and I.Q. test score groups were examined by one way analyses of variance. There were significant differences between I.Q. test score groups on the number of correct responses (F=16.824 df=2,91 P<0.01), number of incorrect responses (F=21.124 df=2,86 P<0.01), and incorrect response time (F=3.817 df=2,86 P<0.02). There was no main effect of I.Q. test score group on correct response times (F=2.2662 df=2,91 P<0.10).

Planned comparisons between the middle and low I.Q. test score groups and high and middle groups were used to identify the position of the effects. There were significant differences between the low and middle groups for the number of correct responses (F=18.469 df(1,91) P<0.001) and number of incorrect responses (F=22.616 df(1,86) P<0.001). Differences between the middle and high groups were insignificant for all measures i.e. (F=2.414, 2.114, 3.3) for number correct, incorrect response time, and number of incorrect responses respectively. The significant main effect for incorrect response time is caused by the difference between the high and low I.Q. test score groups (F=7.632 df(1,86) P<0.01). All other indices where the effect of I.Q. test score was significant also have this difference, as indicated by the significant difference between the middle and low I.Q. test score groups, and the group mean trends.

A significant effect of age was only found with correct response time (F=10.489 df=2,91 P<0.01 .). All other one way analyses of variance on the remaining indices were insignificant (F=2.067, 1.680, 2.8386) for number of correct responses, incorrect response time and number of incorrect responses respectively. The planned comparisons on correct response time showed that the effect was due to a difference between the middle and old subjects in the range (F=10.597 df(1,91) P<0.01), and by inference the old and younger subjects also differ significantly. However, the comparison between the middle and young subjects was insignificant (F=2.20 df(1,91) P>>0.05).

The use of a graded task, once again allowed an analysis of the effect of item difficulty on correct response time. Median correct response times were computed for each subject over each progressingly more difficult set of four slides. This produced a total of eight median scores, corresponding to eight levels of increasing difficulty. A multiple regression analysis was used to establish the gradient and intercepts on the X and Y axes of the linear function of response time against increasing difficulty for each age and I.Q. test score group separately (Fig 6.5, 6.6).

In addition, an analysis of variance was computed with age and I.Q. test score group as between factor variables and difficulty with eight levels, as the within factor variable. There were significant main effects of age group (F=11.141 df=2,78 P<0.001) and I.Q. test score group (F=3.740 df=2 P<0.028), both of these effects were in the expected direction, with the older subjects, and subjects with low I.Q. test scores taking the most time to produce correct responses. The main effect of difficulty validates the scale (F=40.304 df=7,546 P<0.001), 'easy' sets of pictures generate faster responses than 'difficult' sets. There were significant interactions between age group and difficulty (F=1.98 df=14,546 P<0.017) and I.Q.test score and difficulty (F=2.457 df=14,546 P<0.002) (see Figures 6.7, 6.8). The interaction between age group and difficulty occurs because the younger subjects take longer than the middle group in the range to respond to the most difficult slides, but the younger group respond faster than the middle group to the easier picture slides. For the I.Q. group X difficulty interaction, the effect of the most difficult pictures is more prononced for the low and middle I.Q. test scorers with there response times increasing relatively more than those with high I.Q. test scores. Also, the low scorers respond faster than the middle group

Fig. 6.5

Regression Lines For Age Groups With Mean Scores.

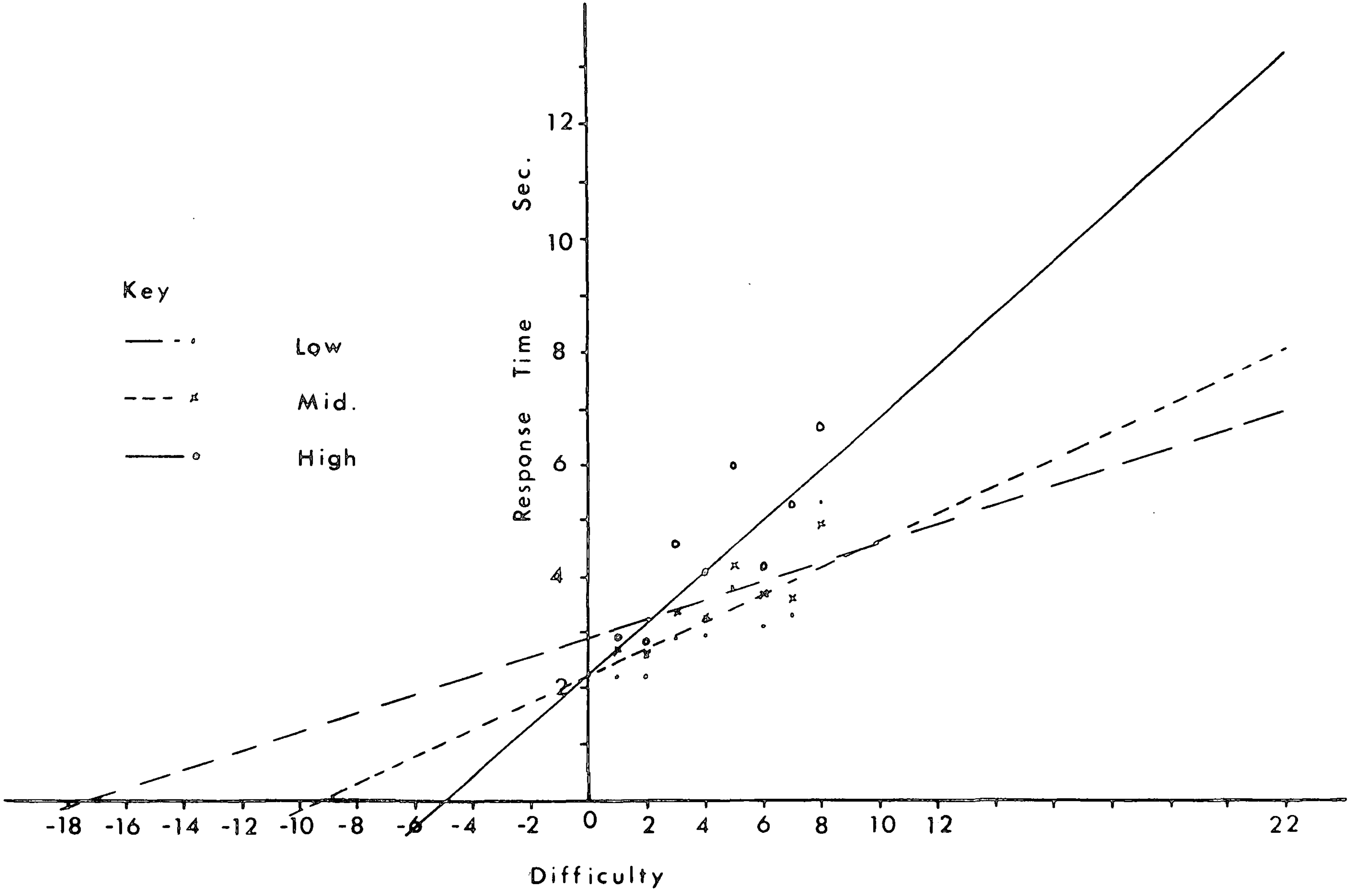


Fig. 6.6

Regression Lines For I.Q. Test Score Groups With Mean Scores.

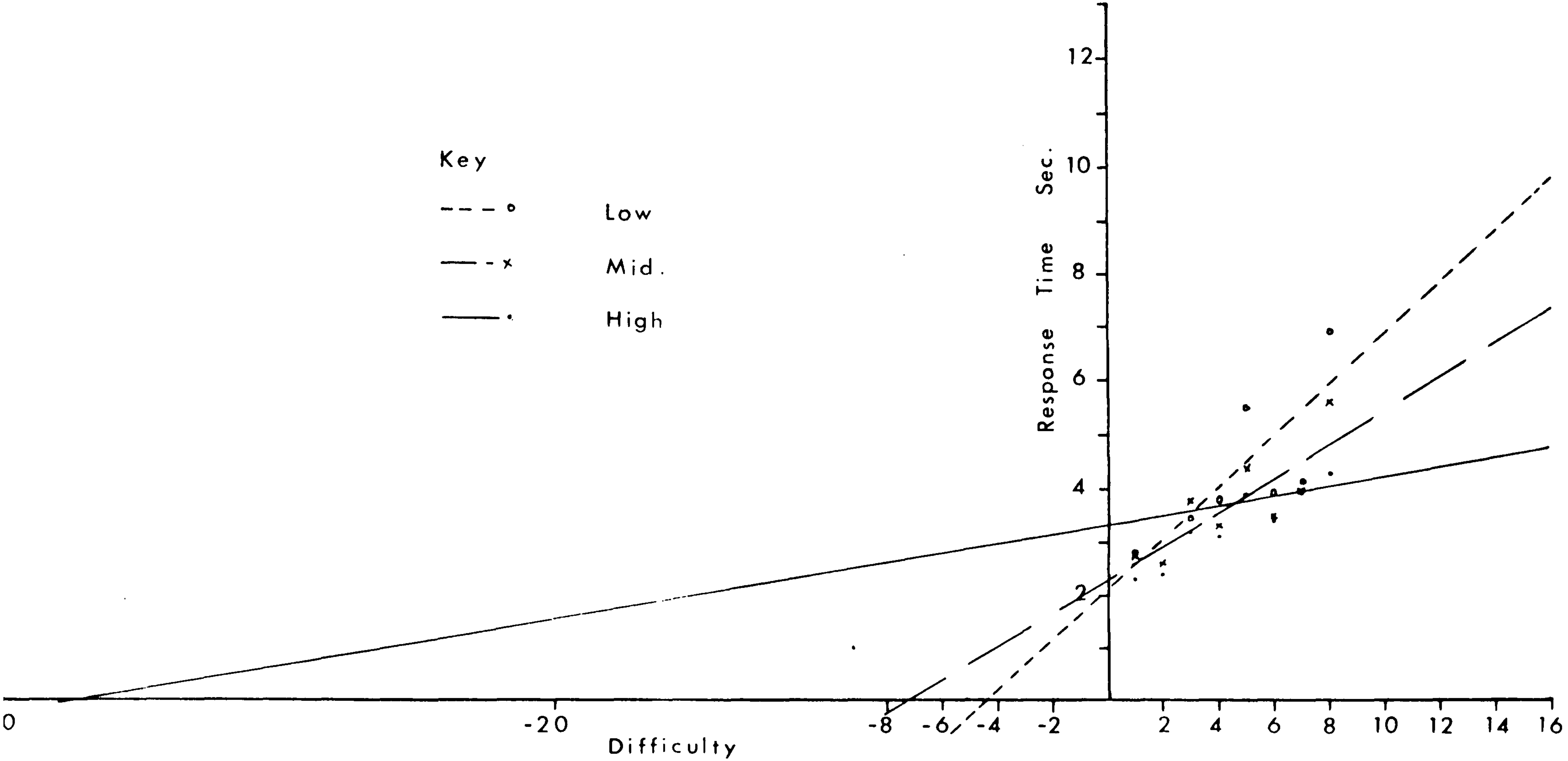


Fig. 6.7

Interaction Between Age Group & Difficulty.

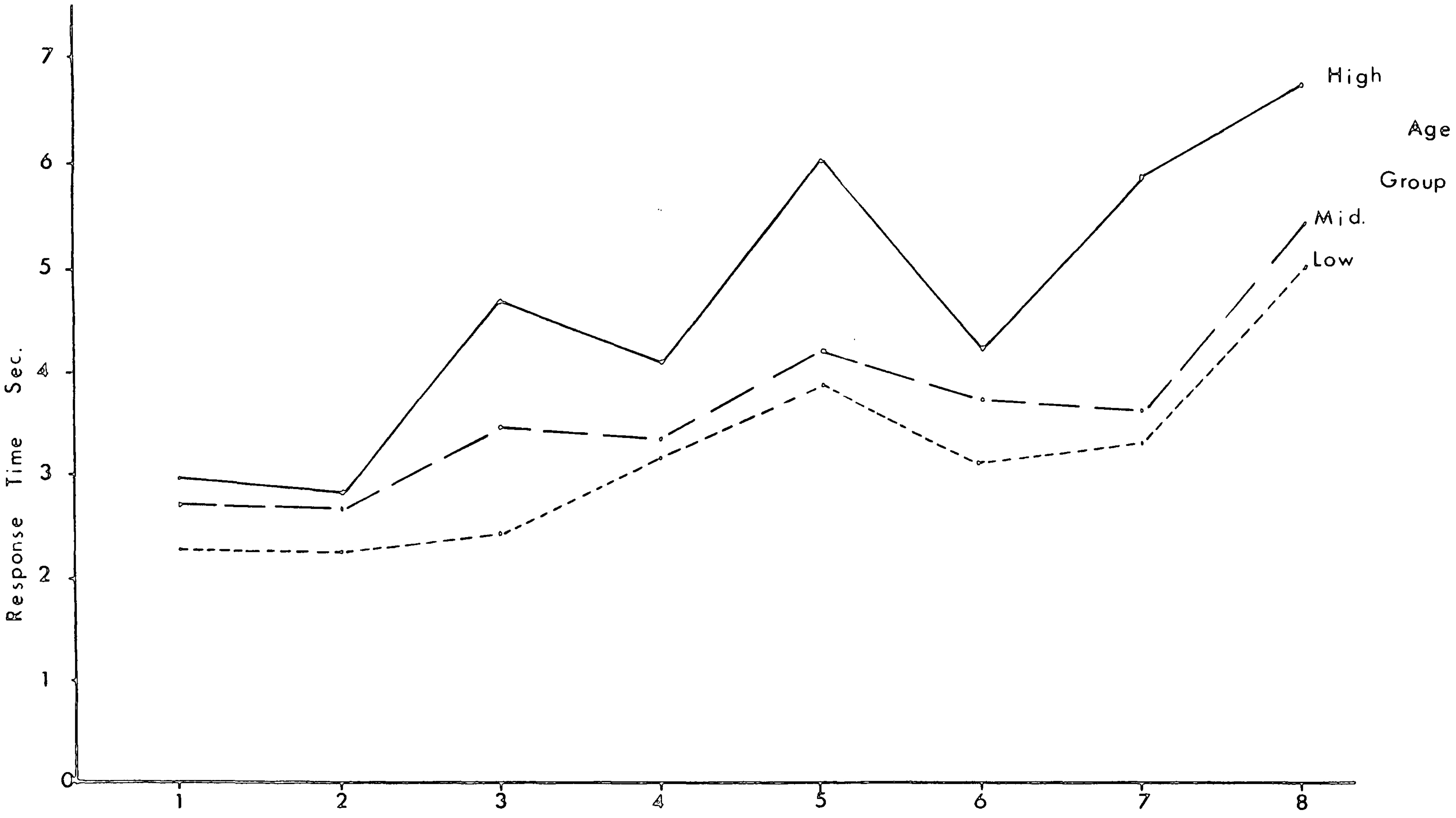
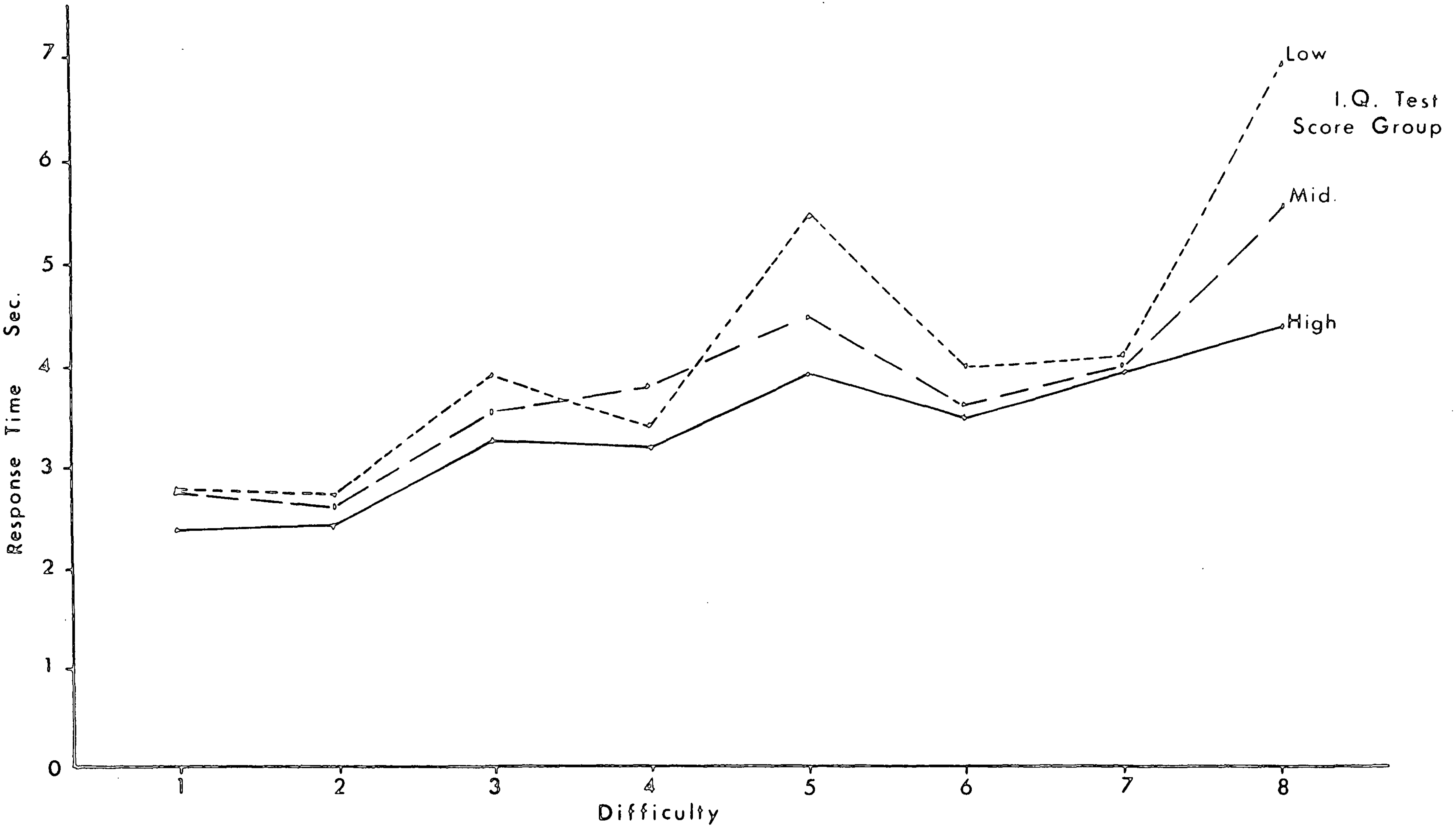


Fig. 6.8

Interaction Between I.Q. Test Score Group & Difficulty.



on the second level of difficulty which is peculiar to their performance on other levels of difficulty where they are generally slower than the middle I.Q. test scorers (Table 6.8).

Both age and I.Q. test score were assessed as predictors of correct response time for each level of difficulty. Linear regression analyses using forward selection were computed for each difficulty level (Table 6.9). The pattern is less distinctive than that found for the naming latency data in which AH4 accounted for more variance than age on every level except level 3. In this task, AH4 part one test score accounts for a higher percentage of the variance on levels 1 and 2, but age picks up some of the remaining variance increasing the R-Square value by 0.1046 and 0.081 when it is added to the equation. However on subsequent levels, age accounts for a large portion of the variance up to level 8, the most difficult level, for which AH4 accounts for 19.09%, and the addition of age to the regression equation does not significantly increase this. There is one exception to this sequence with age accounting for the most variance, level 6 show AH4 test score as the better predictor. Again, it is difficult to account for this, but it may reflect the change from 'item' names to 'action' names. The strength of AH4 over age in the initial part of the task, may indicate that I.Q. test score accounts for more of the variance when a task is unfamiliar, and subjects are still in the phase of grasping the task and devising appropriate task strategies. Levels 3-7 indicate that on this relatively simple task of matching names to pictures on which high levels of performance are achieved, age accounts for more of the variance.

Table 6.8

Experiment 6.2

Analysis Of Variance I.Q. Test Score Group / Age Group X Difficulty

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	63.355	2	31.678	3.740	0.001
Age	188.710	2	94.355	11.141	0.001
I.Q. X Age	14.567	4	3.642	0.430	0.032
S-Within	660.570	78	8.469		
Difficulty	569.536	7	81.362	40.304	0.001
I.Q. X Difficulty	69.453	14	4.961	2.457	0.232
Age X Dificulty	55.955	14	3.997	1.980	0.698
I.Q. X Age X Difficulty	80.831	28	2.887	1.430	0.898
Difficulty X S-Within	1102.230	546	2.019		

Table 6.9

Experiment 6.2

Regression Analysis For Each Level Of Difficulty Using Forward Selection

Level 1

....

Step	R-SQR	STD ERROR	VARIABLE		PARTIAL	SIGNIF
1	0.2123	0.64726	AH4	IN	-0.4608	0.000
2	0.3179	0.60566	AGE	IN	0.3661	0.000

Level 2

....

Step	R-SQR	STD ERROR	VARIABLE		PARTIAL	SIGNIF
1	0.2106	0.58562	AH4	IN	-0.4589	0.000
2	0.2916	0.55787	AGE	IN	0.3203	0.002

Level 3

....

Step	R-SQR	STD ERROR	VARIABLE		PARTIAL	SIGNIF
1	0.0874	1.9659	AGE	IN	0.2956	0.004
2	0.1181	1.9431	AH4	IN	-0.1837	0.079

Level 4

....

Step	R-SQR	STD ERROR	VARIABLE		PARTIAL	SIGNIF
1	0.2147	1.1452	AGE	IN	0.4634	0.000
2	0.3163	1.0745	AH4	IN	-0.3596	0.000

Level 5

....

Step	R-SQR	STD ERROR	VARIABLE		PARTIAL	SIGNIF
1	0.1984	1.9019	AGE	IN	0.4455	0.000
2	0.2965	1.7918	AH4	IN	-0.3497	0.000

```
Level 6
.......

Step      R-SQR         STD ERROR     VARIABLE           PARTIAL      SIGNIF

 1        0.1225         1.3036          AH4        IN    -0.3500       0.000

 2        0.1825         1.2652          AGE        IN     0.2614       0.012

Level 7
.......

Step      R-SQR         STD ERROR     VARIABLE           PARTIAL      SIGNIF

 1        0.1059          2.1228         AGE         IN    0.3254       0.002

 2        0.1541          2.0764         AH4         IN   -0.2322       0.027

Level 8
.......

Step      R-SQR         STD ERROR     VARIABLE           PARTIAL      SIGNIF

 1        0.1909          2.8963         AH4         IN   -0.4369       0.000
```

However, when the task reaches a certain level of difficulty, where subjects are likely to make errors, AH4 test score becomes the better predictor.

6.4 DISCUSSION

Both experiments provide evidence for group differences in retrieval from long term memory and memory search, and will therefore be considered together.

The McKenna and Warrington pictures used in Experiment 6.1 had been especially selected to reveal difficulties in retrieval by neurological patients. This may be the reason why the effects of age and I.Q. found with this material were much greater than with the material from the B.P.V.S., which was selected to allow even people with very modest vocabularies to attain some correct responses. In Experiment 6.1 the low I.Q. test score group and older subjects made fewer correct responses and made slower responses than the younger subjects and subjects with high I.Q. test scores. Experiment 6.2 showed the same effects with the matching task. Both of sets of results replicate the findings of Thomas, Fozard and Waugh (1977), in which older people took longer to name pictures of objects than their younger counterparts.

The documentation available on our volunteers allowed group and individual differences to be examined in terms of models of the structure and search processes available in long-term memory. Both correlation matrices show strong positive relationships between the number of correct responses, AH4 test score and Mill Hill vocabulary sores. In addition, the latency of these correct responses correlates negatively with both these indices. It appears that the high I.Q. test scorers and subjects with high vocabulary scores have high levels of performance on both matching and naming latency tasks in terms of speed and accuracy. Age has a stronger relationship with response latency than with accuracy: (R=0.416, 0.457, -0.167, -0.221) for response latency and accuracy respectively. This trend has been found in previous studies, where age effects are greatest when response time is the index of performance on which the comparisons are made (Chapter5).

Waugh and Barr (1980) have argued that the speed factor is responsible for age effects found on memory performance. They suggest that speed of decisions not differences in accuracy as a function of decision difficulty account for many of the age differences found. This implies that if materials are presented slowly enough and subjects are allowed sufficient time to carry out the necessary cognitive processses for a task, older people will perform as accurately as younger controls. The data from these studies show lower correlations between age and accuracy than between age and the speed of correct responding. In both of the tasks subjects were requested to respond in as little time as possible, but there was no time limit imposed upon them. In this way both speed and power factors were measured. The picture presented from the correlation matrix is that chronological age affects speed rather than accuracy.

If the increasing difficulty factor of the tasks is considered, it is clear from both the multiple regression analysis and the analysis of variance that, as the items become more difficult, the older subjects and those of low I.Q. test score take significantly longer to search through memory and to retrieve a name or match a word to one of four pictures presented. From the correlation matrix, it is clear that AH4 test score and Mill Hill vocabulary scale are highly correlated with each other. Hence, those of high I.Q. test score are likely to have large vocabularies. It is interesting to consider memory search time in relation to size of vocabulary. If we consider only the intercept effects i.e. the times for retrieval of names for the easiest items administered in the task we find differences between the groups. This is shown by the difference in intercepts on the X axis of the plotted regression lines, so that even on the easiest items, the old and the low I.Q. test scorers show a lag in retrieval time relative to the younger subjects in the sample and those with high I.Q. test scores. Even those subjects with small vocabularies show frequency effects, i.e. their idiosyncratic rare items take longer than their idiosyncratic frequent items, but the slope of the latency/difficulty function is steeper for the old and low I.Q. test scoring subjects than for the young and high I.Q. groups. If serial search was operating, the prediction would clearly be that since the low I.Q. group hold a smaller number of words in memory, they should be able to find any be word in store faster than the high I.Q. group, who must, presumably search through a larger number of words which they hold in store. This was not found. Thus differences in retrieval time can not reflect differences in the relative sizes of the sets of words through which our subjects conducted a simple search. Differences in retrieval time must rather reflect differences in the organisation of items (words) in the memory store. It seems that individuals with the largest number of items in store also have the best possible organisation of these items. It is the organisational structure of semantic memory, not merely the volume of information it contains, which determines the efficiency of access and retrieval to items held within it.

Of the models suggested in 5.1, that of Norman and Rumelhart (1970) is clearly inappropriate, since they suggest serial search through the lexicon. By developing the Shiffrin (1970) model, in association with a more structured model of semantic memory such as that of Collins and Loftus (1975), group differences can be explained. Shiffrin suggests that latency should be a linear function of the number of draws into a search set prior to response output. The search sets can be considered as a collection of nodes set in a network, as suggested by Quillian (1969). When a picture is presented for naming, it is firstly encoded by the visual coding system, generating a transient visual code based on abstracted information from the visual stimulus. This visual code is used to identify relevant information in the semantic network and the lexicon. If the semantic network is well structured, with many associations between nodes, then the search through semantic memory and the lexicon for the correct name will proceed efficiently. The more associations available, and the better defined the structure of the system, the more pointers and leads available to direct the memory search operations and draws into the set. This model can easily be applied to both matching and naming latencies.

The data suggest that the individuals with a high I.Q. test score, and younger subjects, are able to retrieve names and match them to pictures at a faster rate than the other groups, even though their search set size is larger. This can be explained by their memories having a better organised structure, and the system being richer in terms of the number of associations within it. This feature means that memory search is guided efficiently and effectively. Those of low I.Q. test score, and the older subjects, are likely to operate a more haphazard search, with fewer pointers in a less organised system. Thus age does not simply imply merely a reduction in the number of items available in memory, nor a simple increase in the time required to access any desirable item. Rather it implies some overall lag in memory retrieval time or in response production (since the older subjects took longer on even the simplest items). But, more importantly than this, old age seems to imply a gradual disintegration of organisation of material held in memory so that items rarely accessible become disproportionately harder and so slower to retrieve than items which are, perhaps in daily use suggesting a "use it or lose it" model of the lexicon.

If the relative contributions of age and I.Q. test score to the variance on these tasks are assessed, it is found that on the matching task, where the retrieval difficulty is minimised since subjects are presented with the name of the item or action and have to match the name to one of the four pictures presented, I.Q. test score accounts for more of the variance than age on the initial trials, but once the simple task is grasped, age is the better predictor of correct response time up to the final level of difficulty. On the most difficult slides AH4 part one test score accounts for the largest percentage of the variance with a partialled correlation of -0.4369. For the naming latency task (Experiment 6.1), in which fewer cues are provided, and the task can be described as being more dependent on strategic processing and search, AH4 part one test score accounts for more of the variance than age on all but one of the stages of difficulty. This gives rise to an interesting point, since when a task minimises the retrieval difficulty (as in the matching task where a large number of cues are provided) age is a better predictor of response latency than I.Q. test score, except when the task is found to be very difficult. However, if few cues are provided and the task requires the generation of cues and a greater use of cognitive resources to guide search through long-term memory, I.Q. test score is a better predictor of performance. It also appears that when a task is unfamiliar, and the subject is attempting to devise appropriate strategies for performance, I.Q. test score accounts for more of the variance than age. In the initial trials of both tasks more of the variance was accounted for by the two factors than at any other stage (R-SQR=0.3439, 0.3179).

Group differences were found on both tasks. This emphasises the previous finding in the television news recall studies, where the differences between I.Q. test score groups persisted in recognition as well as free recall. These data are very similar, since the matching task involves similar processes to a recognition memory task. However, the increase in the number of cues does not bring the performance of the low I.Q. test score group to the level of the high group. Not only are the high group able to search and retrieve names from long-term memory at a faster rate, but they can also match names to pictures faster. The benefits of a richly-associated memory base which is well structured come through in both tasks.

The implication of these studies is that the old and the low I.Q. test scorers require more time to retrieve information from long-term memory. Even though their long-term memory contains less information, it is assumed to be less well structured and more impoverished in terms of the number of associations between different concepts, and within concepts. Therefore, fewer clues and pointers are available to select search sets and guide search within the set. These deficits can be seen to affect comprehension and memory in the elderly. They are likely to be slow to recognise events and items, and therefore prone to confusion as they are unable to maintain a pace of identification necessary for comprehension. In a pastime such as watching television this is a serious handicap; scenes change rapidly and slowness of identification means that the theme of the plot may be easily lost. If the elderly are slow to retrieve semantic information as well as lexical information, this may well tend to be responsible for the difficulties they encounter when attempting to keep up with a conversation in a group, or in understanding instructions. These effects have been illustrated in the previous experiments in which subjects with low I.Q. test scores and elderly subjects found difficulty recalling television news broadcasts and drawing inferences from pictorial information. Naming latencies are just one measure of retrieval of information from long-term memory the removal of samentic information found in scripts and frames (Shank and Abelson, 1977) is explored in the following chapter.

CHAPTER SEVEN

7 ASSOCIATIONS TO SCENES

7.1 INTRODUCTION

In everyday life the main use of information stored in memory is to identify what is happening in the present, and predict what is likely to happen in the future. In order to comprehend sentences and visual scenes the presented information needs to be matched with previously stored representations of actions and events.

Schank and Abelson's (1977) 'script theory' has already been referred to as a possible model of the knowledge system. This provides a solution to such problems as how the previous knowledge is structured in long term memory so that inferences can be drawn in the process of elaboration. In the studies reported here, this model is used to explain group differences. A plausible explanation for the elderly's slowness in drawing inferences and deficit in inferential reasoning is found by considering the richness of the scripts or frames in which the knowledge in long-term memory is stored.

The Schank and Abelson model is a specific elaboration of the frame theory proposed by Minsky (1975). It proposes that our knowledge is organised around stereotypic situations with routine activities. Through both direct and indirect experiences, hundreds of cultural stereotypes with some idiosyncratic variations are acquired. These are called scripts or frames. A script is a structure that describes an appropiate sequence of events in a particular context. It is made up of slots and requirements about what can fill the slots. Information is represented in one of two formats either propositional or analogical, and it is either of a primarily perceptual, linguistic or conceptual nature. The structure can be described as an interconnection of slots with what is in one slot affecting what can be in another. Within the model the scripts are organised in conceptual bundles which form a hierarchical system. A concept is a network of propositions or procedures which may exist at several levels of abstraction from single objects and events to scenarios and episodes. A frame of a single object for instance, may have propositions or procedures which indicate what the object looks like, e.g. the propositions will have arguments or 'slots' for visual features and their structural relations, and procedures for detecting these relations and features. Propositions and procedures will also be present to represent what its function is, where it is likely to be found, and how one interacts with it. Frames can also be more global. The frame of a football match will consist of arguments and slots which are frames in themselves, but in this structure they function as headings for the actors of the scenario, the actions which are likely to occur, and the course of events which one can expect.

Scripts or frames have several uses. One function is their use in anticipating and planning for what is going to happen next. With a specific goal in mind, a script is selected from memory whose outcome matches this goal. The relevant conditions for the prior events to occur can then be identified and contrived; in this role scripts are used to plan the future. In a predicting role, the present circumstances are identified by matching events with those of a script and inferences can then be drawn to anticipate the following outcome. It is the role of scripts in comprehension which is of interest here. When information is presented by any channel, long-term memory is searched to find a relevant script or frame. The script retrieved will contain differing levels of abstraction which will depend on the nature of the present information. It follows that the retrieved information from the script is used to fill in the gaps which frequently occur. A gap is missing information in a particular context which is filled in by inferential processing, in this case a 'script' is a mechanism which describes one or several ways in which the inference could take place. In this way script retrieval and the structure and richness of scripts play a fundamental role in comprehension, since within them is stored associations and typical formats of everyday events.

It seems a plausible suggestion that some of the difficulties which the elderly experience in comprehending and drawing inferences from everyday scenes and events, can be explained in terms of differences in the efficiency of their maintenance of 'scripts', which guide their interpretations of everyday life situations.

Light and Anderson (1983) explored age differences in the representation or utilisation of the generic knowledge contained in scripts. They could find no evidence for age-related differences in the way the two groups represented stereotypical action sequences in semantic memory. When they investigated memory for similar actions, the young were better able to recall and recognise new instances of the script than the old. However, both young and older subjects recalled more atypical than typical actions and were more accurate in their discrimination of the explicitly stated from unstated atypical actions. This is a well documented feature of memory for scripts (Bower, Black & Turner, 1979; Graesser, Woll, Kowalski & Smith, 1980; Hasno & Kumar, 1979). Highly typical actions are recalled or falsely recognised even when they are not stated explicitly, due to their activation as part of the generic script. Light and Anderson conclude from the evidence on inferential reasoning that older adults encounter difficulty primarily when inferential reasoning requires storage and manipulation of information, i.e. when working memory is involved. Reasoning in situations that require inferences based on generic knowledge is seen to be reasonably constant across adult years.

One factor that must be taken into account when considering the inferential ability of the elderly is the speed at which they can infer events. Everyday life is composed of dynamic, complex events, which require efficient processing for comprehension to be achieved. Previous evidence suggests that the rate at which information can be retrieved from long-term memory declines with age and I.Q. test score. This is also the case for the speed and accuracy with which inferences are drawn. Both deficits will affect comprehension, and may be associated with the structure and richness of the scenes stored in long-term memory. For instance if the elderly are slow to retrieve information, then in the situations where information is presented quickly by the time they have identified one theme, they may have missed the next, and a third may have already begun. An impoverished script will also hinder comprehension since inferences will be made even more difficult to construct and will take longer to elucidate.

In the studies reported here, both the speed of script retrieval and the richness of the associations in the elderly's scripts were investigated. Group differences were examined in terms of age and AH4 test score. In both studies everyday scene titles were used, for a word generation task and forced-choice reaction time task involving the acceptance or rejection of items for a stated scene. These were used to determine these features of the system. The young and high I.Q. group were expected to have richer scenes and to be faster at identifying members of the presented scenes. The findings were used to interpret group differences in comprehension from previous studies.

7.2 EXPERIMENT 7.1

7.2.1 Introduction

The main purpose of this experiment was to generate normative data from an elderly population, to be used in Experiment 7.2. The task simply involved generating nouns to eleven scene titles (e.g. The School, The Church, The Seaside). From the overall frequency counts of the words generated, a measure of the associative strength of each word was obtained.

The data derived from this task are by themselves of interest. It has been suggested in Chapters 4 and 5, that a possible explanation for the deficits in comprehension and recall of pictorial material and picture sequences, can be based on the structure and nature of the schemata within long-term memory. By asking subjects to generate items in response to titles of everyday scenes, an indication of the richness of the associations stored is obtained and the fluency with which associations are retrieved.

Light and Anderson (1983) found little evidence for age-related differences in the way that stereotypical action sequences were represented in memory. In this study subjects generated scripts for daily routine activities e.g. grocery shopping, going to the doctor, writing a letter to a friend. To date this is the only research which compares the structure of information in semantic memory in younger and older adults using stimuli which are more complex than single words or word triads in word association tasks (eg., Howard, 1980; Lovelace & Cooley, 1982; Perlmutter, 1978; Riegel & Birren, 1966; Riegel & Riegel, 1964). The data from these studies are inconclusive, as some researchers report greater variability of response by older subjects on word association tasks (e.g., Lovelace & Cooley, 1982; Perlmutter, 1979; Riegel & Riegel, 1964), whereas others have not found this to be the case (e.g., Howard, 1980; Perlmutter & Mitchell, 1982). A more general finding from these studies is that older adults are less fluent, producing fewer responses in timed tasks (Howard, 1980; Riegel & Birren, 1966; Schaie & Parham, 1977; Schonfield & Stones, 1979; Stones, 1978). However, Perlmutter (1978) found no effect of age on the number of associates produced when subjects worked at their own pace.

The data generated from this study allow for a much wider range of associations than other more constrained tasks have encouraged or allowed. For example in Light and Anderson's experiment subjects were required to use general everyday descriptions of routines, which means that the scripts described did not indicate the richness of those in memory, merely the type of ordinary routines stored. However, by using a relatively easy, unconstrained task, which allows subjects to generate even very remote associations in which there is no explicit criterion for adequacy, the task should maximise the number of associations produced by all groups. This means that if the older subjects make their own idiosyncratic associations in comparison with the other groups they are not penalised for it. Hence the richness and diversity of the schemata stored in memory should be revealed.

7.2.2 Method

7.2.3 Subjects

150 elderly volunteers from the North East Age Research project were selected on the basis of their age and AH4 part one test scores. The same three AH4 test bands were constructed as in previous studies, and these were filled with 50 subjects in each, ensuring an even distribution of age within each band. All subjects were paid for their participation in this study. The structure of the bands is shown in Table 7.1. Mean age was 64.55 years, std. dev.=7.28 and mean AH4 score 31.53/65, std. dev.=13.19.

7.2.4 Procedure

Subjects were tested in groups of approximately twelve. Each subject was given a booklet containing the eleven scene titles shown below:

The School
The Birthday Party
The Church
The Picnic
The Restuarant
The Fairground
The Bathroom
The Park
The Post Office
The Doctor's Surgery
The Seaside

Instructions were printed on the cover of each booklet, and also fully explained by the experimenter. They were strictly instructed not to turn over pages until told to do so. On each page, the title of an everyday

Table 7.1

3x3 Matrix Of Subjects For Age and AH4 Test Score Groups

AH4 Test Score Group Age Group	Low Group 1	Middle Group 2	High Group 3	Total Mean (x) Std. Dev. Range (r)
Young Group 1	13	17	16	N=46 x=56.02 std.dev.=2.02 r=52-59 years
Middle Group 2	16	19	18	N=53 x=64.07 std.dev.=2.58 r=60-69 years
Old Group 3	17	18	16	N=51 x=72.74 std.dev=3.47 r=70-85 years
Total Mean (x) Std. Dev. Range	N=46 x=16.65 std.dev.=3.89 r=2-20/65	N=54 x=30.15 std.dev.=5.88 r=21-39/65	N=50 x=46.68 std.dev.=5.58 r=40-61/65	N=150 x=31.53 64.55 std.dev.=13.19 7.28 r=2-61/65 53-83 years

scene was printed. Subjects were told to generate as many items as possible associated with each scene, in two minutes. They were told that the time available was limited. Following each generation, subjects were given a short break before commencing the next scene title.

7.2.5 Results

The number of words generated by each subject to each scene title were counted, and then summed over the eleven scenes to give a grand total of words generated for the entire task. The descriptive statistics for the grand totals are shown in Table 7.2 according to age and I.Q. test score groups. These data were used in an analysis of variance with age and I.Q. test score group as betweensubject factor variables. There were significant main effects of age and I.Q. test score group (F=19.841, 32.9 df=2,8 P<0.001) respectively. The two way interaction was also significant (F=3.361 df=4,8 P<0.012) see Figure 7.1 and Table 7.3. The young and the high I.Q. test scoring volunteers generated more words overall than the other groups.

Frequency counts were computed for each word generated to give an index of the how many subjects generated a specific word in association to a particular category. For each subject the number of words which only they produced i.e. with a frequency count of one, to particular category were counted, and then a grand total of number of 'original' words for each subject was calculated. These scores were used in an analysis of variance with age and I.Q. test score group as

Table 7.2.

Descriptive Statistics Of Total Number Of Words Generated For I.Q. Test Score Groups & Age Groups

Total Scores Over All Scene Titles

I.Q. Test Score Group	N	MIN.	MAX.	MEAN	S.E. MEAN
Low	46	54.00	217.00	129.587	04.821
Middle	53	67.00	266.00	165.302	06.040
High	50	98.00	343.00	200.040	07.511
Age Group					
Low	46	54.00	343.00	182.543	09.314
Middle	53	85.00	267.00	165.453	06.613
High	51	78.00	242.00	150.451	05.718

Table 7.3

Experiment 7.1

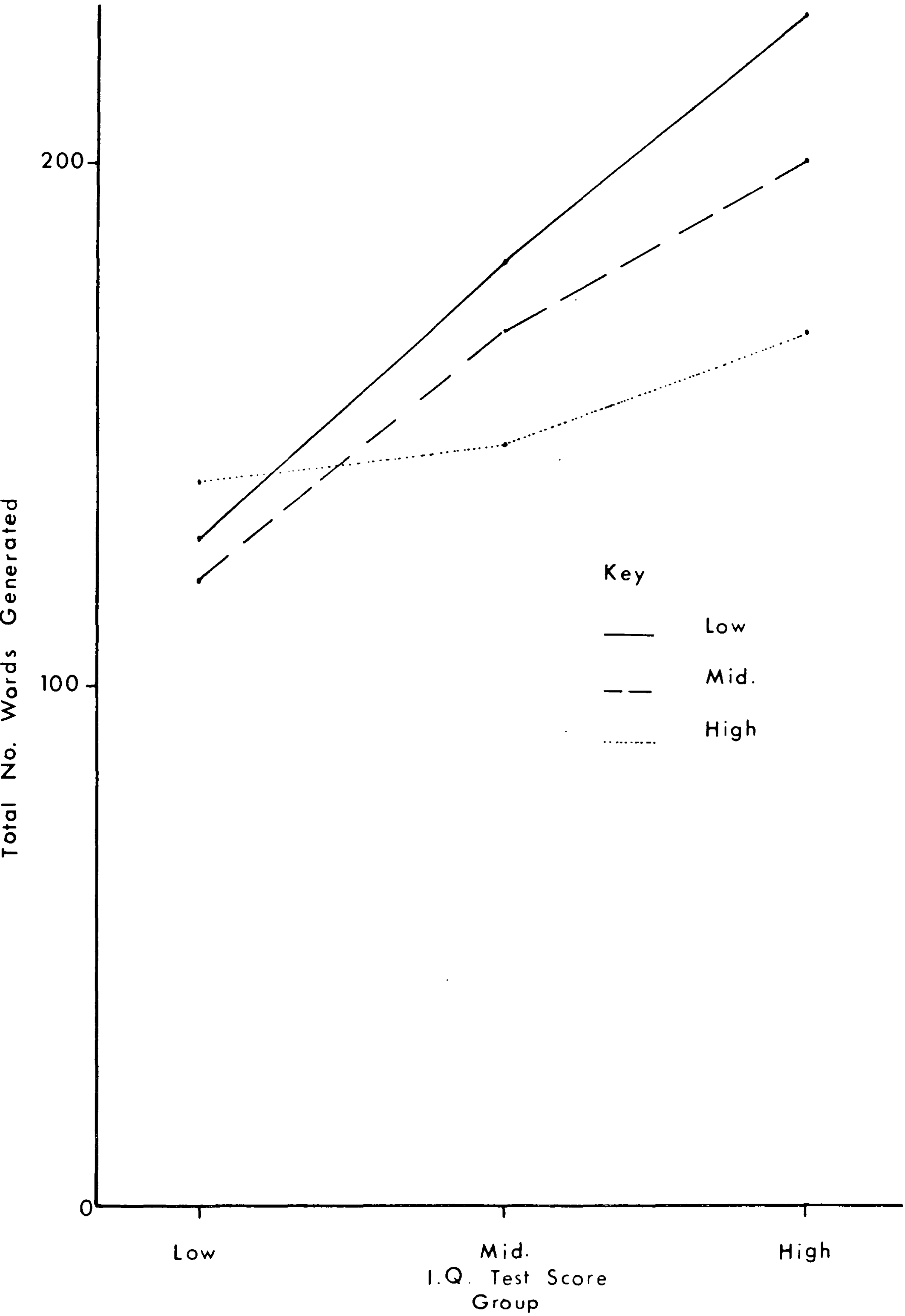
. . . .

Analysis Of Variance   I.Q. Test Score Group / Age Group X Total Number Of Words Generated.

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Main Effects	138340.545	4	34585.136	19.841	0.000
I.Q.	114724.481	2	57362.241	32.907	0.000
Age	19387.458	2	9693.729	05.561	0.005
I.Q. X Age	23437.548	4	5859.387	03.361	0.012
Explained	161778.093	8	20222.262	11.601	0.000
Residual	244039.236	140	1743.137		
Total	405817.329	148	2742.009		

Fig. 7.1

Interaction Between Age & I.Q. Test Score Groups.



between-subject factor variables to examine group differences. Again there were strong effects of age and I.Q. test score group (F=3.00, 9.813 df=2,4 P<0.05 0.01 .) This means that not only do the young and high I.Q. test score groups generate more associations to everyday scene titles, but thay also generate more 'original' associations which no one else in the sample produced. However, when the proportion of 'original' words to total number of words generated was calculated for each subject and placed in a similar analysis of variance, there were no significant main effects (F=2.217, 0.752 P<0.113 P<0.474). See Tables 7.4, 7.5 for full analyses of variance.

Separate analysis of variance were computed for each scene title for number of words generated, and number of 'original' words, with age and I.Q. test score group as between subject variables. A complete summary table of these analyses can be found in Appendix 7. There were main effects of age and I.Q. test score group on number of words generated for every scene title, and for the number of 'original' words for all but two titles.

Appendix 8 gives a sample of words generated and their frequency counts for each scene title. A cross tabulation of frequency and category, showed that the largest count of 'original' associations was generated by 'The School' and the least by 'The Birthday Party'.

Table 7.4

Experiment 7.1

....

Analysis Of Variance I.Q. Test Score Group / Age Group X Total Number Of 'Original' Words.

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Main Effects	7347.280	4	1836.820	6.580	0.000
I.Q.	5478.590	2	2739.295	9.813	0.000
Age	1676.362	2	838.181	3.003	0.052
I.Q. X Age	2054.086	4	513.521	1.840	0.125
Explained	9401.366	8	1175.171	4.210	0.000
Residual	39082.098	140	279.158		
Total	48483.463	148	327.591		

GROUP(10)>

Table 7.5

Experiment 7.1

....

Analysis Of Variance I.Q. Test Score Group / Age Group X Proportion Of Original Words

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Main Effects	262.852	4	65.713	1.487	0.209
I.Q.	195.859	2	97.929	2.217	0.113
Age	66.413	2	33.206	0.752	0.474
I.Q. X Age	34.333	4	8.583	0.914	0.941
Explained	297.185	8	37.148	0.841	0.568
Residual	6185.290	140	44.181		
Total	6482.475	148	43.801		

7.2.6 Discussion

The two main findings from this experiment are that the elderly and the low I.Q. test scorers generate fewer associations in two minutes to everyday scene titles, and of the words they generate fewer are likely to be 'original' i.e with a frequency count of one, from a frequency count of all subjects for that particular scene. Previous studies have found that the elderly are less fluent at producing responses in timed tests (Howard, 1980; Reigel and Birren, 1966; Schaie and Parham, 1977; Schonfield and Stones, 1979; Stones, 1978). There are two explanations for this result; first the elderly are less efficient at accessing and retrieving scripts or scenes in semantic memory, which may be due to a poorly organised knowledge structure, or second, the elderly store impoverished scenes which contain fewer 'slots' (Minsky, 1975) so that less data and fewer associations are stored in memory. However, a further consideration can not be ignored. Since this task was a timed task, it is possible that the low I.Q. test score group and the elderly may have been restricted by writing speed leading to fewer associations.

Experiments 6.1 and 6.2 have already suggested that the elderly are likely to possess a poorly organised semantic memory, because retrieval of words present in their vocabulary is slower than retrieval for the young and high I.Q. test score group. The lag in retrieval time found in these previous experiments could explain why the elderly retrieve fewer words in two minutes than the younger and high I.Q. test scores. Hence it is plausible to suggest from both Experiments 6.1 and 6.2 and this study, that the organisation of semantic memory deteriorates with age which leads to a lag in access and retrieval time. These data suggest that the elderly and low I.Q. test score group can not retrieve information from memory as fast as the young and high I.Q. test scorers. This can be seen to affect comprehension of everyday events. If an elderly person is less able to retrieve associations and words quickly because of a disorganised knowledge base, when scenes and events are encountered familiar to daily living, it will take longer for identification and comprehension to occur. Further, since their semantic memory is impoverished, the elderly have fewer expectations to guide the sensory search of the scene, and to form a basis for interpretation and comprehension. This means that more inferential reasoning is generally necessary, and processing is less efficient.

7.3 EXPERIMENT 7.2

7.3.1 Introduction

Experiment 7.2 examines the speed of search by the elderly for associations in long-term memory. Further, a possible description of the associative network in which the search occurs is also investigated. A forced-choice reaction time task is used to provide an index of search time, and the number of associations accepted as belonging to each scene. From these data, it is possible to gain useful insight into the structure and organisation of the knowledge system which can be used effectively as a basis for the discussion of group differences in comprehension and recall of everyday events.

There are several parallels between this study and Experiments 6.1 and 6.2. Both consider search processes in long-term memory; the latter concerns the retrieval of names of objects and the former items associated with scene titles. As in Chapter 6, the size of the search set is a necessary consideration in such studies. The speed of both spoken and written associations to word stimuli has been shown to decline with age (Birren, Riegel and Robbin, 1962; Howard, 1982; Perlmutter, 1978; Riegel and Birren, 1966; Riegel and Riegel, 1964). This effect is associated with a rise in the number of words known (Riegel and Birren, 1966). However, Experiments 6.1 and 6.2 have clearly shown that a large vocabulary is associated with faster search times and an increase in the number of correct responses. It was suggested that group differences in this task could be explained by the high I.Q. test score group and the younger subjects having better organised memory nodes, and also having a richer network of associations between conceptual nodes which direct the search for names efficiently. There is no reason to suppose the use of more semantic material in this experiment would not show similar effects.

The task used here involves activating the relevant schemata in long-term memory in response to the presented scene title, and then searching the schema to find or derive an association between the presented noun and the stored information. It is predicted that the high I.Q. test score group store more associations within their schemata, and are better able to form associations. They possess a much richer knowledge structure with a denser network of associations and connections. When information of any kind is presented, the nature of this structure means that relevant information of any kind is rapidly retrieved, because the search is closely directed by the numerous associations available.

7.3.2 Materials

The words generated in response to each of the eleven scenes just referred to in Experiment 7.1 were analysed for their frequency of occurrence within the pool of those generated by all volunteers. For each scene, a high association set of words was constructed, containing the twenty words which occurred with the highest frequency. Similarly, a second, low association group was also constructed, containing twenty words which had each received a single citation by an individual subject. A third group of twenty words with no association was also constructed for each scene. For this purpose, the words selected were those which had received a single citation from one of the other ten scenes. All sixty words for each scene were balanced for phonemic length and frequency of everyday usage using Postman and Keppel (1970). Ten of the eleven scenes were used in the experiment, while the remaining one was used for a practice trial. Appendix 9 contains lists of the words used.

7.3.3 Subjects

94 volunteers were selected according to their age and AH4 part one test score, as in all previously reported experiments. They were aged between 49 and 85 years, mean age 63.319 years. A 3X3 matrix of three age groups and three AH4 part one test score groups was filled as shown in Table 7.6. All subjects were paid 2 pounds to cover travelling expenses. None of these volunteers had participated in Experiment 7.1.

7.3.4 Procedure

A baseline of the subjects' reaction times in a semantic forced-choice procedure was obtained by giving them a short preliminary task. Fifty words were presented one at a time on the BBC microcomputer screen, and volunteers simply had to classify each, in turn, as the name of a 'living' or 'non-living' item. They were presented with 10 trial words to familiarise themselves with the task and the position of the 'yes' and 'no' response keys.

The actual experiment followed after a short break. Full instructions were written into the BBC programme, and both response-type and response latencies were stored on disc. The subjects were told that the title of a scene would appear at the top of the screen (e.g. 'The Supermarket') and underneath it a word. The task was to decide whether there was any association between the scene title and the word. Pressing the 'Yes' button indicated that some association was realised, and the 'No' button

Table 7.6

3x3 Matrix Of Subjects For Age and AH4 Test Score Groups

AH4 Test Score Group Age Group	Low Group 1	Middle Group 2	High Group 3	Total Mean (x) Std. Dev. Range (r)
Young Group 1	8	12	11	N=31 x=54.87 std.dev.=2.25 r=49-58 years
Middle Group 2	12	12	8	N=32 x=62.72 std.dev.=2.75 r=59-67 years
Old Group 3	10	10	11	N=31 x=72.39 std.dev=4.62 r=68-85 years
Total Mean (x) Std. Dev. Range	N=30 x=14.27 std.dev.=3.90 r=7-20/65	N=34 x=30.59 std.dev.=5.37 r=21-39/65	N=30 x=47.23 std.dev.=5.03 r=40-58/65	N=94 x=30.69 63.32 std.dev.=14.08 7.90 r=7-58/65 49-85 years

indicated that there was no connection seen between the two. For all trials the scene title remained on the screen so no memory component was necessary in order to establish which scene title the words were to be tested against. On all trials, subjects were asked to press the response keys as quickly as they could. After each set of words was presented for a scene, subjects were allowed to take a short break. There was also a break of at least ten minutes at the half-way stage. The experimenter began with a practice session of 30 words, ten from each association group. During the practice trials, subjects had an opportunity to ask questions and ensure that they fully comprehended the task. It was the experimenter's impression that all subjects found the instructions clear, and the task straightforward. The ten scene titles were presented in four different random sequences which were randomly allocated to testing sessions, and the presentation of the sixty words for each scene was randomised for each subject. Response type and response times were recorded on disk along with order of word presentation. There was no 'correct' response in this task, in each case the subject's choice was considered to be valid and the response type and time recorded.

7.3.5 Results

The preliminary study, to establish if group differences were present in a simple semantic forced-choice reaction time task (i.e. whether the presented word was living or non-living) was analysed first. In each case, the first response for each subject was eliminated, since this was found to be inconsistent with the subject's subsequent performance.

Median response times were computed for correct and incorrect responses, and the number of correct responses and number of errors counted for each subject. The overall mean number of correct responses was very high (47.56/49) and hence mean error rate was low (2.5/49). The mean correct response time overall, computed from the median scores, was 0.82 seconds.

An analysis of variance was used to compare the performance of the I.Q. test score groups and age groups, for response times and numbers of correct responses. There were no main effects in either analysis (see Table 7.7). This means that any differences found in the second experiment will be due to the additional processing required for this task, and not to differences in force-choice reaction times.

The analysis of Experiment 7.2 centred on the effect of strength of association on response latency. Median response times were computed for each subject, for the three association strength categories. The overall means of these are shown in Table 7.8. As expected, responses were slowest to words of no association, and fastest to words of greatest association. In addition, the number of positive responses was counted for each subject, again in relation to the associative strength category of the words: overall means for these are shown in Table 7.9.

A 2X1 analysis of variance was computed on subjects' response times, with I.Q. test score group and age as between subject factors and associative strength as the within subject factor. There were significant main effects of all three factors, and a significant

Table 7.7

1X1 Analysis Of Variance Age Group X Number Of Correct Responses

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Age	6.49	2	3.2475	1.161	0.3179
S-Within	254.62	91	2.7980		
Total	261.11	93			

1X1 Analysis Of Variance Age Group X Correct Response Time

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
Age	142.05	2	71.023	0.158	0.8537
S-Within	40799.0	91	448.340		
Total	40941.05	93			

Table 7.7 (Continued)

1X1 Analysis Of Variance   I.Q. Test Score Group X Number of Correct Responses

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	10.202	2	5.1010	1.850	0.1631
S-Within	250.92	91	2.7573		
Total	261.12	93			

1X1 Analysis Of Variance   I.Q. Test Score Group X Correct Response Time

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	781.15	2	390.58	0.885	0.4162
S-Within	40160.0	91	441.310		
Total	40941.15	93			

Table 7.8

<u>Descriptive Statistics For Age Groups</u>

Age Group 1

. . . .

	N	MIN.	MAX.	MEAN	STD. DEV.
HIGH ASSOCIATION WORDS (MSEC)	31	539.00	860.00	672.84	77.551
MEDIUM ASSOCIATION WORDS (MSEC)	31	633.00	1150.00	834.94	143.990
LOW ASSOCIATION WORDS (MSEC)	31	659.00	1524.00	886.00	209.040

Age Group 2

. . . .

	N	MIN.	MAX.	MEAN	STD. DEV.
HIGH ASSOCIATION WORDS (MSEC)	32	514.00	1293.00	757.50	169.280
MEDIUM ASSOCIATION WORDS (MSEC)	32	570.00	1534.00	921.00	211.230
LOW ASSOCIATION WORDS (MSEC)	32	588.00	1362.00	897.00	190.240

Age Group 3

. . . .

	N	MIN.	MAX.	MEAN	STD. DEV.
HIGH ASSOCIATION WORDS (MSEC)	31	534.00	1405.00	755.32	151.740
MEDIUM ASSOCIATION WORDS (MSEC)	31	608.00	1739.00	920.45	207.680
LOW ASSOCIATION WORDS (MSEC))	31	671.00	1701.00	959.29	237.680

Table 7.8 (Continued)

DESCRIPTIVE STATISTICS I.Q TEST SCORE GROUPS

I.Q. Test Score Group 1

	N	MIN.	MAX.	MEAN	STD. DEV
HIGH ASSOCIATION WORDS (MSEC)	30	564.00	1405.00	790.03	189.510
MEDIUM ASSOCIATION WORDS (MSEC)	30	658.00	1739.00	969.83	254.400
LOW ASSOCIATION WORDS (MSEC)	30	657.00	1701.00	966.83	233.860

I.Q. Test Score Group 2

	N	MIN.	MAX.	MEAN	STD. DEV
HIGH ASSOCIATION WORDS (MSEC)	34	571.00	1058.00	735.44	105.410
MEDIUM ASSOCIATION WORDS (MSEC)	34	661.00	1186.00	915.38	134.460
LOW ASSOCIATION WORDS (MSEC)	34	654.00	1572.00	979.88	220.110

I.Q. Test Score Group 3

	N	MIN.	MAX.	MEAN	STD. DEV
HIGH ASSOCIATION WORDS (MSEC)	30	514.00	886.00	660.23	91.268
MEDIUM ASSOCIATION WORDS (MSEC)	30	570.00	1076.00	789.73	126.700
LOW ASSOCIATION WORDS (MSEC)	30	588.00	1045.00	787.10	105.800

Table 7.9

Descriptive Statistics For Number Of Positive Responses Overall Trials

AGE GROUP	N	MIN.	MAX.	MEAN	STD. DEV.
LOW	31	298.00	495.00	318.19	48.231
MIDDLE	39	281.00	468.00	360.05	38.549
HIGH	24	162.00	465.00	371.54	59.027

I.Q. TEST SCORE GROUP	N	MIN.	MAX.	MEAN	STD. DEV.
LOW	30	162.00	465.00	350.73	57.408
MIDDLE	34	298.00	495.00	383.56	45.810
HIGH	30	299.00	431.00	373.77	32.986

interaction between I.Q. test score group and associative strength. This arose because the high I.Q. test score group are significantly faster on all three associative strength categories, and the middle I.Q. test score group show a significant difference in response times between the low association and no association words, which the other groups do not show (Figure 7.2 and Table 7.10)

The numbers of positive responses were analysed separately for each associative strength group. There were significant effects of I.Q. test score group for both the high and low association groups. This effect arose because the low group in each case gave fewer positive responses than the other groups. All other effects were insignificant (see Table 7.11).

7.3.6 Discussion

These data are consistent with the findings in Chapter 6, and can easily be interpreted using the model of long-term memory suggested there. As expected, the high and middle I.Q. test score groups formed or recognised more associations between the scene title and the presented words than the low I.Q. test score group for both the high and low association words. There were no age significant age effects, but since one would expect this measure to have some association with vocabulary size, it is not surprising. From these data, the suggestion that the high I.Q. test scorers have a richer knowledge structure and a denser network of associations is confirmed.

Table 7.10

Experiment 7.2

. . . .

Response Times

Analysis Of Variance I.Q. Test Score Group / Age Group X Word Association Strength

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	1234455.00	2	617227.50	08.310	0.001
Age	525246.12	2	262623.06	03.536	0.003
I.Q. X Age	79770.125	4	19942.531	0.268	0.897
S-Within	6313568.00	85	74277.250		
Association	1879118.00	2	939559.00	135.182	0.001
I.Q. X Association	125126.56	4	31281.64	4.501	0.002
Age X Association	45832.16	4	11458.04	1.649	0.164
I.Q. X Age X Difficulty	29021.75	8	3627.72	0.522	0.839
Association X S-Within	1181552.00	170	6950.30		

Fig. 7. 2

Interaction Between I.Q. Test Score Group & Word Association Strength.

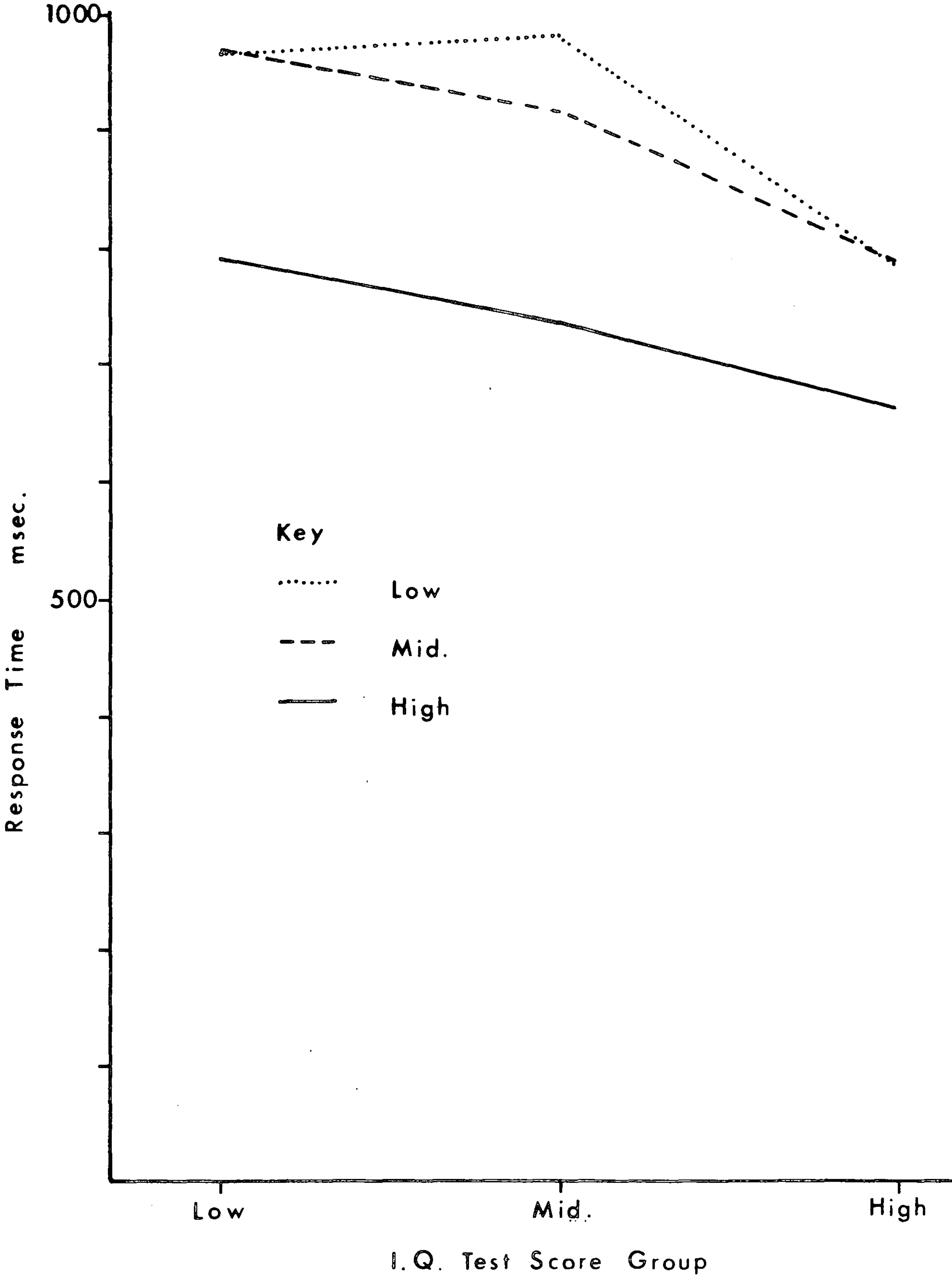


Table 7.11

Experiment 7.2

....

Number Of Positive Responses

Analysis Of Variance   I.Q. Test Score Group / Age Group X Word Association Strength

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
I.Q.	5075.46	2	2537.730	03.541	0.033
Age	2080.86	2	1040.429	01.452	0.240
I.Q. X Age	2870.70	4	717.676	01.001	0.411
S-Within	60915.00	85	716.647		
Association	1202798.00	2	601399.00	3097.647	0.001
I.Q. X Association	2981.59	4	745.39	03.839	0.005
Age X Association	1875.81	4	468.95	02.415	0.051
I.Q. X Age X Difficulty	2824.86	8	353.12	01.819	0.077
Association X S-Within	33005.00	170	194.15		

The second finding, of the effect of age and I.Q. test score on speed of response latency with respect to the rejection or acceptance of the associations, is again associated with the results in Chapter 6. Decision times for the presence or absence of an association were significantly faster for all associative strengths in the high I.Q. test score group. This is most pronounced for the low and no association words. Again, the effect of search set size appears to go in the opposite direction to that predicted by serial search or an interference theory. It appears that a rich associative network connecting nodes in a Quillian (1969) manner gives rise to a well directed and efficient search through the relevant schemata. This point is emphasised particularly by the search speed differences for the high association words. In this instance, the majority of the words presented will be near the top of a hierarchical system within the node, or close to the node if a Collins and Loftus (1975) model of semantic relatedness is considered. However, in spite of this high degree of relatedness being present in a smaller set, the low I.Q. test score group are slower to identify the presence of an association.

An explanation of these findings is made in terms of the structure and organisation of schemata in long term memory. There is undoubtedly a general speed factor accounting for some of the variance in this task, but it cannot sufficiently explain the group differences presented. The model suggested in Chapter 6 can easily be adapted to explain these findings. When a scene title is presented, the relevant schema is activated in long-term memory. The following presentation of an item initiates a search through the schema and, if necessary, a search through associated schemata if the original search is unsuccessful. If the network of associations within a schema is dense, the search for the item will be well directed and the decision made quickly, whether positive or negative. A negative response in all cases was associated with a longer search time, since the search through the schema can be assumed to be exhausted before this response is given. As Shiffrin (1970) suggests, the latency of response should be a linear function of the number of draws into the search set. If the draws are well directed by the richness of the information present in the set, fewer draws are likely to be needed and the response latency will be reduced.

The data from both of these experiments suggest that the elderly with low I.Q. test scores are slow to access and retrieve information concerning everyday scenes. This lag in retrieval time is unlikely to be due to the scene titles being unfamiliar to subjects or the task being difficult in any way. A possible explanation for the slow access and retrieval is that their semantic memory is poorly organised and contains fewer associations to guide search. Further, the diminished number of associations both within and between scenes has implications for the speed and accuracy of comprehension for everyday events. Fewer expectations will be generated concerning the contents and likely development of the event encountered, hence comprehension and analysis of the sensory stimuli is poorly directed and will be forced to rely on more inferential reasoning. These decrements lead to resource expensive processing and less efficient and accurate comprehension.

In summary, from these experiments we can not claim that the low I.Q. test score group and older people have fewer associations to scenarios held in long-term memory, but these results suggest that they do not generate or retrieve them as rapidly as younger subjects or those with a high I.Q. test score. From Experiment 7.2, it is clear that the old and low I.Q. test score group accept as valid fewer associations than the young and high I.Q. test score group. The final point from these studies is that the high I.Q. test score group and the young are much faster at validating associations, this has implications for everyday cognitive performance. Correct recognition and appropriate anticipation are necessary cognitive skills for the comprehension of fast changing events as portrayed on television or encountered in everyday living.

CHAPTER EIGHT

8 DISCUSSION

8.1 INTRODUCTION

This thesis has discussed how advancing age affects the everyday task of comprehending complex, dynamic events. In doing so, it has also considered which of a number of indices best predicts changes in performance accompanying advancing age, and attempts to develop a better understanding of the ways in which the elderly are handicapped in complex situations in everyday life. The assessment of changes in performance with increasing chronological age will be addressed first, since it has bearing on the design and interpretation of all the experiments reported.

8.2 WHAT PREDICTS THE CHANGE IN PERFORMANCE WITH INCREASING AGE?

One of the most striking aspects of these data is that few effects of chronological age per se could be found. This should not be surprising on reflection since from neurobiology it is evident that people 'age' at different rates, and there could well be several different patterns of ageing relating to various changes in pathology as well as social and environmental factors. In addition, many researchers have found an increase in inter-subject variance on different performance measures with elderly populations. This also suggests that age per se is a poor predictor of changes in performance.

From these studies, the best index of the cognitive effects of ageing appears to be I.Q. test score. In many of the performance measures taken, the cognitive effects of increasing age have not been separated from those of a decline in I.Q. test score. This will be in part associated with the particular I.Q. test chosen. It could be that if a different I.Q. test had been used, more or less of the variance would have been accounted for by this factor, and this would have either increased or reduced the likelihood of isolating apparant age effects. What remains is the fact that although there are evidently gross changes with age, most of those which affected performance in our tasks, and which therefore probably affect performance in the comprehension of daily events and television broadcasts are all picked up by the single, simple I.Q. test we used. There seems to be rather little residual variance which chronological age accounts for - atleast in the measures we have used.

Age and I.Q. test score effects have been separated in the data, but their independent effects have not been found to be associated with one specific component of the tasks. Waugh and Barr (1980) have suggested that it is the speed factor as opposed to the power factor of a test which accounts for the age effects on memory. However, our experiments found age effects on both (see Chapter 4 and Chapter 6).

The selection of subjects on the basis of age and I.Q. test score allowed both age and I.Q. test score to be assessed as indices of changes in cognitive performance. When individual differences in performance were examined, the score on AH4 part one picked up most of the variance on the measures taken. Hence the data show the relationship between this single intelligence test and performance on a number of different cognitive tasks including memory for complex everyday events, inference construction, and retrieval from long-term memory. The fact that the AH4 part one test is a good predictor of performance on these tasks means that it may be a very useful index of change in cognitive performance with age. The shape of the function of change in cognitive performance with increasing chronological age is still debated. There are two plausible models i) suggests that a maximum level of functioning is reached at some point in young adulthood, which is followed by a linear or linear accelerating decrement, and ii) suggests a period of stability from maturity to the terminal drop (Riegel and Riegel, 1972). Therefore, it is important to develop tests which are first, sensitive to changes in performance with increasing chronological age, and second, tests that can detect which processes change first and are affected most by the ageing process. In addition, tests need to illuminate to what extent age changes are related to one or a number of factors.For example: are changes related to speed and power of processing, memory capacity, or computational power in reasoning as shown by Cohen (1981) in the interpretations of inference deductions from propositions in text? It is easy to see how a change in any one of these factors may bring about apparent changes in other systems which would be otherwise unaffected by ageing.

8.3 COMPREHENDING AND RECALLING EVERYDAY EVENTS

On a global level, the studies of memory for television broadcasts demonstrated that elderly people with low I.Q. test scores have difficulty recalling facts and details from such an information source. In a free recall task they recalled on average 6.8% of the details given and in a multiple choice task 38.5%. The experiments which followed have attempted to identify the processes which explain the group differences in performance on this everyday memory task.

8.3.1 Differences In Strategic Processing

Baron (1978) presents evidence that many differences in performance between groups can be accounted for by differences in the use of strategies. In the television recall studies there were two opportunities for differences in strategic processing to appear. First, free recall offers fewer cues for retrieval than a multiple choice task. Hence, it is plausible to expect a greater employment of strategical processing in a free recall task. If group differences could be explained by the use and sophistication of strategies, then group differences should be greater in free recall than multiple choice tasks. This was not found: Differences between the high and low I.Q. test score groups were significant for both recall conditions, and indeed the high I.Q. test score group showed greatest improvement in percentage recall for the multiple choice recognition task.

Throughout these studies the main interest has been the presence or absence of an age effect. By using a wide range of elderly subjects with a wide range of AH4 part one scores it was found the AH4 part one accounted for more of the variance than age per se.

There is a mean decline of AH4 score with age (Rabbitt 1986 personel communication). However, the selection procedure used according to AH4 score and age may be critised for being biased towards AH4 effects. This method of selection may have failed to included the very old and the very poor AH4 test scorers. But since we know that AH4 test score declines with age then the effects found are indeed age effects however, here they are seen to equate to a decline in AH4 test score with AH4 test score picking up most of the variance on the tasks used.

In the few instances where age is the better predictor of performance there is a strong speed component to the task which suggests a speed of processing or motor control deficit. Generally, it was difficult to isolate the effects of ageing from those of a decline in AH4 test score. By using a balanced design the relative contributions of these variables have been examined.

The second opportunity for group differences in the use of strategies to be illustrated was in the analysis of the details recalled under free recall conditions. A useful strategy to employ when presented with such a large amount of information for retrieval would be to attend selectively to the thematically essential details. The stored themes could then be used as cues for the retrieval of additional details. If this strategy was used by the high I.Q. test score group and not by the low I.Q. test score group then the proportion of correctly recalled thematically relevant to thematically irrelevant should be greater in the high scorers than the low scorers. However, the high I.Q. test score group recalled a significantly higher percentage of all detail types than the low I.Q. test score group, and showed no significant difference between the percentage recall of thematically essential and of irrelevant details. So although this strategy may have been implemented it has not been isolated as a source of group differences in this analysis. An equally plausible explanation can be made in terms of differences in memory capacity and structure, or the efficiency of the memory system; both would result in higher recall of all detail types and higher recall under both recall conditions.

Experiments 3.1 and 3.2 investigated the employment of strategies in a more rigorous fashion. The task required subjects to categorise presented words but at the same time attempt to memorise the words belonging to an emphasised category. Strategic processing beyond grouping the words semantically is obviously required. In both a free recall and a recognition test the high I.Q. test score group recalled significantly more of the words from the emphasised category than the low I.Q. test score group. However, the same was true of the remaining category words. Thus here again, no evidence was found that the extent of 'strategic focussing' was any greater in the high I.Q. test score group, i.e. that high I.Q. scorers were employing strategic focussing more effectively than the low I.Q. test scorers. So there is no evidence that the higher performance of the high test score group was due to other factors than differences in the efficiency of their information processing system, and in the structure of their long term memory data bases. If the difference was due to strategic processing poorer recall should have occurred for the remaining category words by the high I.Q. test score group.

It was disappointing that these studies did not capture differences in strategic processing by the high and low I.Q. test score groups, but this by no means implies that it does not contribute to group differences in the memory and comprehension of everyday events. It is probable that the design of Experiments 2.1 and 2.2 were insensitive to this effect since all subjects knew that a recall test would be given, and these task demands could have changed the nature of the strategies they adopted. In Experiments 3.1 and 3.2 the lack of conclusive evidence was more puzzling. All groups recalled a significantly higher percentage of the words from the stated category than the remaining categories. This suggests that they were all able to employ some effective means of emphasising these words beyond classification. Nonetheless, the fact that the high I.Q. test score group could recall a significantly higher percentage of the remaining words than the other groups, again implies that processing efficiency or memory structure are involved in group differences.

8.3.2 Group Differences In Theme Construction and Inferential Reasoning

In the previous section (8.2), it was suggested that the high I.Q. test score group may adopt a strategy of attending selectively to thematically essential details in a television broadcast as a means of providing cues for subsequent recall. A possible source of group differences in the task of retrieving and comprehending news broadcasts is the identification of thematically relevant material. Elderly people with low I.Q. test scores may have problems identifying thematically essential material and drawing plausible inferences from the presented information in order to construct the relevant theme. These hypotheses were rendered plausible by the results of Experiments 4.1 and 5.1. The low I.Q. test score group were less accurate in their decisions as to whether a target slide was a plausible conclusion to a theme shown in previous slides, and also recalled fewer plausible inferences from black and white photographs of events. In the latter case they failed to respond to more questions than the high I.Q. test score group and gave more implausible inferences.

Two explanations for this impairment require consideration. First, the difference may be due to a reduced working memory capacity (Light & Anderson, 1982). Experiment 4.1 has a large working memory component subjects are required to hold previously presented slides in working memory and the inferences associated with them until the next slide confirmed or negated the alternatives constructed. However, Experiment 5.1 had less of a working memory component, testing the retrieval of both factual and inferential information from a previously shown black and white photograph. In this case, the decrement in inferential reasoning associated with the low I.Q. test score group is likely to be the result of difficulties with comprehending the photographs and identifying the depicted event. Both these processes depend upon the efficient use of knowledge stored in long-term memory. The processes of identification and inference construction with the use of information stored in long-term memory is also an important element in the theme construction task of 4.1, and will be discussed in detail in the following section.

8.3.3 Differences in the Structure and Organisation of Long-Term Memory

The elderly with low I.Q. test scores are poor at comprehending scenes and drawing inferences from pictorial stimuli. These findings can be considered in relation to the structure and nature of the long-term memory system, since a key component process of these cognitive skills is the efficiency and speed of activation and retrieval of information from this long-term store. It is proposed here that one of the properties associated with a high I.Q. test score is a richer, more elaborate data-base or knowledge structure. The network of associations between concepts and attributes is dense, which means that information is automatically processed to a deeper level because of the number of associations present. This feature of the system enhances retrieval processes markedly (Craik and Tulving, 1975), as well as increasing the efficiency of the search and encoding processes. In contrast, the low I.Q. test score group have long-term memory structures which are relatively impoverished in terms of the number of associations between and within concepts. A consequence of this is that the structure is less well organised and less adhesive to information. Greater effort is required for deep level processing to occur, and more cognitive capacity is necessary in order to encode and retrieve information.

Data from Experiments 6.1, 6.2, 7.1 and 7.2 support this description of the long-term memory system and illustrate some characteristics associated with it. Rabbitt (1984b) suggested that low I.Q. subjects may have impoverished networks of associations between items held in memory compared with young or clever people. They find it more difficult to learn new associations and may forget useful associations they had previously mastered. It follows that the elderly with low I.Q. test scores may have a reduced 'bandwidth' of available associations for 'memory driven' priming. Experiment 7.1 supports this idea. Subjects who were elderly and those of low I.Q. test score generated fewer associations to everyday scene titles than the young and high I.Q. test score group, indicating that the representation of scenes in memory is more impoverished with fewer associations within the concept and between concepts. Additional support was found in Experiment 7.2, where the low I.Q. test score group realised fewer associations between target words and titles of scenes. They appear to adopt a stricter boundary between concepts and to hold fewer associations within concepts.

The density of the network of associations in long-term memory has an effect upon the speed at which information can be accessed and retrieved. Experiments 6.1, 6.2 and 7.2 examined group differences in matching and retrieval of information in long-term memory. Any investigation of search and retrieval must consider the size of the set in which the search occurs. Experiments 6.1 and 6.2 both involved a search through the verbal store, whereas Experiment 7.2 used the semantic store. In each case the high I.Q. test score group were required to search through larger sets, since they have larger vocabularies and generate more associations to semantic scene titles. Retrieval times in all experiments were faster for the high I.Q. test score group and for young subjects. In spite of the increase in set size, the clever and the young operate a faster search strategy than the elderly with low I.Q. test scores. There are two possible explanations for this, either the high I.Q. test are considerably faster at retrieval because they operate all processes at a faster rate with greater efficiency, or they are faster because their data-bases are better organised and search strategies operate both quickly and efficiently.

An explanation for search can be derived from the structure and organisation of the associative network. Clearly, if serial search was taking place, subjects with large search sets would be expected to take longer to retrieve words and facts, and the opposite result would have been found. Instead, the rich network of associations enhances the retrieval of information by giving rise to well directed and efficient search. When an object or scene is presented for identification, the relevant schema requires activation in memory. The large number of associations between and within concepts means that there are more pointers in memory to guide the search, and more chance of initially activating an associated schema. If the incorrect schema is activated, there are likely to be sufficient associations between it and the target schema for the correct location to be reached. The latency of the response is a linear function of the number of draws into the search set (Shiffrin, 1970). In the highly associated network the draws into the set are well directed so that fewer are needed, thereby reducing the number required.

8.3.4 Differences in Comprehension and Recall of Events

The nature of the association network in long-term memory affects the rate and likelihood of comprehension. Its effect on the rate of search and identification have already been discussed. However, once a scene has been recognised it generally requires interpretation, which depends upon the amount of information extracted from the stimulus, and the amount of associated knowledge stored in long-term memory.

Experiments 4.1 and 5.1 have illustrated that the elderly with low I.Q. test scores are slow to comprehend scenes and events. When presented with pictorial story sequences, they required longer viewing times for each picture, and were slower to make a decision on whether the last slide in each sequence was consistent or inconsistent with the theme illustrated by the previous slides belonging to that sequence. Further, when they reached a decision, it was more likely to disagree with the consistency judgements made by the 42 raters than the decisions made by the high and middle I.Q. test score groups. In Experiment 5.1, the same group recalled fewer inferences from photographs of events, and the inferences they recalled were more implausible than those recalled by the other groups. They also retrieved less factual information from the stimuli. These group differences in comprehension and inference construction can be related to the structure of the data-base in long-term memory.

Evidence suggests that the schemata held in long-term memory by the low I.Q. test score group contain fewer 'slots' (Minsky 1975) that are filled with specific instances and data accumulated through past experience, than those stored by the high I.Q. test score group (Rabbitt 1984b; Chapter 7). When an event or scene is encountered there is less information and knowledge available for interpretation and comprehension. The comprehension of a scene or an event involves the simultaneous operation of two processes - 'top-down' analysis and 'bottom-up' analysis. 'Top-down' analysis starts with hypotheses and then attempts to verify them by looking at the stimulus, and attempting to determine which structural description applies to a given input. On the otherhand, 'bottom-up' analysis refers to stimulus driven activity. These two interact, since schemata stored in long-term memory can be used to determine hypotheses for 'top-down' processing and these direct the stimulus search in 'bottom-up' processing.

Palmer (1975) has proposed how the facilitation of stored knowledge and experience might take place in a limited capacity system, in which processing resources (or attention) can be differentially distributed. This is based upon Norman and Bobrow's (1975) analysis of data-limited and resource-limited processes, and can be used to explain group differences in comprehension and memory for complex events.

If we consider each schema to contain nodes whose function is to compute the concept it represents, and then derive a measure of the goodness of fit between the sensory data presented and its concept, each schema's success can be described as depending on the goodness of fit score it obtains. For each an asymptotic level of success will be reached when all pertinent data have been analysed. We can consider the rate at which this level is reached to depend upon the amount of processing resource given to a particular procedure. The greater the amount of resource allocated to a particular schema within a superordinate schema, the faster and more completely the concept will be computed and the goodness of fit measure made available for other procedures.

Expectations determine the allocation of processing, so that strongly expected schema receive large allocations of resouces. In a relational network 'expectation' is described as a request for data from an associated node. Obviously, the more resources a procedure has, the more rapidly it can compute its fit to the available sensory data. The facilitating affect of expectations is to increase the rate at which predicted information is analysed. Intelligent behaviour by the system can be seen as the system concentrating its efforts on computing information, which from experience is likely to be present, given what has already been identified. However, if expectations are not specific or exact, then the resource will be distributed over more alternatives, and these will require more resources than contained in the limited pool, so that the efficiency of testing for alternatives will be lowered.

Group differences between low and high I.Q. test scorers in memory and comprehension of complex events can be explained in terms of this model. The low I.Q. test score groups generate fewer expectations from their data-base, which has fewer associations and is less well organised. This results in the operation of less efficient analysis of the scene or event, since processing resources will be distributed over more alternatives, because there is less information available in long-term memory to form specific expectations. Both 'top-down' and 'bottom-up' processing is therefore impaired in the low I.Q. test score group. The high I.Q. test score group have better 'top-down' and 'bottom-up' systems, based upon a well organised and richly associated data-base from which more specific expectations can be generated. The advantages of this structure extend to improved recognition and comprehension with the use of better internal and contextual cues. In addition learning is more rapid, since the data-base is able to encode and retrieve information easily by association. The main advantage for the high I.Q. group of rapid identification and comprehension is that they are probably able to automatise more subprocesses and free more resources for additional processing tasks. The low I.Q. test score group has a more impoverished data-base which leads to less efficient recognition and comprehension, because they are unable to generate as many specific expectations. Hence, they have have fewer excess resources to enable them to operate effect learning and retrieval strategies.

The effects outlined above explain the data from the television recall studies. Both Experiments 2.1 and 2.2 found that the high I.Q. test score group recalled more facts in both free and cued recall conditions. In addition Experiment 2.2 showed that this group retrieved a higher percentage of all detail types in the broadcast than the other groups. The high global recall performance is due to their more absorbent and highly associated data-base benefitting search and comprehension processes, and releasing more resources for storage and retrieval strategies.

8.3.5 Differences In The Efficiency Of Processing

There are two possible sources of individual and group differences associated with the memory system itself. The effects of the structure and organisation of the data-base have already been discussed. However, the effects of the efficiency of cognitive processing also require discussion. There is no doubt that as age increases, the speed at which mental operations take place decreases. Waugh and Barr (1980) have reviewed the evidence for the age-related decrement in processing in relation to memory tasks, and have concluded that most observed decrements can be accounted for by an age-related decline in the speed of performance.

An explanation of the decline in memory performance based upon the decline in the efficiency of processing is so general that a mass of evidence can be interpreted in support of it, and very little evidence appears to contradict it. Unsurprisingly, the group and individual differences seen in the presented data can be explained easily by differences in the efficiency of processing. Clearly the high performance achieved by the high I.Q. test score group on recall tests may be due to their superior efficiency in encoding, storage and retrieval. The differences in naming and matching latencies illustrating the high rates of retrieval by the high I.Q. test score group reinforce this interpretation. Speed of cognitive processing can also be seen to affect the ability to comprehend material. A working memory which can handle information efficiently will be able to manipulate and transform information. Faster activation and retrieval of information in long-term memory will free more cognitive resources for inferential reasoning and remembering the presented information.

8.3.6 Comments On Models Of Group And Individual Differences

The preceding sections have discussed the interpretation of the data in relation to three models of individual differences: differences in strategical processing, differences in cognitive structure and differences in the efficiency of cognitive processing. In spite of the lack of support for differences in strategical processing from the data presented, one might say that strategical differences require a much greater and more complex 'support' structure of other skills. In showing the evidence of differences in this support structure perhaps we have implicitly shown differences in the capabilities for strategic processing. However, there is sufficient evidence for age-related changes in the use and sophistication of strategies to suggest that this warrants further investigation. Both of the remaining models have been applied successfully to the data, which leaves the problem of distinguishing which is responsible for the differences shown.

Differences in the rate of cognitive processing can just as well be the result of differences in the structure of the memory system as its efficiency. The model of differences in the efficiency of cognitive processing can be criticised for being too general, and hence difficult to test. It presents a very simple explanation of differences, but fails to question the cause of the reduction in efficiency. Efficiency can be reduced by several factors, e.g. a reduction in the capacity of working memory, a disorganised data-base, or poor strategical processing. Isolating these effects presents a difficult problem. If a group of subjects are found to be poor at performing simple mental arithmetic problems on a sequence of single digit numbers, this could be interpreted as due to a reduced working memory capacity. An alternative explanation which has not been eliminated is that this group are unable to process the presented numbers as efficiently, hence less cognitive capacity is available for memorising the digits involved in the mental transformations.

An obvious route to follow when exploring differences in rates of processing is the effect of changing presentation rates. At slow presentation rates, the low I.Q. test score group's processing rate will be sufficient to keep pace with the presentation rate, and hence their error rate will not be significantly different from the high I.Q. test scorers. However, an alternative explanation would be that the difference in the cognitive structure between the groups does not impede the performance of the low I.Q. test score group when presentation rates are slow. At fast presentation rates, differences in error rate could be due to differences in processing efficiency or differences in the cognitive structure both would slow the rate of processing in the low I.Q. test score group leading to a higher error rate.

It is easier to consider each model in isolation. However, the most plausible explanation of group differences is probably a combination of all three models. An interactive and interdependent model of individual differences makes it increasingly more difficult to identify the processes which contribute to the effects found. This difficulty has been avoided in this thesis, and several processing differences have been identified. The models suggested can always be criticised for being too general. Clearly, this is a continuing problem in ageing research, yet investigating changes in cognitive structure and efficiency has proved to be a profitable approach to studying the performance of the elderly on an everyday memory task. The need for more specific models of change in cognitive performance with age is a familiar plea (Rabbitt 1981a), and it is hoped that this thesis has contributed to our knowledge of changes in performance on a more complex cognitive task.

8.4 SUGGESTIONS FOR FURTHER RESEARCH

An important point raised by this research is that the study of ageing can no longer consider chronological age as the primary factor associated with changes in performance. The evidence suggests that the effects of chronological age are weak in comparison with other effects such as I.Q. test score. If changes with age are to be monitored then tasks which correlate highly with I.Q. test score will provide a sensitive measure of changes in performance.

The experiments reported in this thesis have shown the practical implications of the decline in speed of retrieval of information from long-term memory and the use of inferential reasoning in comprehension and memory for events. The comprehension of what is happening in the course of everyday living is a persistant difficulty for the elderly, and these studies suggest that their main deficit lies in processing speed and the organisation and structure of information in long-term memory. Both of these factors affect the elderly's ability to construct inferences and comprehend events. Further, these studies offer some advice to those who are concerned with the assessment of the elderly. Performance on these tasks generally correlates highly with AH4 test score rather than age. A simple test of general intelligence seems to provide a sensitive index of changes in cognitive performance with increasing age.

The need for models of change in performance with age is clearly something which the study of the cognitive effects of ageing should pay more attention to. Frequently, the models applied are inappropriate or not specific enough to be adequately tested. Further, the demand to consider the effects of ageing on everyday cognitive functioning is becoming increasingly important, with pressures from the changing composition of the population at large, and the challenge of analysing performance on more complex everyday skills. The research presented has attempted to meet this challenge. The investigation of individual and group differences in the comprehension and memory of events has provided insight into the components of this complex task and offered an explanation of changes in the cognitive structure of the memory system associated with ageing.

APPENDICES

9 APPENDIX 1

9.1 EXAMPLES OF CUED RECALL QUESTIONS EXPERIMENT 2.1

a) Who left Britian under order?

b) Where has President Reagan been visiting?

c) Where was the oilslick?

d) Which vessels were exchanging oil?

e) What was England's first innings score?

f) What do protestors advocate instead of turkey?

g) Who have Spanish Police killed?

h) What viewpoint was a protestor giving to President Reagan?

i) Who's home was raided by police?

j) Where was the cricket match?

9.2 EXAMPLES OF PROBE QUESTIONS EXPERIMENT 2.2

a) Which fleet was inspected?

b) In which county was the fossil found?

c) Where were British troops remaining?

d) What has delayed the birdwatchers' trial?

e) What were the kidnappers of the Italian girl demanding?

f) Who says public spending must be cut?

9.3 EXAMPLES OF MULTIPLE CHOICE QUESTIONS EXPERIMENT 2.2

a) Where was Caroline Hogg last seen?

A. On a beach B. At the zoo C. At a gala D. At a funfair E. At a park

b) Where was her body found?

A. Edinburgh B. Coldstream C. Stafford D. Leicester E. Twycross

c) Who is withdrawing troops from the Lebonan?

A. Israel B. Britain C. Palestine D. America E. Syria

d) Where were birdwatchers arrested?

A. Ankara airport B. Istanbul C. Turkish military zone

D. Ankara E. Turkish Coast

e) Who is embarrassed by the case?

A. The judge B. The birdwatchers C. The British Government

D. The Turkish Police E. The Tirkish Army

f) Where were the auxillary fleet inspected?

A. Portugal B. Southampton C. Devonport D. Plymouth E. Dover

g) Where was the fossil found?

A. In a quarry B. In the cliffs C. In a sandpit D. In a claypit

E. On a beach

f) Where was the letter bomb intercepted?

A. The Head Post Office B. The House Of Commons C. Westminster Police Station D. 10 Downing St. E. Conservative Central Office

10 APPENDIX 2

10.1 STIMULUS WORDS EXPERIMENT 3.1

10.1.1 Four Footed Animals

dog
cat
horse
cow
lion
tiger
elephant
pig
bear
mouse
rat
deer
sheep
goat
zebra

10.1.2 Clothing

shirt
socks
pants
shoes
blouse
skirt
coat
dress
hat
sweater
tie
jacket
gloves
belt
underwear

10.1.3 Colour

blue
red
green
white
pink
brown
gold
black
yellow
indigo
purple
grey
beige
mauve
tan

10.1.4 Kitchen Utensils

knife
spoon
fork
pan
can opener
stove
bowl
mixer
cup
dish
beater
rolling pin
toaster
strainer
ladle

10.1.5 Vehicles

wagon
car
bus
aeroplane
train
bicycle
van
boat
scooter
ship
tractor
taxi
jet
motorbike
rocket

10.1.6 Trees

magnolia
elm
cherry
sycamore
cedar
oak
holly
spruce
ash
beech
pine
birch
walnut
poplar
weeping willow

10.1.7 Vegetables

carrot
potato
celery
beetroot
radish
pea
lettuce
cabbage
onion
parsley
bean
spinach
cauliflower
turnip
sprouts

10.1.8 Sport

football
hockey
rugby
squash
boxing
tennis
lacrosse
bowling
fishing
judo
golf
skiing
polo
badminton
horseriding

10.1.9 Furniture

chair
sofa
couch
stool
cabinet
table
desk
dresser
bureaux
wardrobe
bed
lamp
television
bookcase
sideboard

11 APPENDIX 3

11.1 STIMULUS WORDS EXPERIMENT 3.2

11.1.1 Birds

robin
eagle
parrakeet
wren
swallow
seagull
pheasant
peacock
warbler
magpie
sparrow
crow
hawk
parrot
dove
thrush
finch
swan
bluetit
pelican
jay
canary
blackbird
pigeon
owl
housemartin
lark
stork
vulture
swift

11.1.2 Part of a Building

window
wall
room
furniture
chimney
attic
kitchen
cellar
entrance
front

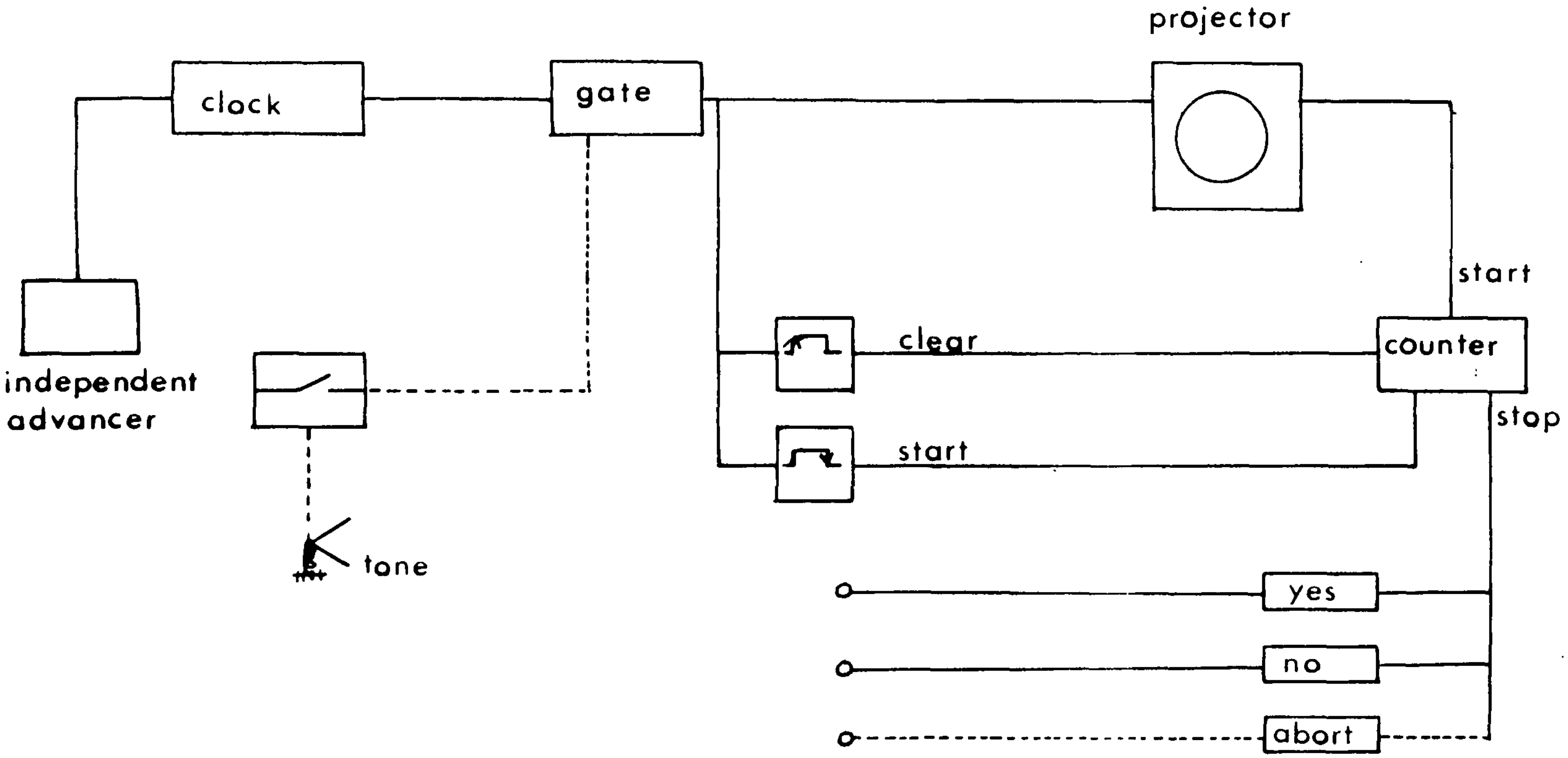
door
floor
basement
stairs
foundation
corner
lights
office
cornerstone
side

roof
ceiling
closet
doorway
lobby
beams
cement
porch
rafters
corridors

11.1.3 Weather Phenomenon

hurricane
snow
storm
clouds
typhoon
blizzard
drizzle
tidalwave
ice
gale
tornado
shower
wind
sunshine
thunder
fog
humidity
drought
rainbow
warm
rainstorm
sleet
cyclone
lightening
cold
monsoon
frost
hailstorm
windstorm
mist

12 APPENDIX 4



Appendix 4

Circuit For Apparatus In Experiments 4.1, 6.1 & 6.2

13 APPENDIX 5

13.1 PICTURE SEQUENCES EXPERIMENT 4

Twenty four picture sequences were presented to the subjects. There were eight sequences of each length i.e. two, three, and four pictures, of the eight four were rated as requiring postive responses and four negative, and four were rated as easy and four as difficult. The rated response and ease with which the consistency descision was rated are indicated at the beginning of each sequence.

13.1.1 Sequence 1

Positive, easy

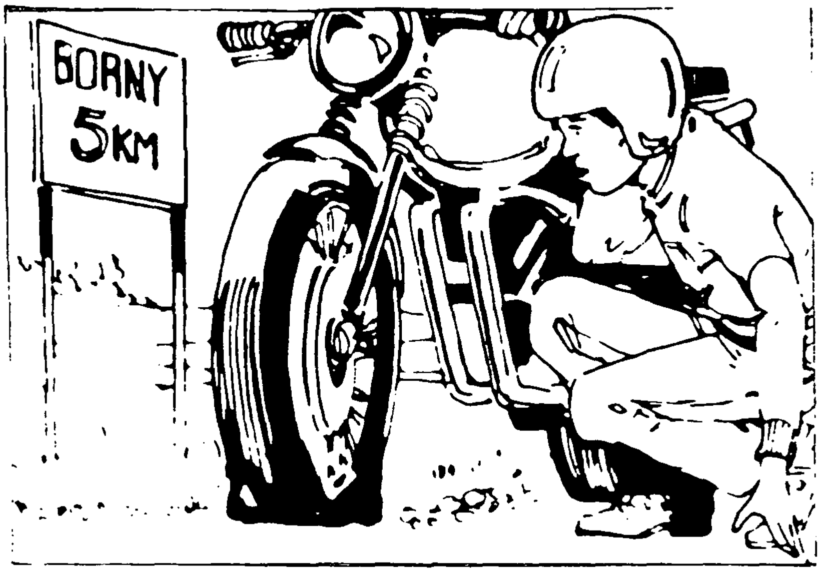
1

2

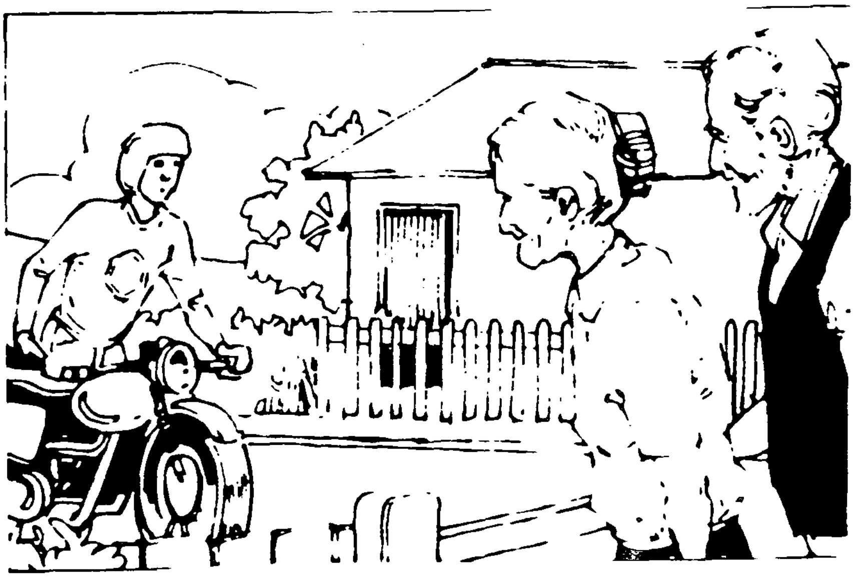
13.1.2 Sequence 2

Positive, difficult

1



2



13.1.3 Sequence 3

Negative, easy

1

2

3

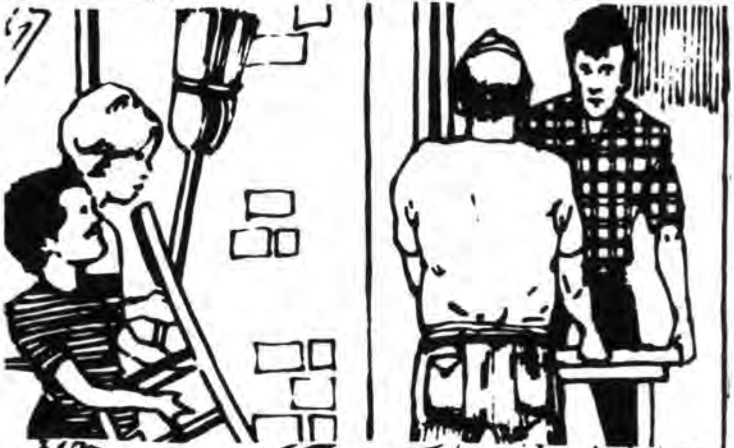
4

13.1.4 Sequence 4

Negative, difficult

1

2



3

14 APPENDIX 6

14.1 EXAMPLES OF PICTURES & QUESTIONS EXPERIMENT 5.1

14.1.1 Picture 1

1. How many items was the youth carrying?
2. How were they held together?
3. How many people were behind him?
4. In which country was the photograph taken?
5. What was the youth carrying?
6. Who was he employed by?

14.1.2 Picture 2

1. How many people were there in the photograph?
2. What were the men on the platform wearing?
3. What gesture were they making?
4. What were they facing?
5. What event were they attending?
6. What was significant about their gesture?

14.1.3 Picture 3

1. What letters could be seen on the shop?
2. What was the boy leaning against?
3. How many people were there in the street?
4. What was the boy doing?
5. What speed were the cars moving?
6. Was the boy successful in his task?

14.1.4 Picture 4

1. What were most of the women wearing on their heads?
2. What was the woman in the foreground doing?
3. How many crosses were there on the banner in the background?
4. In which building was the photograph taken?
5. Why were the people there?
6. Which country was the photograph taken in?

14.1.5 Picture 5

1. How many people were there in the photograph?
2. What were they wearing for protection?
3. By what means was the apparatus suspended from the roof?
4. Where was the photograph taken?
5. What were the men doing?
6. What was the substance they were working with?

15 APPENDIX 7

15.1 2X1 ANOVAS FOR EACH SCENE TITLE EXPERIMENT 7.1

2X1 ANOVAs For Scene Titles Experiment 7.1

....

The Post Office

N=150

NUMBER OF ORIGINAL WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	85.009	4	21.252	4.370	0.002
I.Q.	79.426	2	39.713	8.166	0.000
AGE	4.319	2	2.160	0.444	0.642
2-WAY INTERACTIONS	11.997	4	2.999	0.617	0.651
I.Q. X AGE	11.997	4	2.999	0.617	0.651
EXPLAINED	97.006	8	12.126	2.493	0.015
RESIDUAL	680.820	140	4.863		

TOTAL NUMBER OF WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	704.778	4	176.194	13.581	0.000
I.Q.	607.607	2	303.804	23.417	0.000
AGE	79.138	2	39.569	3.050	0.051
2-WAY INTERACTIONS	106.810	4	26.703	2.058	0.090
I.Q. X AGE	106.810	4	26.703	2.058	0.090
EXPLAINED	811.588	8	101.448	7.820	0.000
RESIDUAL	1816.305	140	12.974		

2X1 ANOVAs For Scene Titles Experiment 7.1

The Birthday Party

N=150

NUMBER OF ORIGINAL WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	14.502	4	3.626	1.014	0.402
I.Q.	10.400	2	5.200	1.454	0.237
AGE	3.823	2	1.911	0.535	0.587
2-WAY INTERACTIONS	25.560	4	6.390	1.787	0.135
I.Q. X AGE	25.560	4	6.390	1.787	0.135
EXPLAINED	40.062	8	5.008	1.401	0.201
RESIDUAL	500.583	140	3.576		

TOTAL NUMBER OF WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	710.900	4	177.725	8.463	0.000
I.Q.	603.583	2	301.792	14.371	0.000
AGE	86.551	2	43.276	2.061	0.131
2-WAY INTERACTIONS	220.350	4	55.087	2.623	0.037
I.Q. X AGE	220.350	4	55.087	2.623	0.037
EXPLAINED	931.250	8	116.406	5.543	0.000
RESIDUAL	2939.972	140	21.000		

2X1 ANOVAs For Scene Titles Experiment 7.1

....

The Fair Ground

N=150

NUMBER OF ORIGINAL WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	102.576	4	25.644	4.366	0.002
I.Q.	64.125	2	32.062	5.458	0.005
AGE	35.977	2	17.988	3.062	0.050
2-WAY INTERACTIONS	30.377	4	7.594	1.293	0.276
I.Q. X AGE	30.377	4	7.594	1.293	0.276
EXPLAINED	132.953	8	16.619	2.829	0.006
RESIDUAL	822.349	140	5.874		

TOTAL NUMBER OF WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	962.570	4	240.643	12.914	0.000
I.Q.	638.829	2	319.414	17.141	0.000
AGE	283.155	2	141.578	7.598	0.001
2-WAY INTERACTIONS	233.551	4	58.388	3.133	0.017
I.Q. X AGE	233.551	4	58.388	3.133	0.017
EXPLAINED	1196.122	8	149.515	8.024	0.000
RESIDUAL	2608.845	140	18.635		

2X1 ANOVAs For Scene Titles Experiment 7.1

....

The Doctor's Surgery

N=150

NUMBER OF ORIGINAL WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	30.954	4	7.738	2.814	0.028
I.Q.	29.213	2	14.607	5.312	0.006
AGE	1.175	2	0.588	0.214	0.808
2-WAY INTERACTIONS	7.972	4	1.993	0.725	0.576
I.Q. X AGE	7.972	4	1.993	0.725	0.576
EXPLAINED	38.926	8	4.866	1.770	0.088
RESIDUAL	384.940	140	2.750		

TOTAL NUMBER OF WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	784.071	4	196.018	11.769	0.000
I.Q.	616.091	2	308.045	18.495	0.000
AGE	141.821	2	70.911	4.258	0.016
2-WAY INTERACTIONS	167.239	4	41.810	2.510	0.045
I.Q. X AGE	167.239	4	41.810	2.510	0.045
EXPLAINED	951.310	8	118.914	7.140	0.000
RESIDUAL	2331.737	140	16.655		

2X1 ANOVAs For Scene Titles Experiment 7.1

. . . .

The Church
N=150

NUMBER OF ORIGINAL WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	55.333	4	13.833	3.599	0.008
I.Q.	41.309	2	20.654	5.374	0.006
AGE	12.898	2	6.449	1.678	0.190
2-WAY INTERACTIONS	3.536	4	0.884	0.230	0.921
I.Q. X AGE	3.536	4	0.884	0.230	0.921
EXPLAINED	58.869	8	7.359	1.915	0.062
RESIDUAL	538.070	140	3.843		

TOTAL NUMBER OF WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	1603.052	4	400.763	15.299	0.000
I.Q.	1443.112	2	721.556	27.544	0.000
AGE	123.978	2	61.989	2.366	0.098
2-WAY INTERACTIONS	228.270	4	57.067	2.178	0.074
I.Q. X AGE	228.270	4	57.067	2.178	0.074
EXPLAINED	1831.322	8	228.915	8.739	0.000
RESIDUAL	3667.444	140	26.196		

2X1 ANOVAs For Scene Titles Experiment 7.1

....

The Bathroom
N=150

NUMBER OF ORIGINAL WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	59.182	4	14.796	4.233	0.003
I.Q.	47.278	2	23.639	6.763	0.002
AGE	9.963	2	4.982	1.425	0.244
2-WAY INTERACTIONS	16.885	4	4.221	1.208	0.310
I.Q. X AGE	16.885	4	4.221	1.208	0.310
EXPLAINED	76.068	8	9.508	2.720	0.008
RESIDUAL	489.355	140	3.495		

TOTAL NUMBER OF WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	1602.823	4	400.706	24.228	0.000
I.Q.	1127.537	2	563.769	34.088	0.000
AGE	418.340	2	209.170	12.647	0.000
2-WAY INTERACTIONS	243.661	4	60.915	3.683	0.007
I.Q. X AGE	243.661	4	60.915	3.683	0.007
EXPLAINED	1846.485	8	230.811	13.956	0.000
RESIDUAL	2315.408	140	16.539		

2X1 ANOVAs For Scene Titles Experiment 7.1

....

The Picnic

N=150

NUMBER OF ORIGINAL WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	29.300	4	7.325	1.889	0.116
I.Q.	18.098	2	9.049	2.333	0.101
AGE	10.576	2	5.288	1.364	0.259
2-WAY INTERACTIONS	11.900	4	2.975	0.767	0.548
I.Q. X AGE	11.900	4	2.975	0.767	0.548
EXPLAINED	41.201	8	5.150	1.328	0.234
RESIDUAL	542.907	140	3.878		

TOTAL NUMBER OF WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	798.191	4	199.548	8.470	0.000
I.Q.	619.843	2	309.922	13.155	0.000
AGE	149.765	2	74.883	3.179	0.045
2-WAY INTERACTIONS	310.736	4	77.684	3.297	0.013
I.Q. X AGE	310.736	4	77.684	3.297	0.013
EXPLAINED	1108.927	8	138.616	5.884	0.000
RESIDUAL	3298.241	140	23.559		

2X1 ANOVAs For Scene Titles Experiment 7.1

....

The Park
N=150

NUMBER OF ORIGINAL WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	98.153	4	24.538	6.255	0.000
I.Q.	62.171	2	31.086	7.924	0.001
AGE	32.918	2	16.459	4.195	0.017
2-WAY INTERACTIONS	31.858	4	7.965	2.030	0.093
I.Q. X AGE	31.858	4	7.965	2.030	0.093
EXPLAINED	130.011	8	16.251	4.142	0.000
RESIDUAL	549.251	140	3.923		

TOTAL NUMBER OF WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	1455.362	4	363.840	18.026	0.000
I.Q.	1148.002	2	574.001	28.438	0.000
AGE	256.541	2	128.271	6.355	0.002
2-WAY INTERACTIONS	341.859	4	85.465	4.234	0.003
I.Q. X AGE	341.859	4	85.465	4.234	0.003
EXPLAINED	1797.221	8	224.653	11.130	0.000
RESIDUAL	2825.813	140	31.237		

2X1 ANOVAs For Scene Titles Experiment 7.1

....

The Restaurant

N=150

NUMBER OF ORIGINAL WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	98.205	4	24.551	5.447	0.000
I.Q.	52.558	2	26.279	5.830	0.004
AGE	43.780	2	21.890	4.857	0.009
2-WAY INTERACTIONS	40.103	4	10.026	2.224	0.069
I.Q. X AGE	40.103	4	10.026	2.224	0.069
EXPLAINED	138.309	8	17.289	3.836	0.000
RESIDUAL	631.020	140	4.507		

TOTAL NUMBER OF WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	1723.956	4	430.989	17.202	0.000
I.Q.	1585.152	2	792.576	31.634	0.000
AGE	113.003	2	56.501	2.255	0.109
2-WAY INTERACTIONS	199.458	4	49.865	1.990	0.099
I.Q. X AGE	199.458	4	49.865	1.990	0.099
EXPLAINED	1923.415	8	240.427	9.596	0.000
RESIDUAL	3507.592	140	25.054		

2X1 ANOVAs For Scene Titles Experiment 7.1

The School

N=150

NUMBER OF ORIGINAL WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	115.503	4	28.876	4.034	0.004
I.Q.	77.466	2	38.733	5.411	0.005
AGE	34.122	2	17.061	2.383	0.096
2-WAY INTERACTIONS	52.168	4	13.042	1.822	0.128
I.Q. X AGE	52.168	4	13.042	1.822	0.128
EXPLAINED	167.671	8	20.959	2.928	0.005
RESIDUAL	1002.168	140	7.158		

TOTAL NUMBER OF WORDS GENERATED

Source	SUM SQRS	DF	MEAN SQR	F-STAT	PROBABILITY
MAIN EFFECTS	1424.287	4	356.072	12.773	0.000
I.Q.	1328.599	2	664.300	23.830	0.000
AGE	66.807	2	33.404	1.198	0.305
2-WAY INTERACTIONS	154.498	4	38.624	1.386	0.242
I.Q. X AGE	154.498	4	38.624	1.386	0.242
EXPLAINED	1578.785	8	197.348	7.079	0.000
RESIDUAL	3902.786	140	27.877		

16 APPENDIX 8

16.1 HIGH FREQUENCY WORDS GENERATED TO SCENE TITLES

Frequency counts are shown in parentheses.

16.1.1 The School

desks	(128)
teachers	(111)
pencils	(106)
blackboard	(106)
books	(104)
pens	(103)
chalk	(94)
pupils	(64)
chairs	(53)
headmaster	(36)

16.1.2 The Restaurant

tables	(126)
waiters	(92)
food	(90)
chairs	(84)
tablecloths	(83)
waitresses	(69)
knives	(67)
forks	(65)
wine	(60)
plates	(58)

16.1.3 The Post Office

stamps	(143)
postal-orders	(79)
counter	(69)
parcels	(59)
pens	(50)
pensions	(47)
money	(46)
scales	(44)
letters	(43)
envelopes	(38)

16.1.4 The Birthday Party

presents	(112)
games	(92)
candles	(81)
cake	(61)
children	(58)
sandwiches	(51)
birthday cake	(49)
food	(44)
tables	(43)
jellies	(41)

16.1.5 The Fairground

roundabouts	(103)
bingo	(54)
swings	(50)
fortune tellers	(44)
side-shows	(43)
candy-floss	(43)
music	(41)
people	(35)
coconut shies	(35)
ice-cream	(35)

16.1.6 The Doctors Surgery

doctors	(93)
receptionists	(91)
chairs	(86)
patients	(51)
prescriptions	(50)
magazines	(49)
nurses	(44)
waiting room	(42)
telephones	(42)
stethoscope	(40)

16.1.7 The Church

pews	(118)
altar	(107)
hymn books	(79)
organ	(68)
pulpit	(66)
font	(65)
flowers	(65)
choir	(54)
bible	(51)
vicar	(49)

16.1.8 The Bathroom

bath	(139)
soap	(138)
towels	(132)
shower	(93)
toothbrush	(86)
bath mat	(84)
sponge	(74)
toothpaste	(68)
talcum powder	(67)
washbasin	(60)

16.1.9 The Seaside

sand	(133)
sea	(110)
spades	(81)
deck-chairs	(76)
boats	(70)
children	(69)
buckets	(61)
ice cream	(59)
rocks	(53)
seaweed	(46)

16.1.10 The Picnic

sandwiches	(78)
food	(68)
cups	(59)
grass	(46)
tea	(44)
plates	(39)
car	(37)
hamper	(35)
picnic basket	(33)
balls	(33)

16.1.11 The Park

trees	(113)
flowers	(103)
grass	(74)
seats	(67)
swings	(66)
bowling green	(62)
lake	(58)
tennis courts	(53)
bandstand	(50)
paths	(49)

17 APPENDIX 9

17.1 STIMULUS WORDS EXPERIMENT 7.2

Association strengths are indicated.

17.1.1 The Seaside

High	Middle	Low
....	......	....
sand	shorts	eggs
sea	park	manager
children	view	money
boats	papers	hills
rocks	gardens	nuts
donkeys	theatres	taps
shells	hats	ties
spades	seals	nuns
waves	bay	breakfast
pebbles	taxis	chalk
tents	tar	shelf
seaweed	rods	wasps
fish	fruit	patients

ships
crabs
bucket
pier
dogs
promenade
towels
foam
pavements
huts
nets
horses
blankets
booths
pope
pens
statues
blinds
calls
trailer
ruler

17.1.2 The Park

High	Middle	Low
....	......	...
trees	cars	health
flowers	speaker	milk
grass	stream	rooms
lake	models	bible
paths	couples	gospel
children	wall	terms
dogs	families	plays
birds	soldiers	cross
ducks	shed	puplit
people	sun	hall
boats	hills	breakfast
bandstand	weeds	stumps
shrubs	snacks	bows
toilets	bushes	pier
swans	bins	seales
pond	woods	magazines
gates	lawn	oils
roundabouts	broom	rods
prams	hoes	weights
bowls	bulbs	chalk

17.1.3 The Post Office

High	Middle	Low
....	......	...
stamps	posters	playground
counter	newspapers	foam
parcels	packets	donkeys
money	government	music
pens	tokens	aromas
letters	waiting	trees
scales	raids	puppets
envelopes	labels	pebbles
pensions	boxes	cups
postbox	prams	tombstones
queues	clips	prescriptions
forms	services	sun
pencils	seales	theatres
postmaster	airmail	hedges
leaflets	pamphlets	stings
clock	overseas	lather
assistants	weights	caberet
licences	slots	serviettes
clerk	van	gardens
telephone	official	dust

17.1.4 The Bathroom

High	Middle	Low
....	......	...
soap	tray	mothers
bath	warmth	tea
towels	handles	cheese
shower	rail	lawn
sponge	cabinets	candles
toilet	shelf	organ
mirror	cap	pencils
taps	tablets	tinkers
shampoo	lacquer	cigars
toothpaste	first aid	mixers
wash basin	beakers	table cloths
toothbrush	make-up	altar
water	feet	people
flannel	lather	shrubs
scales	ripples	bandstand
razor	oils	hats
curtains	crystals	bells
radiator	detergents	sandwiches
carpet	tub	maps
plug	bowl	plates

17.1.5 The Picnic

High	Middle	Low
....	......	...
sandwiches	castles	stamps
food	summer	paper
cups	hats	tub
grass	lake	bay
tea	meal	warmth
plates	rocks	lift
car	field	feet
hamper	stumps	raids
rug	maps	shells
fruit	eggs	clerk
children	family	wall
cakes	pram	queues
knives	jug	circus
trees	farm	choice
flies	plough	couch
spoons	oranges	monster
saucers	hedges	lacquer
milk	sports	waves
coffee	dust	telephone
wasps	stings	crabs

17.1.6 The Restaurant

High	Middle	Low
....	......	...
tables	eggs	bus
food	party	trees
chairs	nuts	plug
knives	brandy	hymns
forks	receipts	ink
plates	tip	rocks
waiter	screens	maps
wine	sweet	official
waitress	aromas	dwarf
spoons	linen	bishops
glasses	meal	skills
menu	biscuits	ripples
cups	ties	razor
table cloths	mixers	graphs
napkins	cigars	pavements
saucers	stores	airmail
chef	cheese	handles
flowers	breakfast	bay
serviettes	caberet	acrobats
kitchen	choice	letters

17.1.7 The Church

High	Middle	Low
....	......	...
pews	vestments	blazer
altar	deacon	scales
pulpit	nuns	whistle
organ	gospel	tar
flowers	rooms	breakfast
choir	palms	crystals
vicar	brocade	entertainment
candles	bishops	toothbrush
prayer books	anthem	showmen
cross	wood	ships
priest	singers	gum
congregation	pope	calls
aisle	carving	counter
statues	amplifier	prizes
bells	sanctuary	huts
stained glass window	tombstones	chatter
people	father	gypsies
hymns	canon	party
cassocks	manse	stockings
bible	buildings	gymnasium

17.1.8 The Doctor's Surgery

High	Middle	Low
....	......	...
doctor	health	cars
chairs	lift	bowl
desk	smoke	animals
patients	newspapers	fun
nurse	lamp	priest
magazines	mothers	pond
stethoscope	spectacles	caravans
prescriptions	heating	competitions
telephone	calls	horses
books	paper	services
table	staff	farm
couch	overall	tar
seats	needle	paths
scales	pumps	hamper
thermometer	documents	waiter
pen	iodine	awards
medicine	apparatus	tents
people	help	government
notices	charts	sanctuary
pills	pillow	choir

17.1.9 The School

High	Middle	Low
....	......	...
books	play	summer
blackboard	bullies	beggars
chalk	alphabet	stethoscope
pencils	bus	van
pens	graphs	cassocks
teacher	classes	coffee
desks	plays	soldiers
pupils	speeches	shorts
chairs	swing	soap
headmaster	stationary	stalls
ruler	blinds	lather
ink	stockings	bucket
playground	whistle	flannel
paper	student	doctor
children	letters	land
gymnasium	blazer	fortune tellers
pictures	houses	horses
maps	gum	rug
hall	terms	waiting
rubber	awards	needle

17.1.10 The Fairground

High	Middle	Low
....	......	...
roundabouts	competitions	saucers
swings	dwarf	taxis
music	land	help
sideshows	acrobats	theatres
people	manager	car
noise	skills	pencils
coconuts	corkscrew	seaweed
stalls	bows	brocade
fortune tellers		
	showmen	stained glass window
prizes	monster	spoons
children	parents	father
caravans	beggars	pamphlets
gypsies	kiosks	serviettes
slides	puppets	cabinets
hotdogs	tinkers	toothpaste
chips	entertainments	promenade
circus	chatter	receipts
crowds	trailer	flies
barkers	bells	knives
animals	fun	cross

APPENDIX 10

THE AH4 TEST OF INTELLIGENCE

The main considerations involved in selecting an intelligence test for screening elderly subjects were practical ones. All subjects used in the studies reported were taken from a pool of two thousand volunteers participating in a longitudinal study.

The test had to be easy to administer and of a reasonably short duration. The AH4 test has two parts each of ten minutes. Further, the test should permit group testing and not require verbal responses since individual testing would be too time consuming given the large number of volunteers to be screen tested. The AH4 was considered suitable for this criterion, responses are made to multiple choice questions by writing the appropriate number corresponding to the choice of answer on a response sheet.

Each part of the test contains 65 questions with ten trial questions to ensure that subjects comprehend the structure of the questions. The first part contains questions concerning verbal and numerical intelligence and the second part is exclusively concerned with spatial intelligence. Examples of questions are given below:

a) Easy means the opposite of 1 problem 2 simple 3 difficult 4 always 5 cannot

b) SEED is to PLANT as EGG is to 1 tree 2 bird 3 pollen 4 oats 5 aspidistra

c) 1, 2, 4, 8, 16 What number comes next?

d) 0.9, 1.1, 1.3, 1.5, 1.7,... What number comes next?

REFERENCES

18 REFERENCES

Anders, T. R., Fozard, J. L. & Lillyquist, T.D. 1972. Effects of age upon retrieval from short term memory. Developmental Psychology. 6:214-217.

Baddeley, A. D. & Hitch, G. J. 1974. Working memory. In: G.H. Bower (Ed.), The Psychology of Learning & Motivation. Vol. 8, New York: Academic Press pp 47-90.

Baddeley, A. D. & Hitch, G. J. 1976. Recency re-examined In: S. Dornic (Ed.), Attention and Performance VI Hillsdale, New Jersey: Erlbaum Associates.

Baddeley, A. D. 1981. The cognitive psychology of everyday life. British Journal of Psychology. 72:257-269.

Baltes, P.B. 1968. Longitudinal and cross-sectional sequences in the study of age and generation effects. Human Development. 11:145-171.

Baltes, P. B. & Willis, S. 1979. The critical importance of appropriate methodology in the study of aging: The sample case of psychometric intelligence. In: F. Hoffmeister & C. Muller (Eds.), Brain Function in Old Age. Berlin: Springer-Varlag.

Baltes, P. B., Reese, H. W. & Lipsitt, L. P. 1980. Life-span developmental psychology. Annual Review of Psychology. 31:65-110

Baron, J. 1978. Intelligence and general strategies. In: G Underwood (Ed.), Strategies of Information Processing. London: Academic Press.

Battig, W. F. & Montague, W. E. 1969. Category norms for verbal items in 56 categories: A replication and extension of the Conneticut Category Norms. Journal of Experimental Psychology. Monograph 80 (3 pt. 2).

Belmore, S. M. 1981. Age-related changes in processing explicit and implicit language. Journal of Gerontology. 36:316-322.

Bennett-Levy, J. & Powell, G. E. 1980. The subjective memory questionnaire (SMQ). An investigation into self-reporting of 'real-life' memory skills. British Journal of Social Clinical Psychology. 19:177-188.

Berry, C., Gunter, B. & Clifford, B. 1981. Memory for televised information: a problem for applied and theoretical psychology. Current Psychological Reviews. 1:171-192.

Biederman, I. 1972. Perceiving real world scenes. Science. 177:77-80.

Biederman, I., Rabinowitz, J. C., Glass, A. L. & Stacy, E. W. 1974. On the information extracted from a glance at a scene. Journal of Experimental Psychology. 103:597-600.

Bird, C. 1927. The influence of the press upon the accuracy of report. Journal of Abnormal and Social Psychology. 22:123-129.

Birren, J. E., Reigel, K. F. & Robbin, J. S. 1962. Age differences in continuous word associations measured by speech recording. Journal of Gerontology 17:95-96.

Birren, J. E. 1968. Psychological aspects of aging. Gerontologist. 1:16-20.

Birren, J. E. & Schaie, W. (Eds.) 1977. Handbook of the Psychology of Aging. New York: Van Nostrand, Reinhold.

Botwinick, J. 1973. Aging And Behaviour: A Comprehensive Integration of Research Findings. New York: Springer.

Botwinick, J. 1967. Cognitive Processes In Maturity And Old Age. New York: Springer.

Botwinick, J. 1977. Intellectual abilities. In: J. E. Birren & K. W. Schaie (Eds.) Handbook Of The Psychology Of Aging. New York: Van Nostrand, Reinhold.

Botwinick, J. & Storandt, M. 1980. Recall and recognition of old information in relation to age and sex. Journal of Gerontology. 35:70-76.

Bower, G. & Glass, A. 1976. Structural units and the redintegrative power of picture fragments. Journal of Experimental Psychology: Human Learning and Memory. 2:456-466.

Bower, G. H., Black, J. B. & Turner, T. J. 1979. Scripts in memory for text. Cognitive Psychology. 11:177-220.

Briggs, G. E. & Swanson, J. M. 1969. Retrieval time as a function of memory ensemble size. Quarterly Journal of Experimental Psychology. 21:185-191.

Broadbent, D.E., Cooper, P. F., Fitzgerald, P. & Parkes, K. R. 1982. The Cognitive Failure Questionnaire (CFQ) and its correlates. British Journal of Clinical Psychology. 21:1-16.

Bromley, D.B. 1958. Some effects of age on short term learning and remembering. Journal of Gerontology. 13:398-406.

Burke, D. M. & Light, L. L. 1981. Memory and aging: The role of retrieval processes. Psychological Bulletin. 90:513-546.

Burke, D. M. & Yee, P. L. 1984. Semantic priming during sentence processing by young and older adults. Developmental Psychology. 20:903-910.

Cady, H. M. 1924. On the psychology of testimony. American Journal of

Psychology. 35:110-112.

Canestrari, R. E. 1968. Age changes in acquisition. In: G. A. Talland (Ed.), Human Aging and Behaviour. New York: Academic Press.

Cavanaugh, J. C. 1983. Comprehension and retention of television programs by 20- and 60- year olds. Journal of Gerontology. 38:190-196.

Charness, N. 1982. Problem solving and aging: Evidence from semantically rich domains. Canadian Journal of Aging. 1:21-28.

Clifford, B. 1978. A critique of eye witness research. In: M. M. Gruneberg, P. E. Morris & R. N. Sykes (Eds.) Practical Aspects of Memory. London: Academic Press.

Cofer, C. N. 1977. On the constructive theory of memory. In: F. Weizman & I. C. Uzgiris (Eds.) The Structuring of Experience. New York: Plenum.

Cohen, G. 1979. Language comprehension in old age. Cognitive Psychology. 11:412-429.

Cohen, G. 1981. Inferential reasoning in old age. Cognition. 9:59-72.

Cohen, G. & Faulkener, D. 1981. Memory for discourse in old age. Discourse Processes. 4:253-265.

Cohen, G. & Faulkener, D. 1984. Memory for text: Some age differences in the nature of the information that is retained after listening to texts. In: H. Bouma & D. Bouwhis. (Eds.) Attention and Performance X. Hillsdale New Jersey: Lawrence Erlbaum Associates.

Cohen, J. 1960. A coefficient of agreement for nominal scales. Education Psychology Measurement. 20:37-46.

Collins, A. M. & Loftus, E. F. 1975. A spreading activation theory of semantic processing. Psychological Review. 82:407-428.

Collins, W. A. 1979. Childrens' comprehension of television content. In: E. Wartella (Ed.) Children Communicating: Media and Development of Thought, Speech, Understanding. Sage Beverly Hills: C. A.

Comfort, A. 1957. The biological approach in the comparative study of ageing. In: G. E. Wolstenholme & C. M. O'Connor (Eds.) Methodology of the Study of Ageing. London: J. & A. Churchill.

Cowgill, D. O. & Baulch, N. 1962. The use of leisure time by older people. The Gerontologist. 2:47-50.

Craik, F. I. M. 1968. Short-term memory and the ageing process. In: G. A. Talland (Ed.) Human Aging and Behaviour. New York: Academic Press.

Craik, F. I. M. 1971. Age differences in recognition memory. Quarterly Journal of Experimental Psychology. 23:316-323.

Craik, F. I. M. & Lockhart, R. S. 1972. Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behaviour. 11:671-684.

Craik, F. I. M. & Tulving, E. 1975. Depth of processing and the retention of words in episodic memory.Journal of Experimental Psychology: General. 104:268-294.

Craik, F. I. M. 1977. Age differences in human memory. In: J. E. Birren & K. W. Shaie (Eds.) Handbook of the Psychology of Aging. Von Nostrand Rheingold.

Craik, F. I. M. & Simon, E. 1980. Age differences in memory: The roles of attention and depth of processing. In: L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg & L. W. Thompson (Eds.) New Directions in Memory and Aging. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Craik, F. I. M. & Rabinowitz, J. C. 1984. Age differences in the acquisition and use of verbal information: A tutorial review. In: H. Bouma & D. G. Bouwhins (Eds.)Attention and Performance X Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Daneman, M. & Carpenter, P. A. 1980. Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behaviour. 19:450-466.

Dooling, J. D. & Christiaan, R. E. 1977. Levels of encoding and retention of prose. In: G. H. Bower (Ed.) The Psychology of Learning and Memory. Vol. II. New York: Academic Press.

Doppelt, J. E. & Wallace, W. L. 1955. Standardization of the Wechsler Adult Intelligence Scale for older persons. Journal of Abnormal and Social Psychology. 51:312-330.

Drachman, D. A. & Leavitt, J. 1972. Memory impairment in the aged: Storage versus retrieval deficit? Journal of Experimental Psychology. 93:302-308.

Dunn, L. M., Dunn, L. M., Whetton, C. & Pintille, D. 1982. British Picture Vocabulary Scale. NFER-Nelson: Windsor.

Ellis, H., Shepherd, J. & Bruce, A. 1973. The effect of age and sex upon adolescents' recognition of faces. Journal of Genetic Psychology.

123:173-174.

Erber, J. T. 1974. Age differences in recognition memory. *Journal of Gerontology.* 29:177-181.

Eysenck, M. W. 1974. Age differences in incidental learning. *Developmental Psychology.* 10:936-941.

Eysenck, M. W. & Eysenck, M. C. 1979. Processing depth, elaboration of encoding, memory stores and expended processing capacity. *Journal of Experimental Psychology: Human Learning and Memory.* 5:472-484.

Fairbanks, G., Guttman, N. & Miton, M. S. 1957. Auditory comprehension in relation to listening rate in selective verbal redundancy. *Journal of Speech and Hearing Disorders. 27:23-32.*

Findahl, O. 1971. The effect of visual illustrations upon perception and retention of news programs. Swedish Broadcasting Corporation Audience & Program Research Report.

Findahl, O. & Hoijer, B. 1975. *Man as a receiver of information: On knowledge, social privilege and the news.* Stockholm:Sveriges Radio Audience & Programme Research Department.

Findahl, O. & Hoijer, B. 1977. *How important is presentation? A review of experimental research.* Swedish Broadcasting Corporation Audience &

Program Department Report.

Flavell, J. H. & Wellman, H. M. 1976. Metamemory. In: R. V. Kail & J. W. Hagen Memory in Cognitive Development. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Foulke, E. 1968. Listening as a function of word rate. Journal of Communication. 18:188-206.

Frederiksen, C. H. 1975. Effects of context-induced processing operations on semantic information acquired from discourse. Cognitive Psychology. 7:139-166.

Friedman, A. 1979. Framing pictures: The role of knowledge in automatized encoding and memory for gist. Journal of Experimental Psychology: General. 108:316-355.

Funkhouser, G. R. & McCombs, M. E. 1971. The rise and fall of news diffusion. Public Opinion Quarterly. 35:107-113.

Gaesser, A. C., Woll, S. B. Kowalski, D. J. & Smith, D. A. 1979. Memory for typical and atypical actions in scripted activities: Tests of a script pointer and tag hypothesis. Journal of Verbal Learning and Verbal Behaviour. 18:319-332.

Genova, B. K., Bradley, L. & Greenberg, S. 1979. Interest in news and the knowledge gap. *Public Opinion Quarterly.* 43:79-91.

Goldstein, A. G. & Chance, J. 1964. Recognition of children's faces. Child Development. 35:129-136.

Goldstein, A. G. & Chance, J. 1965. Recognition of children's faces II. Perceptual and Motor Skills. 20:548-549.

Goodman, G. S. 1980. Picture memory: How the action schema affects retention. Cognitive Psychology. 12:473-495.

Gregory, R. L. 1974. Concepts and mechanisms of perception. New York: Charles Scribners Sons.

Grether, W. F. & Baker, C. A. 1972. Visual presentations of information. In: H. P. Van Cott & R. G. Kinkade (Eds.) Human Engineering Guide to Equipment Design. Washington D. C.: U. S. Government Printing Office.

Gruneberg, M. M., Morris, P. E. & Sykes, R. N. (Eds.) 1978. Practical Aspects of Memory. London: Academic Press.

Gunter, B. 1979. Recall of brief television news items: Effects of presentation mode, picture content and serial position. Journal of Educational Television and Other Media. 5:57-61.

Gunter, B., Clifford, B. R. & Berry, C. 1980. Release from proactive interference with television news items: evidence for encoding dimensions within televised news. Journal of Experimental Psychology: Human Learning and Memory. 6:216-223.

Hastie, R. & Kumar, P. A. 1979. Person memory: personality traits as organising principles in memory for behaviours. Journal of Personality and Social Psychology. 37:25-38.

Havilard, S. E. & Clark, H. H. 1974. What's new? Acquiring new information as a process in comprehension. Journal of Verbal Learning and Verbal Behaviour. 13:512-521.

Hermann, D. & Neisser, U. 1978. An inventory of everyday memory experiences. In: M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.) Practical Aspects of Memory. London: Academic Press.

Hintzman, D. L. 1976. Repetition and memory. In: G. H. Bower (Ed.) The Psychology of Learning and Motivation. Vol. 10 New York: Academic Press.

Howard, D. V. 1980. Category norms: A comparison of the Battig and Montague (1969) norms with the responses of adults between the ages of 20 and 80. Journal of Gerontology. 35:225-231.

Hulicka, I. M. & Grossman, J. L. 1967. Age group comparisons for the use of mediators in paired-associate learning. Journal of Geronotlogy. 22:46-51.

Hultsch, D. F. 1969. Adult age differences in the organization of free recall. Developmental Psychology. 1:673-678.

Hultsch, D. F. 1971. Adult age differences in free classification and free recall. Developmental Psychology. 4:338-342.

Hultsch, D. F. 1974. Learning to learn in adulthood. Journal of Gerontology. 29:302-308.

Hultsch, D.F. 1975. Adult age differences in retrieval: Trace dependent and cue dependent forgetting. Developmental Psychology. 11:197-201.

Jenkins, J. J., Wald, J. & Pittenger, J. B. 1978. Apprenhending pictorial events: An instance of psychological cohesion. In: C. Wade Savage (Ed.) Perception and Cognition Issues in the Foundations of Psychology. Minnesota Studies in the Philosophy of Science, Vol. IX.

Kagan, J., Klein, R. E. Haith, M. M. & Mornson, F. J. 1973. Memory and meaning in two cultures. Child Development. 44:221-223.

Katz, E. Adoni, H. & Parness, P. 1977. Remembering the news: What the picture adds to recall. Journalism Quarterly. 54:231-239.

Kintsch, W. 1976. Memory for prose. In: C. N. Cofer (Ed.) The Structure of Human Memory. San Francisco: Freeman.

Korchin, S. J. & Basowitz, H. 1957. Age differences in verbal learning. Journal of Abnormal and Social Psychology. 54:64-69.

Lachman, R., Lachman, J. L. & Thronesbery, C. 1979. Metamemory through the adult life span. Developmental Psychology. 15:543-551.

Lachman, J. L. & Lachman, R. 1980. Age and the actualization of world knowledge. In: L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg & L. W. Thompson (Eds.) New Directions in Memory and Aging. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Laurence, M. W. 1967. Memory loss with age: A test of two strategies for its retardation. Psychonomic Science. 9:209-210.

Light, L., Zelinski, E. M. & Moore, M. 1982. Adult age differences in reasoning from new information. Journal of Experimental Psychology: Learning, Memory, and Cognition. 8:435-447.

Light, L. & Anderson, P. A. 1983. Memory for scripts in young and older adults. Memory and Cognition. 11:435-444.

Lipton, J. P. 1977. On the psychology of eyewitness testimony. Journal of Applied Psychology. 62:90-93.

Loftus, E. F. & Palmer, J. C. 1974. Reconstruction of automobile destruction: An example of the interaction between language and memory. Journal of Verbal Learning and Verbal Behaviour. 13:585-589.

Loftus, E. F. 1975. Leading questions and the eyewitness report. Cognitive Psychology. 7:560-572.

Loftus, E. F., Attman, D. & Geballe, R. 1975. Effects of questioning upon a witness's later recollections. Journal of Police Science and Administration. 3:1-2;165.

Loftus, E. F., Miller, D. G. & Burns, N. J. 1978. Semantic integration of verbal information into visual memory. Journal of Experimental Psychology: Human Learning and Memory. 4:19-31.

Loftus, G. R. 1972. Eye fixations and recognition memory. Cognitive Psychology. 3:525-557.

Loftus, G. R., Nelson, W. W. & Kallman, H. J. 1983. Differential acquisition rates for different types of information from pictures. Quarterly Journal of Experimental Psychology. 35A:187-198.

Lorge, I. 1956. Gerontology (later maturity). Annual Review of Psychology. 7:349-364.

Lovelace, E. A. & Cooley, S. 1982. Free association of older adults to single words and conceptually related word triads. Journal of Gerontology. 37:432-437.

Mandler, J. & Johnson, N. 1976. Some of the thousand words a picture is worth. Journal of Experimental Psychology: Human Learning and Memory. 2:529-540.

Mason, S. E. 1977. The effects of orienting tasks on the recall and recognition memory of subjects differing in age. Unpublished doctoral dissertation, Georgia Institute of Technology.

McKelvie, S. 1976. The effects of verbal labelling on recognition memory for schematic faces. Quarterly Journal of Experimental Psychology. 28:459-474.

McKenna, P. & Warrington, E. K. 1983. Graded Naming Test Manual. Windsor: NFER-Nelson.

McNichol, D. 1972. A Primer of Signal Detection Theory. London: Allen and Unwin.

Minsky, M. 1975. A framework for representing knowledge. In: P. H. Winston (Ed.) The Psychology of Computer Vision. New York: McGraw-Hill.

Mitterer, J. & Rowland, G. 1975. Picture memory: Role of figure-ground organization. Perceptual and Motor Skills. 40:753-754.

Monge, R. H. & Hultsch, D. F. 1971. Paired-associate learning as a function of adult age and length of the anticipation and inspection intervals. Journal of Gerontology. 26:157-162.

Neisser, U. 1978. Memory: What are the important questions? In: M. M. Gruneberg, P. E. Morris & R. N. Sykes (Eds.) Practical Aspects of Memory. London: Academic Press.

Neisser, U. (Ed.) 1980. Memory Observed: Remembering in Natural Contexts. San Francisco: Freeman.

Neuman, R. W. 1976. Patterns of recall among television news viewers. Public Opinion Quarterly. 40:115-123.

Newcombe, F., Oldfield, R. C. & Wingfield, A. 1965. Object naming by dysphasic patients. Nature. 207:1217-1218.

Norman, D. A. & Rumelhart, D. E. 1970. A system for perception and memory. In: D. A. Norman (Ed.) Models of Human Memory. New York: Academic Press.

Oldfield, R. C. & Wingfield, A. 1965. Response latencies in naming objects. Quarterly Journal of Experimental Psychology. 17:273-281.

Oldfield, R. C. 1966. Things, words and the brain. Quarterly Journal of Experimental Psychology. 18:340-353.

Palmer, S. E. 1975. The effect of contextual scenes on the identification of objects. Memory and Cognition. 3:519-526.

Parkinson, S. R. 1982. Performance deficits in short-memory tasks: A comparison of amnesic Korsakoff patients and the aged. In: L. S. Cermak (Ed.) Human Memory and Amnesia. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Perfetti, C. A. & Lesgold, A. M. 1979. Coding and comprehension in skilled reading and implications for reading instruction. In: L. B. Resnick & P. Weaver (Eds.) Theory and Practice In Early Reading (Vol. 1). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Perfetti, C. A. 1983. Individual differences in verbal processes. In: R. F. Dillon & R. R. Schmeck (Eds.) Individual Differences in Cognition (Vol. 1). New York: Academic Press.

Perlmutter, M. 1978. What is memory aging the aging of? Developmental Psychology. 14:330-345.

Perlmutter, M. 1979. Age differences in the consistency of adults' associative responses. Experimental Aging Research. 5:549-553.

Perlmutter, M. & Mitchell, D. B. 1982. The appearance and disappearance of age differences in adult memory. In: F. I. M. Craik & S. Trehub (Eds.) Aging and Cognitive Processes. New York: Plenum.

Pezdek, K. & Miceli L. 1982. Life-span differences in memory integration as a function of processing time. Developmental Psychology. 18:485-490.

Postman, L. Keppel, G. (Eds.) 1970 Norms Of Word Association. New York : Academic Press.

Quillian, M. R. 1969. The teachable language comprehender. A stimulation program and theory of language. Communications of the A.C.M. 12:459-476.

Rabbitt, P. M. A. 1964. Age and time for choice between stimuli and between responses. Journal of Gerontology. 19:307-312.

Rabbitt, P. M. A. 1965a. An age decrement in the ability to ignore irrelevant information. Journal of Gerontology. 20:233-238.

Rabbitt, P. M. A. 1965b. Age and discrimination between complex stimuli. In: A. T. Welford & J. E. Birren (Eds.) Behaviour, Ageing and the Nervous System. Biological Determinants of Speed of Behaviour and its Change with Age. C. C. Thomas.

Rabbitt, P. M. A. 1968. Age and the use of structure in transmitted information. In: G. A. Talland (Ed.) Human Aging and Behaviour. New York: Academic Press.

Rabbitt, P. M. A. 1979. Some experiments and a model for change in old age. In: F. Hoffmeister & G. Muller (Eds.) Brain Function In Old Age. Berlin: Springer-Varlag.

Rabbitt, P. M. A. & Vyas, S. M. 1980. Selective anticipation for events in old age. Journal of Gerontology. 35:913-919.

Rabbitt, P. M. A. 1980. A fresh look at changes in reaction times in old age. In: Stein. (Ed.) The Psychobiology of Aging: Problems and Perspectives. Elsevier.

Rabbitt, P. M. A. 1981a. Cognitive psychology needs models for changes in performance with old age. In: J. Long & A. D. Baddeley (Eds.) Attention and Performance IX. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Rabbitt, P. M. A. 1981b. Breakdown of control processes in old age. In: G. Finley et al. (Ed.) Human Lifespan Psychology Vol. 3.San Francisco: Wiley.

Rabbitt, P. M. A. 1982. Development of methods to measure changes in activities of daily living in the elderly. In: S. Corkin et al. (Eds.) Alzhimers Disease: A report of progress (Aging, Vol. 19.).

New York: Raven Press.

Rabbitt, P. M. A. & Abson, V. 1984. Where did I put it? Does old age and I. Q. predict the frequency with which objects are mislaid. Paper presented at Meeting of the British Psychological Society, December 1984, City University, London.

Rabbitt, P. M. A. 1984a. Investigating the grey areas. Times Higher Education Supplement. 1.6.84.

Rabbitt, P. M. A. 1984b. How people prepare themselves for events which they expect. In: H. Bouma & D. G. Bouwhuis (Eds.) Attention and Performance X. Hillsdale New Jersey: Lawrence Erlbaum Associates.

Rabinowitz, J. C., Craik, F. I. M. & Ackerman, B. P. 1982. A processing resource account of age differences in recall. Canadian Journal of Psychology. 36:325-344.

Riegel, K. F. & Riegel, R. M. 1964. Changes in associative behaviour during later years of life: A cross sectional analysis. Vita Humana. 7:1-32.

Riegel, K. F. & Birren, J. E. 1966. Age differences in verbal associations. Journal of Genetic Psychology. 108:153-170.

Riegel, K. F. & Riegel, R. M. 1972. Development, drop and death. Developmental Psychology. 6:306-319.

Robinson, J. F., Davis, D., Sahin, H. & O'Toole, T. 1980. Comprehension of television news: how alert is the audience? Paper presented at the Radio and Television Division of the Association for Education in Journalism Annual Convention, Boston, M A, August.

Rochford, G. & Williams, M. 1965. Studies in the development and breakdown of the use of names part IV: The effects of word frequency. Journal of Neurology, Neurosurgery,and Psychiatry. 28:407-413.

Rubin, A. M. & Rubin, R. B. 1982. Older persons' T. V. viewing patterns and motivations. Communication Research. 9:287-313.

Salthouse, T. A. 1984. Effects of age and skill in typing. Journal of Experimental Psychology: General. 113:345-371.

Salthouse, T. A. 1985. Transcription typing: A case study in age and expertise. Paper presented at the meeting of the International Society for the Study of Behavioural Development, Tours, France. July 1985

Sandford, A. J. & Maule, A. J. 1971. Age and the distribution of observing responses. Psychonomic Science. 23:419-420.

Sandford, A. J. & Maule, A. J. 1973a. The allocation of attention in multistore monitoring behaviour: Adult age differences. Perception. 2:91-100.

Sandford, A. J. & Maule, A. J. 1973b. The concept of general experience: Age and strategies in guessing future events. Journal of Gerontology. 28:81-88.

Schaie, K. W. & Parham, I. A. 1977. Cohort-sequential analyses of adult intellectual development. Developmental Psychology. 13:649-653.

Schaie, K. W. 1977. Methodological problems in the design of gerontological experiments. In: J. E. Birren & W. Schaie (Eds.) Handbook of the Psychology of Aging. New York: Van Nostrand, Reinhold.

Schank, R. & Abelson, R. 1977. Scripts, Plans, Goals and Understanding. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Scheidt, R. J. Schaie, K. W. 1978. A taxonomy of situations for an elderly population: Generating situational criteria. Journal of Gerontology. 33:848-857.

Schonfield, D. & Robertson, B. A. 1966. Memory, storage and aging. Canadian Journal of Psychology. 20:228-236.

Schonfield, D. 1967. Memory loss with age: acquisition and retrieval. Psychological Reports. 20:223-226.

Schonfield, D. & Stones, M. J. 1979. Remembering and aging. In: J. F. Kihlstrom & F. J. Evans (Eds.) Functional Disorders of Memory.

Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Schramm, W. 1969. Aging and mass communication. In: M. W. Riley et. al. (Eds.) Aging and Society, Vol. 2: Aging and the Professions. New York: Russell Sage Foundation.

Shepard, R. 1967. Recognition memory for words, sentences and pictures. Journal of Verbal Learning and Verbal Behaviour. 6:156-163.

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

Simon, H. A. 1979. Information processing models of cognition. Annual Review of Psychology. 30:363-369.

Smith, A. D. 1975. Partial learning and recognition memory in the aged. International Journal of Aging and Human Development. 6:359-365.

Smith, A. D. 1977. Adult age differences in cued recall. Developmental Psychology. 13:326-331.

Smith, A. D. & Winograd, E. 1977. Age differences in remembering faces. Paper presented at the South Eastern Psychological Association, Hollywood.

Smith, A. D. 1980. Age differences in encoding, storage and retrieval. In:

L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg, and L. W. Thompson (Eds.) New Directions In Memory and Aging. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Spilich, G. J. 1983. Life span components of text processing: structural and procedural differences. Journal of Verbal Learning and Verbal Behaviour. 22:231-244.

Standing, L. Conezio, J. & Haber, R. N. 1970. Perception and memory for pictures: single trial learning of 2500 visual stimuli. Psychonomic Science. 19:73-74.

Standing, L. 1973. Learning 10,000 pictures. Quarterly Journal of Experimental Psychology. 25:207-222.

Stern, A. 1961. Cited by J. Robinson in F. G. Kline and P. G. Ticher (Eds.) Mass Communications Research. vol.1 Beverly Hills: Sage.

Stones M. J. 1978. Aging and semantic memory: structural age differences. Experimental Aging Research. 4:125-132.

Sunderland, A., Harris, J. E. & Baddeley, A. D. 1983. Do laboratory tests predict everyday memory? A neuropsychological study. Journal of Verbal Learning and Verbal Behaviour. 22:341-357.

Tannenbaum, P. H. 1954. Effect of serial position on recall of radio news stories. Journalism Quarterly. 31:319-323.

Talland, G. A. (ed.) 1968. Human aging & behaviour. New York: Academic Press.

Taub, H. A. & Grieff, S. 1967. Effects of age on organization & recall of two sets of stimuli. Psychonomic Science. 7:53-54.

Taub, H. A. 1968. Age differences in memory as a function of rate of presentation, order of report and stimulus organization. Journal of Gerontology. 23:159-164.

Taub, H. A. & Walker, J.B. 1970. Short-term memory as a function of age and response interference. Journal of Gerontology. 25:177-183.

Taub, H. A. 1978. Recall of prose as a function of age and input modality. Journal of Gerontology. 33:725-730.

Taub, H. A. 1979. Comprehension and memory for prose materials in young and old adults. Experimental Aging Research. 5:3-13.

Tenney, Y. T. 1984. Ageing and the misplacing of objects. British Journal of Developmental Psychology. 2;43-50.

Thomas, J. C., Fozard, J. L. & Waugh, N. C. 1977. Age-related differences in naming latency. American Journal of Psychology. 90(3):499-509.

Till, R. E. & Walsh, D. A. 1980. Encoding and retrieval factors in adult memory for implicational sentences. Journal of Verbal Learning and Verbal Behaviour. 19:1-16.

Transkell, A. 1972. Reliability of evidence. Stockholm: Bechmans.

Treat, N. J. & Reese, H.W. 1976. Age, pacing & imagery in paired-associate learning. Developmental Psychology. 12:119-124.

Tulving, E. & Thompson, D. M. 1973. Grading specificity & retrieval processes in episodic memory. Psychological Review. 80:352-373.

Underwood, O.S. 1979. Quelle Histoire! Practice in French composition. (Illustrated by Bob Harvey.) Wheaton, Pergamon Press.

Vickers, D., Nettelback, T. & Wilson, R. J. 1972. Perceptual indices of performance: the measurement of 'inspection time' and 'noise' in the visual system. Perception. 1:263-295.

Waagaaner, W. A., Varey, C. A. & Hudson, P. T. W. 1984. Do audio-visual aids aid? A study of bisensory presentation on the recall of information. In H. Bouma & D. G.Boulwhius (Eds). Attention & Performance. Hillside, N.J.: Lawrence Erlbaum Associates.

Waugh, N. C. & Norman, D. A. 1965. Primary memory. Psychological Review. 72:89-104

Waugh, N. C. & Barr, R. A. 1980. Memory & mental tempo. In: L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg & L. W. Thompson (Eds.) New Directions in Memory and Aging. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Welford, A. T. 1957. Methodological problems in the study of changes in human performance with age. In G. E. W. Wolstenhole & C.M. O'Connor (Eds) Methodology of the study of ageing. London: J. & A. Churchill.

Welford, A. T. 1958. Ageing & human skill. Oxford University Press.

Welford, A. T. & Birren, J. E. 1965. Behavior, Aging & the nervous system. Springfield, Ill.: Charles Thomas.

Welford, A. T. 1980. Keynote. Memory & age: A persective view. In: L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg & L. W. Thompson (Eds.) New Directions in Memory and Aging. Hillsdale, New Jersey: Lawrence Erlbaum Associates. Tallard Memorial Conference. Hillside, N.J.: LEA.

Wells, J. E. 1972. Encoding & memory for verbal and pictorial stimuli. *Quarterly Journal of Experimental Psychology*. 24:242-252.

Witryol, S. & Kaess, W. 1957. Sex differences in social memory tasks. *Journal of Abnormal & Social Psychology*. 54:343-346.

Zelinski, E. M., Gilewski, M. J. & Thompson, L. W. 1980. Do laboratory tests relate to self-assessment of memory ability in the young and old? In Poon, L. W., Fozard, J. L., Cormak, L. S., Arenberg, D. & Thompson, L. W. (eds). *New directions in memory and aging*. Hillside, New Jersey: Lawrence Erlbaum Associates.

Zipf, G.K. 1935. *The psychology of language*. Boston: Houghton-Mifflin.