The Use Of Physical Context Information In Psychological Processing: An Investigation Into The Environmental Context Reinstatement Effect.

Andrew Rutherford B.Sc.

Thesis submitted for the Degree of Doctor of Philosophy

in the Faculty of Science, University of Glasgow, June 1986.

DEDICATION

To Magdelene Combe Rutherford

CONTENTS

3 EXPERIMENT ONE: THE DEMONSTRATION OF AN ENVIRONMENTAL CONTEXT EFFECT

4 THEORETICAL PSYCHOLOGY: PERSPECTIVE AND CRITERIA

9 EXPERIMENT FOUR: AN ATTEMPT TO MANIPULATE THE STRATEGIC USE OF ENVIRONMENTAL CONTEXT INFORMATION AT RETRIEVAL

10 EXPERIMENT FIVE: THE INFLUENCE OF RECALL DELAY ON THE ENVIRONMENTAL CONTEXT REINSTATEMENT EFFECT

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor Tony Sanford, for his perspective and helpful comments.

Useful discussion and support was provided by Tony Anderson, Michael Harker, Robert Henderson, Susan Jandoo, Bob Leiser and Kathy Moar.

Thanks are due also to Jim Mullin who wrote the slide presentation programs, Stuart Oliphant who wrote the word/colour presentation programs (as well as an extremely useful data collation program which saved me from having to spend many hours calculating modes, medians and means) and Tony Cullen who constructed much of the necessary hardware.

SUMMARY

A great deal of anecdotal evidence indicates that people employ context as an aid to memory. Psychological accounts of memory also have made much use of context as a theoretical construct. An examination of the psychological literature revealed that the term "context" has been used to refer to four classes of information; process, semantic, physiological and environmental. Studies investigating the use of process, semantic and physiological information in psychological processing were reviewed generally, while a much more detailed review was carried out with respect to studies investigating the use of environmental context information. The review of environmental context studies reported numerous investigations obtaining such effects and identified several factors that may have been responsible for the reputed unreliability of such phenomena. Consideration of the reviews of the four types of context also indicated that there was greater similarity than difference in their effects suggesting a possible equivalence of psychological function.

To ensure the tenability and validity of the proposed research in the light of the potential influence of those factors identified in the review of environmental context effects, a partial replication of Smith's (1979, expt.1) demonstration of an environmental context reinstatement effect was attempted. Subsequent to the successful attainment of such an effect, but prior to an assessment of various potential and actual models of environmental context phenomena, the emerging metatheoretical criteria that psychological accounts should comply with were discussed. The concept of different terms of abstraction was introduced, with the three forms identified by Marr (1982) defining the requirements of an information-processing

explanation. The requirement of ecologically valid process models for animate systems was emphasised and convergence on the mapping of information into knowledge structures as an initial step in processing was identified. Pre and post cognitive accounts; including Anderson and Bower's (1972; 1974) FRAN and HAM based accounts, Norman and Rumelhart's (1970) system for perception and memory, Tulving's (1982; 1983) GAP system, Glenberg's (1979) component levels account, Smith's (Smith, Glenberg and Bjork, 1978; Smith, 1979; 1982; 1984) cueing account, and Godden and Baddeley's (1980) relative cueing account, were found to be unable to accommodate the full variety of environmental context effects, and/or failed to provide an adequate account in terms of the former psychological requirements. As an alternative, an account derived from Norman and Bobrow's (eg. 1976; 1979) schema characterisation of psychological processing was proposed. Apart from being able to accommodate the variety of environmental context effects, the account also fulfilled the metatheoretical criteria. In addition, the nature of the account provides an insight into the relationship between the different contexts and establishes a basis for the prediction and exploration of environmental context and other psychological phenomena.

Experiment two examined the prediction that low level processing would facilitate environmental context effects and the tenability of the familiarity account of these effects through the use of a reinstatement paradigm. Using the same basic paradigm, experiment three examined the prediction that high level processing would reduce the effect of environmental context. The experimental methodology was considered when an alternative strategy of statistical analysis based on the linear model underlying ANCOVA was introduced and employed to assess the predicted interaction between the form of processing and the environmental context effect. The significant interaction supported the account of environmental context effects and the general utility of the schema characterisation of psychological processing.

Experiment four focused on retrieval, rather than encoding operations. An attempt was made to control and manipulate environmental context interference, and the self-generation of environmental context as a retrieval strategy. However, examination of subjects' recall protocols and questionnaire responses revealed that while the former variable was ineffective, the experimental control and manipulation of the latter variable had failed miserably. Nevertheless, the self-generation of environmental context at retrieval by the vast majority of subjects not only demonstrates that people employ environmental context information at retrieval, but that the potential utility of such information is common, although primarily implicit knowledge.

Experiment five examined the hypothesis suggested by some previous research that with a longer recall delay, environmental context information would increase in utility at retrieval. In such a situation the observation of an environmental context effect with high level processing conditions would indicate that environmental context information had been encoded during the presentation period. However, no effect of environmental context was observed with either high or low level processing conditions. These results were explained by considering the probable form of environmental context representations and the effect of time on such representations.

Experiment six had several goals. One was to demonstrate the applicability of the schema based account to environmental context defined as a function of the stimulus materials. Another was to examine the prediction that under appropriate conditions, an environmental context effect could be obtained with recognition procedures. Also, an attempt was made to determine whether the lack of environmental context effect with high level processing conditions was a consequence of encoding limitations, or the style of retrieval enabled by the form of memory representation produced by such processing. The replication of the influence of processing on the environmental context reinstatement effect indicated the general applicability of the account and also the tenability of its description of recognition and recall which was discussed at some length. The pattern of results obtained also indicated that although both encoding limitations and the style of retrieval enabled by the memory representations produced by high level processing conditions could influence the manifestation of an environmental context effect, the latter was a more likely account of the non-effect observed when high level processing was operated.

Experiment seven explored the often suggested relationship between intentional action and environmental context. Recent psychological accounts of such intentional activities also have been schema based and have acknowledged the role of environmental context in maintaining and/or eliciting actions. An attempt was made to produce a corollary of the intentional tasks reported as being affected by environmental context that was independent of general environment and that could be examined formally. The results obtained supported the theoretical accounts of environmental context use and supplemented the anecdotal evidence previously available.

The final chapter considered the particular style of the approach employed in the present series of studies. In addition, the main findings were reiterated and their implications discussed. It was concluded that the context effects reported and observed were in accord with a sophisticated processing system that strategically employs information in a practical and effective manner.

CHAPTER ONE

THE USE OF CONTEXT: GENERALLY AND IN PSYCHOLOGY

1. Evidence Of A General Appreciation Of The Influence Of Context On Memory.

1.1. Folk applications.

In murder investigations the police are increasingly utilising reconstructions of events (sometimes filmed), particularly of the victim's last known movements, in an attempt to obtain more information about the crime by jogging some potential witness's memory. This method has been successful at reminding people that in fact they did see the events which had been described so many times before.

A common suggestion to someone who has forgotten something is to tell them to go back to the place and/or position where the thought originally occurred. Frequently this helps them to remember whatever it was that they had forgotten.

Returning to an old haunt can produce a variety of memories which seemed to have been totally forgotten. Not only is it possible to remember such things as how a place looked and who or what used to be there; indeed the amount of apparently trivial detail that can be recalled is often surprising, but old feelings and motivations frequently return to mind as well. The type of memory where one feels the internal and external sensations which seem to epitomise a past situation is very often induced by a return to such a familiar place, or by some aspect of a place which is similar to one from the past.

All students, or for that matter anyone who has taken some sort of exam, will probably have experienced the ability to remember exactly in which book or set of notes and on which page, indeed the place on the page, a vital and desired piece of information occupies. Unfortunately, this type of remembering seems to manifest itself most strongly when the sought after information refuses to reveal itself.

These anecdotes have at least two things in common. They are all events which everyone I have spoken to in the past three years has had experience or knowledge of, and secondly, they are all to do with memories which are tied up in some way with the context in which they occurred. However, it is not only in Folk knowledge that demonstrations of an appreciation of some tacit relationship between context and memory can be found.

1.2. Literature and philosophy.

In both of these traditional areas of psychological enquiry, there is evidence of an awareness of some link between context and memory. Authors of the classics to authors of "pulp" magazine stories have used context as a reminder.

Proust is probably the most renowned for his contextual memories. His book, A La Recherche Du Temps Perdu, was translated to "Remembrance of Things Passed". Unfortunately, this title fails to capture the essential meaning of the French version and indeed was a translation that Proust himself was not happy with. The original title conveys the notion of attempting to re-obtain; rather than remember, passed times. This ability to obtain a period from the past from memory in any way approaching the reality it is concluded, is only possible via involuntary memory. This type of remembering is totally cue dependent, and operates at a physical level rather than any other.

A few examples from the novel itself should illustrate the phenomenon.

> No sooner had the warm liquid mixed with the crumbs touched my palate than a shudder ran through me and I stopped, intent upon the extraordinary thing that was happening to me... And suddenly the memory revealed itself. The taste was that of the little piece of madeleine which on Sunday mornings at Combray (because on those mornings I did not go out before mass), when I went to say good morning to her in her bedroom, my aunt Leonie used to give me, dipping it first in her own cup of tea or tisane (vol.1, p.48-50).

> ...I had entered the courtyard of the Guermantes mansion and in my absentminded state I had failed to see a car which was coming towards me; the chauffer gave a shout and I just had time to step out of the way, but as I moved sharply backwards I tripped against the uneven paving-stones in front of the coach-house. And at that moment when, recovering my balance, I put my foot on a stone which was slightly lower than its neighbour, all my discouragement vanished and in its place was that same happiness which at various epochs of my life had been given to me by the sight of trees which I thought that I recognised in the course of a drive near Balbec, by the sight of the twin steeples of Martinville, by the flavour of a madeleine dipped in tea, and by all those other sensations of which I have spoken ... almost at once I recognised the vision: it was Venice ... the sensation which I had once experienced as I stood upon two uneven stones in the baptistery of St. Mark's had ... restored to me complete with all the other sensations linked on that day to that particular sensation (vol.3, p.898-900).

> A servant ... chanced to knock a spoon against a plate ... the

> sensation was again of great heat, but entirely different: heat combined with a whiff of smoke and relieved by the cool smell of a forest background ... I seemed to be in the railway carriage again, opening a bottle of beer ... so forcibly had the identical noise of the spoon knocking against the place given me ... the illusion of the noise of the hammer with which a railwayman had remedied some defect on a wheel of the train while we stopped near the little wood (vol.3, p.900-901).
>
> ...I wiped my mouth with the napkin which he had given me; and instantly... a new vision of azure passed before my eyes ... I thought that the servant had just opened the window on to the beach ... for the napkin ... had precisely the same degree of stiffness and starchedness as the towel with which I had found it so awkward to dry my face as I stood in front of the window on the first day of my arrival at Balbec (vol.3, p.901).

Physical context information, which "by chance" makes contact with, and produces in the person a reinstatement of all feelings, thoughts and sensations of that time and situation is concluded to be the only method by which the true experiences of that time and situation can be remembered. Through the reinstatement of all that a person had experienced at some time past, in the present, these experiences are revealed to be outside time. It was this facet of these occurrences which had given rise to what was described in the first example as the "extraordinary thing" and that removed Proust's "discouragement" and replaced it with "happiness". Part of the reason for this transition was that while he experienced such an occurrence, he escaped the present and so had no worries of the future. Another reason for his new feelings was the realisation that his assessment of life's tedium had been based on incorrect evidence: his voluntary memories. The real remembrance of the past in the present was an event which had exciting consequences in relation to a person's real being; and that in time.

However, this is not the place to delve into the Bergsonian type philosophy of Proust's novel. What should be obvious is the explicit and important role of context in remembering.

The Russian author Vladimir Nabokov (1967) provides another autobiographical example of context aiding memory. In his book, Speak, Memory, he tells of the time, when in Cambridge again, he went to visit his old tutor. The tutor was having some difficulty in remembering who Nabokov was, until he clattered into some tea things as he had done the first time they met. At that point the tutor looked at him and said,

> Oh, yes, of course, I know who you are (p.273)

Another author who has used context information to produce some memory is George Eliot. In The Mill On The Floss, the opening chapter describes Dorlcote Mill viewed from a bridge over the river Floss. This episode is actually a dream, which is confused with reality, as the author's elbows press into her chair causing her to believe that she is resting against the stone bridge.

In the psychological literature, both Carr (1925) and Waters (1934) report two separate incidents which could have come from some comedy. Both tell of a case of a specific language capability (Chinese and Dutch respectively), which was lost out of context, that is when the people left China and Holland, and then reappeared when they returned to the respective countries.

In philosophy, the Seventeenth Century empiricist and associationist John Locke, acknowledged the psychological effect that context could exert. In his An Essay Concerning Human Understanding, he describes several instances of the link that can arise between context and memory. The lightest anecdote, also cited by Godden & Baddeley (1975) and Baddeley (1976), is of

> ... a young gentleman, who, having learnt to dance, and that to great perfection. There happened to stand an old trunk in the room where he learnt. The idea of this remarkable piece of household stuff had so mixed itself with the turns and steps of all his dances, that though in that chamber he could dance excellently well, yet it was only whilst that trunk was there; nor could he perform well in any other place, unless that or some such other trunk had its due position in the room (Locke, 1690).

1.3. Miscellaneous media.

Many less "respected" adventure novels, comics, films and television programmes provide examples of contextual reminders. Loss of memory is often repaired in the most dramatic situations, if not by a blow to the head, then more subtly by a tune, a noise, a sight, a smell, or some combination of these.

In a genre assumed by many of these categories; the detective story, many cases have been solved by having key witnesses return to the scene of the crime or action. Something about being back in the same place helps their memory.

The familiarity of such examples as these really makes specifics redundant, but for good measure one could cite the film Random Harvest, where Ronald Coleman's memory is prodded by the sound of a factory whistle and finally returned by the familiar sound of a squeaky gate, or the television series Tales Of The Gold Monkey, where the forgetful ex-alcoholic Corky had his recall aided by some odd contextual event nearly every week.

1.4. Assessment.

The original purpose for giving all of these examples was to support and illustrate the claim that there is a general awareness of a relationship between context and memory. This awareness manifests itself not only in stories about remembering, but also in some of the practical methods people utilise to aid their ability to remember.

Yet despite all this, when it is explicitly stated that context; and especially the physical or apparently non-semantic context, can help remembering, still it seems a bit strange. A possible reason for this strangeness will be forwarded after some of the empirical evidence regarding the effect of context manipulations on memory has been discussed (see chapter two, section 7.). However, with the first real separation in the types of context which may exist suggested, it seems appropriate to consider what exactly is meant when the term context is used.

2. Definitions Of Context.

2.1. Common usage.

The term "context" is quite difficult to define. In normal language its most common use is to restrict and identify one of several possible meanings. For example, saying that euthanasia is the accepted and normal practice with the vast majority of general practitioners would doubtless cause some consternation, until it was pointed out that of course we were talking within a veterinary context.

Another use of context is to refer to the physical surroundings of some activity, event, or item. Distracting contexts are often held to blame for the failure to achieve some academic goal.

I am sure that without much effort many more examples of both semantic and physical context can be thought of. Yet as this is done and more thought is devoted to the exercise of delineating between the two types of context, it seems that the exercise gets more and more difficult. The problem seems to be that physical contexts carry with them a semantic influence and likewise semantic contexts can be identified with particular physical contexts, or situations. This aspect of context will be returned to later (see chapter thirteen, section 2.3.), but for the moment it may be worth considering why the normal use of such a seemingly ambiguous term as context does not cause any problem in comprehension.

2.2. In psychology.

The problem of defining context does not only extend into, but is compounded in the psychological literature. Context must be one of the most ubiquitous and possibly vague terms presently in use in psychology. Smith, Glenberg & Bjork (1978), quite poetically have described context as,

> ...a kind of conceptual garbage can that denotes a great variety of intrinsic or extrinsic characteristics of the presentation or test of an item ...

while Nadel & Willner (1980), in a more down to earth manner, state that

> The term "context" has been used in such a bewildering variety of ways that one is tempted to conclude that it defies accurate definition.

Their opinion is that context conveys the notion of something which surrounds and influences. The problem is determining what the "something" is. However, for some theorists this "something" is almost everything. Bower & Anderson have defined context as,

> ...background external and introceptive stimulation prevailing during presentation of the phasic experimental stimuli ... [plus the subject's] mental set (Bower, 1972a, p.93).
>
> ...the subject's general mood or attitude, his physical posture and his physiological state as well as any conspicuous external cues prevailing during presentation (Anderson & Bower, 1972, p.101).

and

> ...physical characteristics of an item's presentation, implicit associations to the items and some cognitive elements representing the list in question (Anderson & Bower, 1974, p.409).

The term context appears as one of the key words in a plethora of research covering many areas and situations. Unfortunately from many of these brief mentions, it is impossible to ascertain what the term is actually being used to refer to. Most workers use "context" within their own particular context and as often as not this meaning will differ from that intended to be conveyed when other authors use the term.

3. Origins Of Psychological Definitions Of Context.

3.1. Theoretical.

Really there are two sources of context definition. One source is the theorist who utilises context in the explanation of some psychological function. It seems from the beginning of psychological enquiry and certainly since the application of associationism by behaviourists, that context has been regarded as one of the major cues involved in initiating any set of learned responses (eg. Carr, 1917; 1925; Hull, 1945; McGeoch, 1942; Skinner,1938; Smith & Guthrie, 1921; Tolman, 1932) and although some things may have changed with a more cognitive psychology, the use of context as a theoretical construct has not (eg. Anderson & Bower, 1972; 1974; Barnes, 1964; Estes, 1955; Norman & Rumelhart, 1970; Underwood, 1969). Such uses of context will be returned to later, when the ability of the models in which they are incorporated are considered as providing an explanation for the results of the second source of context definition.

3.2. Operational.

This other source of context definition, both chronologically as well as figuratively comes from those studies which have actually manipulated context in order to determine its effect upon a variety of tasks. This type of context definition is operational. But this does not mean it is made too explicit. Usually the context has to be discovered by examining the description of the experimental materials and procedure. Examples of manipulations which have been called changes in "context" could include physical environment (eg. Godden & Baddeley, 1975), the perception of stimulus relations (eg. Jones, Rana & McGonigle, 1980), semantic distinctions (eg. Light & Carter-Sobell, 1970), physiological state (eg. Spear, Smith, Bryan, Gordon, Timmons & Chiszar, 1980), nature of task (Underwood, 1977) and the background colour of stimulus cards (eg. Weiss & Margolius, 1954).

Although these examples seem to cover a whole host of experimental manipulations, a closer look at the different context studies in the psychological literature suggests that basically there are only four different types of circumstance:-

(i) when the types of tasks performed are altered.

(ii) when the interpretive semantics are altered.

(iii) when the subject's physiology is altered.

(iv) when the physical environment is altered.

To what extent the subjective and physical circumstances can be viewed as independent depends upon one's philosophical stance, while the subjective-physical dichotomy itself suggests another way in which the types of context could be separated. A distinction very similar to this is made by Hewitt (1977) and will be considered later (see chapter two, section 6.3.).

The type of separation made here is based on the fact that it is possible to vary the manipulations involved in producing these four types of context independently of each other. But perhaps of greater interest to the psychologist, is whether or not the information provided by each context is utilised in a similar manner. Here we are talking of a possible equivalence in psychological function.

In the remaining sections of this chapter, the findings from studies which have investigated the first three types of circumstance will be reviewed. A more extensive and detailed review of those studies which have examined the effect of the fourth circumstance: altering the physical environment, will be presented in the following chapter.

4. Experimental Review Of The Influence Of Context.

4.1. When the types of tasks performed are altered.

Falkenberg (1972) used the Brown-Peterson paradigm (Brown, 1958; Peterson & Peterson, 1959), to investigate the effect of a context produced by a task prior to learning and prior to recall of a consonant trigram. He found that the greater the resemblance of the task prior to recall to that carried out before learning, the greater the subject's ability to recall the target information. The measure which was most sensitive in detecting this effect was the number of errors made when the subjects recalled the consonant trigrams. The way in which these errors were defined is best illustrated by Falkenberg's own example. If subjects were given the trigram QSX, but recalled QXK, this would be regarded as two errors. So in order to recall correctly subjects had to remember order information as well as which items were presented.

Apart from demonstrating context effects with this type of manipulation, Falkenberg's paper demonstrates context effects of equivalent size with presentation-recall groups ranging from 5.4 to 18 seconds. This suggests that the advantage provided by a reinstatement of context is not restricted to attempts to retrieve information from long-term memory; although the fact that such short-term memory tasks as the Brown-Peterson paradigm have been shown to have quite large long term memory components should not be forgotten (eg. Baddeley & Scott, 1971).

A barrage of other experiments reported in the Falkenberg paper, tested alternative explanations of the context effect. Explanations such as its being due to proactive or retroactive interference, varying degrees of rehearsal, some artifact of a within subject design, or an increase in between trial interval were all found to be inadequate accounts of the effect.

Falkenberg explains the phenomenon as being an example of stimulus generalisation. He contends that the distinction between stimulus and context is arbitrary and that if the context is regarded as part of the "overall functional stimulus", such effects as he reports would be predicted by stimulus generalisation. This analysis is based on McGeoch's (1942) explanation of forgetting (which will be discussed in chapter five, section 3.1.). Falkenberg however, attributes McGeoch with a four factor account of forgetting, the extra item being stimulus generalisation, while McGeoch would seem to regard this as more of an effect of some circumstances upon which the three factors depend (eg. stimulus similarity). The adequacy of this type of explanation will be considered in chapter five.

As part of an attempt to discover how people can remember and identify the temporal order of information previously presented to them, Underwood (1977) carried out a similar manipulation. Subjects were required to carry out different types of memorisation tasks on a series of four word lists. He found that subjects who did this improved their subsequent ability to identify a word's temporal position (ie. which list, 1,2,3, or 4, did the word belong to?) compared to subjects who carried out the same type of task with all lists. This finding was independent of the degree of learning which occurred with any list.

Similarly Block and Reed (1978) found that subjects' perception of a time period was not influenced by the specific nature of the task (semantic or structural) they were required to carry out, but it was affected when they had to alternately carry out both types of task within this period.

When subjects carried out the two tasks alternately, the same time period was regarded as being longer than when they were required to carry out only one type of task.

Although the latter experiments do not attempt to examine any reinstatement of the original context on the subjects' ability to remember target information (as the majority of experiments to be reported do), still they provide evidence that the context produced by the operation of psychological processes can serve to differentiate items in memory.

Underwood explains his results in terms of associations formed between list numbers (ie. 1,2,3 and 4), the process context and list words, while Block and Reed emphasise the aspect of contextual change to account for their results and refer to Anderson and Bower's (1972) model of memory to provide a mechanism to accommodate their explanation (for an account of Anderson & Bower, 1972, see chapter five, section 4.).

4.2. When the interpretive semantics are altered.

Of all the types of context, this is the one which has received the most intensive and extensive investigation. It would be virtually impossible and totally impractical to provide a complete review of all of this work. Instead a general outline of semantic context studies and their results will be given. Hopefully this will manage to convey some of the importance that these studies have had in changing opinions and explanations of learning and memory, as well as providing an account of the basic findings.

The study of semantic context bridges the gap between the Ebbinghaus and Bartlett traditions of investigation. Generally the study of memory was propagated by adherents to the Ebbinghaus tradition (from Ebbinghaus, 1885). Stimuli presented to be memorised were devised to contain as little information as possible. In this way it was hoped that a person's prior knowledge would not influence the learning of the stimulus materials and so a clear picture of the fundamentals of learning and memory would emerge. To some extent this methodology had success. Functions such as the rate of learning and forgetting determined by these methods have been elucidated by subsequent research rather than invalidated. Yet as this type of methodology and the almost atheoretical approach that aligned with it continued, it became more and more obvious that it was not going to provide a comprehensive explanation of memory function.

Bartlett (1932) disagreed with the general acceptance of the Ebbinghaus approach, claiming that the attempt to eliminate meaningful information from the stimulus materials created a situation which was not representative of normal memory. Bartlett developed his own techniques to investigate more natural memory, using pictures, passages, stories, etc. as stimuli. On the basis of these investigations he claimed that one of the most important aspects of learning and memory was the subject's attempt to extract some meaning from the material they were presented with.

Eventually as psychology reoriented and again regarded mental activity as being within the realm of acceptable theory, the role of meaning began to be considered in models of memory. But even until the late Sixties and early Seventies the method of investigation was almost exclusively a direct descendant of the Ebbinghaus tradition. The presented stimuli were either logically unassociated paired-associates, or word lists. Using such methods, the role of the specific semantic context within which an item was presented began to be investigated.

Since then, the extent of the semantic context investigated has increased. Investigations into the effect of supersentential or thematic context; comparable with those of Bartlett, have been carried out producing similar effects and developing similar models of explanation.

Many studies of semantic context have shown that the ability to remember is best when the original semantic context is reinstated. Although some may be expressed differently, explanations of this phenomenon assume that information is encoded within a particular context and is best retrieved when a similar context is invoked. A variety of systems have been constructed which attempt to model this aspect of behaviour. Some of these will be considered in chapter five. But exactly what evidence is there that information is encoded within a particular context and is best remembered when this is reinstated?

Probably the most illustrative and convincing evidence of the claim that information is encoded within a particular context comes from studies by Dooling and Lachman (1971) and Bransford and Johnson (1972). Both sets of researchers have shown that the provision of an appropriate thematic context, approximately doubles subjects' comprehension and recall of a passage. This doubling of comprehension and recall was in comparison to the situation where no-context was provided, even although each individual sentence was logically meaningful.

Another experiment, by Bransford, Barclay and Franks (1972), emphasised that the context used in forming a representation does not have to be specified explicitly in the information presented. The usual example taken from their study is of the following sentences:

(1) Three turtles are sitting beside a log. A fish swam under them.

(2) Three turtles are sitting on a log. A fish swam under them.

The important aspect of the two pairs of sentences is what they allow the reader to infer about what the fish swam under. It is only in the second pair of sentences that it is logical to infer that the fish swam under the log as well as the turtles.

Bransford et al. tested subjects' ability to recognise sentences saying that the fish had swum "under the log". Subjects who had received the first pair of sentences did not identify other sentences which said that the fish had swum "under the log", but subjects who had received pair (2) were as confident that they had seen "under the log" as they were confident that they had received the actual "under them" wording. This type of error in recognition could only occur if subjects were including information in their representation of the event, beyond that provided by the stimulus materials.

It is this tendency for subjects to manifest a context for the presented information beyond that actually presented to them which explains why the word JAM has a different "meaning" when coupled with STRAWBERRY than when it is paired with TRAFFIC.

Light and Carter-Sobell (1970) presented subjects with a list of words each accompanied by another which invoked a particular contextual representation. The words strawberry/traffic-JAM are taken from their study. Later subjects had to identify words which had been presented to them earlier. The words presented to them were either the original pair, the second word alone or this word preceded by another which invoked a different context. Recognition performance was best when the original pair were presented again. The next best performance was when the second word (eg. JAM) was presented alone and the worst performance was obtained when another word, invoking a different context, was presented along with one of the originals (eg. strawberry-JAM, then traffic-JAM). The same effect has been demonstrated when more subtle variations in encoding context have been presented, for example, train-BLACK, then white-BLACK (Tulving & Thompson, 1973).

A whole host of studies which have investigated semantic context have replicated these effects. One study by Woodall and Folger (1981) has demonstrated that the benefit of reinstating the original semantic context which was present at encoding when subjects were trying to remember, was not restricted to verbal material. They used meaningful gestures to accompany speech and found that subjects recall of what had been said was better if they were re-presented with the gestures which accompanied it.

Unlike Woodall and Folger's experiments, the majority of studies which have investigated the effect of changing or reinstating the original encoding context at the time of remembering have used verbal material as both stimuli and context, and recognition procedures to measure memory. The use of verbal material obviously allows the greatest and easiest manipulation of semantics. Recognition measures also tend to be used for convenience. The typical type of paradigm, as illustrated in the discussion of the Light and Carter-Sobell, and Tulving and Thompson studies, involves the presentation of two words of which one is the nominal stimulus. Together these words produce a certain context. The use of a recognition procedure at test allows the re-presentation of the nominal stimulus either alone, with another word so altering the context, or with the same word as before so reinstating the original context. With this type of procedure the experimenter has a large degree of control over the semantic context at presentation and test. Control of the semantic context at test is not so easy to obtain with recall procedures.

Recall procedures in investigations of semantic context are most often observed in the situation in which some particular semantic context is, or is not provided at the encoding stage (eg. Dooling & Lachman, 1971; Bransford & Johnson, 1972). Here there is no manipulation at test. This type of experiment primarily examines encoding, while the reinstatement-nonreinstatement of context at test experiments focus on the encoding-retrieval interactions.

Despite the greater awkwardness of using recall as a measure of memory performance in a reinstatement-no reinstatement paradigm, it has been done and similar results to those obtained with recognition procedures have been reported. In order to achieve recall within a prevailing semantic context, several studies have used cues (eg. Barclay, Bransford, Franks, McCarrell & Nitsch, 1974; Ley & Huba, 1980), so the type of remembering is actually cued-recall. However some studies have managed to obtain free-recall measures as well as cued-recall measures with reinstatements and changes in semantic context. Marcel and Steel (1973) included a free-recall (no-cue) condition in their investigation of semantic context. Not surprisingly, they found that an appropriate (ie. same semantic context as learning) cue produced better remembrance than the no-cue conditions (immediate and delayed), but they also found that these no-cue conditions produced better remembrance than the provision of an inappropriate (ie. different semantic context to learning) cue. A clever experiment by Reddy and Bellezza (1983) demonstrates why this is. They recorded the overt verbalisations of subjects as they learned a list of words. There were two learning conditions; one where subjects had to construct a story from the words presented to them, and one in which subjects had to produce associations between words and describe the visual images they had of the words. Their ability to remember the list of words was tested in three different ways. In one condition subjects free-recalled, but again they had to overtly verbalise what they were thinking as they did so, and as before this was recorded. Another condition provided subjects with a written record of their own learning verbalisations, but without the list words. The third condition again provided a learning verbalisation transcript without the list words, but each subject received one which had been produced by some other subject.

In accord with the other studies mentioned, Reddy and Bellezza found that the subject's remembrance was related to the degree of semantic context reinstatement. But what they were able to show was that the recorded free recall verbalisations also took the form of semantic context reinstatements. Those subjects in the condition which had to construct a story seemed to be able to recreate the context more easily than the subjects in the other learning condition and as a result were able to recall more. Reddy and Bellezza attribute this easier re-creation of a story to its greater formal structure.

Overall these studies demonstrate the importance of semantic context in memory. Studies such as those by Dooling & Lachman (1971) and Bransford and Johnson (1972) show that comprehension and memory performance are dependent on having a context within which the information can be represented. This ties in with Bartlett's concept of the "pursuit of meaning", although it would seem that meaning is a consequence of the context. The studies involving semantic context reinstatements and changes as memory is tested suggest that in order for information to be retrieved, it has to be in a context. This in turn suggests that an efficient "memory system" should operate by encoding information within a context. Attempts to encode in this manner are equivalent to a "pursuit of meaning". This is something which has plagued attempts to obtain meaningless stimuli and has produced the paradox of degrees of meaningfulness with non-sense syllables.

4.3. When the subject's physiology is altered.

Experiments of this type are usually termed investigations of state-dependent learning. However, this title can be misleading. The characteristic effect observed in such studies is a function of the congruity of the learning and memory test states. The phenomenon is therefore a result of encoding and retrieval interactions. For this reason both Wickelgren (1975) and Eich (1977) prefer to re-label the phenomenon state-dependent retrieval. Although this shifts the emphasis, it could be just as misleading.

This type of analysis of the state-dependent effect is typically psychological. The general model applied is of information-processing that over time can be conceived of having at least three stages; encoding, storage and retrieval. On the basis of the behaviour observed in particular circumstances, certain conclusions can be drawn in terms of the model.

Another approach which seems to be encouraged by the physiological nature of the state-dependent variable is to try and identify the physiological-biochemical changes which produce the effect. These two approaches are not mutually exclusive, but it seems that the latter approach still has to explain its results in terms of the former.

The reason for this is that the two types of analysis describe the same events, but in different terms of abstraction. It is just that the form of abstraction which seems to convey with most meaning the psychological-physiological-biochemical-chemical-physical etc. processes is the psychological information-processing analysis. This topic will be returned to in chapter four, but for the moment it may be well to emphasise that "form of abstraction" has nothing to do with distinctions between wholism and reductionism.

The usual state dependent experiment has four conditions. Subjects are tested under two different physiological states: one drug induced, and one usually non-drugged. Memory for some task is tested with learning in state 1 and recall in state 1, learning in state 1 and recall in state 2, learning in state 2 and recall in state 2, and learning in state 2 and recall in state 1. A significant interaction between learning state and recall state indicates a state dependent effect.

Unfortunately not all experiments follow this methodology. The area of state dependency is one which has many contributors from different disciplines. Physiologists, biochemists and psychiatrists, as well as psychologists have produced research in this area. Any person intent on delving into this large literature should be ready to search out a good number of obscure journals.

Fortunately, Eich (1980) has provided an impressive review of this field including all work published or presented at conferences between January 1965 and December 1978. Eich's purpose in doing this was to try and discover why state dependent effects seemed to be so unreliable. Given the variety of approaches; some of which border on the methodologically inept, the variety of sources and variety of drugs and dosages used, this was an extremely worthy task.

Several hypotheses suggesting why and where the state dependent effect should and should not appear were considered. They were (a) the type of psychoactive drug administered, (b) the dosage of drug dispensed, (c) the nature of the items to be remembered, (d) the "level" of item analysis and (e) the nature of retrieval cues available to the rememberer.

Eich presents a very persuasive case for the hypothesis that the state dependent effect is actually a cue-dependent phenomenon. The reason for the number of studies which could not replicate the state-dependent effect was their use of procedures which are not sensitive to the effect. Specifically, Eich claims that failures to demonstrate state-dependence are restricted to situations in which the utilisation of stored information is tested in the presence of discreetly identifiable retrieval cues. The greatest retrieval cue is what Eich calls a copy cue. Here the same stimulus is presented again and the subject has to say whether or not it was presented originally: in other words a recognition test. Of the fifty-seven studies covered in the review, only two studies using a recognition paradigm found a state-dependent effect. The other eleven studies using this procedure did not obtain a significant result. A second class of cues available to subjects are what Eich calls list cues. These include units of information (other than copy cues) which originally appeared as part of the experimental list. Situations in which this class of cues are available are generally cued recall paradigms, such as paired associate or category cueing. Of this kind of study, ten out of eleven did not produce a state-dependent effect. Of all the other studies (minus seven with negative outcomes previously ascribed to the administration of ineffective drug doses), only three instances did not produce a state dependent effect. Twenty three studies in which only "invisible" cues were available; that is unidentifiable cues which for theoretical convenience are assumed to be used, produced the state-dependent effect.

In conclusion, Eich makes the point that although the nature of the cueing condition seems to be a prime factor determining the appearance of state-dependent effects, there may be other factors which also contribute to the occurrence of the phenomenon. Three of these are explicitly mentioned by Eich. They are; the degree of original learning, the population from which the subject sample is drawn and the extent to which the memory task requires the remembrance of serial order information.

The first of these, the degree of original learning, would seem to be intimately related to one of the possibilities already considered by Eich, namely the "level" of item analysis. Eich reports only one study which investigates such a manipulation in conjunction with changes in state; an experiment carried out by himself. In this study he found an improvement in recall with deeper (semantic) processing in both same and different state recall conditions. There was no interaction between depth of processing and recall condition. However, Eich had subjects process words semantically by asking them to give synonyms. It is possible that this procedure could cause large interference effects. Subjects perhaps would be unable to distinguish between the presented items and their generated synonyms at recall. If so, this would cause problems in scoring and if only presented items were scored after recall, this procedure could actually reduce the recall scores in comparison with some other method of achieving a "deep level" of processing. Any consequence this interference would have in relation to the state dependent effect would depend on whether or not it would exert a systematic effect and whether or not it would produce an increase in variation which could mask the interaction between level of processing and recall condition. Only further research will provide an answer.

The possibility that subjects' ability to remember serial order information is affected by changes in state from learning to test has been suggested by several researchers (eg. Eich, 1977; Hill, Schwin, Powell & Goodwin, 1973; Stillman, Weingartner, Wyatt, Gillan & Eich, 1974), but this hypothesis goes no further than describing an aspect of some tasks which have been affected by the change in state. There has been no suggestion of how, or why a change in state should affect memory for tasks with such requirements.

The suggestion that the type of subject could influence the occurrence of state-dependent effect arises from studies which have examined alcoholics' performance (eg. Weingartner & Fallice, 1971; Lisman, 1974). However, as many studies have obtained state-dependent effects with "normal" subjects, it is questionable to what extent the findings of studies using alcoholics as subjects should be regarded as contradicting any account of "normal" state-dependent effects. A criticism pertinent to this issue is detailed in the following paragraph in relation to the study of mood state with depressive patients.

Eich's (1980) review dealt solely with research which had used drugs to induce changes in state. In addition to this research, there are studies which have investigated the effects of changed states such as are caused by alterations of the subject's mood. These studies are not always so orientated toward a memory perspective, but instead are often carried out to investigate some aspect of psychopathology. Subjects are usually suffering from some clinical syndrome which may influence their performance. Consequently, any perceived changes in psychological function may have more to do with particular aspects of the general syndrome, rather than just the altered mood state.

However, several studies using normal subjects have been able to achieve changes in state by manipulating subjects' mood (eg. Bower, 1981; Bower, Gilligan & Monteiro, 1981; Bower, Monteiro & Gilligan, 1978; Leight & Ellis, 1981). In accord with the drug state studies, both Bower and Leight and Ellis have found an advantage for remembering in a similar mood-state to that of learning. Leight and Ellis induced moods in their subjects by using a procedure developed by Velten (1968). Subjects read 60 statements which refer to the reader (ie. themselves). The examples of a depressive statement and an elated statement given by Leight and Ellis are: "Every now and then I feel so tired and gloomy that I'd rather just sit than do anything", and "If your attitude is good, then things are good, and my attitude is good". After assessing the effect of the mood inducement, they presented subjects with a perceptual grouping task. In the experiment which investigated reinstatement effects only depressive moods were induced, but two input conditions were used: constant and varied. Letter sequences of a particular structure were used as stimuli. An example of one of the sequences would be BONKID. Each letter sequence consisted of four consonants and two vowels presented in this type of order which potentially could create two pronounceable trigrams, but each sequence was actually divided at different points, eg. 2,2,2, giving BO NK ID; or 1,4,1, giving B ONKI D, etc. Each letter sequence was presented four times and could be presented constantly with a particular structure, eg. 2,2,2, or they could be presented with a different structure each time, eg. 2,2,2; 1,4,1; 2,3,1; 1,3,2. Contrary to Eich's (1980) argument, Leight and Ellis found asymmetrical reinstatement effects only with recognition in the constant input condition, and not with free recall. One factor contributing to this result may be that the recognition and recall statistical analyses differed in power. If the error terms are compared, it can be seen that the recognition error term (MSe = 1.75) is almost half that of the recall error term (MSe = 3.21). Another explanation of the difference between the recognition and recall effects in the constant input condition is that subjects in the free recall group used their memory of the mood they had learned in to aid recall, while subjects in the recognition group did not. This would amount to a self-generated reinstatement of mood and is not without a parallel in the environmental context literature (Smith, 1979; described in chapter two, section 2.2.) Unfortunately, little information is provided regarding the free recall and recognition procedures. In the usual recognition procedures, subjects have to press a key to indicate "yes" or "no" as quickly as they can, or they are given a list within which they have to identify items. In both cases attention is focused on the specific items and quick decisions are made through choice or instruction. In recall, subjects' attention can shift from aspect to aspect. Retrieval is much less directed by external influences and more by the self-generation of (sometimes complex) strategies.

The other main finding as far as the perspective of this review is concerned, was that the reinstatement effect was only observed in the constant input condition. Leight and Ellis accounted for this by suggesting that it was due to the greater degree of learning which occurs in the varied input condition. It has already been mentioned in relation to the drug-state studies that greater learning may influence the state-dependent effect (the effect exerted by this variable will be returned to in relation to environmental context). The asymmetry of the reinstatement effect is considered to reflect the lack of distinctiveness provided by the association of the to-be-remembered item to a neutral or normal mood-state. As so much information is acquired in this state, the state will not particularly discriminate the to-be-remembered information from any other information.

In the investigations carried out by Bower, hypnosis and imagination were used to induce mood-states. Bower (1981) reports the initial difficulty they had in obtaining reinstatement effects. Eventually they discovered that the effect could be secured by presenting subjects with two word lists to learn: one in mood A, the other in mood B. Recall of the first list learned was then tested in either the same or in a different mood to that within which it had been learned (a similar type of paradigm will be seen in some environmental context studies). Bower presents an explanation of state-dependent effects which is based on Anderson and Bower's (1974) propositional model. In essence, the explanation of the reinstatement effects and why they only occur when two lists are presented in two separate moods provided by Bower is similar to the account forwarded by Eich (1980). Both contend that state-dependent effects only will be observed when the to-be-remembered information cannot be retrieved except by its distinctive association with the particular state. These models will be reviewed in chapter five.

In an investigation which relates to the "state" literature: in that the source of information accompanying the nominal stimuli arises primarily, although not exclusively, from the subject's physiology, Rand and Wapner (1967) examined the effect of posture changes on the ability to relearn a list of six non-sense syllables. Subjects learned in both erect and supine postures, and relearned in all four reinstatement and change conditions. Four separate lists were presented with each item being shown for 3 seconds. Relearning occurred after subjects had spent fifteen minutes doing sums.

Relearning continued until subjects had reached a criterion of two consecutive correct trials. The ratio of the number of trials to relearn to the number of trials to learn originally, to the criterion of one errorless trial was calculated in all conditions for each subject. Two such calculations were made. One was for a weak criterion: the number of trials to one errorless trial, and the other was for a strong criterion: the number of trials to two errorless trials.

Analysis revealed a significant benefit for congruent posture at learning and recall, but only with the weak criterion. No effect was observed with an analysis of the number of trials to the strong criterion. So, whatever factor is distinguished by the change in criterion has influence on the manifestation of a posture effect.

CHAPTER TWO

A REVIEW OF STUDIES INVESTIGATING THE EFFECT OF MANIPULATING ENVIRONMENTAL CONTEXT

1. Introduction.

1.1. Theoretical orientation.

In most psychological texts concerned with memory there is usually a section that mentions a token number of studies which have obtained some effect on memory by varying the physical environment between learning and recall. Generally, these accounts seem to be included to achieve a certain breadth of coverage, rather than as examples of important empirical results provoking serious theoretical consideration. In all likelihood this is due to the general opinion that such effects are not only unreliable, but are of little consequence in the development of an understanding of memory.

The theoretical viewpoint and purpose of most of the research using a physical environment manipulation carried out between the early 1930s and late 1960s compounded this latter opinion. These experiments set out to test interference explanations of forgetting. Their reasoning was that if associations were established between the nominal stimuli; usually verbal material, and the physical environment, then the physical environment should be able to evoke the verbal material. If at some later point the subject tried to remember some other information within this physical environment and the former stimuli were evoked by their association with the environment, interference between the two sets of information would be observed.

As the main purpose of these experiments was to alter the levels of interference observed by varying the degree of association between the physical environment and some nominal stimulus set, it was assumed that the physical environment exerted its effect by means of little more than accidental (contiguous) association. Hopefully a proper review of the literature reporting such alterations of physical environment in terms of the effects observed and later consideration of theoretical accounts of these effects will persuade that physical environment phenomena are not quite as simple as first suggested, and are of importance in developing an understanding of the psychological operation of memory.

1.2. Definitions and layout.

Prior to the review of these studies it seems appropriate to provide some explanation of the meaning of the term "physical environment". Physical environment refers to all the physical parameters within which a stimulus is represented, from the shape and hue of the stimulus materials, to the lighting, any sounds or smells, the temperature, the size of the room or wherever the subject is, where it is, etc. Time is not included in this definition as it is assumed that any variables associated with a change in time will be accommodated by this, or one of the previous sections, eg. over the course of a day, the light level changes (physical environment), and (state) changes associated with circadian rhythms occur.

Due to the size of this category of variables, it seems sensible to divide it in some way to allow easier reading. The division is based on a simple methodological distinction. The physical environment variable can be some information contained within the stimuli materials, such as specific additional items, the colour background which the nominal stimuli items are presented upon, or even the colour of the nominal stimuli items themselves.[1] Alternatively, the physical environment variable can be what ordinarily one would comprehend as being physical environment, that is an area, place, or room.

2. When The Physical Environment Is Altered.

2.1. Physical environment as a function of the stimulus materials.

One of the first, if not the first published study of context and its effect on human memory was reported by Shuh Pan in 1926. Pan investigated the influence of several types of context with a paired

1 In such a situation it makes sense to consider the experimentally defined physical environment as the nominal environmental context for the same reasons as the experimentally defined stimuli items are termed nominal stimuli items.

associate paradigm. Each pair of words had presented along with them words which were either logically related to the stimuli or response word, words which were logically related to both stimuli and response words, or words which were logically unrelated to stimuli and response words. There were also conditions where no stimuli were presented except for the stimuli and response words (no-context condition) and where numbers accompanied the two stimuli and response words at presentation. In another experiment a face and name were presented on top of a picture card of some well known place in Chicago. Pan investigated the effect of all these contexts at learning and also variations of the contexts at recall.

The type of context manifested by the provision of logically related and unrelated words is likely to be semantic. As such the results of these manipulations are in accord with those already discussed in chapter one (section 4.2.). The context variations which are of concern here are those involving numbers and places.

Each stimulus pair and their context was presented for 3 seconds, but while the word pairs and number contexts were presented only three times, the face name pairs and place contexts were presented until memory for these items was perfect. In comparison to the no-context condition, the learning of the word pairs with a number context was hampered. However, if the data are analysed in terms of Z scores (t-values cannot be calculated as the number of subjects per group is not given), this difference has a probability value of only 0.192, and is therefore not significant. A comparable analysis of learning was not carried out with the place context data.

One further difference separates the number context study from that of the place context. At recall the intervening gap between learning and test for the number context study is twenty four hours, while the same gap for the place context study is forty eight hours. A comparison between recall when the original learning context is reinstated and when no-context is provided is made with the number context data. This result again is not significant (probability associated with Z score = 0.145).

There is no report of a manipulation involving a test for recall when a different number context is present. In contrast, in the place context experiment, there is no report of recall when no-context is presented. Here the comparison is between the original learning context and a new context. A significant advantage for the reinstatement condition is found when the data are re-analysed using a related samples t-test ($t = 8.024$, $p < 0.001$).

Pan also investigated the relationship between the effect of a change in place context and the degree of original learning. As learning had continued until perfect memorisation, those pairs learned first could be assumed to be better learned by the end of the list of face-name and place context presentations. Pan observed that the detrimental effect of context change varied with the degree of learning. The face-name pairings least well learned were most affected by a change in context.

Dulsky (1935) presented subjects with a pair of non-sense syllables on a coloured card every 5 seconds. Ten pairs of non-sense syllables were presented three times, then recall was tested. Learning and recall was then alternated until perfect recall was attained. After five minutes recall was tested again and if it was not perfect, it was once more alternated with learning until it was perfect. Two types of measure were recorded; the total number of correct and incorrect items recalled and the number of relearning trials to the criterion of perfect recall. Eleven different learning and recall conditions were investigated. The main finding was that changes in background colour from learning to test decreased subjects' ability to remember. This was observed with both types of measure. Changing the background colour of the response term side of the card decreased remembrance more than changing the stimulus side colour. This type of change in background colour reduced the ability to remember more than a change from the single background colour for stimulus and response at learning, to another single colour at test.

Some changes in colour background were from colours such as red, yellow, etc., to grey, while others involved swapping colours from one non-sense syllable to another throughout the test. In this latter situation Dulsky assumed that if the response term was associated with its colour and its stimuli term, then in the situation where the colours are interchanged there will be response competition between the term cued by the colour and the term cued by the stimulus term. If this happened one would expect a reduction in the ability to remember in this condition as compared to the condition in which there was a change in colour to the non-term associated grey. Dulsky obtained this difference between the two conditions and so reasoned that the deleterious effect of a change in colour background has more to do with a lack of cues normally provided by the colour than with the novelty of the new background. The last two conditions Dulsky examines are the only ones which are similar to the usual context/no-context reinstatement paradigm. The whole set of stimulus and response terms were presented with the one background colour and are tested with this reinstated or not. Dulsky found slight differences between these two conditions in the expected direction, but claimed it was statistically unreliable. This result is in accord with his claim that the lack of (colour) cues is the most potent factor, as in this situation the response term's colour background was reinstated at test.

There is however, one conspicuous problem with the Dulsky paper. Throughout its entirety there is not one statistical comparison reported. Only total recall and presumably the average number of trials to relearn are reported per condition: no variance measures are given so it is not possible to re-analyse the data. Subject numbers per group are given, but as no explanation of the make up of the totals is provided; with respect to the number of items obtained per recall, any calculation of an average would be meaningless. It is perhaps best to concentrate on the numbers of relearning trials as these are independent of recall, while recall is not independent of trials. But of course this does not overcome the problem of attaining an indication of statistical significance.

Weiss and Margolius (1954) set out to investigate the support that background stimuli could provide in both the learning and recall of what they termed primary stimulus items. In five conditions subjects were presented with nine paired-associate non-sense syllables, each on a different coloured card. The stimulus item was first shown to the subjects for 3 seconds before a shutter lifted to reveal the response item also. Stimulus and response items were then presented together for 3 seconds. After twenty-four hours, subjects' ability to recall the response items was tested.

Two factors were investigated in conditions one to four; the effect of changes in the primary stimulus and changes in background colour on the ability to remember. Conditions five and six examined the effect of colour background on learning.

Group one was the reinstatement condition where the same stimuli items were presented again on their original colours. The group two situation was similar except that all the stimuli items were presented on grey backgrounds. In group three, subjects were presented with stimuli items in which the fourth and fifth letters of the six letter stimuli items had been changed. These represented generalised primary stimuli. All the colour backgrounds remained constant from learning to test, but group four subjects were presented with generalised primary stimuli items on grey background coloured cards. Group five subjects were presented with the original colour of cards, but not with the primary stimulus items. A sixth condition had subjects learn the same paired-associates, but they were presented on grey cards. At recall this situation was similar to condition five in that the different colour cards were presented, but without the stimuli items. In all conditions the number of trials required to learn the list to a criterion of one perfect recall was recorded. At test the number of items correctly recalled on the first trial was recorded, as was the number of relearning trials required to obtain the previous criterion.

A significant benefit in terms of the number of trials to criterion was found for those subjects who had learned with the different background colours compared to those who had learned only with a grey background colour (ie. groups 1-5 cf. 6). For groups one to four, the type of background colour presented at recall was found to be significant. There was no significant effect of the type of stimuli item presented at recall, nor was there any interaction between these two factors. This indicated that there was an advantage for remembrance when the colour background was reinstated even when the nominal primary stimuli items were slightly changed. The average relearning score for group five was significantly less than that for group six, but was not significantly different to those of groups one and three. Another comparison between the group five and group two relearning scores (which is not reported in the article, but which can be calculated from the published data), shows a significant relearning advantage for group five ($t = 3.25$, $df = 14$). These results and perhaps the latter by itself, which compares remembrance of the response, when only the colours are presented, with that when the nominal primary stimuli are presented without the learning colours, suggest that the colour background is the most effective response cue.

Underwood (1963) explained the Weiss and Margolius data in terms of cue selection, where the subject selects aspects of the stimulus compound as the functional cue. Saltz (1963) claims that although this explanation suggests that both the nominal stimulus term and the background colour could both cue the response, it does not unambiguously explain why there is superior performance when both the primary stimulus and the background colours present at learning are presented at test (ie. the reinstatement condition). Saltz suggests that this is due to the colour background making the stimulus more distinctive at learning and so affecting the type of representation: cognitive differentiation. This is distinguished from any perceptual differentiation which would not affect the type of retention of the information, but only the way it was perceived. In an attempt to ascertain which of these hypothesised processes produced the effects observed, Saltz presented ten paired-associate (non-sense syllable-word) items either each on a different background colour, or without any background colour. At test (4 seconds after each learning trial), the nominal stimuli (non-sense syllable) could be presented on its original presentation colour or not, and subjects had to recall the response word. Each paired-associate was presented for 2 seconds at learning and each subject's total recall after fifteen learning and recall trials was recorded.

An analysis of variance indicated that both colour during learning ($F_{1,148} = 19.14$, $p < 0.001$) and colour at test ($F_{1,148} = 12.02$, $p = 0.001$) significantly aided recall of the response item. As there was no interaction between these two factors it suggested that both cognitive and perceptual differentiation aided recall to a similar extent. The recall in both the colour-no colour and no colour-colour conditions were both significantly greater than the condition in which no colour was presented at learning or test (ie. no colour-no colour). As better recall is obtained in situations where the colour was present only at learning <u>or</u> at test and therefore could not directly cue the response term, any cue selection hypothesis would have to involve an association between the stimuli term as well as the response and the background colour. It could be assumed that in the colour-no colour condition, subjects could form associations between the stimuli items and the colour. At test the presentation of the stimuli item could evoke the colour which could then elicit the response. In the no colour-colour condition it would need to be assumed that an association between the stimuli item and colour would be formed at test. On subsequent interleaved learning trials the stimuli item could then evoke the colour which in turn could then be associated with the response.

To reduce the likelihood of such a strategy being operated by subjects, Saltz presented stimuli items which were easier to associate with the response items than they were to the colours. Results similar to those before were obtained and subjects interviewed after the experiment reported that the conscious use of colour as a cue was too confusing to maintain.

Despite all this, the result that when the stimuli and colours presented at learning were reinstated at test the best recall performance was obtained was still not unambiguously accounted for. The reinstatement condition is the only condition in which colour could act as a cue, so it is possible that this in addition to differentiation could have produced the superior recall. However, this condition is also the only one in which both cognitive and perceptual differentiation can occur. Therefore the result could also be explained by their simultaneous operation.

In line with the traditions of experimental psychology, another study was carried out by Birnbaum (1966). This looked at the possibility that in the colour-no colour condition associations did develop between the nominal stimuli item and colour, and the colour and response. Replicating Saltz's (1963, expt.2) procedure, Birnbaum found that both such associations existed between these elements ($p < 0.001$ in both instances). In addition, in a transfer of learning task where subjects had to learn to provide the response term when presented with either one of the previously associated colours, or a different colour, it was found that there was easier learning, as measured by correct response term recall, with the colour-response pairs which were consistent with the original learning ($F_{1,48} = 17.22$, $p < 0.001$). This result is not predicted by Saltz's stimulus differentiation hypothesis. No investigation or mention of the tenability of the more involved cue-mediation hypothesis as an explanation of the no colour-colour result was made.

Elio and Reutener (1978) brought the debate between stimulus cuing and differentiation explanations of the effect of colour context almost into the eighties. However, one important change was that the revived application of the two concepts took the form of possible psychological operations which need not be mutually exclusive. They set about investigating the method of effect colour context had by applying the hierarchical organisation of words paradigm, the influence of which had been researched previously by Bower, Clark, Lesgold and Winzenz (1969). This paradigm involves the presentation of a set of words which have a common conceptual link in an inverted tree like structure. At the top of the set is one word which subsumes the two words below it. These two words subsume two words below each of them and in turn each of these four words are the nominal headings of a four word list. In the Elio and Reutener study this hierarchical clustering and relationships between words also could be manifested in the background colours of the stimuli words.

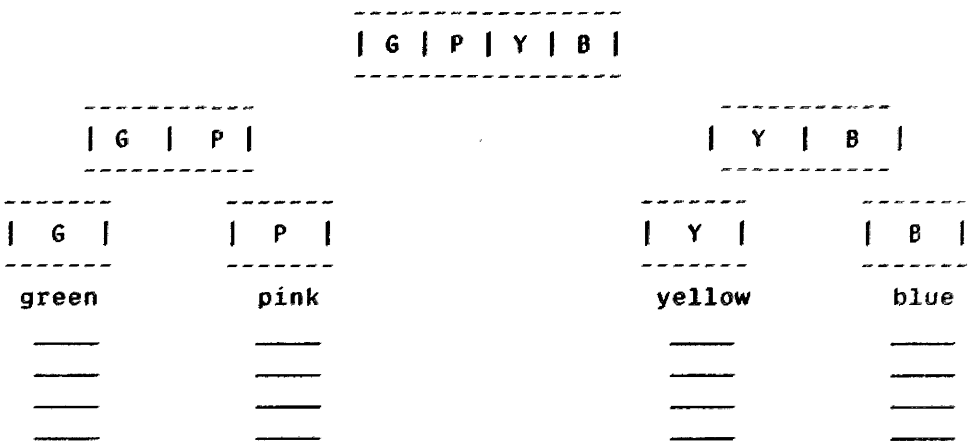

Figure 2.1. Background colour hierarchy (from Elio & Reutener, 1978).

The top word could have four bands of colour, the two words below it would each have two colours from the top word's four colours, and the four word lists and their headings would each be on one background colour (see figure 2.1.).

In all, six conditions, or six types of word hierarchy were presented. Coloured or black and white hierarchies could be presented and this was crossed with conceptually organised word sets or randomly organised word sets. In all conditions the same spatial layout was maintained. The conditions were (i) blocked (ie. conceptually/hierarchically organised) words/blocked colour, (2) blocked words/no colour, (3) blocked words/random colour, (4) random words/blocked colour, (5) random words/no colour, (6) random words/random colour.

Subjects were each presented with four hierarchies from one of the six conditions. Each hierarchy was presented for a period of 13 seconds and immediately after all four had been shown successively, subjects were required to recall as many words as they could from all the hierarchies.

The data analysis revealed that recall was aided by the blocked, ie. organised colour conditions, irrespective of whether or not the words were in a blocked or random arrangement ($F2,114 = 17.98$, $p < 0.001$). As recall was higher in the condition with randomly arranged words with the blocked colour than the random word/random colour condition ($t = 2.5$, $df = 38$, $p < 0.05$), it did not appear that the colour had simply emphasised the word conceptual structure. The result that only the blocked word/blocked colour condition subjects recalled more than subjects in the equivalent no-colour condition ($t = 3.42$, $df = 38$, $p < 0.01$) suggested to Elio and Reutener that the colour context of each word was not being used to cue the words. If this had been the case then subjects in the blocked word/random colour condition also should have had significantly higher recall than subjects in the blocked word/no colour condition. Presumably the rationale behind this is that if any word is recalled it would have an association with its background colour. As only four colours were used, the colour evoked by this association also should have an association with another 1/4 of the set words, minus one (ie. the word already recalled). Overall it should be expected that the same number of words would be remembered in blocked and random colour conditions if a cuing process alone was in operation, as the same number of words would be associated in the same manner to the colours. It is also worth noting that over the four trials each subject experienced, the magnitude of the effects reported increased (eg. random word/blocked colour condition vs. random word/random colour condition on the fourth and last trial, $t = 5.75$, $df = 38$, $p < 0.001$).

The explanation of these results forwarded by Elio and Reutener is based on Anderson and Bower's (1972) model of recognition and retrieval processes in free recall and regards the colour background as a list marker. However, as this model will be discussed in chapter five, it is enough to note that these results indicate that colour context can be used as an organisational factor at encoding and recall.

About a decade before the Elio and Reutener study and in common with many room environmental context studies, Gottlieb and Lindauer (1967) reported an investigation into the result of changing or reinstating a colour shape context on retroactive interference (RI). An A-B, A-C paradigm was applied with the stimuli being number-nonsense syllable paired-associates. At the initial or original learning phase (OL), subjects were presented with eight of the paired-associates at a rate of 2:2 seconds, to learn to a criterion of one errorless trial. At OL all the paired-associates were presented on a common set of coloured shapes. Two groups of subjects received the A-C, interpolated learning phase (IL). The paired-associate list presented here was on a different set of coloured shapes to that of OL. Twenty minutes after the end of OL, all three groups in the experiment were tested for recall. The number stimulus was presented on the same coloured shape background as the IL to what was deemed to be the strong RI group, and was presented on the same coloured shape background as original learning to what was deemed to be the weak RI group. The third group; the control group, who had received no IL, recalled with the number stimuli presented on the same coloured shape background as original learning. It was assumed therefore, that the control group should experience little or no RI. The effect of RI as measured by the correct number of responses on the first recall trial was as predicted, with significant differences between the RI strong, RI weak and no RI groups ($F2,57 = 40.89$, $p < 0.001$). The mean number of trials to relearn the lists to criterion revealed a significant difference between the two RI groups and the control group ($F2,57 = 4.35$, $p = 0.017$), but not between the two RI groups. On the second recall test after the first relearning trial, similar results to those already obtained with recall were found. This suggests that cued-recall is a more sensitive measure of environmental context effects than the relearning measure is.

Based on a similar theoretical viewpoint, Lehr and Duncan (1970) used coloured stimuli letters in a study of the spontaneous recovery of prior learning. Subjects were presented with two paired associate lists in an A-B, A-C paradigm. Stimuli items were CVC trigrams, ie. non-sense syllables, presented at a 2:2 second rate and the response items were high frequency adjectives. In the list one priming condition, the first list stimuli items were coloured red, the second list stimuli items black, and the stimuli items were coloured red at recall. In list two priming, the first list stimuli items were presented coloured black, the second list stimuli items red, and at recall the stimuli items were red. The priming control condition had black stimuli items throughout all learning and recall phases. All three of these conditions had recall tested immediately after learning (six minutes after learning list one), and also twenty minutes after learning list one. In a learning control condition subjects learned only list one, which was coloured black. In all conditions the response words were coloured black. Both lists were taken to a learning criterion of eight out of twelve responses before recall began.

A significant increase in recovery was found only in those lists which had been primed. This was indicated by significant interactions (list one priming, $F_{1,76} = 4.67$, $p = 0.034$; list two priming, $F_{1,76} = 7.96$, $p = 0.006$), between the retention intervals and learning conditions (ie. one vs. two lists learned). This result demonstrates that the presentation of a matching stimulus colour at recall (cf. learning) can reduce the amount of information forgotten from immediate to delayed recall. Lehr and Duncan explained the results as being due to the colour reducing the generalised response competition, ie. interference via its association with the responses.

2.2. Physical environment as a function of the subjects' location.

2.2.1. memory for presented stimuli.

Smith and Guthrie; in their 1921 psychology text, mention what must be the first of the physical environment studies using humans as subjects. They informally report two experiments carried out by W.R. Wilson. He found that there was a greater saving in learning when subjects were tested in the same place as they had learned: indoors or outside, than when learning acquired outside was tested inside, or vice versa. In another experiment he found that performance benefited similarly when either the presence or absence of the smell of pepperment was reinstated.

Abernethy (1940) reported a study of the effect of altering classroom and/or instructors at examinations. This vigorously controlled experiment extended over a period of five years, the first three of which were devoted to obtaining various exam tests of equal difficulty. Two student classes were used as subjects (N = 181), with each divided into four groups. Each group then performed four examinations in different controlled orders at different times throughout the four weeks of actual experimentation. Each group therefore was tested at each level of test conditions, ie. same class, same instructors; same class, different instructors; different class, same instructors; different class, different instructors.

Abernethy found that changing instructors did not affect performance as much as changing classroom, and that changing both was associated with the worst performance. She also comments that the best students, ie. those highest in scholastic rank, showed least effect of change, while those lowest in scholastic rank demonstrated the greatest deleterious effect of change from learning to test.

Metzger, Boschee, Haugen and Schnobrich (1979) also carried out another examination of classroom environmental context. Two classes of (25) university students were given weekly tests in their usual classroom on three occasions. On the fourth week, one group of subjects was tested in a different classroom. On the fifth week, the other group of subjects was tested in a different classroom, while the first group was tested in its usual classroom. On the sixth week of the study, both groups were tested in their usual classrooms.

The data analysis took the form of comparing the number of errors made by subjects in each group, on each test. Over the first three tests, no significant differences were observed between groups. In the fourth test however, the change group made significantly more errors than the non-change group, $t(48) = 33.85$. On the next test; when the groups reversed conditions, again the (new) change group made more errors, $t(48) = 2.97$. On the final test; when both groups were tested back in their usual classroom environmental contexts, no significant differences were observed.

Metzger et al. also analysed the subjects' error scores in terms of high and low error groups. An average error score over the first three tests was calculated for each subject. A median split provided two groups. Subsequent to this, a deviation score was calculated by subtracting each subject's average number of errors (over the first three tests) from the number of errors made when they were tested in a changed classroom environmental context. A comparison of the high and low error groups, indicated that the "best" students suffered most from a change in classroom environmental context, $t(48) = 77.22$.

Unfortunately, Metzger et al. do not provide any details of the type of materials and test involved in the study, apart from implying that the tests were on the topic of physical geography. Although the form of analysis carried out is probably appropriate given the type of questions asked by educationalists, the discrepancy in the size of effects suggests that from a theoretical perspective, a rigorous analysis of the data would have produced more interesting results.

Paradigms developed to investigate interference effects were applied to investigations of room type environmental context before they were applied to colour context studies and state dependent studies (eg. Bower, 1981). Bilodeau and Schlosberg (1951) carried out such a study to investigate the effect on retroactive interference (RI). Subjects were presented with three lists, each consisting of ten pairs of disyllabic adjectives to learn in a particular room. Subjects were given eight trials to learn, within which each pair of adjectives was presented for 4 seconds and were separated by an inter-stimulus gap of eight seconds. An interpolated task which could involve doing long division sums, or instead learning a similar three lists, occurred prior to the relearning of the originally presented list. There were three basic experimental conditions. In condition one, subjects received the three separate phases of the experiment in the same room (AAA). Condition two was identical to condition one, except that where subjects in condition one had had a long division interpolated task, subjects in condition two learned the three similar lists. Condition three was as condition two, except that subjects learned the interpolated task lists in a different room (ABA). In all conditions, rooms were counter-balanced. Relearning in the final phase of the experiment involved the anticipation of each response after seeing the stimulus word. The relearning took place eight minutes after the original learning.

Bilodeau and Schlosberg reported significant differences[2] between all three conditions on the first relearning trial. Subjects in condition one could anticipate most words and subjects in condition three performed better than subjects in condition two (ie. ABA > AAA). The results demonstrated that performance on the recall task was best when the interpolated activity is dissimilar to that of original learning, but if the two activities are similar, carrying out the interpolated activity

---

2 Bilodeau & Schlosberg do not provide a detailed account of their statistical analyses in this paper, preferring to refer any interested reader to Bilodeau's thesis.

in a different environment will benefit performance. One other aspect of the phenomenon reported by Bilodeau and Schlosberg was that although females' performance was superior overall, the experimental effects were equal across both sexes.

Greenspoon and Ranyard (1957) repeated the experiment reported by Bilodeau and Schlosberg, but with several alterations and additions. Subjects were presented with two lists of ten non-sense syllables. Each syllable was presented for 3 seconds with three seconds between each stimulus item. Subjects were required to serially learn the presented lists such that they could anticipate each successive item. In all conditions subjects learned the second list in the second phase of the experiment. Therefore, there was no comparison of the effect of type of interpolated activity. Despite this deletion, Greenspoon and Ranyard employed four basic conditions. Subjects could receive all three experimental phases in one room (AAA), they could try to relearn in a different room (AAB), they could learn in a different room to that in which all subsequent phases took place (ABB), or they could receive only the interpolated learning in a different room (ABA). Both original learning and interpolated learning continued to a criterion of two consecutive errorless trials. Relearning of the original list occurred three minutes after original learning and continued until the learning criterion was reattained.

The first part of the data analysis was to check that neither lists nor rooms (which had both been counter-balanced) had differentially affected learning. On the basis of the number of trials to criterion, no differences were found. An analysis of both the number of trials to relearn to criterion ($F_{3,129} = 16.23$, $p < 0.001$) and the number of items correctly anticipated on the first trial ($F_{3,129} = 32.49$, $p < 0.001$), indicated that the best performance was obtained when interpolated learning occurred in a separate room to that of original and relearning, and with the latter two occurring in the same place. By including the two conditions AAB and ABB Greenspoon and Ranyard were able to show that it was not just a change in environment from original learning to interpolated learning or vice versa, but also the reinstatement of the original learning environment which benefited performance.

Dallett and Willcox (1968) replicated part of Greenspoon and Ranyard's experiment to ensure that their environmental change was an effective manipulation before they carried out their own investigation of such a change on proactive interference. In the partial replication, a significant difference was found between the AAA and ABA conditions ($F_{1,20} = 14.44$, $p < 0.001$), but rather than use another room, the second environment was provided by having subjects put their heads into a specially constructed box.

In their investigation of the effect on proactive interference, Dallett and Willcox presented subjects with lists of twenty five high frequency nouns. Each list was presented at a rate of 3 seconds per item. Subjects were required to master the serial recall of the first list presented to them with thirty trials (however, in a five day experiment reported, this criterion was relaxed). After the first list, subjects learned a second list in the same environment. Subjects returned on the next day and in the same environment had to try and recall as many of the words from the two lists as they could, in any order. Another two lists were then presented for the subjects to learn, but this could occur in the same environment or in a different environment. This procedure was repeated each day until the final day, which consisted only of a recall session. Three separate experiments were carried out: a three day experiment, a four day experiment, and a five day experiment. In the three and four day experiments, half of the subjects experienced a change in environment on days two and three respectively, and then continued to learn and recall in that environment. In the five day experiment both sets of subjects experienced a change in environment, but at different times: one group changed environment on day three, and the other on day four.

In all three experiments, separate analyses of the number of trials to the learning criteria indicated that there were no differences between groups which experienced an environmental change and those that did not. In the four and five day experiments, separate analyses of the number of words correctly recalled revealed significant interactions between groups and days ($F_{2,36} = 3.18$, $p = 0.053$: this is incorrectly quoted by Dallett and Willcox as being < 0.05). This indicated that in these situations overall recall is better when the environment is changed than when it remains constant. To account for the lack of such an effect in the three day experiment, Dallett and Wilcox suggest that it may be necessary for proactive interference to build up before a change in environment is effective. The failure to find any such effects with relearning again would suggest that the recall procedure is a more sensitive measure (cf. Gottlieb & Lindauer, 1967).

Also based on an interference model of forgetting is a developmental study of learning and recall in conjunction with a room environment reinstatement/change paradigm carried out by Jensen, Harris and Anderson (1971). Children from six different age groups were presented with a list of eight non-sense syllables to learn. Syllables of ninety per cent meaningfulness were selected to allow the younger children to master the task within a reasonable time. Each syllable was presented for 3 seconds with another three seconds inter-stimulus gap. Each subject was represented with the list until they could make five correct anticipations. Each age group was divided into two basic experimental groups: one group recalled in the same room as they had learned in (AA), and the other group recalled in a different room (AB). After twenty four hours recall was obtained by a relearning by anticipation procedure similar to original learning.

An analysis of covariance, controlling for the number of learning trials required for the different age groups, and using the number of errors made on the first relearning/recall trial as the dependent variable, indicated that across all age groups there was a significant advantage for performance in the same environment ($F_{1,263} = 4.42$, $p = 0.036$). A subsequent analysis found that females' performance was better overall, but that all other effects acted equally across both sexes (cf. Bilodeau & Schlosberg, 1951).

With a rather more extreme manipulation, Godden and Baddeley (1975; 1980) have investigated the effect of reinstating or changing underwater and on-land environments. In each condition, subjects were presented with thirty six unrelated words at a rate of three words per 2 seconds, with four seconds between each set of three. Subjects were tested in all four conditions which involved: learning on land, recalling on land (DD); learning on land, recalling underwater (DW); learning underwater, recalling underwater (WW); and learning underwater, recalling on land (WD). At test, approximately four minutes after the presentation of the last item, subjects wrote; in any order upon a formica board, as many items as they could remember.

An analysis of the correct recall indicated a highly significant benefit for the reinstatement conditions ($F_{1,12} = 22.0$, $p = 0.001$). Wilcoxon tests revealed that while recall on land was better than recall underwater, for those subjects who learned underwater their recall was best underwater.

One possible explanation of these results is that the apparent deleterious effect of an environment change is due to the movement or disruption inherent in this act, ie. going from one environment to another. The lack of such movement in the DD and WW conditions would explain the superior performance of subjects in these conditions. A variant of this explanation is that rehearsal would be differentially disrupted by the activity between presentation and the test. To test both of these hypotheses, Godden and Baddeley ran another experiment involving two DD conditions. In one of these, subjects entered the water between learning and test while in the other condition the procedure was as undisrupted as before. The recalls of the subjects in the two conditions differed only slightly and in the opposite direction to that which would be predicted by either the disruption or differential rehearsal hypotheses.

It would seem then that free recall performance is best in the reinstated conditions because there is a match between the learning and test environments. Godden and Baddeley (1980) also looked at recognition performance under the same conditions. Surprisingly, especially considering the magnitude of the free recall effect, the only difference found with recognition was that words learned underwater were not identified as well as those learned on land, irrespective of where they were tested. The lack of an interaction between learning and test environments was also reflected in the subjects' d' scores.

Mayes, Meudell and Som (1981, expt.2) reported a study which added a temporal factor to the investigation of the effect of environmental context reinstatement/change on the ability to remember. Subjects were presented with two lists each of thirty highly associated word pairs, with each pair presented at a rate of one every 2 seconds. Subjects learned in only one room, but were tested for recall of one list in a different room, while the other list was tested in the same room as learning. Recall was tested after one minute and after one week. A ten minute gap was placed between the learning of the two lists. In the one minute retention interval condition, subjects left the learning environment to go to a waiting area and read a passage from a book. At recall, subjects were required to produce a response to every stimulus item as well as provide a confidence rating of their response correctness.

An analysis of recall with a correction for guessing revealed a large effect of the reinstatement/change manipulation, but only in the one week delayed recall condition ($F_{1,60} = 36.3$, $p < 0.001$). Another analysis of recall, but only using the data from the reinstatement conditions indicated that subjects who learned and were tested in what Mayes et al. termed the "unusual" room forgot at a faster rate than comparable subjects in the "ordinary" room ($F_{1,60} = 4.2$, $p = 0.045$).

The most extensive research into room type environmental context effects to date, has been carried out by S.M. Smith (1979; 1982; 1984; Smith, Glenberg & Bjork, 1978). In the room type environmental context reinstatement/change paradigm, apart from the context match/mismatch variable, two other factors have been mentioned which are generally confounded in the experimental procedure: the disruption involved in moving from one room to another, and the novelty of the new environment.

Godden and Baddeley (1975) and Dulsky (1935) have determined that these factors, although they would exert an influence to bias the results in favour of a reinstatement effect, cannot account for the phenomenon. However, Smith (1979) has argued that because of the way in which such determinations have been made, it is not clear what proportion of the typical difference between the environmental context reinstatement/change conditions is actually due to the matching of test with learning context. To assess this, Smith operated three basic conditions, each involving five separate phases. Subjects learned a list of eighty high frequency nouns presented at a rate of 3 seconds per item. After this presentation they were given a recognition test with ten stimuli and ten filler items. This first phase took ten minutes. Subjects then returned to a waiting area for three minutes. In the third phase subjects entered a different room to the learning room and spent ten minutes drawing it. They then returned to the waiting area for another three minutes. In the final phase of the experiment subjects spent ten minutes free recalling as many of the words that they had been presented with as they could remember. The recall room was determined by experimental condition. Three separate rooms were used. The three main conditions are represented in figure 2.2., with the letters A, B, and C identifying the different rooms.

ABA	ABB	ACB

Figure. 2.2. Room order for the three main conditions in Smith (1979, expt.1).

In the ABA conditions subjects returned to the same room to recall in as they had spent ten minutes in learning the words. In condition ABB subjects recalled in a different room to that in which they had learned, but it was familiar to them as they had spent ten minutes in this room drawing it. Indeed, they may have been more familiar with this room than the learning room. In the ACB condition, subjects not only recalled in a different room to that in which they had learned, but also to that which they had spent ten minutes drawing. Consequently, subjects in this condition would be unfamiliar with their recall room. Equally in all conditions, subjects had to move between rooms. In other words, disruption between learning and recall was held constant across all conditions, and the novelty or familiarity of recall context was varied between conditions. If the novelty of a recall context does exert any influence on subjects' performance, it should be expected to manifest as a difference between conditions ABB and ACB. If the matching of learning and test context exerts an effect, a difference between conditions ABA and ABB should be observed.

Orthogonal analysis of the recall data revealed significant differences between conditions ABA and ABB ($F_{1,27} = 5.78$, $p = 0.023$), but not conditions ABB and ACB ($F_{1,27} = 0.08$, $p = 0.779$). These results demonstrate the benefit to recall of a match between learning and test context, and together with the Godden and Baddeley (1975) experiment and Dulsky's (1935) results and rationale, suggest that this is the only important factor in the reinstatement/change paradigm.

Another aspect of this experiment was that subjects in the counter-balance conditions (eg. for ABA, BAB) received auditorily presented words, while the other subjects received visually presented words. The modality of input was found not to exert any effect of the subjects' ability to recall in any of the conditions.

In the same paper Smith describes another study (expt.2) in which subjects learned the list of words and completed the partial recognition test exactly as before. Twenty four hours later, subjects returned either to the same room or a different room to recall. The majority of subjects were placed in one of four different context recall conditions. Before recalling, some were given pictures of the original learning room. In another condition, the subjects were told to try and imagine the original learning environment. In a third condition, subjects were told to imagine a room at home while they recalled, and in the last of these different conditions subjects were given the usual instructions just to recall as many words as they could.

Smith found that in the former two conditions, the deleterious effect of a change in environment was eliminated, with recall increasing to equal that of the reinstatement recall condition. The instruction to imagine the room that learning occurred in was as effective as presenting the subjects with pictures of the learning room. The only significant recall differences were found between the subjects in the two unaided conditions (normal instructions in different condition, and imagine room at home different condition) and the rest ($F_{4,45} = 5.13$, $p = 0.002$).

In a third experiment Smith investigated the robustness of the ability to self-generate the learning environment. Subjects were presented with five different tasks. Of these tasks only the list learning, which was carried out exactly as before, was a verbal task. List learning always occurred in one particular room, but the other four tasks could take place in one other room, or each in a separate room. Twenty four hours later, recall of the list was tested either in the same room as it had been learned in, or in a new room. Subjects recalling in a new room again were told to imagine the original learning room. The orders of the four conditions are illustrated in figure 2.3., below.

	Conditions			
	1	2	3	4
Task 1	A	A	A	A
List learning	B	B	B	B
Task 3	A	C	A	C
Task 4	A	D	A	D
Task 5	A	E	A	E
Recall	B	B	F	F

Figure. 2.3. Room orders for expt. 3 (from Smith, 1979).

The worst recall performance was found in the condition where six different rooms were used (condition 4 in fig. 2.3.). The other three conditions differed significantly from this condition ($F_{1,36} = 6.35$, $p = 0.016$), but not from each other. This suggests that self-generated learning environment is not as efficient in promoting recall when the number of intervening or experienced environments is increased.

In an attempt to identify the origin of better performance in reinstated environmental conditions, Smith, Glenberg and Bjork (1978, expt.3) presented subjects with eighty words which had to be sorted into ten categories. Category names were provided and indeed the stimuli words were chosen from these ten categories in equal numbers. The next day subjects were taken to either the same, or a different room and were asked to recall in any order as many of the words that they had sorted the day before as they could.

The results showed that more categories ($F_{1,16} = 5.65$, $p = 0.030$), as well as more words per category ($F_{1,16} = 8.93$, $p = 0.009$) were recalled in the same condition. This suggests that whatever influence environmental context exerts, it affects each individual item.

In the fourth experiment reported by Smith et al., a view of memory similar to that advanced by Kintsch (1970) was adopted. On the basis of this model, an attempt was made to assess where the environmental context exerted its effect. Free recall and recognition tasks were examined. It was reasoned that if recognition involves the accessing of a particular meaning of a word before the decision process is operated and if environmental context influences the semantic encoding of a word, one would expect to see poorer identification of words which have more than one possible meaning (eg. balanced homographs, high frequency words) after a change in environmental context. If only the decision process is influenced by environmental context, then all items to be remembered should be affected by a context manipulation. If only the retrieval process is influenced then one would expect the effect of environmental context to be restricted to free recall situations.

Subjects were presented with each of the ninety-six stimuli items for 3 seconds each. Immediately after the completion of the presentation a recognition test consisting of sixteen stimuli and sixteen filler items was presented. The next day subjects returned to either the same or a different room. Half of the subjects in each condition received a free recall test and the other half received a recognition test. Subjects in the recognition test condition had to identify the eighty untested stimuli items from the eighty filler items which were presented in a ten page booklet.

In recall the expected advantage for subjects performing in a reinstated environmental context was found compared to those subjects performing in the changed context condition ($F_{1,24} = 12.79$, $p = 0.002$). In agreement with previous studies (eg. Geis & Winograd, 1975; Kintsch, 1970), high frequency words and polarised homographs were better recalled than their low frequency and balanced homograph counterparts ($F_{1,120} = 45.49$, $p < 0.001$ and $F_{1,120} = 4.22$, $p = 0.042$, respectively). Analysis of the recognition data, in the form of d' scores, revealed no effect of environmental context on the recall of homographs or non-homographs. However, an interaction between word frequency and environmental context in the expected direction was obtained ($F_{1,72} = 4.12$, $p = 0.046$). It was also found that homograph d' scores in the environmental context change condition had significantly greater variation than in the reinstatement condition ($F\text{-max}_{4,40} = 2.77$, $p < 0.05$). The word frequency effect with recognition was also replicated (cf. Winograd & Geis, 1974).

Overall the results demonstrate that the environmental context influences subjects' ability to recall information, but hardly influences their ability to recognise information. Smith et al. suggest that the poorer recognition of high frequency words indicates an influence of environmental context on the accessing of a semantic sense of a word. However, another interpretation of this result is that as high frequency words are more difficult to recognise on the basis of some record of familiarity, so the additional information provided by environmental context reinstatement provides a comparative advantage at retrieval and therefore exerts an effect on recognition.[3]

In the final experiment reported by Smith et al., the semantic variability at test and input was controlled by presenting a cue word along with the nominal stimulus. The purpose of this was to try and determine if environmental context had any influence on processes other

---

3 The background to this interpretation is built up through chapters five (sections 6. and 8.3.), six (section 1.5.), ten (section 6.1.) and eleven (section 1.3.).

than the derivation of a semantic sense. In the terms of the model adopted by Smith et al., differences in recognition decision processes would be reflected. Changing the cue word from presentation to test has been found to adversely influence recognition of high frequency words, but not low frequency words (Reder Anderson & Bjork, 1974). It was assumed that this was due to the greater potential for different meanings available with high frequency words. With the result of cuing so predicted, any effect of environmental context should be due to an influence on the decision process of recognition.

The analysis of the d' scores showed that the environmental context reinstatement/change factor had not exerted any effect. Smith et al. concluded that this meant that the decision process is not influenced by environmental context.

In the second experiment reported in this study Smith et al. looked at the effect of an environmental context manipulation on cued-recall. Two lists of forty-five weakly associated word pairs were created. Fifteen of the stimuli items were common to both lists, but were paired with different responses in list one to what they were in list two. The word pairs were presented aurally at a rate of one pair in 4 seconds, or visually at a rate of one pair in 3 seconds in the two different learning contexts. In the first session subjects were presented with one list four times, after which they were tested for cued-recall on fifteen of the thirty unique pairs. In the second session on the following day, subjects were similarly presented and tested with the other list in the other learning context. On the third day, subjects were tested either in the day one or day two context, or alternatively in a neutral context. The recall test involved the presentation of the fifteen common stimuli pairs and then the sixty unique pairs from the two lists. Prior to

	Day 1	Day 2	Day 3
Env.1	Learn list 1 (or 2) *. Test unique pairs.	*	Test: 15 common and then the 60 unique pairs.
Env.2	*	Learn list 2 (or 1)*. Test unique pairs.	As above.
Env.N			As above.

Figure 2.4. Diagram of Smith, Glenberg and Bjork's (1978) 2nd experiment (the *'s represent counter balance conditions).

beginning the test, subjects were told that some items may require more than one response (see figure 2.4.).

If learning environment reinstatement benefits recall in comparison with a change in environment at test, then Smith et al. claim that subjects tested in their day one environment should recall more of the information presented there than the information presented in the day two environment. Likewise subjects tested in the day two environment should recall more of the information presented there, than that learned in the day one environment. In both situations where recall is tested in a reinstated environment, more information learned in those environments should be recalled than would be if recall was in the other learning or neutral environments.

Analysis revealed a significant main effect of cue type (greater recall with unique cues; $F_{1,18} = 60.6$, $p < 0.001$) and response type (day two learning recalled better than day one learning; $F_{1,18} = 13.2$, $p = 0.002$). However, the pertinent interaction between test context and response type was not significant. Smith et al. claim that this is due to an inflated error term in the neutral conditions. A subsequent analysis excluding this group did reveal a significant interaction between test context and response type ($F_{1,12} = 4.8$, $p = 0.049$). It may have been noticed that this paradigm also controls for disruption and novelty of test environment. However, in this experiment the defined context also involved a time factor. Subjects only used a particular room at a particular time, ie. 8am, 12am, and 4pm. It is very likely that what have been described here as state variables (ie. physiological changes associated with circadian rhythms) contributed to the result.

The last experiments to be reviewed which have been carried out by S.M. Smith involve the effect on recall of varying learning environment. In the first experiment reported by Smith et al., subjects were presented with forty unrelated high frequency nouns at a rate of 2 seconds per word slide, or 3 seconds per word in the auditory condition. The list was then presented again, with subjects being allowed 10 seconds per word so they could rate each one as either "good" or "bad". Three hours later in a second session, subjects were presented with the same stimuli and procedure. However, the second session could occur either in the same environment or in a new environment. Three hours after the second session subjects were taken to yet another environment. There they were asked to write down as many of the words that they had been presented with that they could remember.

The results showed that subjects who had been presented with the words in two different environments could recall significantly more words than those subjects who had been presented with the words in one environment ($F_{1,12} = 50.63$, $p < 0.001$).

Smith (1982, expt.1) carried out a similar type of experiment, but rather than repeat the presentation of the same list in different environments, he had subjects learn different lists in different environments. Four lists of twenty five high frequency nouns were presented aurally at a rate of one every 3 seconds. Subjects had the lists presented in either one room, two rooms (two lists in each room), or four rooms (one list in each room). After each list presentation subjects returned to a waiting area. The time between each list presentation (ILI) was varied across all groups. This could be either thirty seconds or five minutes. In addition the gap between last list presentation and the recall test (RI) could be thirty seconds or five minutes. Subjects first carried out a free recall test and then were presented with a recognition test consisting of all the stimuli items and an equal number of filler items.

Analysis of the recall scores showed a significant effect of the number of rooms ($F_{2,72} = 3.96$, $p = 0.023$). As the number of rooms increased, so did recall. Those groups receiving the five minute gaps between lists also recalled significantly more than subjects receiving the thirty second ILIs ($F_{1,72} = 4.17$, $p = 0.045$). As would be expected, as the RI increased, recall decreased ($F_{1,72} = 3.98$, $p = 0.050$). Although no interactions were significant, the only group which did not obtain any benefit from learning in different rooms was the thirty second ILI and RI group. Clustering in recall (ie. the adjacent recall of some list words) significantly increased as the number of learning environments increased ($F_{2,72} = 9.0$, $p < 0.001$).

The analysis of the recognition data (cast as the number of hits per subject) revealed no significant effects at all. The only groups showing any similarity of trend were those groups with both an ILI and RI of five minutes.

In the second experiment Smith (1982) reports, subjects were divided into three groups. One group received the four lists in four rooms, another group received all the lists in the one room, and the other group received all the lists in one room, but did not move to the waiting area between list presentations. All groups had ILI and RIs set at five minutes and list presentation was as described for the previous experiment. Two tasks were administered in the new test environment. The first was a forced choice recognition task. Pairs of words were presented and subjects had to select which of the two was a previously presented stimulus item. A recognition test was presented again in case the free recall task presented first in the previous experiment had contaminated the results. ILI and RIs of five minutes were used as it seemed from the results of the previous recognition test that if any effect was to be found, it would probably appear under such conditions. The second test was a list differentiation test.

One possible explanation of the increased recall with a larger number of learning rooms is that room changes allow a greater distinction between lists in memory. For certain models (eg. Anderson & Bower, 1972; 1974; see chapter five, section 4.) list differentiation is a prerequisite of efficient memory performance. If differentiation is affected by changes in learning environment, this could account for the superior memory performance. However, the consequence of superior list differentiation should be evident in a recognition test also.

Analysis of the recognition data conformed with that of experiment one. No differences were found between any of the groups. A similar result was obtained with the list differentiation data, although again this test could have been contaminated by the preceding recognition task.

In the final experiment reported by Smith (1982) the effect of learning in a number of rooms in relation to the reinstatement phenomenon was examined. Three lists each of thirty-two high frequency nouns were created. Presentation was as previously described for the other multiple room learning experiments. One or three rooms could be used at learning and in all conditions five minute ILI and RIs were used. Recall could occur in the same (or one of the same) room(s) as learning, or in another neutral environment. As with all the environmental context studies reported, counter-balancing controlled for room specific effects (see figure 2.5.).

	Number of input rooms	
Test context	One	Three
Same	AAA-A BBB-B CCC-C	ABC-A ABC-B ABC-C
Different	AAA-D BBB-D CCC-D	ABC-D ABC-D ABC-D

Figure 2.5. Room orders for list presentations (from Smith, 1980).

The analysis of the subjects' recall scores found the normal reinstatement/change effect with one learning room groups ($F_{1,116}$ = 3.99, p = 0.048), but not with the three learning room groups. To check that the lack of reinstatement in the three room learning condition was not due to only one of the three rooms being reinstated, a comparison of the number of items recalled from each of the three lists was made. In this situation one might expect the list learned in the test context to be best remembered. However it was found that all three lists in this condition were equally recalled. Together these results would suggest that the deleterious effect of recalling in a different environment to that in which learning occurred can be eliminated by learning in several environments.

Smith (1984) has also compared multiple encoding environments with self-generation of encoding environment in terms of their ability to reduce the environmental context dependent effect. All subjects were presented with three, thirty-two high frequency noun lists. Each word was presented auditorily, at a rate of one every 3 seconds. Between each list presentation was a five minute break which subjects spent in a hall. Half of the subjects returned to the same room each time to learn the next list, while the other subjects went into a different room for each list. Five minutes after the presentation of the third list, all subjects were taken to a novel environment to make their attempt to recall the words. Half of the subjects in the one room presentation condition were asked to try and remember the presentation room as an aid to their recall, while the other subjects received no instructions. A similar division was made with respect to the three room presentation subjects, but with the instructed subjects being asked to try and remember all three rooms as an aid to their recall.

The results indicated that the subjects instructed to remember the presentation environments recalled most ($F1,112 = 7.03$, $p < 0.01$). However, the interaction between the instruction and number of presentation rooms factors ($F1,112 = 3.75$, $p = 0.055$) and subsequent Newman-Keuls tests revealed that the former effect was only present for subjects in the one room presentation condition.

Dolinsky and Zabrucky (1983) investigated the environmental context reinstatement effect with "immediate" recall, in terms of the serial position of best remembered items. Subjects wore welders' goggles, equipped with flip opaque shades. The two environments were created by having subjects engage, or not engage the opaque shades. Each subject was auditorily presented with two lists of eighteen (noun) words, presented at a 2 second rate. Immediately after each list's presentation, subjects recalled verbally for forty seconds. Recall was made with either the opaque shades in the same position as at learning, or in a different position to that of learning.

A slight effect of environmental context reinstatement was observed, $F(1,15) = 3.84$, $p = 0.069$. However, analysis of the serial position curve revealed a significant advantage for different environmental context subjects for the first three positions (primacy), $F(1,15) = 6.25$, $p = 0.025$, but an advantage for reinstated environmental context subjects on recall of the items from position four onwards (recency), $F(1,15) = 8.56$, $p = 0.010$.

Nixon and Kanak (1981) also examined the effect of environmental context reinstatement and change on free-recall serial position. However, at presentation they also instructed half of their subjects to attend to the overall environment; explaining that it could help learning. All subjects were visually presented with a list of thirty-five adjectives, at a rate of one every 3 seconds. After spending fifteen seconds moving to and from an adjacent hall, the two minute free recall test was initiated in the reinstated or a changed environment.

An analysis of the subjects' overall recall scores revealed an advantage for those subjects instructed to attend to the encoding environment $F(1,60) = 10.41$, $p = 0.002$, but no environmental context reinstatement effect. However, an analysis of the serial position of recalled items revaled an interaction between serial position and environmental context reinstatement at recall, $F(1,60) = 3.56$, $p = 0.064$. The first five items were best recalled in the changed environment, while the last five items were best recalled in the reinstated environmental context.

As Falkenberg (1972) demonstrated with process context, the two studies by Dolinsky and Zabrucky, and Nixon and Kanak, demonstrate that environmental context reinstatement effects are not restricted to long term memory. Exactly what effect environmental context has on long term memory in these situations is more equivocal.

Eich (1985) also investigated the effect of instructions directing subjects to employ environmental context in their encoding of nominal stimuli items. After presentation of each list word, subjects either generated an image of the (noun) item conjoined with a particular feature of the environmental context, or generated an image of the item in isolation. After reporting each image, subjects gave a rating of its clarity. The rate of presentation of the twenty-four item list was determined by the subjects' speed of image generation and description. Free recall was tested two days later in either the original or a changed environmental context. Immediately after free recall, subjects were presented with a two-alternative, forced-choice recognition test, with the two alternatives being conceptually related nouns.

Apart from finding a general recall advantage for items imagined to be conjoined with the encoding environment, $F(1,60) = 25.49$, $p < 0.001$, Eich also found that an environmental context reinstatement effect was manifest with such conjoined image items, but not with the isolated image items (interaction effect, $F(1,60) = 4.74$, $p = 0.033$). However in contrast, the recognition analysis (probability of hit) demonstrated a general advantage for isolated items, but no environmental context reinstatement effects.

In common with Eich, Devane and Parkman (1978) also instructed half of their subjects to form a mental picture of two objects. The other half of the subjects were told to associate each pair of nouns in the eighteen pairs list, so that when provided with the first, they would be able to respond with the second. The pairs of words were presented visually at a rate of one every 7 seconds to one group of subjects and presumably at a similar rate auditorily, to the other group of subjects. Subjects were given two presentations of the list before exiting to the corridor. On returning to either the same or a different environment, subjects carried out a cued-recall task (visually or auditorily cued), which in all cases was supervised by a different experimenter.

An analysis of cued-recall indicated an advantage for subjects instructed to use imagery, $F(1,72) = 14.22$, $p < 0.001$. An interaction between this factor and recall environmental context was observed also, in that only those subjects instructed to employ imagery manifested an environmental context reinstatement effect, $F(1,72) = 4.98$, $p = 0.029$.

Given the results reported by Eich, the suggestion is that the images constructed at encoding by subjects in the latter experiment employed environmental context features.

2.2.2. memory for elapsed time.

The effect of changing environmental context on the ability to make temporal judgements also has been investigated. To do this, Block (1982) set up a paradigm involving three parts. First, subjects were given some task to do for a set period of time (D1). After this they are given another task which lasted the same time as the first (ie. D1 = D2). Subjects were then tested on their ability to estimate the length of each time period and then on some aspect of the tasks carried out.

In the first experiment reported, the effect of a change in context from D1 to D2 and the effect of making the judgement in the same (or one of the same) context(s) as compared with making the judgement in a different context was examined. In all conditions subjects were disrupted in the same manner between D1, D2 and judgement.

Block discovered that subjects erroneously regarded D1 as being larger than D2, if between the two periods there was no change in context. If there was a change in context, this mistake was not made. The type of context that the judgement was made in did not influence these results. In each time period subjects had rated fifteen occupations on how well suited they were to each other. Each occupation name was presented for 10 seconds. In the test period the thirty occupation names plus five filler occupations were presented on a sheet of paper. Subjects had to identify if the word had appeared in the two presentations, if it was in D1 or D2, and its relative position within its presentation set. The only effect of a change in environment with the task data was that D2 items were less likely to be regarded as D1 items if there was a change in environment between D1 and D2.

In the second experiment reported, Block specifically investigated the effect of disruption. Three conditions were compared: no disruption, disruption only and disruption and environment change from D1 to D2. The tasks given in each period were identical to experiment one.

It was found that subjects' time estimates in the no-disruption condition demonstrated a larger positive error than subjects in the the disruption only condition. In the disruption and context change condition there was no time error. Analysis of the occupation identification test(s) revealed that a change in context at D2 again decreased the chances that a D2 item would be regarded as a D1 item.

In the final experiment Block looked at the relationship between process context (see chapter one, section 4.1.) and environmental context. To create different process contexts, subjects had to determine if a word matched with a description presented along with it. The description could be in terms of structural features, ie. the sort of type set, or it could be in terms of semantic category membership.

Thirty two words, each presented for 5 seconds, were judged in each time period. From D1 to D2 half of the subjects received mixed and then unmixed judgements. In addition to this, from D1 to D2 the environments could be changed or subjects could be disrupted only. One other difference from the method of experiments one and two was that the memory test involved the presentation of four words randomly selected from each quarter of the presented lists and four filler items (making a total of thirty six items).

The duration judgement data analysis revealed a main effect of process context, and an interaction between environmental and process context. Normally (eg. Block & Reed, 1978, expt.2), a positive time order error appears in the mixed-unmixed process context condition, but here the only condition revealing such an effect was the disruption only, mixed-unmixed condition. In the list differentiation task again it was found that a change in environmental context reduces the probability than an item from D1 will be assigned to D2.

3. Context Studies With Non-Human Subjects.

3.1. Introduction.

Although all the experiments reported in the previous sections have been carried out with human subjects, a sizeable literature also exists with respect to context effects with animal subjects. A great deal of this research is related to work on hippocampal function (see chapter six, sections 2.2. to 2.5.), but there are studies which have looked at the influence of context in normal animals.

3.2. Overview.

Watson (1907) found that rats' performance on a maze was influenced by an environmental context alteration. A decade later Carr (1917) reported an extensive investigation into the effect of varying degrees of environmental context alteration at learning and test. He concluded that the animals (rats) adapted to the whole sensory environment. He found their learning of a maze to be best when environmental conditions were kept constant. Relearning or subsequent performances; after criterion attainment, benefited most from a reinstatement of the original learning conditions. Carr also showed that a variety of contextual changes could differentially influence performance.

More recently, Zentall (1970) has shown that a parallel exists between human RI effects with a change of context and those observed in a similar paradigm with rats as subjects. Fewer errors were made when original learning was tested in the same environment when intermediate learning (IL) took place in another environment, than when IL occurred in the reinstated environment. A similar effect with PI was found not to be significant, but this was attributed to masking by a floor effect.

Deweer, Sara and Hars (1980) demonstrated that a ninety second exposure to the background stimuli of the testing room significantly improved the performance of maze running by rats previously trained on the task.

Spear, Smith, Bryan, Gordon, Timmons and Chiszar (1980) also demonstrated the benefit rats obtain from the reinstatement of environmental context and state variables, while Jones, Rana and McGonigle (1980) have argued that rats' (as well as monkeys' and peoples') performance given prior learning is determined by the context within which information is presented. This type of animal performance is comparable to that observed with human semantic context variations.

4. Ineffective Context Manipulations.

4.1. Introduction.

Of course, not all studies have been able to demonstrate an effect of context. Little is known about "process context", but there seems to be a general agreement regarding the consequences and to a lesser extent, the explanation of alterations in semantic context. With respect to state dependency, Eich (1980) has done much to put in order the enigmatic literature, giving a reasonable account of why failures to produce such effects have occurred. In the environmental context literature fewer failures to produce context effects have been reported, but due to the bias or requirements of publication, this does not necessarily mean that a smaller number of studies have been unsuccessful. However, at least six studies have reported no effect of environmental context manipulations. It may be worth considering these studies in relation to those which have obtained such effects, in an attempt to discover why no influence was observed and consequently what the determining conditions are.

4.2. Review.

Farnsworth (1934) carried out an investigation into the practical effect of changing the classroom at test from that normally taught in. Three classes were used, and pairs of subjects from each were matched on grades obtained prior to the pertinent exam. No significant difference was found between those groups tested in a different environment to that in which they were normally taught and those tested in their usual environment.

In a later study, Farnsworth (1937) had subjects practise a pursuit rotor task and a number task as a seventy decibel chord sounded out. One group of subjects practised for one session and another group practised the tasks for two sessions. These groups were then split in two and half of the subjects carried out the tasks again with noise, while the other half of the subjects carried out the tasks without the accompanying noise. Again, Farnsworth found no benefit to those subjects carrying out the tasks in the same noisy conditions, nor did he find any difference in their performance when he stopped the noise.

In a study using a similar type of environmental context as a variable, Pessin (1932) examined subjects' ability to learn four lists each of seven non-sense syllables. Each item was presented for 1.5 seconds and each list was learned to a criterion of one errorless trial. Each subject participated in four conditions determined by the crossing of quiet and loud conditions at learning and relearning. Loud conditions were produced by accompanying the nominal stimuli items with flashing lights and a buzzer. An analysis of the number of relearning trials to obtain one errorless trial and the number of errors made in doing so showed no significant benefit for the learning reinstated conditions.

Nagge (1935) carried out an experiment where an interpolated list of twelve non-sense syllables was learned in a different room to that of original learning and relearning. Each item was presented at a rate of approximately one every 3 seconds and each list was learned to a criterion of one errorless trial. Analyses of savings in relearning and anticipation scores revealed no difference between those subjects learning the interpolated list in the same room and those learning the interpolated list in a different room to that in which original learning and relearning took place.

Strand (1970) reported a study where the disruption, which is normally part of the environmental change (but not the reinstatement) condition, was included as a separate condition. Using a paradigm designed to examine the influence of RI, she presented subjects with one word list in the control conditions and two lists in the experimental conditions. Each list consisted of twenty words each of which was presented for 1.5 seconds. List presentation was repeated until subjects could correctly recall twelve words. Three conditions (crossed with experimental and control manipulations) were distinguished by the activity engaged in by subjects between learning list one and two. In the no change (N) condition, subjects remained in the one environment throughout. In the disrupted (D) condition subjects moved out of the learning environment, but returned to it to learn list two. Only in the change (C) condition did subjects move into a different environment to learn list two. In all conditions the final recall test was carried out in the list one learning environment.

A comparison of the number of trials to criterion revealed no significant differences between the groups. A comparison of control and experimental groups showed a significant effect of RI ($F_{1,75} = 80.54$, $p < 0.01$). However, although there was superior recall in the D and C conditions compared to the N condition, the D and C conditions did not differ significantly from each other.

On the basis of these results Strand claims that the previous findings of a reduction in RI with environmental context change has nothing to do with the change in context. Instead Strand argues that the apparent memory decrement is a consequence of the disruption involved in moving from one environment to another in the different test context condition.

Of all the failures to obtain effects of environmental context, the result of the experiment carried out by Strand is of most concern. If Strand's conclusion that disruption is responsible for the decrements in memory rather than any match/mis-match between presentation and test environments was accepted, then the whole theoretical perspective from which these effects have been viewed would have to be altered. Rather than reflecting encoding and retrieval interactions in memory, such effects would be more likely to demonstrate the interference of "non-memory" processes on memory retrieval.

It is for this reason that both Godden and Baddeley (1975) and Smith (1979) have attempted to demonstrate that the environmental context effects observed in their experiments could not be attributed to the disruption of moving from one environment to another in the different context condition. Although Strand's finding of an effect of disruption has been used to question the validity of environmental context reinstatement effects, it is worth pointing out that Godden and Baddeley could not replicate this result, even with the degree of disruption involved in moving from land to underwater and vice versa. However, a more detailed discussion of Strand's experiment will have to be postponed until further evidence, regarding those factors which may determine environmental context effects, has been presented and assessed (see sections 5. and 6.2.).

Fernandez and Glenberg (1985) presented the results of a series of eight experiments. After the first experiment the series took the form of an attempt to determine the cognitive factors responsible for the lack of an environmental context reinstatement effect.

The basic learning paradigm employed by Fernandez and Glenberg was to have subjects construct (write) a sentence relating two words, which were presented at a subject paced rate on punch cards. The responses of subjects were collected in specially prepared booklets.

In the first experiment the familiarity of words was manipulated by altering the number of presentations. Each subject was tested after one, or seven days and performed; in order, free recall and then alternately, cued-recall and recognition tasks on all the presented items in the original, or a changed environmental context. In all measures, retention interval and the number of presentations exerted a significant effect, but no effect of environmental context was observed.

In the second experiment, the common presence of the experimenter at presentation and test was eliminated. Two experimenters re-ran the former experiment (one experimenter taking subjects through the presentation phase and the other experimenter carrying out the test phase), but with a sub-set of the stimuli items previously employed and requiring only free-recall after a presentation-test gap of five minutes. Again, no effect of environmental context was observed.

Experiment three examined the possibility that the processing demands of the sentence construction task inhibited the encoding of environmental context information. Consequently, task difficulty was manipulated by providing subjects with associated word pairs or with randomly assigned word pairs, to construct sentences with. However, after a retention interval of one day, the only significant effect obtained was that of task difficulty.

Experiment four tested the hypothesis that the absence of the environmental context reinstatement effect was due to subjects employing non-environmental context information as retrieval cues. To test this, subjects were provided with a sentence within which the two to-be-remembered words were present, or were provided with the sentence generation task as before. It was reasoned that the use of environmental context information as a retrieval cue would be facilitated by the former method of presentation. But again after a retention interval of one day, the only effect was that of the newly introduced factor: type of orienting task.

Experiment five demonstrated a relationship between serial position and the environmental context reinstatement effect, but rather than as Nixon and Kanak, and Dolinsky and Zabrucky had found, Fernandez and Glenberg observed an advantage to same environmental context recall in both primacy, t(90) = 2.67, and recency, t(90) = 2.02, positions. However, the difference in results is probably a consequence of the longer retention interval (five minutes) employed by Fernandez and Glenberg.

Experiment six attempted to replicate the results of experiment five and determine if the environmental context effect is associated with the beginning and end of a session, or a list. Subjects were presented with two lists of randomly paired nouns as had been used in the previous experiment, separated by two minutes of arithmetic. As in all previous experiments, the subjects task was to produce a sentence relating each word pair. However, no significant (main or interaction) effect of environmental context was obtained, although a significant interaction between lists and serial position was observed.

Experiment seven investigated the possibility that the manifestation of an environmental context effect was being inhibited by the generation of sentences. Rather than having to create a sentence with the word pairs, subjects were asked to provide judgements regarding the size of the two objects referred to by the nouns. This task also was self-paced. Five minutes later, subjects performed a free-recall task. Again no significant effects of environmental context were observed.

After such inability to obtain effects of environmental context, Fernandez and Glenberg attempted to replicate the methodology employed by Smith (1979, expt.1). The only differences between experiments concerned the use of two different environments cf. Smith, and the exclusion of recalled filler and recognition items from the analysis of recall scores. In addition, the serial position recall was also analysed. An overall recall advantage for reinstated conditions was found, $F(1,46) = 3.31$, $p = 0.075$, but was reported as not meeting the criteria of significance. An analysis of items in terms of their serial position also failed to obtain significant results.

5. An Assessment Of Factors Influencing Environmental Context Effects.

5.1. Introduction.

As can be seen from the presentation of the two sets of studies: those which have succeeded in obtaining an effect of environmental context and those which have not, quite a variety of methods have been employed in the attempt to investigate such phenomena. Essentially these different methods are functions of the variation of a set of factors which are; the nominal type of environmental context, the degree of difference between environmental contexts, stimulus presentation time, presentation-test gap and the nominal task requirement at presentation and test. In addition, one other variable has emerged, namely the degree of learning of the nominal task. The question which arises next is, what influence do each of these factors have in the production of an environmental context effect?

Fortunately for the purposes of assessment several studies have included comparisons between levels of some of these factors. However, to a large extent comparisons will have to be made between experiments differing on more than one factor. Obviously this reduces the importance that can be attached to any conclusion drawn in such a situation, but as with most reviews this is as much as can be achieved prior to a formal experimental test of the ideas obtained from such an analysis.

5.2. Nominal type of environmental context.

At the beginning of this chapter a distinction was made between environmental context which was presented as part of the stimulus materials and the room type of environmental context. Consideration of the environmental context effects obtained when these two types of environmental context are manipulated suggests that this is an influential factor.

One reason for the greater effects obtained when the environmental context manipulated is a function of the stimulus materials could be that subjects consider this environmental context as more important or significant than room type environmental context. Every activity in life takes place in some sort of environmental context, so unless there is something special or peculiar about the place, it is unlikely to be regarded with much importance. However, when an experimenter presents stimuli words on a variety of different coloured cards it quickly prompts consideration of why different coloured backgrounds are being used.

Implicit on the part of the experimenter is a similar type of assumption. This is evident from the way in which black (or white) backgrounds are often regarded as constituting the neutral stimuli condition.

If the saliency of an environmental context is a determining factor in the size of such an effect, then one would expect studies with room type environmental context that emphasise the use of such environments also to obtain larger effects. Certainly from those studies using many rooms (eg. Smith, 1982, expt.1) and those employing distinctive environments (eg. Godden & Baddeley, 1975; Mayes, Meudell & Som, 1981) there would seem to be some support for this notion.

Of course it is naive to think that any type of environmental context difference would give rise to an effect. This is demonstrated by two of the "non-effect" studies which employed noise as a means of creating an alternative environmental context. Pessin (1932) concluded that the reason he had failed to detect an effect of the loud versus quiet manipulation was because subjects in the loud (and presumably distracting) conditions were increasing their work rate to maintain their performance level. Dulsky (1935) also supported this view, claiming that if these motivational factors had been constant across all conditions, the reinstatement/change effect would have been observed. Farnsworth's (1937) experiment could be criticised in a similar manner to Pessin's study. Farnsworth himself suggests that

> ...a true decrease in [the] score due to the change in Nebenreize was masked by an increase in score due to the elimination of a major source of distraction. (Farnsworth, 1937, p.278)

In both cases it would certainly seem sensible to try and block out such stimuli when one was performing a task. If this was done, it is unlikely that a reinstatement of these conditions would aid performance, as there would be little match with this information in memory.

It is not enough to say that at test a change from the learning environment will be to the detriment of the performance of some task. Consideration must be given to the nature of the task and the nature of the environmental context, and to any possible effect that the latter may have on the former.

5.3. Degree of difference between environmental contexts.

It was not just by coincidence that the same paradigm was used by Nagge (1935), Bilodeau and Schlosberg (1951) and Greenspoon and Ranyard (1957). Bilodeau and Schlosberg's rationale for carrying out their study was to determine if with a greater degree of environmental change, an environmental context effect could be established. Greenspoon and Ranyard attempted to replicate Bilodeau and Schlosberg's study to reduce the dubiety of the situation. The two later successes would seem to indicate that the failure to observe any influence of environmental context in the Nagge experiment, was due to an insufficient separation between the two levels of the independent variable.

One of the early studies of the influence of environmental context by Carr (1917) showed that varying the degree of environmental context reinstatement affected the level of performance of rats. More recently and with human subjects, Godden and Baddeley (1975) obtained an extremely large environmental context effect with what must be one of the largest differences between environmental contexts in the present literature. However, despite the evidence (and presumably most theoretical conceptions of the phenomena) which would suggest a relationship between the degree of difference between environmental contexts and the size of effects observed, very little attention seems to have been paid to this. One particular indication of this is the brevity with which most writers report the environmental contexts that have been used in the study.

These experiments serve to illustrate the point that if environmental context effects are to be observed, it must be ensured that significantly different environments are employed. Unfortunately, a problem arises with respect to the a priori determination of effectively different environments. This issue is considered in chapter three, where the adoption of a practical approach to resolving the problem is described.

5.4. Stimuli presentation time.

Weiss and Margolius (1954) reported data which indicated that the most effective response cue was not the stimulus word in their paired associate paradigm, but the colour background on which the stimuli words were presented on. They also happened to present this information for the longest period of time. The stimulus word on its background colour was presented first for 3 seconds. After this, stimulus word and response word were presented for another 3 seconds. So the stimuli word and colour were presented for a total of 6 seconds and for half of this time the response word was also present.

Overall it seems that the shortest presentation time that has been used in an experiment that has obtained an environmental context effect is approximately 2 seconds per item. However the determination of some minimum stimulus presentation time that is associated with the observation of environmental context effects is complicated by the rather lax reporting of inter-stimulus gaps. If a stimulus is presented for 3 seconds and there is a 2 second gap before the next stimulus appears, the subjects have effectively 5 seconds presentation time. The stimuli items are not going to become unavailable to the subject as soon as they disappear from sight. Rundus (1971) has provided evidence suggesting that the limit to rehearsal in list learning is not the time available to the subject, but rather the number of items which have to be rehearsed in that time. The real consideration is not how long the stimuli are presented for, but what this time allows the subject to do with the stimuli items.

Although the contrast between the Weiss and Margolius, and Strand studies suggests that there is some relationship between presentation time and the magnitude of environmental context effects, the previous discussion mentions some of the difficulties in drawing such a simple conclusion. However, given that psychological processing does have a temporal dimension, it would seem plausible to assume that one of the basic determining conditions of the effect is adequate time for the appropriate processes to be initiated and completed.

5.5. Presentation-test gap.

Most of the studies reported in sections 2.1. and 2.2. use different presentation-tests gaps. Fortunately however, some experiments have specifically examined the effect of varying the presentation-test gap on the influence of environmental context.

Mayes, Meudell and Som, (1981, expt.2) reported what is probably the most simple and direct examination of this. They found that subjects were unaffected by a change in environment if recall was after one minute, but demonstrated the typical environmental context reinstatement effect if they were tested after one week. Although the magnitude of this effect may be supported by the within subjects design and a high degree of inter-environmental difference, it is still very impressive.

Smith (1982) has also investigated the influence of presentation-test gap in relation to environmental context effects, but rather than placing this within the context of a reinstatement paradigm, it was examined in relation to the effect of multiple learning environments. Smith found that the benefit of learning in several rooms became manifest with a presentation-test gap of five minutes rather than one of thirty seconds.

At this point it is also worth mentioning that Falkenberg (1972) obtained process context reinstatement effects with presentation-test gaps as low as 5.4 seconds.

5.6. Nominal task.

Theoretically the nominal task can be divided into encoding and retrieval phases. These phases can be considered as independent to the extent that the type of encoding does not determine the type of remembrance that may be requested of the subject. In most cases the type of presentation encourages a particular type of encoding which is normally appropriate for the type of retrieval subsequently required of the subject. For example, the ability to remember paired-associates is tested by a cued-recall procedure.

For those investigating environmental context as a function of the stimulus materials, the paired-associate paradigm has certainly been the favourite. All but Elio and Reutener have used this method and all including Elio and Reutener have obtained effects of environmental context.

In the case of place type environmental context investigations, paired-associate paradigms have not been so common. However, Bilodeau and Schlosberg have reported an environmental context effect using this method, as has Smith et al. (1978, expt.2). Unfortunately, the Bilodeau and Schlosberg statistical data could not be obtained, but the size of the reinstatement effect observed by Smith et al. with cued recall, is much smaller than that normally found with free-recall paradigms (eg. Smith et al., expts.1,3,& 4).

Smith (1982, et al., 1978) and Godden and Baddeley (1980) have all failed to obtain environmental context effects with recognition procedures, while they have obtained these effects if subjects were asked to free-recall. Similarly, both Dallett and Willcox (1968), and Gottlieb and Lindauer (1967) have reported failures to detect environmental context effects with relearning measures, but have been able to detect the effects with serial learning free-recall and cued-recall, respectively.

Eich (1985) and Nixon and Kanak (1981) provided one group of subjects with instructions directing them to make use of environmental context information. Although Nixon and Kanak did not find that these instructions significantly interacted with the environmental context reinstatement effect, they did find that subjects so instructed recalled more than other subjects. Eich on the other hand not only found a recall advantage for subjects receiving conjoined environment imagery instructions, but also an environmental context reinstatement effect only with these subjects.

It would seem then that not only is the type of memory test a determining factor in the manifestation of environmental context effects, but so too is the form of encoding carried out as the nominal stimuli are presented.

5.7. Degree of learning of the nominal stimuli.

The last factor to be mentioned which seems to influence the observation of environmental context effects is the degree of learning. Pan (1926), in his study of the paired-associate learning of face-name pairings presented on pictures of well known Chicago places, found that those items most rehearsed, and so presumably better learned, were less affected by a change in context (ie. the Chicago place). However, Reed

(1931) refers to Oberschelp who was unable to obtain any relationship between the paired-associates' degree of learning and the influence of context. Unfortunately, the description of Oberschelp's study is very brief. No information regarding the type of contexts used is provided, nor is it made clear if any context effect was observed.

Abernethy (1940) on the other hand, did notice a similar relationship to that observed by Pan. She found that a change in classroom and/or tutor affected those students highest in scholastic rank least; the assumption being that the highest in scholastic rank had achieved a greater degree of learning.

6. General Assessment Of The Influence Of Context.

6.1. Across the different types of context.

After having looked at the different types of context in this and the previous chapter and despite some fairly obvious gaps in the information available, it is possible to make some attempt at cataloguing and comparing the types of effects observed.

One of the basic findings across all types of context is that if the context prevailing at presentation or learning is reinstated, memory performance on the task is better compared to that observed when the context is altered. It is possible that the degree of difference between contexts is proportional to the detriment in memory performance. Alternatively it may be that some threshold degree of difference is required in order that a reinstatement effect will be observed.

Both semantic and environmental context variation between item set presentations at encoding improves subsequent memory performance. The case may be the same for state and process context, but no examples or refutations are known of.

Differences appear to arise with respect to the types of memory test which detect the reinstatement/change phenomenon. With semantic context the effect can be detected with free recall, cued-recall and recognition paradigms; with state and environmental context the effect can be detected by free recall, less well with cued-recall, and apparently not by recognition, while process context seems only to have been investigated with and detected by free recall paradigms.

A great similarity exists between the effects of environmental context and physiological state, in that the effects of both are not easily detected by tasks which provide the subjects with cues. The more complete the cues are, the less likely it is for an effect to be detected (see Eich's analysis of state-dependent effects in chapter one, section 4.3.).

It seems that the form of encoding may also exert an important influence as to whether an environmental context effect will be observed. This is suggested by those studies which have considered the degree of learning and the effect of instructions to utilise environmental context information at encoding.

6.2. In relation to Strand, Farnsworth and Fernandez and Glenbergs' non-effects.

Now that those factors which seem to influence the appearance of environmental context effects have been discussed and briefly considered in relation to the other types of context effects, the possible reasons for the failure of Strand and Farnsworth's (1934) studies to obtain effects of environmental context can be returned to.

There are several points worth mentioning in relation to Strand's experiment. First of all, Strand did obtain at least one effect of context: the consistency of list two recall order was significantly increased with context change ($F_{2,51} = 3.31$, $p < 0.05$). Context did influence subjects' memory, but it would seem to require a sensitive measure to detect it. Secondly, the study reported by Strand is actually an abstraction of a larger study involving the presentation of categories. Each list contained two categories with ten exemplars of each. Strand claims that as subjects were not informed of the categories, the list can be considered to be unrelated items. This opinion may have seemed valid at the time, but with the results of Smith (1979, expt.1) and the apparent concordance between the cue-dependent nature of physiological state effects and environmental context effects, a shadow of doubt must now be cast over this.

In Farnsworth's (1934) study, although a matched pair control was applied, the experiment was not as carefully constructed as that of Abernethy. Smith (1982) has shown that learning in several environments can reduce the deleterious effect of a context change. Presumably study for the exam would not be confined to the teaching environment, and so this also would reduce the magnitude of any observed effect. However, Farnsworth did find a non-significant advantage for the reinstated conditions in two of the three class comparisons he made (D/6D = 0.54 and 0.84). It is perhaps more surprising that this and Abernethy's results should have conformed at all to the reinstatement pattern, when the other potentially confounding influences existing within these studies are considered.

One of the most obvious differences between the first seven Fernandez and Glenberg experiments and other studies of environmental context effects is the type of task subjects were presented with. Both sentence construction and size judgements (on the basis of visual imagery) would seem to require a style of processing involving the elaboration of the words beyond that expected as a consequence of the casual learning that takes place in most list memorisation tasks.

The suggestion that the cause of the lack of effect is the different sorts of task presented by Fernandez and Glenberg is given further credibility when it is considered that the replication of Smith's (1979, expt.1) carried out by Fernandez and Glenberg did obtain a significant difference between the two recall environmental context conditions, in favour of the reinstated environmental context condition. Although Fernandez and Glenberg reported the F-value obtained as not significant, the probability associated with the value was 0.075. Most other researchers would regard this as a result that at least suggests that some difference exists between groups, even if it has not fully achieved the accepted, yet arbitrary criterion of 0.05. However, even with such a criterion set, it can be demonstrated that Fernandez and Glenberg have actually obtained a significant result. They may have noticed this if they had used a t-test instead of an F-test. The classical probabilities associated with the traditional ANOVA F-values are two-tailed (Keppel, 1982). Therefore, if a directional hypothesis is being tested, these probabilities are not strictly applicable. As with one tailed t-tests, the correct probability is actually half that obtained from the two-tailed tables. In this case, as the superiority of reinstated environmental context conditions obviously is predicted, the difference between means observed is significant at the 0.038 level.

6.3. Different or similar contexts?

After a review of the context literature, Hewitt (1977) suggested that there were two types of context; intrinsic and extrinsic. Intrinsic context was that information which was always represented with the nominal stimulus item, but was not an integral part of its meaning, eg. environmental and state. However, despite the greater similarity between environmental and state effects, there is still a good degree of similarity between all the types of context effects. The question is whether or not there is enough difference between the different context effects; particularly between environmental and state on one hand and semantic and process on the other, to justify making a distinction between them. Presumably, any such distinction would have the consequence of implying some difference in psychological origin. As mentioned in chapter one (section 3.2.), the important theoretical issue is whether or not there is an equivalence in the psychological function of the different types of context.

7. The Oddity Of Physical Context Effects.

7.1. Introduction.

There is a good deal of evidence to support the claim that physiological state and environmental context influence memory performance. Of the two, that physiological state should exert an influence is more intuitively obvious. This intuition predicts that such effects should be either improvements, or decrements in performance. However, these simple predictions are not in accord with the type of interactive effects which are actually observed.

It was mentioned earlier that despite the implicit assumption of a relationship between physical context and memory, it still had an air of oddity about it. Even though the assumed relationships do have an impressive amount of empirical support, it does not stop the ability of physical context to influence memory from seeming odd. Yet it does not seem odd that semantic context; nor physiological state in the simple manner described above, should be able to influence memory.

7.2. Probable cause.

With respect to psychological function, people tend to have a dualistic attitude. Within this there is room for certain physiological changes and semantic alterations to influence memory. The fact that a particular change in a person's or an animal's physiology can affect memory, supports a mechanistic view of brain operation which is distinct from the operation of mind. However, the particular type of state dependent effects reported here disturb this view as they do not conform to the normal concept of an improvement or a deterioration in performance. The fact that semantic context influences memory confirms the view that our minds operate abstractly, imposing meaning or extracting meaning placed there by another mind. The effect of semantic context supports the view that our behaviour is controlled at a satisfying, anthropomorphically abstract level.

The oddity of physical context effects may come from their apparent contradiction of the way in which we believe our behaviour to be controlled. The suggestion is that perhaps such things as shapes, colours and degree of light could influence our intellectual functioning. The logical conclusion of this line of reasoning seems too extreme: it appears to state that the very aspects of our behaviour that were taken to indicate the existence of an anthropomorphic mind, are to a large extent outwith our control and instead depend on a variety of incidental and supposedly insignificant stimuli. This would seem to have repercussions for the dualistic concept of mind and the notion of free-will, by implying a comparatively simple mechanistic account of such intellectual functions.

7.3. Resolution.

Hopefully this thesis will show that this logic is flawed and that such a conclusion is false. Non-semantic context and specifically environmental context do influence psychological processes. This influence however, is not arbitrary. It is the product of a system which utilises information in a practical and effective manner. The apparent oddity of physical context effects is one consequence of applying inappropriate terms of conception in the attempt to understand the use and subsequent influence of context in psychological processing.

CHAPTER THREE

EXPERIMENT ONE

THE DEMONSTRATION OF AN ENVIRONMENTAL CONTEXT EFFECT

1. Introduction.

1.1. Determining the amount of environmental difference.

In the same way that it is difficult to specify with any great deal of precision degrees of semantic difference, it is also difficult to specify degrees of environmental context difference. Exceptions to this are of course situations in which simple environmental contexts are used, such as background colours. However, when the whole room or surroundings are defined as the environmental context, the determination of degrees of difference becomes an intuitive rather than objective exercise.

Research such as that by Baroni, Peron and Salmaso (1980), Brewer and Treyens (1981), Salmaso, Baroni, Job and Peron (1983) and Tversky and Hemenway (1983) is beginning to investigate the way in which such context information is encoded, represented and retrieved. Eventually this should reveal what the important aspects of defined environments are and as a consequence should allow some determination of the degree of difference between environments.

An alternative approach in determining environmental difference would be to assume that if no environmental context effects could be discerned in a memory experiment, then the degree of difference between the environmental contexts is not significant. However, it does not take a great deal of thought to appreciate the flaw in this reasoning. Far too many factors could intervene to produce a "non-effect"; even with a significant difference between environments. Relying on the results of a typical memory experiment would cause the determination of a significant environmental difference to depend on the effect of all the factors that influence memory phenomena; from encoding variations to retrieval strategies. The present level of understanding of the way in which environmental context information is utilised by psychological processes in any particular encoding and retrieval mode does not inspire confidence in the assertion that no effect means no significant environmental difference.

The first requirement to increase our knowledge of the way in which psychological processes utilise environmental context information is the creation of a situation in which an environmental context effect can be replicated and then manipulated. To achieve this, effectively different environmental contexts have to be employed.

1.2. Establishing different environmental contexts.

As the first step in this investigation of environmental context effects, radically different room environments were prepared. The differences between the rooms were based on a combination of logic and intuition. For example, it seemed reasonable that if one was a long, thin, bright room with a window at one end, this would be perceived as quite different to a short, fat, dark room without any windows. However, the problem of specifying the degree of difference still remains. To cope with this problem rather than solve it, detailed descriptions of all the environmental contexts are presented in the appendix (appendix A). It may be that at some point psychological knowledge will be such that these descriptions will be able to be used to provide an estimate of the perceived degrees of difference.

1.3. Experimental paradigm.

The first experiment had as its prime purpose the demonstration that these environments were sufficiently different to give rise to an environmental context effect. To achieve this, one of the simplest and most elegant paradigms constructed to isolate and investigate the environmental context reinstatement effect; reported by Smith (1979, expt.1), was adapted and employed.

As was described in the previous chapter, the experimental session was divided into three parts; a presentation period, an environmental familiarisation period and a recall period. However, in the experiment to be reported, during the environmental context familiarisation period subjects carried out a non-verbal matching task, rather than drawing the room they were placed in. Subjects in the presentation environment do not have their attention so specifically focused on their surroundings as they do when they draw the intermediate environment. By presenting another task at this point it was assumed that the subsequent acquaintance with the intermediate environment would be on a par with that of the presentation environment.

Another change from the procedure described by Smith was the reduction of the between room waiting time from 3 minutes to 30 seconds. The reason for this was to reduce the likelihood of subjects dwelling on the task and rehearsing items during this period.

The only other significant procedural change in the experiment was in the scoring of subjects' recalls. Whereas Smith scored all items correctly recalled from both the stimuli list and the recognition test, here only those correctly recalled items which had not appeared in the recognition test were scored. The reason for this was that for the recognition and filler items, presentation times and the type of psychological processing were liable to be much more varied.

Smith in common with other researchers included in the design of the experiment a counter-balance to control for room specific effects. This aspect of the design was maintained, but rather than just controlling by counter-balancing presentation and recall environments, the number of subjects in each sub-condition was increased to allow a proper analysis for such effects.

1.4. Experimental hypotheses.

The consequences of these alterations to Smith's original paradigm will be considered later in the discussion of the experiment. For the moment it may be worth reiterating the experimental hypotheses. If the experimental procedure and environments are effective and the match between presentation and test environments is an influential variable, there should be significantly higher recall in the same and familiar (ABA-BAB) conditions compared to any other conditions. If the familiarity of the test environment exerts an influence, there should be significantly greater recall in the different, but familiar (ABB-BAA) condition when compared to the different and unfamiliar (ACB-BCA) condition. If previously, the relative familiarity of the test environment falsely had made it seem as though an environmental context effect had occurred, only the latter hypothesis should be supported. If the imbalance in disruption between conditions was responsible for any effect; as it is now equal across all conditions, no effect should be evident. If any one environment is more conducive to encoding, a room effect should be observed, while an interaction between the two factors would indicate some more complex relationship between specific encoding and recall environments.

2. Method.

2.1. Subjects:

48 Glasgow University students; males and females of approximately equal numbers, participated as subjects. All subjects were unpaid volunteers, naive regarding the purpose of the experiment.

2.2. Environmental contexts:

Rooms A, B and C were employed in the experiment. Descriptions of these rooms can be found in appendix A.

2.3. Stimuli:

80 high frequency (> 50 occurrences per million words), three to six letter logically unrelated nouns were selected from the Kucera and Francis (1967) word count.[1] 70 of these items were presented on slides. Each slide consisted of one centrally located word printed in upper case. Two sets of slides were constructed: black print on a clear background and clear print on a black background. The latter slides were used in room B to avoid extra illumination. The remaining 10 words selected from the word count and 10 of the words

---

1 A full listing of all stimuli and filler words is presented in appendix B.

prepared on slides (both randomly selected) constituted the recognition test. These 20 items; randomly arranged in the form of a single column of upper case words, were presented to subjects on a slip of paper.

2.4. Apparatus:

In both presentation environments (rooms A and B) the slides were projected onto white walls by Kodak carousel projectors. Presentation rate was controlled by a Kodak electronic interval timer.

2.5. Design:

A two factor (3 x 2) between subjects design was applied. The first factor was defined by the relationship between the presentation environment and the recall environment; same or different, and the familiarity of the recall environment; familiar or unfamiliar. The second factor was defined by the presentation room environments; A or B.

2.6. Procedure:

Each subject's 31 minute experimental session was divided into three 10 minute periods, separated by two 30 second waiting periods. In the first 10 minutes subjects viewed an automatic presentation of the 70 word slides. Each slide was presented for approximately

4 seconds. 3 seconds exactly would have been a preferable presentation time, but unfortunately the lowest setting on the interval timer was approximately 4 seconds. As a result of the projector mechanics, there was a variable 1 to 1.5 seconds gap between slides as they changed. This break was quite compatible with the presentation of the stimuli. A similar break between stimuli items would have been provided if the projector had not been so helpful. Subjects were told of the slide presentation times and that their task was to try and memorise all the words they would be presented with, but not necessarily in the same order. They were then instructed to turn on the projector using the switch beside them when the experimenter left the room, and to switch off the projector after the last word slide had been presented and black blank slides began to be shown. After the presentation was complete, the subject informed the experimenter who then administered the recognition task. Subjects were given unlimited time to mark those words they recognised, but no subject took more than 2 minutes. When they had completed the recognition task the subjects returned to the waiting area before moving to the next scheduled room. In this environment subjects were presented with a visual, non-verbal matching task. After completing this task subjects again returned to the waiting area for 30 seconds. In the third period, subjects entered the appropriate environment and were asked to try and write down all the words they had been presented with in the first period, including all those making up the recognition test, that they could remember. At the end of the 10 minute period the experimenter returned, terminated the

experimental session and debriefed the subject. The room order for those subjects in the SAME and FAMILIAR condition was ABA/BAB, the room order for the DIFFERENT and FAMILIAR condition was ABB/BAA and the room order for the DIFFERENT and UNFAMILIAR condition was ACB/BCA.

2.7. Scoring:

Each subject's written recall of the word list was examined for the presence of presented items. Mis-spelling was ignored, but the extremely small number of synonyms/approximations produced by some subjects were regarded as incorrect. Three types of item could be recalled; filler (F) items from the recognition test, recognition (R) items from this list and presented (P) items which did not appear in the recognition test. Separate counts of each type of item were made.

3. Results.

3.1. P-item ANOVA.

3.1.1. preliminary analysis.

The familiar analysis of variance (ANOVA) actually consists of two components: a linear model upon which variance is partitioned and a F-test of the variances of interest. Both of these component procedures make certain assumptions regarding the nature of the data presented to them (Kirk, 1982). Most, if not all of these assumptions are met by the application of the normal requirements of experimental design and by the methods of calculation employed. However, some assumptions made of the data are generally outside the experimenter's power of control. Specifically, these assumptions concern the normality and variance homogeniety of each data set.

The F-test assumes that the two variance estimates come from independent normal distributions, while the ANOVA model assumes that the within group error terms are normally and independently distributed with common variance and a mean equal to zero. As the ANOVA model also assumes that the only source of variation within each group is the error terms, the normality assumption for both components is effectively equivalent.

Norton (1952) found that when deviations from normality are heterogeneous across groups, a 2 to 3 point overestimation of the 5 percent significance level could occur ie. when alpha = 0.05 the real value could be as high as 0.08. However, it is likely that this distortion would have been less if sample sizes greater than Norton's 3 or 5 had been used. With regard to the homogeniety of variance assumption, Rogan and Keselman (1977) reported increases in type 1 error of 2 to 4 percentage points above the nominal significance level with large degrees of heterogeneity.

So, although ANOVA with equal sample sizes is robust with respect to departures from these assumptions; particularly the assumption of normality, extreme violations do have effect. As such assumption violations can influence the outcome of the ANOVA, it seems most sensible to know if any such violation exists, in order that appropriate action or interpretation may be initiated. The only way to obtain this information is to carry out preliminary tests of these assumptions. Hartley's (1940; 1950) F-max statistic is a simple test that is sensitive not only to variance heterogeneity, but also non-normality (Box & Anderson, 1955). In situations where non-normality is present, the F-max statistic is inflated. Therefore, the F-max test would seem to be an ideal procedure for identifying violations in normality or variance homogeniety that would be likely to affect the outcome of an ANOVA.

Hartley's F-max test; carried out on the data presented in the top half of table 3.1., revealed no significant deviations from the assumptions of normality and homogeneity of variance, F-max(6,7) = 4.19.[2]

---

2 All alpha levels were set at 0.05.

CONDITION	ABA	BAB	ABB	BAA	ACB	BCA
X RECALL	27.12	26.37	21.75	20.37	21.12	20.00
S.D.	9.25	6.28	9.63	8.17	4.70	4.84
CONDITION	SAME-FAMILIAR(SF)		DIFFERENT-FAMILIAR(DF)		DIFFERENT-UNFAMILIAR(DU)	
X RECALL	26.75		21.06		20.56	
S.D	7.91		8.93		4.77	

Table 3.1. Mean recall and standard deviation of P-items by group.

3.1.2. overall F-tests.

A 3 x 2 (presentation-recall environment match/recall environment familiarity x presentation environment) completely randomised analysis of variance was carried out on subjects P scores. As is reflected in figure 3.1., there was a significant main effect of presentation-recall environment match/recall environment familiarity, $F(2,42) = 3.44$, $MSe = 55.02$. The main effect of presentation environment was not significant, $F(1,42) = 0.26$, nor was the interaction between these two factors, $F(2,42) = 0.01$.

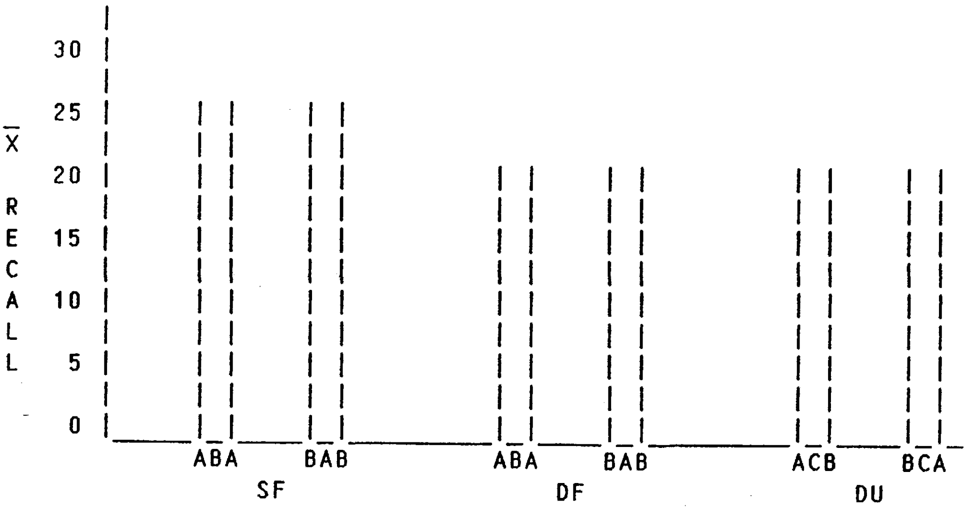

Figure 3.1. Mean recall of P-items per group.

3.1.3. specific comparisons.

Planned orthogonal pair-wise (one tailed) comparisons confirmed that the significant main effect was due to the difference between the means of the SF and DF conditions, $t(42) = 2.168$ and not between the means of the DF and DU conditions, $t(42) = 0.191$ (see table 3.1. and figure 3.1.).

3.1.4. power and degree of association.

A post-hoc analysis of the power (1-B) of the F test identifying the significant main effect revealed it to be operating at 0.50. The degree of association (W^2) between the significant main effect and subjects P scores was also calculated and found to be 0.095.

3.2. ANOVA Of P, R And F-Item Recall.

3.2.1. preliminary analysis.

P, R, and F-item recall was prepared for analysis by calculating the proportion of the total number of items presented that were recalled. For R and F-items the total number presented was ten, while for P-items the total was sixty.

An approximation to the binomial distribution is often a consequence of calculating the proportion scores. An arcsin transformation is usually recommended to establish normality of distribution when proportions are employed (Kirk, 1982, p.83). However, with the present proportion data, an examination of the distributions revealed that the only significant deviation from normality was a large positive skew with R-items. This was removed most effectively by the square root transformation, Y' =

sqrt[Y + 1] + sqrt[Y].

After transformation, Hartley's F-max test on the between subject scores indicated that the assumptions of normally and independently distributed errors were tenable, F-max(6,7) = 1.63.

The means and standard deviations pertaining to this analysis are presented in table 3.2. below.

Condition	ABA	BAB	ABB	BAA	ACB	BCA
X Recall	1.741	1.656	1.651	1.613	1.688	1.694
S.D.	0.226	0.203	0.241	0.189	0.196	0.212

Table 3.2. Average Transformed P, R and F-item recall.

However, in repeated measures designs, further assumptions are made concerning the nature of the relationships between the repeated measures. It is assumed that there is homogeneity of variance of differences between groups. Although variance-covariance matrices reflecting data of this form include matrices with compound symmetry, this is not a necessary condition for the existence of homogeniety of variance of differences.

Unfortunately, the form of the data obtained in most psychological studies is unlikely to conform to this restrictive assumption. Essentially the assumption is that in repeated measure experiments the effect of different variables will increase or decrease the different subjects' scores by similar amounts. Consequently, Keselman, Rogan, Mendoza and Breen (1980) go as far as suggesting that rather than investigating the validity of the assumption, effort is directed to dealing with the problem caused by the assumed violation.

The problem caused by such violation is that the calculated F-statistics are overly optimistic. To cope with this it is possible by altering the degrees of freedom to obtain more stringent probability values associated with the calculated F-values.

Two such corrections are often made. The Greenhouse-Geisser correction which assumes maximal heterogeneity of differences, by dividing the numerator and denominator df's by the df's associated with the repeated factor. Alternatively, the Huynh-Feldt correction alters the df's of the F-statistic numerator and denominator, on the basis of the extent to which the data violates the homogeneity of variance of differences.

A test of the assumption of homogeneity of variance of differences was carried out (symmetry test, Anderson, 1958, p.259). This indicated that the probability of the heterogeneity of variance of differences observed occurring by chance was < 0.0001.

3.2.2. overall F-tests on word type recall.

An analysis of a three factor (3 x 2 x 3) mixed design; with repeated measures on the third factor, was carried out. The first two factors were identical to those described in the previous analysis, while the third factor expressed the type of word recalled (P, R or F-item). Table 3.3. presents a summary of this analysis. Table 3.4. presents the means and standard deviations pertaining to the main effect of word type.

SOURCE	df	mean square	F	p	Huynh-Feldt p
FML	2	0.0638	1.63	0.207	
LE	1	0.0543	1.39	0.245	
FL	2	0.0249	0.64	0.533	
error	42	0.0391			
word type (W)	2	7.1002	148.15		0.000
W x FML	4	0.0465	0.97		0.413
W x LE	2	0.0026	0.05		0.907
W x FML x LE	4	0.0133	0.28		0.845
error	84	0.0479			

Table 3.3. Summary of the three factor mixed ANOVA.

	P	R	F
percentage X	37.88	56.67	8.96
X trans data	1.994	1.247	1.780
S.D. trans data	0.183	0.280	0.153

Table 3.4. Percentage mean, and transformed mean and standard deviation of overall P, R and F-item recall.

3.2.3. specific comparisons.

Pair wise (two-tailed) orthogonal comparisons[3] revealed significant differences between overall R and P-item recall, t(84) = 7.926 and overall P and F-item recall, t(84) = 10.250.

3 As the assumption of homogeneous variance of differences was found to be untenable, separate error terms were calculated for each comparison (Kirk, 1982).

4. Discussion.

4.1. General conclusions.

In common with Smith (1979, expt.1), the pattern of results observed in this experiment support the view that the environmental context reinstatement effect is a consequence of a match between the presentation/encoding context and test/recall context. The establishment of such a result also indicates that one of the research requirements has been fulfilled: effectively different room environments have been created. The lack of any room-specific effect or interaction between factors, suggests that at least with these environments, some aspect(s) common to, or equal across both rooms is involved in the production of the environmental context reinstatement effect.

4.2. Familiarity.

However, the interpretation of the experimental results may be compromised by the alteration of the intermediate task. It could be argued that if the visual matching task was more demanding than the memorisation task, subjects would have comparatively less time or fewer processing resources to attend to the room environment. If this was the case there could be a difference in the familiarity of the recall environments between SF and DF conditions biased in favour of the SF condition. The unfamiliarity of the DF condition could be the variable causing the observed effect rather than the match or mis-match between presentation and test environments. Although argument alone cannot exclude the possibility that the relative demands of the memorisation and visual matching tasks made the familiarisation procedure ineffective, one piece of evidence can be presented which suggests that this is very unlikely to have occurred. After the experimental session subjects were informally debriefed. During this time all persons questioned reported that they found the memorisation task more difficult or demanding than the visual matching task. Providing the greater ease of carrying out the visual task reflects the fact that less time or fewer processing resources are utilised, subjects will not be prevented by the task from becoming as familiar with their environment as subjects employed memorising the words. In such circumstances there will be no systematic familiarity bias to confound interpretation of the experimental results. However, as was implied at the outset, to ensure the unequivocal nature of the effect and its interpretation, familiarity will again be examined in the next experiment.

4.3. Presentation-test gap.

The other procedural alteration to Smith's experiment mentioned at the beginning of this chapter: a decrease in the amount of time spent in the waiting area, may contribute to the explanation of why subjects' recalls were so much higher in this experiment compared with those in Smith's experiment. P-item recalls were scored from a total set of 60 words. The average percentage recall across all conditions was 38%. In Smith's experiment the recall was scored from a total set of 90 words, with the average being 22%. It was said that the five minute difference in the presentation-recall gap between experiments may contribute to the explanation, but from what is known of the rate of forgetting, it does not seem likely that such a difference in recall can be attributed wholly to the difference between presentation-recall gaps of 11 and 16 minutes.

4.4. Presentation time and encoding strategies.

Two other factors could have contributed to the higher recalls in the present experiment. The first of these was the additional presentation time of each word caused by the intransigence of the interval timer. The second factor was the use of mnemonics by subjects. Again during the informal debriefing, many high scoring subjects reported having applied a strategy such as forming a mental image, or composing a story from the words. Subjects with low scores often reported such activities as "just sat and watched the words appear". This variety of different encoding strategies and their consequence on recall is reflected in the within group variance or error term. In Smith's experiment this value was equal to 37.13, while in the present experiment it was 55.02. An F-test was carried out to compare the two values ($F_{42,27} = 1.46$, $p = 0.151$) and although they did not differ significantly, it still represents a sizeable increase in variance.

4.5. The meaning of the indices of power and degree of association.

Despite the increase in error variance, power analyses revealed that the present experiment was operating at a greater power than Smith's experiment ($1-B = 0.4$).[4] In this situation two other factors could have produced such a difference; the magnitude of treatment effects and the sample size. The omega squared value: a measure of the magnitude of treatment effects, suggested a greater effect was observed in Smith's experiment ($W^2 = 0.126$). As a greater treatment effect would serve to

---

4 Although Smith does not report power nor omega squared indices, they can be calculated from the data that is provided.

increase the power of an analysis, it would seem that the greater power of the present experiment is attributable to its larger sample size (48 cf. 30). However, in common with power analysis, the omega squared value is a ratio employing the variation in sample scores. All else being constant, both power and omega squared increase with decreases in sample variation. As a result these indices cannot be regarded as objective measures between analyses. In situations where comparisons are to be attempted care must be exercised in the interpretation of the measures and only a rough guide can be hoped for. In the present comparison it would seem that the experiments are differentiated primarily by the greater variation in the present experiment.

Beyond this, Glass and Hakastian (1969) and Dooling and Danks (1975), have argued that the omega squared measure is only appropriate for random effects models, while Carrol and Nordholm (1975) and Lane and Dunlap (1978) have claimed that omega squared tends to overestimate the treatment effect. Carrol and Nordholm found this overestimation was most pronounced when sample sizes were of the order of 5 to 10 subjects per condition. However, they did feel that the calculation of omega squared is of value when significant F-values are obtained and when the sample size is larger.

The omega value obtained from this analysis is useful in that it gives a rough idea as to the extent of environmental context match/mismatch at recall. Cohen (1977) provides a "scale" by which the size of these values may be judged. He suggests that a small effect in psychology (due to high error terms normally encountered) is reflected by an omega squared value of 0.01. A medium effect is reflected by a value of 0.06 and a large effect is reflected by a value of 0.15 or more. So the environment context effect observed here falls into the medium to large category of magnitude of effect.

4.6. ANOVA of P,R and F-item recall.

The latter ANOVA revealed a main effect of word type with items from the original stimulus list that were included in the recognition test, being best recalled. Next best recall was of stimuli items which were presented only as part of the original list, while worst recall was of recognition test filler items. The average recognition score was 8.06, with an average of 0.42 false alarms. This would suggest that the pattern of word type recall could be explained in terms of a high degree of learning of stimuli items at original presentation, followed by a second presentation of some of these items which because of their high degree of learning subjects could easily recognise, without having to pay a great deal of attention to the filler items. The consequences for recall were that the second presentation of the items in the recognition test incremented their recall above that of P-items, whereas the degree of attention afforded the filler items could not result in an ability to recall them as well as the other two item types.

The effect on word type supports the view expressed in the introduction (section 1.3.) that the processing of P, R and F-items would be more varied than the processing of P-items alone. But perhaps more importantly, this in conjunction with the lack of any other effect approaching significance, sheds some light on the discrepancy in the effect of the presentation-recall environment match/ recall environment familiarity factor, between the two ANOVAs. Specifically, it suggests that the inclusion of R and F items increments the overall experimental variance such that it masks any other effect. Certainly such a determination is not in conflict with the evidence discussed in relation to power and degree of association measures. However, as will be seen in future chapters, there may be another factor present when R and F items are included in the analysis, that exerts an influence to reduce the degree of environmental context effect observed.

4.7. In prospect.

In the next two experiments most of the topics and points touched on in this discussion will be returned to in a more formal manner. For the moment, on the basis of the evidence from the experiments reported in chapter two as well as from this seminal experiment, the view that environmental context effects are distinct psychological phenomena raises the questions of why and how they occur. The next few chapters consider attempts which have been made to answer these questions.

CHAPTER FOUR

THEORETICAL PSYCHOLOGY: PERSPECTIVE AND CRITERIA

1. Introduction.

Accounts of context and accounts incorporating context have been developed throughout the history of psychology. In any assessment of the utility and validity of an account it is bound to be the prevailing criteria of such endeavours that are brought to bear. In what seems to be a period of transition in psychology, it is appropriate to consider what these criteria are and are likely to be, before any of the accounts are presented and assessed.

2. The Theory Of Psychological Information-Processing.

2.1. Origins.

The "information-processing" approach to the study of psychology has been apparent since the mid-nineteen-fifties, but it seems only from the mid-seventies that a proper awareness of the essential nature of this approach became evident. Allport (1980) has summarised the historical route through which the approach has been developed.

The basis of the approach: the analogy between brain and computer, initially focused attention on comparisons involving features which were a consequence of digital computer design. Many ideas in psychology can be traced to this type of comparison. For example, short-term memory is often viewed as the central processor which mediates access to long-term memory. Long-term memory itself is often regarded as a passive representation through which the central processor must search.

Although Allport identifies another source, it seems that one consequence of this structuralist approach would be the view that information processing is carried out in <u>discrete</u> stages. Typically, this sort of analysis has taken the form of flow-charts tracing the path of the stimulus through the system. Neisser (1976), presented a summary diagram of the information processing approach to perception and cognition which; whether intended or not, includes one of the best parodies of such pseudo-explanative accounts (see figure 1.1.).

```
                          ------------           ------------
 ------------            |   More     |         | Still More |
| Processing | ---->     | Processing | ---->   | Processing |
 ------------             ------------           ------------
```

Figure 4.1. Part of Neisser's summary of the information processing approach (from Neisser, 1976).

The basic criticisms of such analyses are that drawing an arrow between boxes does not explain how a certain piece of information causes the appropriate process(es) to operate on it, nor does assigning a label to a box sufficiently explain how information is transformed from one type to another.

As process models were seen to be required to provide explanations in psychology, a greater communality with the discipline of artificial intelligence (AI) began to arise. As AI was involved in developing systems doing

> things that would require intelligence if done by men (Minsky, 1968, p.v)

there had always been a certain overlap in interests. However, with the requirement of process models to provide explanations in psychology and the consequent need to be able to describe processes in detail, the adoption of programming languages; which do this in computing, seems a natural progression. At present this notation appears to be the best candidate in the attempt to unambiguously represent mental activity. Whether or not it will remain so depends on the demonstration of inadequacy in the notational system (rather than in any particular process so described or the hardware it is implemented on) and/or the development of some superior notational system.

2.2. Current conception.

The assumption presently underlying the development of the information-processing approach is that both the brain and computer are essentially symbol manipulators. Two related consequences of considering the brain as a symbol manipulator are that it gives a new perspective to the dualist separation of mind and body (Boden, 1977) and brings mind into the physical universe (Newell, 1981).

However, the analogy between brain and computer does not mean that the two are identical. Too often, a distorted picture of the attempt to comprehend by applying the computational analogy is perceived, if this point is not fully appreciated. It is important to remember that the similarity of symbol manipulation is a function of a particular type of abstraction. Therefore, the real value of the computer metaphor in psychology is not in the computer, but in the nature of the abstracted computational processes that can be implemented on it.

2.3. Forms of abstraction.

David Marr (eg. Marr, 1977; Marr, 1982; Marr & Nishihara, 1978; Marr & Poggio, 1977) in his development of the metatheory of information-processing has concluded that there are in fact three levels; more correctly types of abstraction, at which an information-processing task must be understood before its operation is fully comprehended.

The first type of abstraction is what Marr terms the computational theory. This characterises the underlying nature of the process. The goal of the process is identified, as are the rules that define the constraints that the process must conform to. Marr regards the goal of the process as corresponding to <u>what</u> it does and its constraints <u>why</u> it does it. The next type of abstraction is in terms of the representations and algorithms that constitute a process and concerns <u>how</u> the computational theory of the process may be implemented, while the third type of abstraction is in terms of the physical realisation of the process in the particular information-processing device.

Although the difference between theory and representation and algorithm (sometimes expressed as theory and implementation) has been noted by other writers (eg. Palmer, 1978; Winograd, 1977; see section 2.5.), as yet no one has specified the nature of the different abstractions in such detail as Marr.

In addition to specifying the nature of these three forms of abstraction, Marr (1977) also distinguishes between type 1 and type 2 theories of information-processing. Type 1 theories are obtained from processes which can be abstracted in the terms of a computational theory. Type 2 theories are obtained from processes which can only be abstracted in terms of their representations and algorithms, as their interaction is their simplest description. The determining factor is the modularity of the process: how independent the sub-processes are of the other sub-processes as they operate (Marr, 1982).

Marr has identified modularity as a principle of information-processing design. If a process is not designed in a modular fashion it is extremely difficult to modify. Due to the interactive nature of non-modular processes, small changes can have deleterious repercussions throughout the whole process. Therefore, modification of such a process is a complex and effortful undertaking. Consequently, for the type of development of biological information-processing systems conceived by evolutionary theory, a modular structure would be much preferred.

2.4. Philosophical position.

An important consequence of appreciating that there are different types of abstraction possible in the analysis of information processing tasks is that it reveals the physical realisation (eg. physiological mechanisms in psychology), representation and algorithm, and computational theory as different conceptualisations of the same events. The different forms of abstraction describing these events are imposed by us in our attempt to comprehend and have no other sort of reality. This means that the physiological mechanisms are the manifestation of computational theory or representation and algorithm, when the terms of reference are physiological. When the terms of reference are representations and algorithms or computational theory, the events that an attempt is being made to describe have not changed, only the manner of abstracting these events has altered.

The types of abstraction possible are not restricted to the three mentioned and some types of abstraction may be more or less related. However, it is not always possible to reduce or build one type of description from that of another. As the different types of abstraction describe the same events, but in different terms of conception there is no requirement for a veridical relationship.

2.5. The value of computational theory.

As the form that a representation takes determines the information which is made explicit and that which is made implicit, so the different descriptive conceptualisations of events will have greater and lesser degrees of utility for different purposes. Marr (1982) has argued and experience indicates that for the purposes of understanding, the most important type of abstraction is that of the computational theory. As he points out it is easier to understand if one comprehends the problem while examining the operations which attempt to solve it. An appropriate example of this is the way computer programs tend to be written and certainly are thought of, in high level languages which preserve the goal orientated aspect of the task, rather than in machine code which is closer to the physical realisation of the processes.

Stephen Palmer (1978) is another writer who has attempted to present a metatheory of representation and algorithm[1] in the hope that this will provide a framework within which processes can be understood. He argues that it is necessary to have a conception of what the limits of different representations and their associated algorithms are and consequently what functional differences exist between them. In effect what Palmer is doing is abstracting the process in terms of what it can do, or the goals it can achieve. However, Palmer is not interested in defining the constraints within which a process must operate. He is more concerned with determining the inherent constraints of any particular representation and algorithm. His objective is to be able to identify or eliminate candidates for particular jobs.

Although both Marr and Palmer have utilised similar types of abstraction, the products of their efforts are somewhat different. Whereas Marr's specification of the nature of the particular types of abstraction has contributed to the development of the theory of information processing, Palmers informal; perhaps intuitive, use of such abstractions has achieved what is essentially a categorisation of process ability, rather than a metatheory of representation and algorithm.

In a similar vein, Winograd has discussed the relationship between theory and programming. However, Winograd does not elaborate on the purpose or requirement of theory, but instead focuses on the role of alternative implementations as a means of providing the information from which a theory may be developed.

---

1 Marr and Palmer use slightly different labels to convey the same meaning. Marr's processes require representations and algorithms if they are to be implemented. Palmer's representation depends on processes (Marr's algorithm) to be implemented.

2.6. The need for computational theory.

In terms of Marr's distinctions, it is the information-processing task conceived in terms of the computational theory that is most novel and the terms of the representation and algorithm, and physical realisation which are most familiar. In common with the majority of workers involved with information-processing systems, psychologists' descriptions of processes have been presented mostly in terms of representations and algorithms, with some touching on the underlying physical realisation: the physiology, neuroanatomy, etc. Indeed, prior to the development of information-processing metatheory, process models were defined in terms of representations and algorithms.

The usual argument presented to support and encourage the use of process models as a vehicle for theoretical expression and development is that they provide psychology with three important assets. By adopting the notation utilised in programming, psychological processing can be represented unambiguously, a criterion by which the sufficiency of these ideas can be assessed is provided: if a model of psychological processing is to be a viable possibility it should when programmed be able to carry out the task or set of tasks it claims to describe, and in running a computer simulation, previously unrecognised features of a model's operation are revealed; particularly with less tractable models.

However, despite the obvious theoretical and practical advantages afforded by the construction of this type of model, some concern has been expressed by its very advocates regarding the product of such an approach. One problem associated with information-processing tasks is that a great variety of representations and algorithms can produce the same result. Different representations and algorithms can carry out the same task and any combination which mirrors human performance may, or may not be equivalent to that employed by a human or animal subject. Newell (1973) touched on this when he expressed concern with regard to the dichotomous nature of hypothesis construction and the proliferation of models (in terms of representations and algorithms) to accommodate data previously thought to contradict a hypothesis. Indeed, Anderson (1976) has gone so far as to argue and attempt to demonstrate the impossibility of determining which type of representation and algorithm is employed. The use of computer models would seem to have replaced ambiguity with irreducible variety. But is such a situation really unique?

In all sciences there are alternative explanations of empirical phenomena. The fundamental purpose of the scientific method is to identify the account which accrues most support. That this should be more difficult in psychology is not very surprising. It may be the case that present experimental techniques are not able to discern between serious candidate models, but at the moment it would seem that the range of potential candidates has not been identified. Perhaps in the light of greater knowledge of information-processing attained through the effort of identifying this range, methods of comparison will not only be available, but fairly obvious.

However, along with the need to explore alternative representations and algorithms, there is also the requirement for some method of accumulating what is learned. It is also in this role that the abstraction of computational theory is vitally important. Newell's (1973) main concern was that there was no accumulation of knowledge. In its place was an abundance of alternatives. A similar concern over the lack of "unifying principles" has come from Johnson-Laird (1978). Such unifying principles are a consequence of a different type of conception. They are not to be found in descriptions of representations and algorithms, but rather have to be abstracted from those descriptions. It is the abstraction of computational theory which is the critically important conception of an information-processing task and it is in these terms that scientific principles of such processes will be formulated (Marr, 1982).

2.7. Computational theory and process implementation.

The development of the theory or metatheory of information-processing has affected the argument favouring the expression of theoretical ideas in terms of process models. The most obvious consequence of applying this rationale is that the benefit of revealing hidden features of a model's operation by running it on a computer is seen to be an exposition of an erroneous computational theory, or the use of more general representations and algorithms than are required for the process described by the computational theory.

With the appreciation of the additional abstraction of computational theory, the importance of the programming notation is complemented rather than degraded. It would seem that for comprehension and theoretical purposes the computational theory is most important. For more "practically" oriented purposes, knowledge of the way in which process(es) are implemented is important. In such circumstances a description of the process(es) in terms of representations and algorithms is required. But for a complete understanding of a psychological process it must be able to be described in all terms of abstraction.

As was mentioned earlier, the sufficiency of a process implementation although supporting a correspondence between the model and reality, does not mean that this is the method employed by the system being modelled. However, fewer models should operate satisfactorily within the stipulates of the computational theory. By applying the computational theory and eliminating potential candidates, the task of finding the appropriate implementation is made slightly easier.

2.8. Computational theory and control processes.

Rabbitt (1979) has criticised models of information-processing because of their failure to describe the way in which these processes change with time. Rabbitt attributes the blame for this state of affairs to the omission of control processes in information-processing models. However, part of this problem may be due to the previously unrecognised fact that the process models which have been presented were not whole models. These models have lacked the abstraction of the computational theory and due to the relationship between this and control processes, have made the description of the latter extremely difficult.

It is from the account of the computational theory that the constraints on a process are defined. Knowledge of the limits and requirements of a process is a prerequisite to the determination of effective control processes. Although not often thought of in such terms, control processes attempt to allow the process(es) to fulfill their purpose or attain their goal. When expressed in such terms it can be seen that control processes require a computational theory to provide the detail of their job specification.

3. Biological Considerations In Psychological Theory.

3.1. Introduction.

The evolutionary requirements of biological systems emphasised the value of modular process design. This is one example of the benefit of considering the particular types of requirements of such systems, as attempts are made to model and understand them.

3.2. Purpose of the system.

Donald Norman (1982) has discussed in some detail the differences between animate systems and the type of inanimate systems typically described by cognitive scientists, be they philosophers, computer scientists, or psychologists. Norman points out that while the type of inanimate system designed by cognitive scientists has as its main function the doing of the defined cognitive task, in contrast, the animate system has as its main function the business of living. This can involve a variety of tasks such as obtaining food, protecting from physical injury, establishing relationships, reproducing and looking after the young. Given such different objectives, Norman suggests that the generally held view that cognitive operation is that which biological functioning sustains is actually back to front. Instead he argues that cognitive operation developed as a tool to aid the animate system with the problem of survival. Norman raises the possibility that intellectual thought is an artifact of the production of a larger intelligent component; oriented toward more fundamental objectives such as those described above, attaining some critical mass which provides it with sufficient computational power to determine its own direction.[2]

Irrespective of the extent to which one concurs with Norman's views, the point is made that when considering any task, it is important to regard its psychological performance in relation to the types of problem that animate systems have evolved to cope with. The view that psychological performance is dependent upon, or is derived from processes which share the fundamental aspect of ecological utility rationalises the use of data from a variety of different sources; such as natural and clinical observation and experiments investigating list learning, when similar types of psychological problems are being tackled. A consequence of this view is that the type of models that psychologists should be trying to formulate should reflect this aspect of cognitive functioning. Models which can only carry out one artificial task are extremely unlikely to include processes which replicate those utilised in psychological performance.

4. Some Modelling Considerations Acquired From Experience.

4.1. Structure.

Although primarily a product of cognition conceived in terms of the particular task requirements which was criticised by Norman, much research has indicated the advantage of psychological models which do not rigidly partition aspects of information-processing. When the purpose of a system has been to simulate human behaviour the distinctions in

2 This has similarities with the Colebrook hypothesis (Rutherford & Hamilton, 1980), which was originally presented to account for a great deal of indiscriminate and contradictory human behaviour.

terms of the quasi-independent topics of perception, learning, language and memory have blurred. The benefit of a system which allows information present in one domain to relate to information in another is intuitively obvious and practically demonstrable, yet the value of such an ability is emphasised when the higher order purpose of the system is considered. For animate systems in particular the capacity to utilise information collected from a variety of operational sources has significant consequences for survival.[3]

4.2. Basic process.

One interesting feature of the type of models which unite different types of performance or aspects of information-processing is that they tend to do so by equating a variety of different tasks on some basic problem. This problem seems to be the determination of what is being presented to the system, either as an external stimulus event or as an internally generated entity, such as an intended movement or a puzzle solution. In all cases determination is made on the basis of information present in the system. Items corresponding to the representation are made available to the system and their degree of correspondence is assessed. The basic goal of such process(es) is to map information into existing representational structures.

A variety of different conceptions of information-processing such as proposed by Moore and Newell (1973), Minsky (1975), Bobrow and Norman (1975), Bobrow and Winograd's (1977), Schank and Abelson (1977),

---

3 There is an interesting parallel between this and the comparative lack of interaction of information from different sensory modalities in lower animals (Carlson, 1980, p.7).

Rumelhart and Ortony (1977) and Sanford and Garrod (1981) have converged in agreement on this view of the initial stages of processing. One of the advantages of this type of system is the wealth of information/processes that are potentially available to take part in any operation. There is of course, a direct relationship with context in that semantic context is often regarded as being the representational structure into which the presented information is mapped.

5. Summary Of Criteria And Considerations For A Psychological Account.

5.1. The form of characterisation.

At present influenced by a greater understanding of the nature of information-processing: through the development of computer science in general and AI in particular and its relationship with psychology, slowly psychologists are changing their criteria of what constitutes an adequate explanation of psychological phenomena. It is now appreciated that information-processing can be described in several terms of abstraction. Of greatest value to theoretical development and in eliciting comprehension is the abstraction of computational theory. The abstraction in terms of representations and algorithms is necessary if a description is to be implemented and if what is normally considered to be psychological processes are to be specified. The abstraction of a process in terms of its physical realisation with human and animal subjects involves those descriptions employed by the neurosciences. Accounts in terms of each form of abstraction are necessary for a <u>complete</u> understanding of information-processing, but to answer different questions, different abstractions will be more or less useful.

5.2. The type of system.

In determining the purpose and form of a psychological process, it is important to bear in mind the requirements of an animate system. Processes specifically devised to meet artificial task demands are less likely to model processes used by the animate system, than are modifications of ecologically valid processes which meet the artificial task demands.

5.3. Theory development.

With the framework to abstract information-processing now available, the development of "unifying principles" seems more likely. Although this type of development has not been made explicit, there has been of late a general agreement on the need for large knowledge structures into which information is mapped as an initial process in "intelligent" functioning.

CHAPTER FIVE

PREVIOUS ACCOUNTS OF ENVIRONMENTAL CONTEXT PHENOMENA

1. Introduction.

The notion of context has been utilised in many different accounts of human behaviour, but few of these have actually been attempts to explain the empirical phenomena of environmental context. Most of the accounts which have used the concept of context in their elaboration have done so in an attempt to explain more general aspects of behaviour. However, despite the differences in focus, this latter type of account should still contribute to an explanation of the way in which context exerts its effects, even if that contribution is only in the form of a simple inversion of the account.

2. The Phenomena Of Environmental Context.

2.1. Summary of environmental context effects.

Before considering how well different accounts accommodate and explain the environmental context phenomena detailed in chapter two, it is worth summarising the basic effects which have been observed across the different experiments employing a variety of paradigms and types of environmental context. Table 5.1. summarises these phenomena.

PHENOMENA	LABEL
1. Greater free recall in reinstated learning context.	Free-recall effect.
2. Greater cued-recall in reinstated learning context, but to a lesser extent than 1.	Cued-recall effect.
3. No disadvantage to recognition outside learning context.	Recognition effect.
4. Greater free-recall with more learning contexts.	Multiple context effect.
5. Greater free-recall outside learning context if subject thinks of learning context.	Self-generated context effect.
6. Less interference between information sets if they are presented in different contexts.	Interference effect.
7. Greater ability to judge periods of time when they are distinguished by a change in context.	Temporal-judgement effect.

Table 5.1. Summary of basic environmental context phenomena.

2.2. Interference effects.

Although seven effects have been listed, it is worth pointing out that only six have been demonstrated as empirical phenomena. As yet, interference has been forwarded only as an "explanation" of reductions in correct remembrance. Bilodeau and Schlosberg (1951), Greenspoon and Ranyard (1957) and Dallett and Wilcox (1968) have all carried out experiments purporting to reduce interference by introducing a change in environmental context between information sets (ie. word-lists). However, none of these papers contain any report of the empirical phenomenon of interference: intrusions. Only the number of correct responses and/or the number of relearning trials is reported. Consequently, it remains to be demonstrated that changes in context reduce interference. Therefore, subsequent consideration of a model's ability to account for environmental context phenomena will be restricted to the other six effects, although an account of interference is provided in chapter six (section 2.1.).

2.3. Temporal judgement effects.

As different accounts of the environmental context effects are presented, it should become evident that the temporal judgement effects reported by Block are really a consequence of processes which utilise context information retrieved with nominal entities. As these processes are likely to be particular to this function, it is not intended to examine this aspect of the effect. The main consideration will be the way in which such nominal entities are retrieved, although the ability of a system to provide context information which could input to processes providing temporal judgements will be discussed.

3. Pre-cognitive Accounts.

3.1. Carr and McGeoch.

The first psychological accounts of environmental context effects were presented in the terms of stimulus-response associationism. One of the first investigators of environmental context phenomena Harvey Carr, states this position quite succinctly.

> As a rule any activity can be most readily reinstated in that situation in which it was acquired because the two must necessarily become directly associated (Carr, 1925 p.251).

McGeoch's (1942) explanation of why forgetting occurs, although firmly based in associationism also, began to describe some of the mechanics underlying this facet of memory performance. McGeoch postulated a three factor account of forgetting, the factors being:

i) interference by intervening activities,
ii) altered stimulating conditions,
iii) inadequate set at the time of recall.

The factor of most interest here is that of altered stimulating conditions. McGeoch suggested that any stimulus defined by an experimenter is the nominal stimulus only and that the actual functional stimulus can include other features of; the environment, the subject's state and, mental activity at the time of learning. It is this whole functional stimulus which becomes associated with a response. McGeoch cites Carr (1925) when he states that,

> retention will be higher the more numerous and complete are the stimuli earlier associated with the activity and now present again at the time of measurement of retention... Alteration or removal of these stimuli at the time of recall will be correlated with a failure of recall (McGeoch, 1942 p.501).

One of the main developments in McGeoch's account of the effect of altering stimulating conditions from that of Carr's (1925), was the formal distinction between the nominal and functional stimulus. Although Guthrie (1935) at least, had implied in his discussions of learning that a single response could be associated with a combination of stimuli or stimulus elements (eg. Guthrie, 1935 p.28), McGeoch made this notion explicit.

3.2. Appraisal.

In common with most associationist-behaviourist writers, Carr (1917, 1925) provides little more than detailed and formalised descriptions of situations, rather than descriptions of the psychological processes which give rise to the behaviour observed in such situations. Although McGeoch (1942) makes an important point explicit with regard to the nature of the information encoded as the stimuli, his account still has more to do with the associationist-behaviourist style, than it does with the type of account now regarded as approaching an explanation.

4. Anderson And Bower.

4.1. Introduction.

The idea that a stimulus may have a variety of encoding permutations has recurred in many models of learning and memory. Bower (1972a), Estes (1955; 1959), Madigan (1969), Martin (1968; 1971; 1973), Melton (1970) and Tulving (1968) all utilise the concept of stimulus variability at encoding in their accounts of psychological processes.

In all of these models some particular notion of context is utilised to produce the variation in stimulus encoding. In most cases, this notion of context is most comparable to what has been described as semantic interpretation (eg. Madigan, 1969; Martin, 1968; 1971; 1973; Melton, 1970). However, Bower (1972a) formulated an account of stimulus encoding variability which is less restrictive in its notion of context, regarding it not only as a product of semantic interpretation, but also as what has been termed here as state and environmental information. Bower recognises the contribution of other models of stimulus encoding variability; particularly Estes (1959) and Martin (1968), to his account.

4.2. Encoding variability.

In his exposition of the model, Bower suggests that a nominal stimulus is encoded by a variety of operators. The product of each operator is a stimulus element. An encoded stimulus is represented by a pattern, or particular set of stimulus elements. It is in the determination of the probability of each operator being active and so encoding a stimulus element that the context acts (see figure 5.1.).

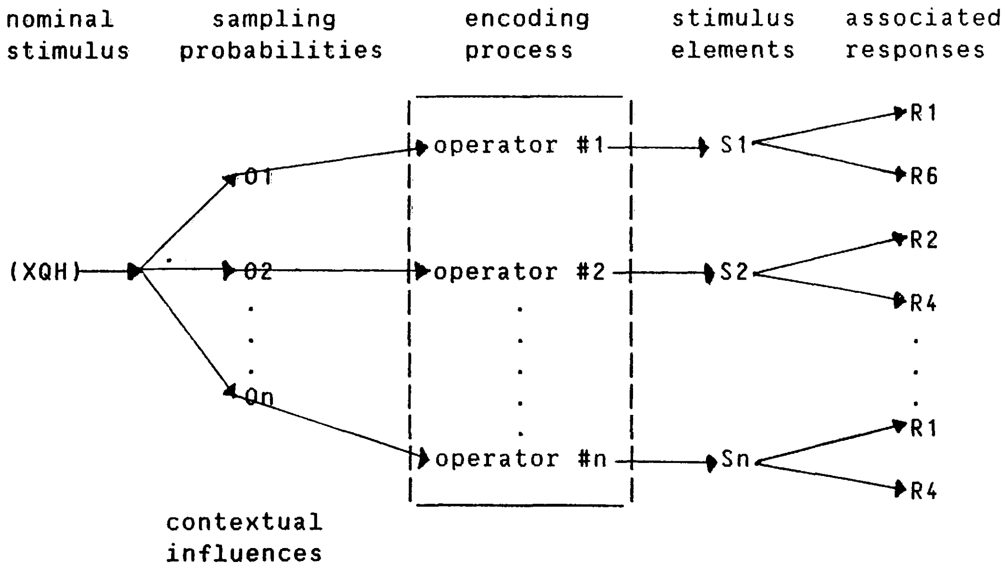

Figure 5.1. Schematic representation of Bower's notion of encoding variability (from Bower, 1972a).

One of the advantages of Bower's model compared to those of Madigan, Martin and Melton is that it details how similar semantic interpretations still can differ in their representation (see Crowder, 1976, p.288). The nature of the encoded representation of the stimulus is not as absolute as in previous accounts of stimulus encoding variability. Bower also suggests that each response item can be another cognitive element such as an idea or concept. The idea or concept could represent membership to a particular set or list. In this latter case, Bower would term the idea or concept a list marker or tag. It is with the association of the group of elements representing the nominal stimulus to the list marker that Bower's (1972a) account meets up with that of Anderson and Bower's (1972; 1974) accounts of list item recognition and recall.

4.3. FRAN (free recall in an associative network).

Anderson and Bower (1972) postulate that at encoding stimuli items are associated with context via list markers. These list markers represent the relationship between each occurrence of an item and the context prevailing at its presentation. The context is assumed to be composed of a variable and almost infinite number of elements with which there is a probability that any particular set of these elements will become associated to the presented item via the list marker (see figure 5.2.).

At recognition, a decision is made as to whether an item was previously presented by assessing how much evidence there is for the item belonging to the list in question. This is done by determining the number of contextual elements associated to the item's list marker which are also members of the prescribed contextual set. This set is identified by special list markers which identify the "prototypical" representation of the list-N contextual elements. It is with these elements that a comparison is made with the possible list-N items' associated contextual elements. A signal detection analysis is proposed to carry out the decision process on the basis of probability distributions which describe the difference between list items and non-list items.

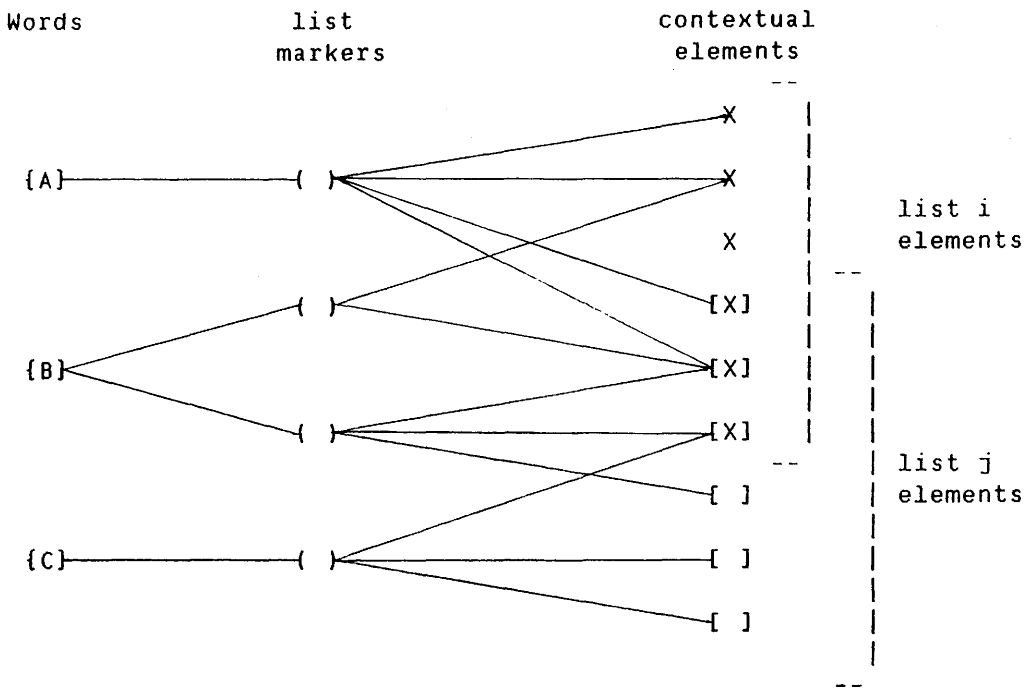

Figure 5.2. Schematic diagram of the associations between words, list markers and context (from Anderson & Bower, 1972).

This describes how potential set items are assessed as members of a prescribed set or list, but one of the more eloquent parts of Anderson and Bower's model is the provision of a method which produces items for assessment. To begin with, they assume that an associative memory network already exists. As items are presented, their representations in this network are activated. From each of these nodes there are a variety of associative paths to other nodes representing other items. As the set of items are presented, any pathways between set items are tagged with a list marker. The consequence of this is that at recall a route can be followed through the associative memory network between nodes representing items. This route identifies items as potential members of the required set. Items so generated are then assessed for set membership at the recognition phase as previously described.

4.4. HAM (human associative memory).

The model just described, presented by Anderson and Bower (1972), was embodied in the computer program FRAN (free recall in an associative network). Their 1974 update of this account was based on HAM (human associative memory; Anderson and Bower, 1973) which had as its main focus memory for language rather than single words or items. HAM utilises a propositional representation and distinguishes between words and concepts as well as specifying the semantic relationship between concepts. HAM parses each proposition into a tree-like structure. This representation can be broken down into two sub-trees; one embodying the context and the other the fact. The context node is further divided into location and time nodes, while the fact node is also decomposable to subject and predicate nodes. The last division necessary to describe is that of the predicate node into relation and object nodes. The nodes are defined by the nature of the links between them which describe the semantic relationship. Near the bottom of the tree are the concept nodes which are connected to particular instance words. These nodes represent the particular idea that each subject has of each concept and are assumed to lie in memory beforehand (see figure 5.3.).

Such a propositionally based model is able to represent explicitly the complex information structures that Anderson and Bower's earlier formalisation could only imply. Previously, the list marker of dog; now represented by the concept node j, would have been associated to a group of contextual elements. In the HAM system the implied contents of this group of contextual elements can be explicitly represented by the contextual propositions such as "stomach gurgling", "I study dog " and "list-N".

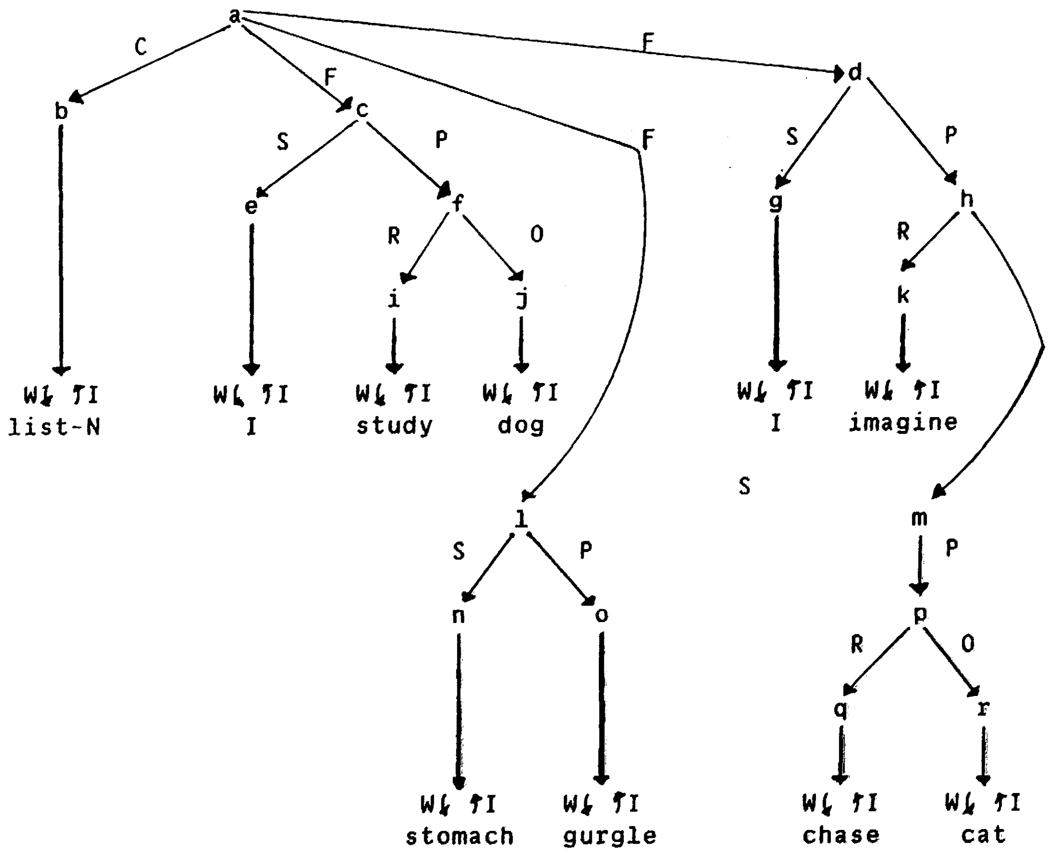

Figure 5.3. Propositions formed in memory after the presentation of the word dog (from Anderson & Bower, 1974).

Anderson and Bower explain how, on the basis of the type of information encoded in the memory representation, different reasons for believing that an item is one of the required set will result. Retrieval in the HAM system depends primarily upon a MATCH process. This process attempts to correlate two information structures; the probe and a piece of the associative memory structure. Essentially the probe is a representation in the system of the question being asked of memory, with its exact structure determined by the nature of the question. The MATCH process identifies terminal nodes in memory similar to those in the probe and attempts to locate associative paths which connect these terminal nodes with the same type of associative relation that those of the probe are connected with. The MATCH process tries to provide the part of memory which gives the greatest match with the probe structure. This type of memory access is effective if a comparison between the probe and some part of the memory system is possible. If such a comparison cannot be made because certain propositions which the probe consists of have not been encoded in memory, an answer may still be able to be given on the basis of an inference from the information that has been encoded in memory. For example, Anderson and Bower (1974) suggest that a persistent noise or any other occurrence or aspect of the experiment could be encoded along with the nominal stimuli items. If the subject encodes that such an event happened during the experiment then an inference can be made that the word associated with this event was a presented nominal stimuli item and so therefore a member of the sought after set. The system can accommodate different strategies to assess the evidence of presentation. The confidence with which a subject asserts that an item was presented during the experiment will be a function of the type and amount of information the subject can obtain with respect to the item in question.

4.5. Appraisal.

Anderson and Bower provide the first account of the effects described in chapter two that approaches the requirements of the "ideal" model described earlier. Although Anderson and Bower's account extends over the life of two models, both are claimed to be functionally equivalent at the level of analysis set by the former (Anderson & Bower, 1974, p.409). The major problem for both the FRAN and HAM models is to accommodate the different environmental context effects with recall and recognition. In both systems the same context information is used in the retrieval and recognition phases. As a result, it is difficult to see how context dependent recognition could be unimpaired, while context dependent recall is impaired when the test environment does not match with the learning environment.

It is this inability of "two process" models (also Kintsch, 1970) to account for the difference between environmental context effects in recall and recognition, that led Smith, Glenberg and Bjork (1978) to conclude that Tulving and Watkin's (1973) view of recall and recognition employing essentially similar retrieval processes could account for the data more succinctly.

It is rather ironic that while Anderson and Bower's model seems as though it could cope with the more complex phenomena of environmental context such as multiple context effects (through the creation of a larger number of prototype list markers) and self-generated context (through the operation of the probe creation part of the MATCH procedure in HAM), it fails on the basis of its fundamental method of deciding which items are to be sought after and when obtained, which are required. However, while HAM can provide a basis for the explanation of temporal judgement effects; as temporality is included as part of the context information recorded, to account for the effects reported by Block would require particular methods of encoding temporal information and assessing it which are quite different to the processes employed in HAM.[1]

---

1 For a criticism of Anderson and Bower's computational approach see Kolers and Smythe (1984) and Allport (1984).

5. Norman And Rumelhart's System For Perception And Memory.

5.1. Introduction.

The LNR research group were also early proponents of propositional representations. Indeed, the model of cognition presented by Rumelhart, Lindsay and Norman (1972) and subsequent research presented by Norman, Rumelhart and LNR research group (1975) has much in common with the HAM propositional system. A forerunner of these models: a model of perception and memory, presented by Norman and Rumelhart (1970) also makes explicit use of context as a component in its operation.

5.2. System description.

As part of the first semantic analysis in this model, context is attached to the components of the individual memories; the attributes. The attributes are already present in the memory store, so only the context cues are actually "novel" introductions. The pattern of attributes attached to by context cues defines the items presented (see figure 5.4.).

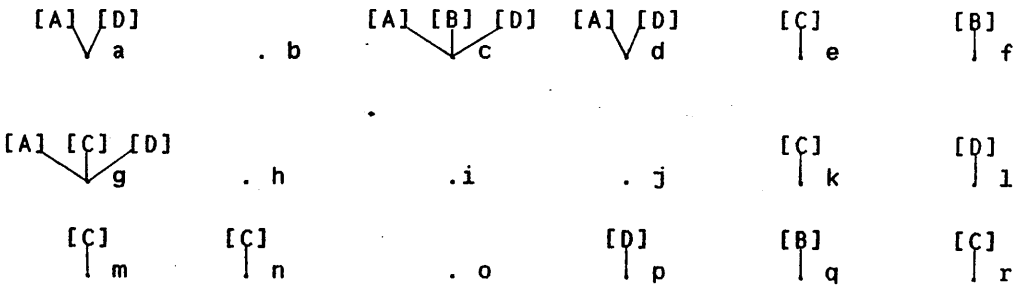

Figure 5.4. Memory representation employing context markers on each attribute (from Norman & Rumelhart, 1970).

At retrieval, using whatever information is presented, the system tries to identify the appropriate "other half": context or attributes. In recognition the subject is given a test item whose attributes are then checked against the context they are associated with. If an item in memory is found with a criterion number of similar attributes corresponding to a consistent context (or contexts), this is taken to be a previously presented item. At recall the subject uses the information about context available and tries to retrieve all the individual attributes associated with the context. After this, an attempt is made to determine what the item must have been.

In their description Norman and Rumelhart make no new suggestions as to what context is, or could be. They leave this question open, but from the apparent influence of the context in defining an item via the attributes, a primarily semantic aspect is implied.

5.3. Compatibility.

This model of perception and memory presented by Norman and Rumelhart provides a good example of the problems created by the separate systems approach criticised earlier. Norman and Rumelhart attempted to link the different systems which had been proposed by researchers investigating different parts of performance. Their model copes with this task quite well, but they themselves admit that due to the requirement of making each component compatible with the others, certain compromises and consequently deficiencies occur in the model's performance. It is perhaps due to the experience of this attempt, that they now favour a more wholistic system approach, with perception, language, memory, etc. regarded as different aspects of the whole system's performance, rather than as discrete systems in their own right.

5.4. Appraisal.

With respect to providing an account of environmental context phenomena, the major problem for the model is that it is far too dependent upon external context cuing. The main conception of context in the model seems to be semantic. In situations where such a context exists and could "accompany" an item at each presentation, the model would be able to operate as described. However, when context is able to be removed from the item as environmental context can be, the fundamental pieces of information against which the attributes are checked are removed, making this process inoperable.

Of course, if context is composed of semantic and environmental information, these processes still could operate. In such a situation different contributions of semantic and environmental information have to be assumed to account for the effects on free-recall, cued-recall and recognition. One way in which this could be done is by an automatic association of predetermined proportions of the different types of context. Alternatively, a process which first identified the item via the attributes, and on the basis of this identification and the current goals of the system, decides the type of context employed, could be implemented. In many models this would be a tenable assumption, but due to the passive bottom-up nature of this model, the requirement of an active top-down process is a significant alteration to the models whole conception.

The data on multiple learning environments would seem to create a few problems also. Presumably the effect of multiple learning environments would be to increase the number of contexts used to associate to the attributes. For environmental context to exert an effect on recall, the recall attempt must occur in an environment with context information suitable to cue the context associated to the attributes. In other words, it has to occur in the learning room(s). Due to this necessity it is not possible for items presented in one room to be more advantageously recalled in another. Indeed the model would predict a decrease in the amount able to be recalled in a multiple learning room paradigm.

For a model to account for self-generated context effects, at a minimum, it needs to be able to do several things. First, it must be able to obtain some record of the relevant context. This means that on the basis of only a nominal cue, a record of context containing information akin to that used in the encoding of the nominal stimuli must be made available. Then it must be able to match the internally generated context information with that "associated" to the nominal stimuli items", or with whatever representation is being used.

Norman and Rumelhart's model would seem to be able to cope with these requirements by utilising various expressions of a context. Presumably the record of the presentation environment would be stored as an item, ie. attributes associated with a context, and would require a similar context to cue these attributes. This may be achieved if the similar context was something relating to the experiment. A context which in nominal terms was equivalent to labeling the attributes as having been used in an <u>experiment</u> should be able to identify both the record of environment and the nominal stimuli items.

Temporal judgements provide another difficulty for the model. The data from Block's experiments suggest that it is the changes in context which provide information about duration. This is particularly problematic for Norman and Rumelhart's model as it operates primarily by utilising constancy of context. If change in context is to be used to provide information it seems as though this would have to be done after a constant factor had identified the entity to be considered via the context associated attributes. One possible method would be to establish two functionally different types of context. The first would allow an entity to be retrieved by context constancy. The other type of associated context could then be used to provide input to those processes which would determine duration. However, again this calls for considerable alteration to the structure of the model.

5.5. Conclusions.

Overall the Norman and Rumelhart (1970) model fails to give an adequate account of the basic environmental context phenomena, mainly through a rigidity in the way the context is associated to the attributes and subsequently cued. It would seem more efficient to adopt or construct some alternative system which operates in a manner compatible with the functions required, rather than to persevere with one which becomes more awkward and ungainly with each modification. However, one interesting feature that does arise in considering how the model could cope with the tasks manifested in the environmental context experiments is that context can be conceived of at different levels. This aspect of context will be returned to in chapter ten, where its relation to similar points raised in chapter one (section 2.1.), will be discussed.

6. Tulving's Encoding Specificity / GAP System.

6.1. Nature of the account.

The encoding specificity hypothesis was first forwarded by Tulving and Osler (1968) to explain their findings from a study of the effect of retrieval cues. The types of cues and context that researchers have generally used to investigate encoding specificity have been semantic. However, the nature of the encoding specificity hypothesis and what Baddeley (1976, p.293) has distinguished it from; the encoding specificity principle, has made them favourite candidates in the attempt to provide an explanation of environmental context effects (eg. Metzger, Boschee, Haugen & Schnobrich, 1979).

Tulving (1974) provides the following description of what he terms the encoding specificity principle.

> the properties of the memory trace of a word event are determined by specific encoding operations performed on the input stimuli, and ... it is these properties, rather than the properties of the word in semantic memory, that determine the effectiveness of any given stimulus as a retrieval cue for the event. The principle suggests that if a stimulus in the retrieval environment renders possible or facilitates recall of the target word T, the retrieval information was appropriate to or compatible with the information contained in the episodic trace of T. Conversely, if a particular stimulus is ineffective in retrieving a particular trace, the conclusion follows that the appropriate relation was lacking (Tulving, 1974, p.778-779).

As Baddeley (1976, p.293) has pointed out, the "principle" is not an experimentally testable hypothesis: it cannot be falsified and is self-fulfilling. What is experimentally testable, is the hypothesis that only information encoded with the to-be-remembered items at input will be useful as retrieval cues for those items. Of course the difficulty is determining exactly what is encoded at input, without employing subjects' remembrance as an indicator.

Tulving (1982; 1983) has incorporated the notion of encoding specificity in his General Abstract Processing System (see figure 5.5.).

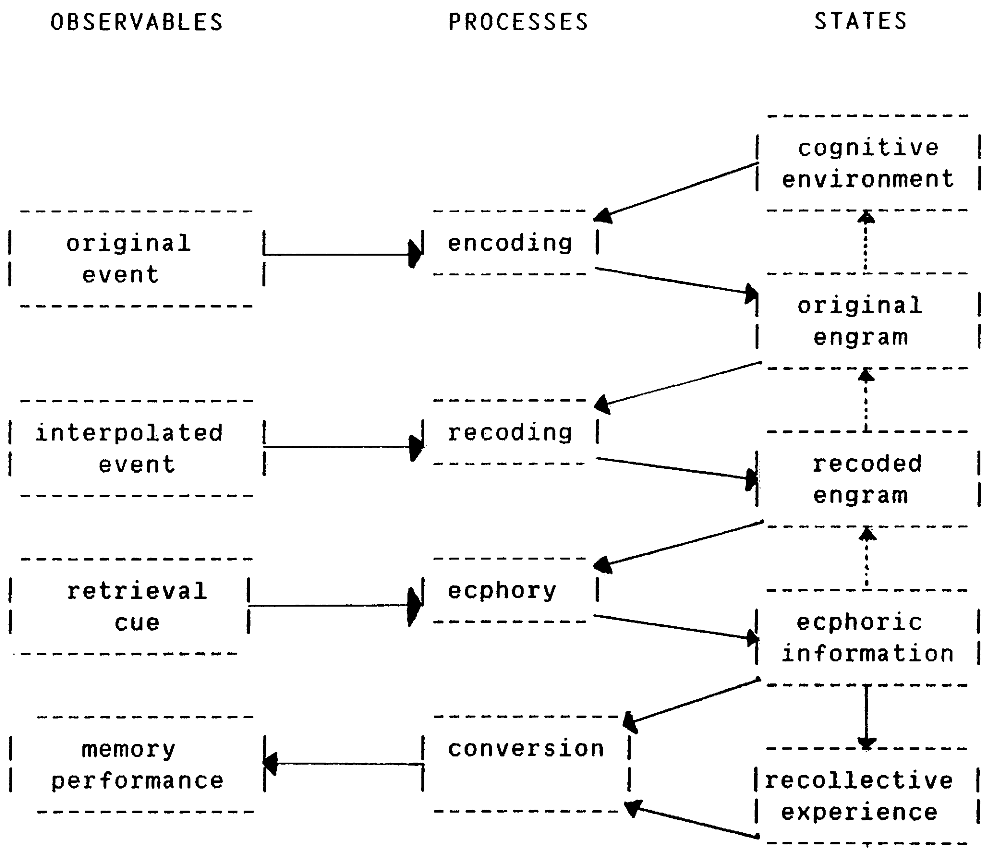

Figure 5.5. The GAP system. Arrows indicate "influences" and broken arrows indicate states that do not influence a current act of remembering, but could influence later ones (from Tulving, 1983).

It can be seen from the diagrammatic representation of the General Abstract Processing (GAP) system that the processes of ecphory are of vital importance for memory performance. Ecphory is the process that produces a particular combination of semantic and episodic information. It is this combination of (ecphoric) information that is used in the comparison with memory traces. Encoding specificity states that the greater the similarity of the ecphoric and memory traces, the greater the chance of a match and consequently, successful retrieval. Presumably retrieval would continue with the result of the match between the ecphoric information and the memory trace being input to the conversion process. Tulving differentiates between the types of performance required in recognition and recall tasks, and relates this to the amount of ecphoric information that is required for the conversion process to elicit the actual memory performance. Less ecphoric information is required for recognition performance (familiarity judgement) than is required for recall performance (identification of nominal items).

6.2. Relation to environmental context.

As the environmental context is part of the overall nominal stimuli context present during the nominal stimuli input, it is possible for it to be encoded with the nominal stimuli items. When memory is tested in the same environment, the environment can act as a retrieval cue providing extra ecphoric information. Recall in such situations should be superior compared to situations where "ineffective retrieval cues" are provided, such as other environments.

6.3. Eich's cue dependency account.

Very closely related to the account of environmental context phenomena that would be expected from Tulving's GAP system is Eich's (1980) account of state-dependent effects. Eich's view of state-dependent effects shares the (seminal) conceptual structure of Tulving's GAP system and applies not only to state-dependent effects, but also to environmental context phenomena.

The basic notion in Eich's account is that the more explicit the cues that are provided, the easier it is to obtain appropriate ecphoric information. In addition, the conversion process requires differing amounts of ecphoric information for the different types of memory performance. As recognition tasks require the least amount of ecphoric information for their performance and provide the greatest intrinsic ecphoric information (a copy cue), they should be least influenced by fluctuations in the ecphoric information of reinstated encoding conditions. On the other hand, the recall task requires most ecphoric information for performance and provides the least intrinsic ecphoric information. Consequently, it should be influenced most by fluctuations in the provision of encoding condition ecphoric information. Cued-recall; falling as it does between the two in terms of the amount of ecphoric information provided by the cue item, should be less influenced by changes in encoding conditions than recall, but more influenced than recognition.

6.4. Appraisal.

Compared to the models presented by Anderson and Bower and Norman and Rumelhart, Tulving's system would seem to provide a comparable account of the effect of a change in environment on recall. The GAP system also can account for differing effects with free-recall, cued-recall and recognition under the common assumption that environmental context is not the only information stored with the nominal item, and that this when cued in conjunction with the different performance conversion thresholds can make the additional environmental context information provided by a reinstatement of the encoding environment less important.[2] However, the GAP system has considerable difficulty in attempting to account for such environmental context phenomena as the multiple context effect, the temporal judgement effect and the self-generated context effect. The latter phenomenon can be accommodated if it is assumed that either the GAP system can provide itself, or can be provided with some equivalent to a retrieval cue, but one that is generated internally.

One of the major problems in accepting Tulving's GAP system as providing an account of environmental context effects is that the GAP system is only an outline of some of the kinds of processes that seem to be be employed to encode and retrieve information from memory described in the manner parodied by Neiser and criticised by Allport (see chapter four, section 2.1.). In terms of the what and how of processing, Tulving's description seems to be a combination of both, without satisfying the criteria of either.

Perhaps such criticisms are unfair as Tulving does describe the GAP system as,

> not a theory.. It's purpose is not to explain, let alone "predict", any specific phenomena. It represents an overall structure within which explanations of various elements of remembering and their interactions can be, and perhaps must be, sought. It is logically compatible with a large variety of specific models...(Tulving, 1983, p.129-130).

---

2 See Godden and Baddeley's (1980) account presented in section 8.3. of this chapter for a similar exposition.

However, it remains to be seen if psychologists will pay attention to such restrictions of use, or if they will be tempted by the apparent sophistication of the system and its presentation to misrepresent it as a model of memory, and attempt to employ it as a means of explaining and predicting memory performance effects, as has been done here.

In terms of the utility of the GAP system as an overall structure, already it would seem to be redundant. Not only are most of the concepts presented in the GAP system to be found in the model presented in chapter six, but because of the nature of this model they are much more clearly identified. The nature and form of Tulving's GAP system is such that the manner of its operation is extremely ill defined. In an attempt to specify the operation of the system's components and their interaction, Tulving goes to the length of presenting new and unique terms of description. However, without a non-ambiguous account of (at least) each process goal, rather than a nominal identification of the processes involved in the GAP system, such new terms achieve little clarity.

With respect to encoding specificity; which often has been suggested as an account of various effects without the accompaniment of the GAP system, it is worth pointing out that neither is this an explanation, ie. a description of a process in any of the terms of abstraction discussed. It is a hypothesis with some support which requires explanation. Unfortunately, these hypotheses; which are a vital and integral part of scientific procedure, can clog up the process of providing explanations when they themselves are misconceived as providing such explanations.

7. Glenberg's Component Levels.

7.1. Introduction.

In the same way that Martin (1968; 1971; 1973), Madigan (1969), Melton (1970) and Bower (1972) have attempted to give accounts of repetition effects, Glenberg (1979) has also presented an explanation of such effects based on what he calls a "component levels theory". The term component stems from the fact that different types of information contribute toward the representation of an entity in mind. The levels tag comes from reference to a "hierarchy" of use; some types of information are used in more representations than others.

7.2. Model description.

Glenberg postulates three types of component: contextual, structural and descriptive. The Contextual component includes not only information which has been previously categorised as environmental and physiological information, but also that which has been roughly described as cognitive or process context (eg. Block, 1980; Underwood, 1977; chapter two, section 4.1). This type of information is automatically encoded as presented items are perceived. However, this information is not only encoded, but it can also influence the encoding of the other information types.

The Structural component involves information which has been related in some manner, such as by association or categorisation. The type of structure identified will depend upon the prevailing (mainly cognitive) context.

The final information type; the Descriptive component, is obtained from the existing semantic memory representation of the item. Two things determine exactly what will be taken from the whole semantic memory representation. First, as with the structural component, control processes will determine the extent and complexity of information process to be entertained and secondly, the context (again mainly cognitive) will guide which aspects of the complete semantic representations will be written into the new representation.

At retrieval, cues provide access to the memory traces. A retrieval cue increases its effectiveness as its encoding resembles that of the memory trace. The ability of the multicomponent retrieval cue to better activate the multicomponent memory trace as the correspondence between the two increases is tempered by a decrease in this ability, as the number of traces which contain any particular component increases. Glenberg also points out that it is possible for traces retrieved on the basis of one information type to act as cues for retrieval by any other information type.

7.3. Relation to environmental context.

This account makes specific provision for different types of information and as a consequence is able to accommodate the data from the investigations of environmental context using recall, cued-recall and recognition measures. The notion that a cue's effectiveness is inversely related to the number of traces it is a component of also provides an account of the benefit of multiple learning contexts.

The data obtained from the self-generation of environmental context presumably would be accommodated by Glenberg's "component levels", in a manner similar to that of Norman and Rumelhart's model. Descriptive information would be associated, or would categorise (structure) the nominal stimuli information as having been presented in the particular environmental context. The provision of a cue corresponding to that descriptive information associated with the nominal stimuli items should make this information available. In common with Norman and Rumelhart's model, context information, but in a more semantic form, ie. descriptive rather than contextual, is used to access the nominal stimuli items. This can be done in two ways. In a manner similar to Norman and Rumelhart's model, the descriptive component could provide access to low level context information about the environment which would be checked against similar information encoded with the nominal stimuli items. Alternatively, it could be assumed that the form of the context associated with the nominal items at encoding, was of a similar high level which would allow a direct comparison with the cued descriptive information. This latter method cannot be applied in Norman and Rumelhart's model as the only way to record an item such as an environment is via the attributes and their associated contexts, and the item can be recomposed only after context has cued the attributes. Always, some context is needed to identify the environment the presentation occurred in, whereas in Glenberg's account the high level context descriptive information can be the cue as well as the information associated to the nominal items in the memory trace.

As Glenberg explicitly states that there is a change in the components stored in the memory trace of nominal items over time, it is quite simple to accommodate the requirement of a rate of change of context to provide the basis for an account of the temporal judgement effect. The only proviso in such a system is that there should be other components in the trace of "doing the nominal activity" or "being in the particular place" that can be activated by a cue containing corresponding components.

7.4. Appraisal.

In common with Tulving's encoding specificity, Glenberg utilises predictive hypotheses rather than explanations to account for observed effects. For example, the assumption that the number of "traces" that a cue is utilised in reduces its overall effectiveness to elicit any of those traces is a hypothesis which has empirical support (eg. Watkins & Watkins, 1975).

However, the hypothesis is presented in terms that require explanation; how does this effect manifest itself? The inclusion of such hypotheses in terms of the result some situation produces, rather than in terms of how such an effect comes about, can increase the ability of the model to predict the outcome of events, but not explain how such outcomes arise.

The model can be criticised on the basis that judged by the criteria of the terms of abstraction of representation and algorithm, it does not describe sufficiently <u>how</u> the processes operate. However, Glenberg's account also can be criticised for failing to suggest <u>what</u> it is trying to achieve. This relates to the point concerning the ecological utility of psychological processes. It is difficult to believe that the purpose of any process(es) is to fulfill the task demands made by experiments on repetition effects, but this is the apparent reason for proposing this theory.

What Glenberg has done is identify a variety of important factors that influence memory encoding and retrieval. What he has not done is to describe the processes that give rise to these factors.

8. Direct Attempts To Explain Environmental Context Effects.

8.1. Introduction.

The experimental investigation of the relationship between environmental context and memory has produced a variety of phenomena to be accounted for. Despite this, very few attempts have been made to provide psychological explanations for these effects, beyond reference to views of forgetting founded in that presented by McGeoch (1942).

8.2. Smith's cueing account.

Although Smith, Glenberg and Bjork (1978) attempted to identify the locus of the environmental context effects in terms of a non-specific generate and recognise model of retrieval, this was but appended to a view of memory and memory formation, as an almost indiscriminate associationist operation; a view little changed since the days of Ebbinghaus. This "theoretical" approach to environmental context effects in memory has continued through Smith's research and although similarities between these effects and other psychological phenomena have been identified, there has been no elaboration beyond this. Overall, the general account of environmental context effects presented by Smith is that environmental context can act as a retrieval cue.

8.3. Godden and Baddeley's relative cueing account.

In their later paper, Godden and Baddeley (1980) present an account of the lack of environmental context influence on recognition which requires assumptions about the way in which information is encoded. Essentially Godden and Baddeley argue that as there is only an arbitrary relationship between environmental context and the nominal stimuli, it will not influence the way in which this information is interpreted. In recognition, as so much information is provided by the to-be-recognised item, the extra information provided by the arbitrary relationship with context will exert no effect. In recall, however, where much less information is provided to identify the to-be-remembered items in memory, even the small amount of information provided by the arbitrary relationship with environmental context will aid retrieval and exert an effect.

Since then, Baddeley (1982a) has refined his ideas, introducing the notion of processing domains (cf. Craik & Jacoby, 1979) and rather than considering context as intrinsic or extrinsic; which he considers to lack the connotation of being a psychological consequence, he now distinguishes between interactive and independent context. Baddeley accounts for the discrepant effects of environmental context on recall and recognition as a consequence of two factors. The first is that nominal items and environment are encoded independently and in parallel. However, the relationship established between the two can aid access to the nominal item and so benefits recall. It is assumed that recognition presents a problem of determining previous occurrence, rather than one of access. This process is not influenced by the relationship between the nominal item and the independent context and so no environmental context effects are observed with recognition measures.

8.4. Appraisal.

The model of memory underlying the account of environmental context effects presented by Godden and Baddeley is superior to that implied by Smith in that explicit acknowledgement is made of the interactive nature of memory performance. All memory effects are by definition a consequence of the encoding, storing and retrieval of information. Any model of information-processing which attempts to account for memory performance will need to reflect this interactive feature.

However, the accounts presented by Godden and Baddeley, and Baddeley fail to meet the criteria required of a process model and in this respect have much in common with the style of Tulving's account and the levels of processing presentations (eg. Craik & Jacoby, 1979; Jacoby & Craik, 1979; Lockhart, Craik & Jacoby, 1976). Baddeley briefly considers the nature of his account and argues that its purpose is to

> provide a coherent framework to assist in the understanding of a wide range of phenomena (Baddeley, 1982, p.726).

A similar view also has been forwarded by Norman (1979), but to support the argument for more process oriented accounts. Doubtless, the type of descriptive accounts criticised do provide a sort of understanding, but the danger is that this is achieved by the application of terms of abstraction that obscure the unique features of the processes described. It is the unique features of a process that makes it possible to determine why that process and not another was implemented.

Although the concept of recollection as introduced by Baddeley (1982) would seem to be useful in accounting for some of the other environmental context effects, it is not specified enough to enable it to be distinguished from alternative methods of achieving similar goals. However, as an identification of a psychological phenomenon that has implication for the types of processes suggested to achieve remembrance, it is well overdue.

One change in the account presented by Baddeley (1982), to that presented by Godden and Baddeley (1980), is the adoption of a "two-process" view of recall. Unfortunately, this may be a detrimental amendment, as is illustrated in chapter eleven, when the implications of such "two-process" views are considered and relevant evidence is presented.

CHAPTER SIX

A SCHEMA BASED ACCOUNT OF ENVIRONMENTAL CONTEXT PHENOMENA

1. Schema Model Of Psychological Processing.

1.1. Introduction.

All of the theoretical conceptions considered in the previous sections have been unable to give complete accounts of the environmental context phenomena identified in chapter two. One of the fundamental criticisms has been that the type of description of psychological mechanisms has been (and would be) insufficient to provide what has been classed as an explanation. The requirement to provide an explanation in the form of a process model is a hard standard to meet. Although there is a difference between formal models; which attempt to describe in terms of the process criteria, and in-formal models; which describe the ideas leading to the formal model, it cannot be expected that either form of the ultimate process model is to be found in the present literature. Although the majority of process models are still at the informal stage of development, what seems to be important is that the science of information-processing is paid heed to and that both formal and informal models are conceived in accordance with this available knowledge.

1.2. The schema system.

In a series of articles, Norman and Bobrow (Bobrow & Norman, 1975; Norman, 1979; 1982; Norman & Bobrow, 1976; 1979) have developed a view of psychological information-processing which conforms to this approach. Of course they are not alone in producing such a view, but their model has developed from a general view of human performance, enabling it to provide a framework for more specific examinations of particular aspects of psychological processing.

They suggest that the processing system is composed of an autonomous collection of schemata. Each schema is an independent processing structure which can communicate with the other schemata. Information is processed in both top-down and bottom-up fashions, with constant feedback regarding the state of schemata processing. Processing is an active operation in that schemata "compete" for information. The schema with slots which best "fit" the data available at that time takes precedence and provides a continuous output of partial results to other schemata as its processes operate. As the whole system operates in this manner there are no stages of processing as envisaged by earlier types of psychological model. The schemata determine the direction of the information processing by virtue of the suitability of subsequent schemata for the outputs of previous schemata.

The notion of schemata is related to other conceptions of knowledge structures mentioned earlier. Although there has been some debate over the fundamental type of psychological representation (eg. analogue vs propositional; place vs additive) and the importance of such distinctions (eg. Palmer, 1978), the point has been made that different types of representation are feasible, provided there are suitable interfaces between these representations (Bobrow, 1975). The terms in which schemata are discussed do not distinguish between the types of representation that could be employed, but rather describes the architecture of the system, and the goals and requirements of those processes involved in particular aspects of psychological performance.

As psychological processes are called from and make continual use of other memory information, an important consideration in any model system should be the manner by which the relevant information is obtained from the vast quantities available. For example, perception depends upon the integration of new sensory information with existing knowledge structures (ie. schemata). However, the structures appropriate to integrate this information must be identified and selected from all the information that constitutes memory. In effect, there is a memory retrieval problem.[1] Norman and Bobrow have suggested that the transfer of information occurs through the formation of descriptions. A description is an intermediary in memory retrieval.

One of the principles of operation of the system is an ability to be vague or precise in the specification (via the description) of an idea, event

---

1 Sanford (1985, p.95-98) provides further examples and details of psychological processes that depend upon memory information.

or item. The degree of specificity necessary is dependent upon the purpose of the description and the form of other descriptions. For example, if there are five boxes of which four are made of wood and one is metal, it is possible to identify the metal box with a "vague" specification. It can be identified as "the metal box". In contrast, if one of the wooden boxes is to be identified then the specification has to be more precise than just "the wooden box". Perhaps colour, shape, size, or location, could be included to identify the intended object. As there is a continuum of specification possible, the task requirements will determine the lowest level of specification that can be usefully employed.

1.3. Memory records.

A description of an entity can be a single, or collection of perspectives. A perspective is a particular view of an event, idea or item. A perspective is formed by identifying the significant fields of a prototype, while a prototype is itself a previously existing schema.

The level of specification of a description is determined by the choice of a prototype (and necessarily its perspective) and further specification of the differences between the entity and the perspectives employed in the description (see figure 6.1.).

In a fashion similar to Minsky's (1975) frames, where each slot in the frame is potentially a sub-frame, so Norman and Bobrow's system is

```
 ------------------- DESCRIPTION -----------------
 |                        |                        |
 |                        |                        |
PROTOTYPE #1          PROTOTYPE #2   ....      PROTOTYPE #k
 |                        |                        |
 |                        |                        |
 |                        |                        |
PERSPECTIVE           PERSPECTIVE              PERSPECTIVE
 |                        |                        |
 |                        |                        |
  further                further                  further
specification          specification            specification
```

Figure. 6.1. Diagrammatic representation of description structure.

organised with each description potentially the component (ie. prototype schema) of another.

This is the structure of what Norman and Bobrow term a memory record. Records are regarded as being the basic unit of memory and have the property that access is to all of the record, or none of it.

Rumelhart and Norman (1978; 1981; 1982) have described different modes of learning which could give rise to the type of memory record outlined above. This is an area requiring much more research and theoretical development. It is also a topic which is outside the scope of this thesis.

1.4. Memory retrieval.

Norman and Bobrow (Norman, 1982; Norman & Bobrow, 1979) identify four aspects of memory retrieval, which they are quick to point out, are not considered to be stages. They regard retrieval as a cycle of continual specification, matching, evaluation and if necessary, fault diagnosis. This cycle refines the description of the desired information with respect to the product of evaluation until the information has been retrieved, or the whole process terminated (see figure 6.2.).

ASPECT	USES AS INPUT	PROVIDES AS OUTPUT	COMMENT
Retrieval specification:	Purposes, needs, descriptions.	Target description, verification criteria.	May require memory retrieval.
Match:	Target description	Memory records.	
Evaluation:	Memory records, verification criteria.	If success, terminate. If failure, diagnose.	May require memory retrieval.
Failure diagnosis:	Information from evaluation process.	Revised retrieval specification.	

Figure. 6.2. Four aspects of memory retrieval (concatenated from Norman, 1982 and Norman & Bobrow, 1979).

(i) retrieval specification: this consists of two types of information; the target description and verification criteria. When a description is formed often it is not known what form other characterisations in memory have. In addition, the characterisation in memory can be desired in order to fill a gap in knowledge, so the specification of that characterisation has to be based on information other than an explicit description of the entity. In such a case the specification has to be formed from information such as why the description is required and what seems to be consistent with that which is already known. The provision of information to form a retrieval specification may itself require memory retrieval.

(ii) matching: this label covers both the access to candidate records and the selection of those which correspond to the target description. At present there are a variety of candidate methods by which this may be accomplished. However, the form of abstraction of this model is not concerned with the detailed description of such processes, rather this model stipulates the goals and restrictions on any implementation of this process.

(iii) evaluation: the verification criteria formed by the retrieval specification are applied here. Williams and Hollan (1981) describe the verification procedure as an attempt to identify those records that are wrongly retrieved and to lower the probability of distortion by inference. As performance differences such as between recall and recognition would suggest, they describe a variety of ways in which retrieved memory records may be verified. One of the phenomena observed by Williams and Hollan was overshoot. This is where a person continues the retrieval process in an attempt to assess the validity of a candidate memory record. This suggests that if necessary, the evaluation procedure may also employ memory retrieval.

(iv) fault diagnosis: if a retrieved record fails to meet the verification criteria it will still be similar to the desired memory record. In such a case the reason(s) for it failing to meet the verification criteria will contribute information to the formation of a new retrieval specification.

There are two factors which determine the success of memory retrieval. Constructability refers to the ability to create an appropriate description to identify a particular memory record. Discriminability refers to the ability of an appropriately constructed description being able to discriminate from all records in memory, its target record. Therefore, the records retrieved from memory depend on the form of the retrieval specification created at the time of retrieval and the form of the encodings of other records in memory.

1.5. Environmental context, memory records and descriptions.

In the terms of Norman and Bobrow's model, what a subject has to do in order to be able to remember some idea, item or event is to construct an unambiguous description to be stored as a memory record. That environmental context influences the retrieval of information suggests that it is incorporated in the memory records formed through descriptions at encoding. There are a variety of ways in which this could occur. Environmental context could be represented by the schema which is used to perceive the whole environment. In other words, a high level schema could identify the nominal stimuli items as having been presented in that environmental context. Alternatively, low level schemata; representing certain features of the environment, could be incorporated in the descriptions of these items. This is comparable with the type of representation of nominal stimuli items suggested by most other models employing the notion of context.

These are the two basic ways in which environmental context could be utilised in the formation of a description of nominal stimuli items. However, the nominal stimuli items themselves have to be described at some level. Items can be described in relation to each other; producing associative effects, and in terms of higher order schemata; producing category effects. Although this raises the possibility that environmental context information could be employed as a component in a nominal item description at a variety of different levels, the results of Smith, Glenberg and Bjork (1978, expt.3) suggest that environmental context is represented (at least) with each individual nominal item.

The environmental reinstatement effect with free recall is accounted for by assuming that whatever the form of context utilised to describe the memory record, it has a greater chance of being used similarly in the retrieval specification when the subject is in a situation where this information is available ie. the learning environment. The smaller context effect with cued-recall and apparent lack of effect with recognition, is accounted for by considering the form of the retrieval specification that could be constructed in these situations. In recognition, virtually all of the information presented at encoding is re-presented. Therefore, the retrieval specification has a greater chance of being appropriately constructed. The explanation of cued-recall is identical, but with the acknowledgement that there is less correspondence between the information presented at learning and test with the cued-recall paradigm, than there is with the recognition paradigm. The greater context effect with cued-recall is what would be predicted.

The data on self-generated context is handled very easily by Norman and Bobrow's model in comparison with all the other models considered. This is because Norman and Bobrow's model includes the translation of external information to a form used by the processing system and details the way in which this information is communicated throughout the system. Consequently, the explanation of the data on the self-generated context is similar to that of context reinstatement with free recall. A retrieval specification to try and identify a memory record containing information about the learning room would be constructed. With a successful retrieval this information can then be used to form retrieval specifications to identify the nominal items in memory.

Multiple environmental context learning effects also are quite easily coped with. If similar descriptions are formed for several items, discriminability is reduced. This could happen if the same environmental context information was employed in the descriptions of many items. The consequence would be a reduction in the number of items able to be recalled. One way of preventing such an occurrence would be to increase the number of contexts, so reducing the likelihood of environmental context information overlap in the descriptions formed. Provided construction of retrieval specifications are appropriate, the increased discriminability should allow an increase in recall. This explanation assumes the same type of processes that will be suggested to explain the inverse relationship between the extent of "cue" use and its effectiveness (see section 2.1.).

The nature of the model described by Norman and Bobrow provides a basis for an account of temporal judgement effects. As context information is contained within a description and descriptions are decomposable, it should be a simple operation to provide the necessary input to the processes which determine duration.

2. Further Applications Of The Model.

2.1. General

One feature of Norman and Bobrow's model is the suggestion that much of the effort involved in psychological processing, such as perception, language and thinking, is in an attempt to obtain and provide for unambiguous specifications of memory records. The initial stage of this process is achieved by mapping information into knowledge structures as discussed in chapter four (section 4.2.). Bobrow and Norman (1975) have already claimed that another principle of psychological processing is that all data must be accounted for. All signals received must be processed at some level. This bottom-up, data driven analysis need not provide specific descriptions. More specific descriptions will be required if the data is unexpected: outside of the default values of the predicted schema to the extent that extraordinary further specifications (excuses) are required, or if the incoming data are determined as relevant to the schemata involved in the central analysis, ie. that currently using most processes.

As Norman and Bobrow's model is process oriented; many of its ideas having been implemented in KRL (Bobrow & Winograd, 1977), it is capable of providing explanations of those predictive hypotheses which were criticised earlier for masquerading as explanations. The data supporting the encoding specificity hypothesis is explained in terms of constructability. In cases where an item that is known to be available to the subject, but has not been recognised, an inappropriate retrieval specification is assumed to have been employed. In contrast, the inverse relationship between the frequency of particular information used in "memory traces" and its effectiveness as a cue is a consequence of the reduced discriminability caused by similar or overlapping descriptions.

In a similar fashion interference effects would be expected if different items are encoded using similar types of description. Interference effects also could be a consequence of poor constructability. If retrieval specifications are not specific enough they will not be able to identify the memory record uniquely.

The later expression of the levels of processing framework (Craik & Jacoby, 1979; Jacoby & Craik, 1979; Lockhart, Craik & Jacoby, 1976) contains many ideas similar to those presented in Norman and Bobrow's model. In particular, there is a correspondence between the notions of level of elaboration and domain, and level of specification. However, again the advantage is with the Norman and Bobrow model because it is process, rather than "descriptively" orientated. This allows the "levels of processing" account to be accommodated by the schema model. One consequence of this is that effects such as the interaction between level of processing and encoding specificity (Fisher & Craik, 1977) can be explained parsimoniously.

More recently, Morton, Hammersley and Bekerian (1985) have employed very many of the aspects of the schema system presented by Norman and Bobrow (1979). However, for some odd reason they precede their model's description by falsely criticising such schema models for being unable to give an account of why almost everything about a person can be remembered, except their name. The model presented by Morton et al. also seems to have a tendency to regard memory records as passive entities. This view is reminiscent of the inappropriate analogy of "memory as a warehouse".

2.2. The amnesic syndrome.

One reason for interest in environmental context effects is the relationship between these effects and the amnesic syndrome (see Weiskrantz, 1978). Both human and animal experiments have shown a striking dependence on environmental context in subjects with bilateral hippocampal damage. However, given the grand nature of the deficit in human performance with such damage, it has been surprisingly difficult to produce similar performance decrements in animals by hippocampal destruction. As a result several reasons have been suggested for the apparent discrepancies. For example, Horel (1978) has argued that hippocampal damage is not the cause of the amnesic syndrome;[2] the possibility that the hippocampi are involved in different psychological operations in humans compared to animals has been raised; and also that the reason for so few deficits in animals' performance is because the type of tasks they are presented with are not equivalent in the nature of their psychological requirement to those presented to humans (eg. Iversen, 1976). The latter argument is usually followed by a demand for more formal analysis of task requirements, but one problem with this is that the result of any such analysis depends upon the particular model of psychological processing applied.

The reasons for the apparent dichotomy between human and other animal performance with hippocampal damage are not mutually exclusive, but if the direction of research in this area has been guided by any one of these arguments it would have to be identified as the analysis of a task in terms of its psychological requirements. A great deal of effort has gone into the identification of tasks which can and cannot be carried out in a normal manner by hippocampal subjects: humans and other animals. On this basis, researchers in the area have developed models which attempt

---

2 Horel (1978) argues that amnesic effects attributed to hippocampal damage are actually a consequence of severing the temporal stem which carries afferents and efferents of temporal cortex and amygdala, but not hippocampus. In human subjects the most detailed psychological examination of deficits has been carried out on amnesic patients who are still alive. Presently therefore, it is not possible to obtain detailed and exact determinations of locus and extent of damage. Animal data is more available, but debate has continued as to whether the hippocampal damage alone causes the effects, or if accidental temporal stem damage could be the cause.

to discriminate between the ability to do some tasks and the inability to perform normally on others. With respect to a desire to understand the psychological processes underlying environmental context effects, these are potentially interesting and important models because they have to account for extreme environmental context dependency.

2.3. Environmental context effects as a consequence of hippocampal damage.

Winocur and Olds (1978) have demonstrated that animals' performance on tasks previously thought to be immune from the effect of hippocampal damage is significantly impaired, if at test the environment is changed. In addition, it was shown that tasks previously affected by the hippocampal damage could be maintained at a normal level if the environmental context in which the task was learned and tested was made distinctive enough (Winocur & Bindra, 1976; Winocur & Olds, 1978; Winocur, 1980).

A similar type of effect has been shown with human amnesic patients. Winocur and Kinsbourne (1978) presented paired-associate words to amnesic subjects in a very distinctive environment. In this condition subjects asked to recall the latter word; given the former as a cue, could perform at the same level of competence as normal subjects. Normal subjects did not show any improvement between normal environment learning and recall, and distinctive environment learning and recall. However, some improvement in normal subjects' performance in the distinctive compared to normal condition may have been obtained if a free-recall task had been employed. It is worth noting that the only type of verbal memory task which amnesics appear to have any degree of success with involves the provision of a (partial) copy cue (see also, Warrington & Weiskrantz 1968; regarding copy cues, see chapter one, section 4.3.)

2.4. Models of hippocampal function

In attempting to account for the possible functions of the hippocampus, a great number of models have been generated. Many have been abandoned over the years including Milner's (1966; 1968) consolidation hypothesis, Warrington and Weiskrantz's (1968; ;1970; 1973) retrieval interference hypothesis, while Gaffans (1972; 1974) familiarity hypothesis seems to have been ignored rather than rejected.

An edition of the nineteen-eighty volume of Physiological Psychology was devoted to hippocampal function. This edition carried papers from many workers researching in this area and provided a good summarisation of the different approaches being applied and the psychological models being developed. Roughly four different views are evident and will be presented briefly here.

Thompson and Berger (Thompson, Berger, Berry, Hoehler, Kettner & Weisz, 1980) apply what they call a "model system" approach to the study of the brain subtrates of associative learning. Employing the intact mammalian brain and selecting a "prototypic" behavioural system having the properties of associative learning which are well defined and characterised that exhibits robust learning, they adopted the approach of recording neuronal unit activity during the course of learning. Their goal is to characterise the activity of various brain systems in learning and memory. They believe that once this is accomplished, the structures and systems that exhibit altered activity with learning will have been identified and analysis of synaptic mechanisms will be feasible. All major brain systems must be explored, but due to the extensive literature on hippocampal involvement in learning and memory, they regarded it as a good place to begin. In order to investigate this structure in the manner described above, they decided to use the classical conditioning of the rabbit nictitating membrane (NM) response to a tone-conditioned stimulus (CS) using a corneal airpuff unconditioned stimulus (UCS), while recordings were taken from chronically implanted micro-electrodes. Both single and multiple unit recordings were made.

This type of work and the data accumulated from it has suggested to the researchers that the hippocampus is involved in registering the temporal sequence of events. Typically, the hippocampal activity measured in the classical conditioning paradigm precedes the behavioural response by 35-49 msecs. This relationship evolves during the first eight CS-UCS pairings (Berger, Alger & Thompson, 1976). Berger, Clark and Thompson, (1980) have recorded from other limbic structures connected to and via the hippocampus and conclude that the electrical activity does not originate in the hippocampus, but from the entorhinal cortex or its afferents (this includes other cortex). What the hippocampus does is to amplify this output from the entorhinal cortex relative to spontaneous activity rates. These results (plus others) are consistent with a view of the limbic system modulating sub-cortical brain mechanisms critical to conditioned behavioural responding.

Solomon (1980) takes the temporal position further. He argues that the hippocampus uses this temporal information to allow the animal to ignore stimuli which do not uniquely predict a change in the probability of the occurrence of the UCS. In turn, this allows the animal to attend to stimuli which do signal such change. This view; shared by Moore and Stickney (1980), is in accord with an experiment by Devenport and

Holloway (1980). They found that hippocampal animals displayed a large amount of superstitious behaviour in comparison with the controls. From this they suggested that the hippocampus is involved in the detection of contingency and without it the animal relies on contiguous relationships. This suggests that the hippocampus does more than register the temporality of events. Indeed it must, as without the structure this function is unhindered. It seems likely in fact, that the observed temporal relationships may have more to do with the paradigm employed than any special function of the hippocampus. It is difficult to think of any function which the hippocampus could fulfill in learning, which would not produce similar temporally related data if examined in this way. So, apart from achieving a greater understanding of the neural pathways conveying information into and out of the hippocampus, all that can be taken from the classical conditioning-electrophysiological studies is that the hippocampus may be involved in identifying relevant stimuli.

A far different approach to the study of hippocampal function is that taken by O'Keefe and Nadel. They take what they call a neuroethological approach. That is, they study electrophysiological recordings and other physiological interventions of the hippocampus and their effect upon behaviour in more natural tasks as compared to the usual experimental paradigm. An example of this arrangement would be an artificially created environment with objects lying in particular locations. The animals (rats) would be taught to obtain water from a particular place. This place would be defined in terms of its relationship with the other objects. Therefore, the place need not always be in the same point in space, but would always be in the same point in relation to the other movable objects. Alternatively, the goal point could be identified by a cue, such as a light above it. O'Keefe and Nadel have demonstrated that hippocampal animals are incapable of locating the goal if it is defined in terms of the spatial relationships between the other objects. They have also recorded increased activity from the hippocampus when the animals utilise this strategy to locate the goal. O'Keefe and Nadel (1978) theorise that the hippocampus forms a cognitive map of space using the objects within that space. Nadel and Willner (1980) take this one step further in claiming that environmental context is superordinate to its components. It cannot be regarded as a compound of component stimuli, but instead is a greater whole. They see the cognitive map of the environment produced by the hippocampus as the receptacle for other information. Nadel and Willner therefore claim that environmental context contains, and presumably through some sort of association, predicts events. This is an interesting view and the thought of environmental context acting as some sort of schema to which other information could be attached has already been mentioned. Unfortunately, it does not explain why humans and animals become more environmentally context dependent when the supposed schema producer is destroyed.

O'Keefe and Conway (1980) reported a new memory task for rats called the "despatch task"; the evidence from which provides good support for the cognitive map theory. The rats are taught a spatial task. The goal is in relation to cues, with a different location depending upon the arrangement of the cues. In the test situation, the animal is allowed to see the cues from a confined position. The cues are then removed and after a delay, often of half an hour, the animal is released to locate the goal. Hippocampal animals (ie. fornix lesion - cuts output from hippocampus) cannot perform correctly on this task while normal controls can. However, hippocampal animals can perform on the cued task, when all the cues are laid beside the goal. However, neither control nor hippocampal animals can seem to maintain the memory over time when this latter cued learning task is tested. Both normal and hippocampal animals can learn to locate the goal when it is cued, as long as the cues remain. If the cues are removed, the animals from both groups cannot locate the goal. It seems that this type of task does not produce an effective "long term" trace. O'Keefe and Conway admit to having no explanation for this.

Another proponent of the spatial function of the hippocampus was David Olton. Using a radial eight arm maze and recording from single units in the hippocampus, Olton accumulated evidence which seemed to suggest that there were cells of the hippocampus which responded only to particular spatial aspects of the environment. However, Olton has changed his mind regarding the function of the hippocampus. He now believes that the behaviours associated with hippocampal damage etc. are indicative of a working memory deficit.

Olton (Olton, Becker & Handleman, 1980) defines working memory to be that used in experimental procedures when information on any single trial is useful only for that trial. Reference memory is tapped for information which is useful for all trials. All behavioural tasks have a reference memory component, whereas only some will have a working memory component. Olton claims that hippocampal animals are impaired on working memory type tasks only. He also claims that impairment is found in working memory types of tasks irrespective of whether the task can or cannot be performed using a cognitive mapping strategy.

Olton's analysis and that of Hirsch (1980) share certain similarities. Hirsch argues that the hippocampus is involved in what he calls the conditional retrieval of information from memory. Conditional operations allow two (and presumably more) highly similar situations to be treated quite differently. Conditional probabilities are an example. The statement that given Y_1, $p_a = 0.5$ is in no way a contradiction of the one that given Y_2, $p_a = 0.02$, despite the fact that p_a is the same entity in both cases. As retrieval is regarded to be the selection of one piece of information from among many possible candidates, and as many things may be known about a given item, retrieval will be to a greater or lesser extent a conditional operation. The process of conditional retrieval is then defined as the result of interaction between a categorical operator representing global factors, and a local operator representing local considerations. Interaction between these two types of operator is said to be necessary for constructing dimensional representations, including maps. Hirsch presents a convincing case for this model, which seems able to account for the majority of the available data. Unfortunately, he does not particularly explain the processes by which the local and global operators work. The similarity with Olton's ideas is in the fact that both writers seem to be describing systems which are either dependent upon or select information which is quite specific in nature.

The final view to be presented is that of Gordon Winocur. He regards hippocampal animals as being limited in their ability to extract information from their environments, and as a result are more vulnerable to the effects of interference when relevant stimuli are spatially or temporally disassociated, or rendered ambiguous by conflicting experiences. Winocur believes that the hippocampal syndrome reflects a basic processing deficit. In this respect his view is more general than others and therefore probably is less incorrect.

2.5. General assessment of amnesic syndrome models and data.

The data presented by the different theorists certainly support their own contentions, but some have little power to account for other hippocampal effects. Thompson and Berger et al.'s work is more of an attempt to describe the physical realisation of processes rather than the representation and algorithm, or the computational theory. Consequently, the other descriptions of processes presented in other terms of abstraction are not in competition. However, when this essentially correlational paradigm is used to provide causal accounts in other terms of abstraction, a certain dubiety must be attributed to their conclusions. One of the difficulties associated with this approach may be appreciated by distinguishing the terms of abstraction involved, and considering an analogous situation: would it be possible by observing the changes of current in a digital computer, and its output of numbers, to ascertain that it was performing an analysis of variance, if one did not already comprehend the underlying nature and purpose of such a process?

Overall the data and models from the hippocampal literature seem to converge on the assessment that tasks which require a large amount of specificity are impaired. Hirsch almost describes the Norman and Bobrow model when he talks of global operators (prototype-perspective) and local operators (further specification). Olton's definition of working memory type tasks is really a statement of tasks which need a fairly high degree of specificity to be carried out correctly. The analysis given by Winocur is what one would expect if the deficit was an inability to properly or adequately specify the task requirements.

The work of Thompson and Berger et al. also could be accommodated by this viewpoint. It was suggested that what their work had shown was hippocampal involvement in some type of attentional process; the identification of pertinent stimuli. This would be necessary in order to specify the task requirements. As was mentioned earlier, it is difficult to attribute any particular functional aspect to the hippocampus on the basis of this data, but the information obtained from these micro-electrode-classical conditioning studies certainly is compatible with the function proposed.

O'Keefe and Nadel's explanation of hippocampal function in terms of cognitive mapping has a great deal of supportive evidence. However, it is likely that if the hippocampus is involved in the formation of unambiguous memory record descriptions, then damage to this system would also effect the ability to form spatial representations. Therefore, the deficit in spatial ability can be regarded as a consequence of impairing the processes which contribute to establishing a cognitive map, rather than being the origin of the deficit.

3. Appraisal Of, And Predictions From The Schema System.

3.1. Framework and process.

Probably the most important aspect of a schema based system is that it provides a structured and powerful framework within which mental activity can be considered. However, it is one thing to accept the notion of "packages of processes" operating on sensory input and schemata output, thereby providing information and knowledge, and another to be able to describe the processes which carry out such tasks. It is likely that by attempting to understand the processes upon which the gross schemata operations depend, our comprehension of the nature of schemata will change. It even may become evident that schemata are, as Schank (1981) has suggested, momentary creations of underlying psychological processes; entities which in themselves are little more than processing artifacts. At present however, our level of theoretical sophistication makes the conceptualisation of schemata a necessary rather than just a useful construct.

3.2. Encoding and retrieval interactions.

One feature of the schema system which contributes to the power of the framework is its emphasis on the variety of different types of information encoding and retrieval that are possible. This is an important feature of psychological processing which must be borne in mind when consideration is made of the way in which any effect is manifested.

The ability to remember; which is a continually ongoing activity serving most other psychological processes with information, depends upon discriminability and constructability. Consequently, the factor that first sets a limit on the ways in which information can be retrieved is the nature of its encoding. Only information which has been incorporated into an entity's description can be effective in retrieving that memory record. However, it is possible that a salient part of the target memory record, which could discriminate the description from others, would be ineffective. This would occur if in the construction of the retrieval specification the pertinent piece of information had been omitted.

3.3. Encoding and retrieval interactions in relation to environmental context effects.

With respect to environmental context effects the obvious and first requirement for an effect is the inclusion of some sort of environmental context information in the memory record. Subsequently at retrieval, similar environmental information must be used in the construction of the retrieval specification and this information must be utilised in the retrieval of the target memory records. It is possible to conceive of a situation where although similar environmental context information was present in both target and retrieval specification, some other information was actually effective in identifying the sought after record.

As the attempt to retrieve information is based on a disambiguation of one memory record from others on the basis of retrieval specification, it could be possible to have more than one disambiguating retrieval specification of the same memory record. The number of different, but unique retrieval specifications of the one memory record will depend on the detail in which the entity was encoded in the attempt to achieve an unambiguous record.

Norman and Bobrow (1979) have argued that with lower "levels of processing" both constructability and discriminability are likely to be low because of the comparatively small number of unique descriptions which can be created from the relatively small number of features (identified by schemata) at this level. With higher levels of processing more schemata are available and so a greater number of unique descriptions should be possible, resulting in increased discriminability and constructability.

Apart from the phenomena of increased recall and decreased interference with "higher level processing", another consequence of the greater number of different, yet unambiguous retrieval specifications is a reduction of dependence on any particular unique retrieval specification for remembrance.

3.4. Predictions.

In relation to environmental context effects, this line of theorising has certain predictions associated with it. If it is assumed that environmental context information's inclusion in memory records is a robust aspect of processing, then a reduction in the alternative detail of descriptions formed at encoding should increase the dependence of retrieval specifications upon disambiguation by environmental context information. Conversely, an increase in the alternative detail of descriptions formed at encoding should decrease the dependence upon disambiguation by environmental context information. In the former case one would expect a comparatively large environmental context effect and in the latter case if the alternative retrieval routes were sufficient, no environmental context effect should be observed.

In the next two chapters experiments which specifically examine these predictions will be presented. Subsequent chapters report studies which continue what is basically an exploration of the processes through which environmental context exerts its influence. As this is carried out, constant reference will be made to the model of psychological processing outlined in this chapter in an attempt to obtain a theoretical, as well as an empirical grip on the phenomena.

CHAPTER SEVEN

EXPERIMENT TWO

THE INFLUENCE OF LOW LEVEL SEMANTIC ELABORATION ON THE ENVIRONMENTAL CONTEXT REINSTATEMENT EFFECT.

1. Introduction

1.1. Environmental alterations.

In chapter three the requirement for effectively different environmental contexts was outlined. The observation of an environmental context effect indicated that effectively different environments had been created. Although there was some discussion of the psychological processes which could have been involved in the production of this effect, ultimately these depend upon there being differences between environments to provide the basis for differential operation.

Unfortunately, just after this experiment was completed, notice was given that the University was to begin major construction work very close to the building where the room environments were located. The major problem this created was noise. In an attempt to eliminate this, alterations to the room environments were necessary. The most significant alteration to the appearance of the room environments was the boarding-up of the window in Room A. All in all, the result of this and other slight changes was to make the two main environments: Rooms A and B more alike.[1]

The problem caused by any change in environmental context is determining the extent to which the findings of experiment one can be assumed to extend to the new conditions. The particular problem caused by these changes is that the increased similarity of the two environments may preclude any effect.

1.2. Aspects of memory performance.

In the preceding chapter it was predicted that "low level" processing of nominal stimulus items would produce greater environmental context dependency. "Low level" processing was taken to be descriptive of operations which did not produce detailed encoding or retrieval specifications. However, the account of retrieval processes presented in chapter six concentrates on only one aspect of encoding and retrieval: the need to uniquely identify a memory record. The way in which the construction of records and the processes that retrieve them, provide and obtain information to allow disambiguation was largely ignored. It is to the relationship between this latter aspect of retrieval and memory performance that attention is now turned.

---

1 Detailed descriptions of the room environments discussed are presented in appendix A.

1.3. Memory organisation.

Mandler (1979) has suggested that there are two independent dimensions of organisation: integrative and elaborative. Integrative organisation refers to intrastructural or within-item organisation and elaborative organisation refers to interstructural or between-item organisation. The two dimensions are illustrated in figure 6.1. below.

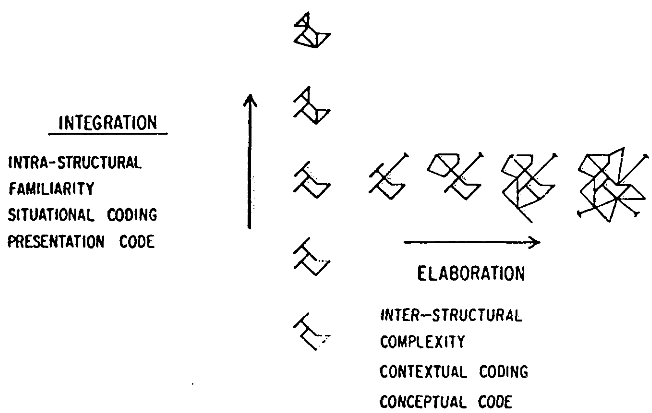

Figure 6.1. A graphic analogy of integration and elaboration (from Mandler, 1979).

The processes which give rise to the types of organisation identified by Mandler seem to be enhanced by different types of rehearsal: maintenance or primary rehearsal increases intrastructural organisation, while elaborative or secondary rehearsal increases interstructural organisation.

In relation to memory performance, it seems that the types of processes underlying maintenance rehearsal which give rise to a highly integrated memory record enhance familiarity judgements (Mandler, 1980), and are involved in what Jacoby and Dallas (1981) have termed perceptual identification. It is the processes underlying elaborative rehearsal that give rise to those memory records that are most easily produced for evaluation in situations which require recall or recognition, although integrative processing benefits recognition performance more than recall performance.[2]

1.4. Descriptions, organisation and retrieval.

In the production of a description therefore, the choice of prototype, perspective and further specification influences the type of organisation that will be manifested by the memory record. In this system it is not possible for the two dimensions of organisation to be wholly independent. Instead, it is better to regard the two types of organisation as describing opposite ends of a continuum of potential memory record structures. Similar views also are advanced by Craik and Jacoby (1979) and Glenberg and Adams (1978).

As was described in chapter six, the retrieval of information may require several retrieval cycles before an entity is retrieved. Initially, the retrieval specification is formed from what little is known of the sought after record, such as its assumed purpose. As retrieval continues, a variety of records are likely to be presented for assessment. From

2 For a review of the evidence pertaining to the different types of processing and their relation to memory performance, see Craik (1981).

these, or from the difference between their content and that deemed to be required, further pertinent information can be acquired. This information can then be utilised in the subsequent retrieval specification formed as part of the next retrieval cycle.

If these processes of retrieval are considered, the advantage(s) afforded the attempt to retrieve elaborately encoded information becomes apparent. With a greater number of schemata contributing toward a description, the chance of initial retrieval specification and subsequent retrieval specifications incorporating pertinent information is increased. As any retrieval specification has to be formed from what is known about the desired entity, the more pertinent information that a retrieval specification begins with and accumulates, the more successful it is likely to be. Although this may seem to be a very obvious fact, its consequences are important. As more pertinent information is uncovered, the retrieval specification can be made more pertinent to the target. As the pertinence of the information increases, so will the pertinence of the records procured for assessment. In other words as the idealised retrieval proceeds, the approach to a successful retrieval specification will be exponential.

In addition, elaborate encoding is likely to increase the number of forms of description that could uniquely specify an entity in memory. If sufficient detail to disambiguate <u>one form</u> of description is not available, elaborate encoding may allow <u>another form</u> of description to identify the entity. The degree of similarity between the descriptions that are formed to identify the entity in memory will determine the usefulness of the information contained in the old-form retrieval specification, for the new-form retrieval specification. Also, it is possible that the same information re-structured; establishing another form of retrieval specification, could achieve disambiguation. An interesting theoretical problem is the extent to which an entity can be retrieved by differing descriptions, before it is considered that these descriptions are retrieving separate representations.

1.5. An illustration.

In relation to the stimuli presented in these experiments, elaborative processing would result in what is normally described as "associations" being established between words, or rather between their representations. What this means in terms of the system just described is that schemata representing the words are being utilised to form descriptions. The schemata employed can be at a variety of levels, but for simplicity in the following example it will be assumed that they are all at the semantic level: a level that is normally available in consciousness after the presentation of a noun.

If the three words KING, RIVER, and SKY are associated together, it is likely that the subject will form some description representing (amongst other things) a king by a river, perhaps looking at the sky. A variety of other features of this scene may be included, depending upon the degree of default value usage by the elaborative processes and any other schemata suggested by those involved in the encoding. In addition, subjects would need to record in some manner, which of the schemata employed were specifically instantiated by the word stimuli.

Elaborative encodings such as this can have their retrieval initiated by a variety of different pieces of information. Probably the first thing incorporated in the retrieval specification is that "independent" words, rather than a story or a description of a pictorial scene, have to be recalled. In addition, the subjects will probably notice that the words were nouns and are likely to be aware of the fact that they had "seen some links" between words.

As the subject builds up a retrieval specification, there is three times the chance of incorporating a piece of information (at the level of the semantic word schemata) pertinent to the scene of the king by the river looking at the sky, than there is of incorporating one of any other three pieces of information at the same level (ie. probability of independent events: P[king] = 1/x, P[river] = 1/x, P[sky] = 1/x, whereas P[king or river or sky] = 1/x + 1/x + 1/x = 3/x, where x = number of possible schemata choices).

Given that a retrieval specification includes a representation of king, the schemata produced for assessment are likely to include the three item description. If they do not, then it is likely that the representation of one of the other items will be presented for assessment (ie. river or sky). If the second schema in a three schema description is then incorporated in the retrieval specification it is very likely that the target description will be uniquely identified. This type of process could begin with any piece of information and in a situation where only three schemata were involved in a description, three separate combinations or "routes" to the three item specification would be available.

Obviously, this is a ridiculously idealised and simplified illustration of the types of retrieval that could be employed after elaborate encoding had been in operation. However, even this example serves to demonstrate the power and flexibility of the proposed system at retrieval when elaborate encoding takes place.

1.6. Experimental promotion of non-elaborative encoding.

To promote subjects' formation of descriptions that are discriminable primarily through the use of environmental context care must be taken to provide an experimental task that does not require, nor encourage, elaborate processing. At the same time however, it would be unsatisfactory if not pointless, to have subjects perform some task that prevented their perception and reading of the stimuli as words. In such a situation there would be doubt over the equivalence of the instruction to "recall the words", between the last experiment and this. Fortunately, word stimuli are unlikely not to be read. Reading is such an automatic process (especially with student subjects) that it would require very extreme circumstances to prevent.

Several different methods of minimising the elaborateness of subjects encoding operations have been employed in experimental studies. As most of these investigations have been inspired by Craik and Lockhart's (1972) views on levels of processing, the form of elaboration that has been controlled has been semantic.[3] One common method is to require subjects to answer questions about each word as it is presented. The nature of the question biases the type of processing carried out on the word. It is also likely that the act of asking separate questions about each word directs attention to each word in turn, rather than to any potential links between words. However, a task which could be carried out by subjects on the basis of a set of instructions provided before any of the stimuli were presented was considered to be more compatible with the task carried out by the subjects in the previous experiment. Although it is likely that subjects attempting to memorise a list of words will employ a variety of different processing strategies to do so; and although this may also have some correspondence with the changes in processing directed by the various questions about the words, it was felt that the subjects' allocation of processing resources over the course of the word presentation was more likely to be affected and altered by a task which imposed its own process-time requirements. As environmental context information must be encoded at some point in time, the allocation of processing resources may be an important feature of environmental context effects.

Eventually a task that fulfilled all of the requirements discussed and that could be accommodated within the time restrictions of the stimuli

---

3 Indeed, human processing is so semantically biased that it is difficult to imagine elaboration in any other way.

presentation rate was decided upon: subjects were asked to count the number of words that had three or more vowels. In addition, the task was explained as an examination of attention and perception.

1.7. The effect of subjects' ideas about the experiment.

In all experiments, the opinions that subjects have regarding the purpose of their activities can influence their behaviour as much if not more than any instructions given by an experimenter. In memory experiments utilising an unintentional learning paradigm, any suspicion on the part of the subject that they are to have their memory for the presented items tested will have serious consequences in terms of the type of processing they are likely to engage in and their subsequent ability to remember. The general effect of such a suspicion is liable to eliminate or substantially reduce the distinction between elaborative and non-elaborative processing.

As has been discussed previously, there is a tendency for people to try and find meaning in what they are presented with (see chapter one, section 4.2.). In other words, people tend to employ elaborate encoding strategies. Continuous non-elaborative processing is therefore an unusual processing strategy. Consequently, when subjects' processing fluctuates under elaborate processing instruction conditions, it is likely to vary in degree, whereas under non-elaborate processing instructions, any suspicion of having to remember the presented items or reversion to the normal mode of processing is liable to produce a qualitative rather than a quantitative change in processing. A qualitative change in encoding mode should produce an increase in recall score well out of proportion to that which may be expected on the basis of inter and intra subject differences within a processing mode.

1.8. Attempts to cope with confounding variables in the experiment.

To reiterate, there are two related problems which could arise and influence the results of this particular experiment. Subjects could form the opinion that their memory for all or some of the words will be tested and so alter their mode of processing. In addition, subjects' incidental changes in processing (to a more elaborate encoding mode) will also increase their ability to recall the words.

To prevent subjects who had formed a suspicion at encoding that their memory would be tested from contributing to and distorting the pattern of results, a questionnaire: designed to provide further information on the types of encoding and recall strategies employed during the presentation and test phases, was administered to subjects at the end of the experimental session. If subjects responded that they had made an attempt to learn any of the presented words they were replaced by another subject whose responses on the questionnaire suggested that they had not attempted to learn any of the words.

To cope with incidental changes in processing, it was decided to invoke statistical control of learning. Fortunately, the experiment was designed with the inclusion of a recognition test as a closure task. This provided all that was required of a control variable for analysis of covariance.[4]

1.9. Recall protocol recording and the questionnaire.

The single questionnaire administered to subjects attempted to provide some insight into what subjects were doing during the 10 minute presentation and test periods. Also, it was decided to alter the method of recall from a version written by the subject to a tape-recorded version. To what extent the verbal output may queue recalled items is not known, but it is certainly less than that required when the written method is employed. One problem overcome by the verbal method of recall is subjects forgetting recently recalled items as they write down one or more remembered just before. From the subjects' point of view this method of data provision is easier, but it also can be more rigorous. Subjects are aware that everything they say and to some extent do, is not only recorded, but recorded in time. This can encourage subjects to maintain their attempt to recall words, whereas if they had to write them down they could abandon their recall attempt as difficulty increased, knowing that the experimenter would be oblivious to how much effort they had expended in trying to remember.

---

4 For an account of the requirements of a control variable in analysis of covariance, see Kirk (1982).

1.10. Recall environment familiarity.

Experiment one was criticised on the grounds that the visual matching task may have prevented subjects from paying attention to the intermediate environment. If this was so, subjects in the DIFFERENT conditions would still be less familiar with their recall environment than SAME condition subjects. This, rather than the environmental context at presentation matching with the environmental context at test, could be the factor causing the difference in recall between conditions.

To ensure that this was not the case, the familiarity of recall environment was again manipulated, but this time Smith's procedure of having subjects pay heed to the intermediate environment by instructing them to draw it was adopted. This procedure is likely to make subjects more familiar with the intermediate environment than with any other environment. In such circumstances any comparative decrement in recall in DIFFERENT conditions could not be attributed to a lack of familiarity.

1.11. Increase in time spent in waiting area.

One other procedural change was introduced in this experiment. Whereas in experiment one subjects spent only 30 seconds in the waiting area betwêen rooms, subjects in the present experiment were allowed 2 minutes. The main reason for this change was to make the transition between environments and the preparation of the environment the subject was to enter next, less frantic. The original purpose of such a short gap as 30 seconds between rooms was to reduce the opportunity subjects had to rehearse any of the presented words. However, the information obtained from subjects in experiment one during their debriefing indicated that the recognition test presented for closure was having the desired effect.

In addition, the nature of the drawing task in this experiment was such that if subjects were going to make an attempt to rehearse any of the presented words, they would have ample opportunity to do so during the 10 minute drawing period. An extra 1 minute and 30 seconds potential rehearsal time was unlikely to exert any significant effect.

1.12. Reasons for paradigm change.

Although these changes bring this experiment (and as will be seen, future experiments) more in line with the paradigm employed by Smith, they move the paradigm further from that employed in experiment one. It would have been convenient if full comparisons could have been made between experiment one, this and future experiments. However, with the alterations in environment being unavoidable and in themselves confounding any comparison, it seemed to be a good time to introduce other changes based on the experience gained from running experiment one and the view of the purpose of the experiments developed since the instigation of the study.

1.13. Experimental hypotheses.

The present experimental hypotheses replicate and extend those of experiment one. The hypotheses are replicated in that the order of best to worst recall is similar and extended in that the degree of difference between SAME and DIFFERENT condition recalls is predicted to be greater. Of course it is exactly this latter hypothesis which cannot be evaluated until some appropriate comparison is made.

In the meantime, for simplicity, this experiment is confined to; testing the tenability of the familiarity account of several environmental context effects, examining the effect of a different psychological processing mode in relation to the environmental reinstatement effect, and investigating the practical and theoretical utility of an alternative analytical procedure in the study of environmental context effects.

2. Method.

2.1. Subjects:

Although all subjects were naive regarding the purpose of the experiment, several subjects behaved in a manner determined by their belief in the purpose of the experiment, rather than as instructed. These subjects were identified by their answers to the questionnaire administered at the end of the experimental session. Any subject who responded that they had made an attempt to learn some, or all of the words as they were presented was replaced by a subject who had not responded in such a manner. All subjects; including those whose data was not analysed, were paid one pound for taking part in the experiment. Indeed, it was not until subjects had completed the session, had been debriefed and paid, that their responses to the questionnaire were seen. Eventually, 48 Glasgow University students; males and females of approximately equal numbers, completed the experiment in a satisfactory manner.

2.2. Environmental contexts:

Rooms A, B and C2 were employed in the experiment. Descriptions of these rooms can be found in appendix A.

2.3. Stimuli:

All the stimuli used in this experiment were identical to those used in experiment one.

2.4. Apparatus:

As in experiment one, the slides were projected onto the white walls in both presentation environments: rooms A and B. However, whereas previously the Kodak projector slide presentation rate had been controlled by an electronic timer, the limitations of this device had prompted the development of a computer controlled slide presentation system (details of this system's operation can be found in appendix C). A footswitch was placed under the desk in both rooms A and B. Slide presentation was initiated by one press on either of these switches. A microphone in each of the two rooms was connected to a JVC stereo cassette deck, and a buzzer controlled by the computer was placed in the waiting area.

2.5. Questionnaire:

A very simple questionnaire consisting of three separate questions was formulated in an attempt to obtain information about what the subjects were doing during the presentation and recall periods, and how similar they thought the room environments A and B were. The two questions regarding encoding and recall were open ended, requiring subjects to write in their own words what they did at presentation and recall, while the third question required subjects to provide a value on a scale ranging from 1 to 100 indicating how similar they thought rooms A and B were (a greater value indicated greater similarity). A specimen copy of the questionnaire can be found in appendix D.

2.6. Design:

As in experiment one, a two factor (3 x 2) between subjects design was applied. The first factor was defined by the relationship between the presentation environment and the recall environment; same or different, and the familiarity of recall environment. The second factor was defined by the two learning environments, rooms A and B.

2.7. Procedure:

Each subject's 34 minute experimental session was divided into three 10 minute periods, separated by two, 2 minute waiting periods. In the first 10 minutes subjects viewed an automatic presentation of the 70 word slides. Each slide was presented for 3 seconds with a variable 1 to 1.5 seconds gap between slides as they changed. Subjects were told of the slide presentation times and were instructed that their only task was to count the number of words containing three or more vowels. They were then told to press the footswitch under the desk in front of them to initiate the presentation when the experimenter left the room. When the presentation was complete, the buzzer sounded in the waiting room informing the experimenter, who returned and administered the recognition task. Subjects were allowed 2.5 minutes to mark the words they recognised. After the recognition task was completed, subjects returned to the waiting area for 2 minutes before moving to the next scheduled room. In the second environment subjects were given a piece of paper and a pencil, and were asked to draw; in as much detail as possible, a view of the room from where they were sitting. If they found that they had completed their drawing before the experimenter returned, they were asked to draw the room again, but from a different angle and to continue drawing until the experimenter did return. After 10 minutes the experimenter returned and directed the subject back to the waiting area. After another 2 minutes, subjects entered the appropriate room environment and were asked to recall all the words they had been presented with in the first period, including those in the recognition test, that they could remember. Subjects were instructed to say each word that came to mind out loud, as their recalls were being recorded. The subjects were asked to begin as soon as the experimenter left the room and to keep trying to remember more words until he returned. After ten minutes had passed, the experimenter returned to stop the subjects' free recall. Those subjects in the SF and DF conditions were then asked to complete the questionnaire. All subjects then returned to the waiting area where they were paid and debriefed. The room order for those subjects in the SAME and FAMILIAR (SF) condition was ABA/BAB, the room order for the DIFFERENT and FAMILIAR (DF) condition was

ABB/BAA and the room order for the DIFFERENT and UNFAMILIAR (DU) condition was ACB/BCA.

2.8. Scoring:

Each subjects' verbal recall protocol was transcribed and scored in the same manner as reported in experiment one.

3. Results.

3.1. ANCOVA Of P-Item Recall.

3.1.1. selection of a covariate.

As was described in the introduction (sections 1.7. and 1.8.), the purpose of employing ANCOVA was to control for the degree of elaborative processing which subjects were assumed likely to engage in. Some assessment of the degree of elaborate processing that subjects have engaged in can be made from the subjects' performance on the recognition test. Apart from scoring subjects' recognition performance in terms of the number of correctly identified items (RHTS) and the number of incorrectly identified items (RFAS), d' scores were also considered.

However, although d' scores might be regarded as a more sophisticated and accurate measurement, there were problems in using this measure with the present data. To calculate d' scores, the probability of a hit and the probability of a false-alarm have to be known. For proper calculation, both of these values should be greater than zero. Unfortunately, in the present experiment many subjects made no false-alarm responses and as a result, many d' calculations would have been extremely unreliable. Excluding subjects that provided no false alarms causes other problems. Apart from reducing the power of the analysis by lowering sample size, the exclusion of these subjects alters the distribution of scores. In this case only potentially high d' scoring subjects would be excluded from the analysis.

Consequently, as a preliminary to the analysis of covariance, a stepwise (forward and backward) multiple linear regression (Dixon, 1983) was carried out to determine the most efficient predictor of P-item recall and to ensure that the overall regression of predicted and predictor variables was linear. As the two predictor variables employed in this procedure were RHTS and RFAS, any advantage in P-item recall prediction afforded by a model incorporating both variables (similar to a d' value) should be detected.

In stepwise (forward or backward) regression; as the variables with the largest F-values are chosen from a set of potential predictor variables, the usual F-tables do not apply. The probabilities associated with the F-values obtained through such procedures are affected by the fact that with a greater number of potential predictor variables there is a greater likelihood of obtaining high F-values by chance. Unfortunately, the appropriate probabilities are a function of not only the number of subjects and the number of variables, but also the correlational structure of the potential predictor variables. The correlational structure of the predictor variables is important as it reflects how much unique information is being provided by each predictor. If all predictor variables were orthogonal, the actual significance level would be given by,

$$1 - (1 - a)^m$$

where **a** is the tabled alpha value and **m** is the number of orthogonal predictor variables.

As the amount of unique information provided by each predictor decreases, so the likelihood of a high F-value occurring by chance decreases. If all predictor variables are perfectly correlated, the tabled probability of the F-value is the actual significance level. In most situations however, the correlational structure is not so simple as all or nothing. Consequently, calculating the probabilities associated with the F-to-enter, F-to-remove and the F-value of the overall regression line, can be a difficult and time consuming exercise. In addition, there is presently no general agreement as to the best method of calculation.

For these reasons, the adjusted probabilities reported with stepwise regressions have been calculated on the basis of the formula presented above. To give some idea as to the degree of conservativism of the adjusted p-values, the correlations between the predictor variables are also presented.[5]

The best prediction of P item recall was obtained using RHTS alone. The F-to-enter value (ie. the F-value of that proportion of the variance accounted for by a variable when the variance accounted for by the other variables is partialed out) for RFAS was less than 1. Table 7.1. gives those values associated with the simple regression equation.

The basic assumptions underlying a regression analysis are that the error terms from the model are normally and independently distributed with mean zero and exhibit homoscedasticity (ie. common Y variance at each X). However, as these assumptions are extremely unlikely to be fulfilled

5 For further discussion of this topic see Draper & Smith (1981, p.311), Forsythe (1983) and Forsythe, May and Engleman (1971).

variable	std reg coeff	R^2	adj R^2	F(1,46)	adj p	predictor corr
RHTS	0.417	0.174	0.156	9.71	0.006	0.101

Table 7.1. Summary of simple regression of P-item recall on RHTS.

unless the type of regression applied to the data is compatible with the relationship between predicted and predictor variables, the assumption reduces to one of linear regression (Keppel, 1982).

A chi-square test of homoscedasticity suggested that this assumption was tenable, $X^2(1) = 2.03$, $p = 0.15$. An examination of the plot of P-item recall against RHTS and the plot of residuals (P-item residuals against predicted P-item recall and the normal probability plot of P-item residuals) also supported the assumptions of the regression analysis and consequently, that linear regression was appropriate.

3.1.2. preliminary analysis for ANCOVA.

Unlike the preliminary data analysis in experiment one, an F-max test prior to the ANCOVA would not be appropriate. Variance estimates in the ANCOVA contributing to the assessment of the F-value are made after the removal of the variance accounted for by the relationship between the covariate and dependent variable. Consequently, it is after this variance has been removed that a test of homogeneity of variance should be carried out (Huitema, 1980). This is most easily done as a final check on the tenability of the assumptions underlying the analysis.

The regression coefficient that removes the variance on the error term explained by the covariate-dependent variable relationship, is obtained by pooling the within-group regression coefficients. However, an important assumption underlying this procedure is that there are no significant differences between within-group regression coefficients. To check this, a test of homogeniety of within-group regression coefficients is carried out. In addition to the assumption that the overall form of regression is linear, it is also assumed that there is linearity of regression within each treatment group. One problem in examining this assumption directly would be determining the significance of any departure from linearity. In the same way that the significance of tests of variance homogeniety are assessed, the probability of obtaining such departures from linearity (with the particular number of groups and sets of scores) would have to be determined in relation to the probability of obtaining such results by chance, from similar sub-sets of linearly related data. At present there are no established procedures for doing this. One way round this problem is to test the linearity of the between groups regression line. This is simply the regression of the dependent variable group means on the covariate group means. Any deviation from linearity of regression within any of the groups will distort the linearity of the between groups regression line and consequently, any departure from linearity within the treatment groups will be assessed by this test.

The most efficient order of tests of the assumptions underlying the ANCOVA is therefore:-

(1) linearity of overall regression line.

(2) homogeneity of within group regression lines.

(3) linearity of between group regression line.

(4) normal distribution and homogeneity of within group variances.

The first assumption has already been checked in relation to the determination of the most efficient predictor of P-item recall. The assumption of homogeneous within group regression lines was not challenged by the results of its test, $F(5,36) = 0.721$. A subsequent test of the linearity of the regression between the group means; given the homogeniety of the within group regression lines, was not significant, $F(4,41) = 0.005$. It can be assumed therefore, that the assumptions of homogeneous and linear within group regression lines are tenable.

The last assumptions to be tested are those of within group variance homogeneity and normality of distribution. Here the usual tests of homogeneity (with appropriate df reduction) can be used with the adjusted dependent variable scores (Huitema, 1980, p.118). Harley's F-max was computed; using the adjusted values presented in the top section of table 7.4., and revealed no significant departures from the assumptions of normality and variance homogeneity, $F\text{-max}(6,6) = 5.95$.

CONDITION	ABA	BAB	ABB	BAA	ACB	BCA
X RECALL	6.12	7.25	3.87	5.37	5.00	4.37
S.D.	3.18	5.23	3.87	3.58	3.25	2.82

CONDITION	SAME FAMILIAR(SF)	DIFFERENT FAMILIAR(DF)	DIFFERENT UNFAMILIAR(DU)
$\bar{X}$ RECALL	6.68	4.62	4.68
S.D.	4.33	3.72	3.04

Table 7.2. Unadjusted mean recall and standard deviation of P-items.

CONDITION	ABA	BAB	ABB	BAA	ACB	BCA
MEAN	5.00	5.87	6.00	7.37	6.50	6.25
S.D.	1.85	1.46	1.07	1.51	1.69	1.83

CONDITION	SAME FAMILIAR(SF)	DIFFERENT FAMILIAR(DF)	DIFFERENT UNFAMILIAR(DU)
MEAN	5.43	6.68	6.37
S.D.	1.67	1.31	1.76

Table 7.3. Covariate mean and standard deviation.

CONDITION	ABA	BAB	ABB	BAA	ACB	BCA
$\bar{X}$ RECALL	7.58	7.61	4.08	3.87	4.58	4.27
S.D.	1.83	4.46	3.05	3.38	3.28	2.37

CONDITION	SAME FAMILIAR(SF)	DIFFERENT FAMILIAR(DF)	DIFFERENT UNFAMILIAR(DF)
X RECALL	7.59	3.98	4.42
S.D.	3.41	3.21	2.86

Table 7.4. Adjusted mean recall and standard deviation of P-items.

3.1.3. overall ANCOVA F-tests.

As in experiment one, a 3 x 2 (presentation-recall environment match/recall environment familiarity x presentation environment) completely randomised design was applied. The dependent variable was subjects' P scores and the covariate was their RHTS scores.

Figure 7.1. illustrates the significant main effect of presentation-recall environment match/recall environment familiarity, $F(2,41) = 5.37$, $MSe = 10.29$. The other main effect of presentation environment was not significant, $F(1,41) = 0.03$, nor was the interaction between these two factors, $F(2,41) = 0.01$. The pooled within group regression coefficient of P-item recall on RHTS was 1.246.

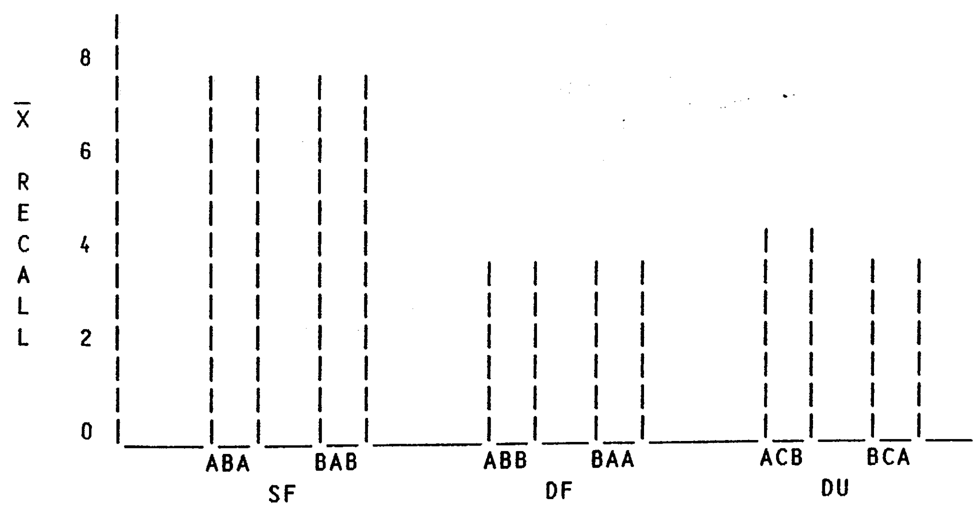

Figure 7.1. Adjusted mean recall of P-items per group.

3.1.4. specific comparisons.

Again, as in experiment one, planned orthogonal pair-wise comparisons; using Fisher's protected LSD (Huitema, 1980; Kirk, 1982), confirmed that the significant main effect was due to the difference between SF and DF conditions, $t(41) = 3.019$. An orthogonal comparison (two-tailed) between the means of the DF and DU conditions revealed the unpredicted superiority of DU condition recalls to be non-significant $t(41) = 0.381$ (see table 7.4. and figure 7.1.)

3.1.5. power and degree of association.

A post-hoc analysis of the power (1-B) of the F-test identifying the significant main effect revealed it to be operating at approximately 0.95. The degree of association (W^2) between the significant main effect and subjects P scores was calculated to be 0.166.

3.2. ANOVA Of P, R And F-Item Recall.

3.2.1. preliminary analysis.

As in chapter three, P, R and F-item recall was prepared for analysis by calculating the proportion of the total number of items presented, that were recalled. No significant departures from the assumption of normally distributed data was found with the proportion recall of P, R and F-items, skew/standard error and kurtosis/standard error < 1.96. Prior to further analyses the proportion values were multiplied by 100 to produce percentage recall. The only reason for doing this was to make the scores more meaningful.

Hartley's F-max test on the between subject scores indicated that the assumption of homogeneous error variances was tenable, F-max(6,7) = 3.19. Table 7.5. contains the means and standard deviations pertaining to this analysis.

Condition	ABA	BAB	ABB	BAA	ACB	BCA
X recall	34.65	29.03	34.65	35.49	33.19	32.43
S.D.	16.95	16.04	18.31	20.38	14.95	11.40

Table 7.5. Percentage mean and standard deviation of total P, R and F-item recall per condition.

As in chapter three, Anderson's (1958, p.259) symmetry test indicated that the probability of the heterogeneity of variance of differences occurring by chance was 0.005.

3.2.2. overall F-tests on word type recall.

A three factor (3 x 2 x 2) mixed design analysis; with repeated measures on the third factor, was carried out. The first two factors were identical to those described in the previous analyses, while the third factor expressed the type of word recalled (P, R or F-item). Table 7.6. contains a summary of the results of this ANOVA, while Table 7.7. contains the means and standard deviations pertaining to the word type comparison.

SOURCE	df	mean square	F	p	Huynh-Feldt p
FML	2	131.732	0.34	0.714	
LE	1	123.457	0.32	0.576	
FL	2	135.783	0.35	0.707	
error	42	387.908			
word type (W)	2	21756.780	99.66		0.000
W x FML	4	468.885	2.15		0.086
W x LE	2	287.346	1.32		0.273
W x FML x LE	4	255.922	1.17		0.329
error	84	218.314			

Table 7.6. Summary of the three factor mixed ANOVA.

	P	R	F
X	8.89	48.33	42.50
S.D.	6.23	17.08	22.23

Table 7.7. Percentage mean and standard deviation of recall of each word type.

However, in addition to this significant effect, the interaction between presentation-recall environment match/recall environment familiarity and word type approached significance. The means and standard deviations of the pertinent comparisons are presented in Table 7.8. and plotted in figure 7.2..

		SF	DF	DU
P	X	11.15	7.71	7.81
	S.D.	7.21	6.22	5.08
R	X	49.37	46.87	48.75
	S.D.	11.76	20.42	17.88
F	X	35.00	50.62	41.87
	S.D.	25.04	25.89	13.59

Table 7.8. Percentage mean and standard deviation of word type by condition.

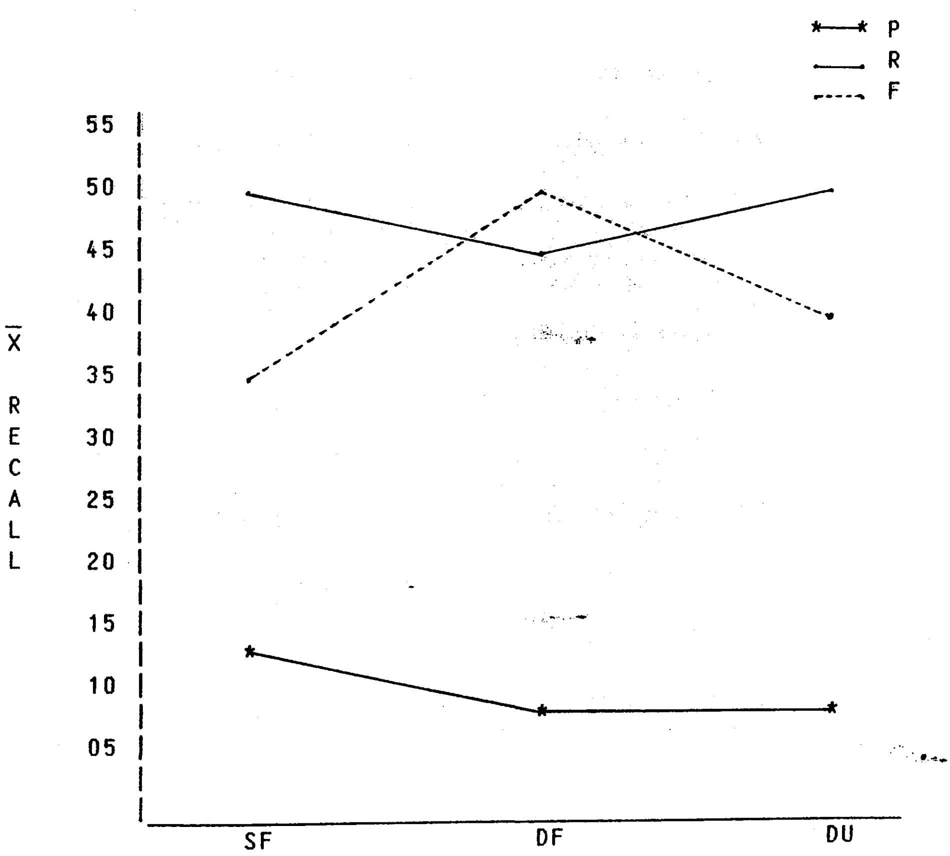

Figure 7.2. Mean transformed recall of word type by condition.

3.2.3. specific comparisons.

Pair wise (two tailed) orthogonal comparisons[5] revealed a significant difference between overall F and P-item recall, $t(84) = 9.630$, but not between overall R and F-item recall, $t(84) = 1.664$.

3.3. Questionnaire Data

3.3.1. encoding and retrieval strategies.

As a consequence of the screening procedures described in section 1. of this chapter, only slight variations on the general strategy of counting the number of words with more than three vowels were observed. These variations seemed to be caused as much by the different forms of expression used by the different subjects, as any substantial differences in processing. As a result no attempt was made to categorise the descriptions of strategies employed at encoding.

Examination of the descriptions of retrieval attempts revealed a greater variety of strategies. One main distinction was between the items on the recognition test and those only presented by slide. 19 out of the 28 subjects who completed the questionnaire made a distinction in their attempt to recall these items. The probability of this occurring by chance is 0.026 (binomial test; assuming that it is equally likely that subjects would, or would not distinguish between the two forms of

---

5 As in chapter three; because the homogeneity of variances of differences assumption was untenable, separate error terms were calculated for each comparison (Kirk, 1982).

presentation).

Of those 19 subjects who distinguished between the two forms of presentation, 8 stated that they had imagined the recognition test list of words on the slip of paper, as they had tried to remember. Of these 8 subjects, 6 were in the SF conditions and 2 were in the DF conditions. 5 subjects (3 SF and 2 DF; 3 of whom were included in the previous 8) also reported imagining the slide presentation, as they tried to remember.

The reasons for the brevity of the open-ended questionnaire data analysis will be presented in section 4.7. of the discussion.

3.3.2. rating of environmental similarity.

The median rating of the similarity of the two rooms; A and B, was 25 (X = 31.18, S.D. = 24.10). Regression analyses were carried out to see if the subjective ratings of environmental similarity (SIM) predicted P-item recall. As the ANCOVA reported in section 3.1. of this chapter has already established a difference between SF and DF conditions in terms of P-item recall, for simplicity it was decided to carry out two separate regression analyses: one for the SF, and one for the DF conditions. Table 7.9. contains summaries of these analyses.

condit'n	variable	std reg coeff	R^2	adj R^2	F	df	p
SF	SIM	-0.281	0.079	0.013	0.86	1,10	0.376
DF	SIM	-0.280	0.078	0.010	1.19	1,14	0.294

Table 7.9. Summary of SF and DF condition regression with SIM ratings predicting P-item recall.

Chi-square tests of homoscedasticity indicated that this assumption was tenable in each analysis (SF, $X^2 = 0.13$, $p = 0.72$; DF, $X^2 = 0.01$, $p = 0.92$). An analysis of residual plots (as described in section 3.1.1.) also suggested the tenability of the assumptions underlying the regression analysis.

3.4. Multiple Linear Regression Model Predicting P-item Recall.

Apart from SIM ratings, RHTS also were found to predict P-item recall (see section 3.1.1.). It was decided to investigate the predictive power of the linear regression model employing both RHTS and SIM variables. For the same reasons as expressed in section 3.3.2., separate analyses were carried out for SF and DF conditions.

3.4.1. linear model for SF conditions.

As in section 3.1.1., a stepwise linear regression was carried out. The best prediction of P-item recall was obtained using RHTS only. The F-to-enter value for SIM was less than 1. Table 7.10. summarises the regression model.

variable	std reg coeff	R^2	adj R^2	F(1,10)	adj p	predictor corr
RHTS	0.723	0.523	0.476	10.98	0.000	-0.413

Table 7.10. Summary of SF condition regression model predicting P-item recall.

A chi-square test of homoscedasticity suggested the tenability of this assumption, $X^2(1) = 0.56$, $p = 0.46$. An examination of the plot types described in section 3.1.1. and obtained with the present data also suggested that the assumptions underlying the regression analysis were tenable.

3.4.2. linear model for DF conditions.

The only significant prediction of P-item recall was obtained using the variables RHTS and SIM. Tables 7.11. and 7.12. summarise this regression model.

variable	R^2	adj R^2	F(2,13)	adj p	predictor corr
RHTS, SIM	0.433	0.346	4.97	0.049	0.172

Table 7.11. Summary of overall regression model predicting P-item recall in the DF conditions.

Variables	std reg coeff	F(1,13)	adj p	p
RHTS	0.605	8.14	0.028	0.014
SIM	-0.384	3.28	0.177	0.093

Table 7.12. Individual contributions of RHTS and SIM to model.

The chi-square test of the homoscedasticity assumption suggested this was tenable, $X^2(1) = 1.61$, $p = 0.21$. An examination of the diagnostic plots described earlier and obtained with the present data, also suggested that the assumptions underlying the regression analysis were tenable.

However, as can be seen from table 7.12., SIM does not contribute significantly to the regression model. The slight increase in the amount of variation accounted for when SIM is incorporated into the model allows the calculated F-value to achieve significance at the (arbitrary) 0.05 level, but this slight increment in R^2 really does not justify the inclusion of an extra predictor variable. Table 7.13. summarises the regression model incorporating RHTS only.

variable	std reg coeff	R^2	adj R^2	F(1,14)	adj p
RHTS	0.539	0.290	0.240	5.73	0.061

Table 7.13. Summary of regression model predicting P-item recall in the DF conditions incorporating RHTS only.

Once again, but with RHTS as the sole predictor of P-item recall in the DF condition, the chi-square test suggested that the homoscedasticity assumption was tenable, $X^2(1) = 0.86$, $p = 0.35$. An examination of the plot types described previously and obtained with the present data, also suggested that the assumptions underlying the regression analysis were tenable.

4. Discussion.

4.1. The use of P-item recall data.

As in chapter three, the data revealing the effect of environmental context was subjects' P-item recall scores. The argument for employing P-item recall as the primary data is the same in this experiment as it was in the last. However, in this situation where an attempt was made to determine the mode of processing that subjects would employ as they encoded the word stimuli, the rational is more obvious through its seemingly greater pertinence. A single score; the total of P, R and F-item recall, would represent not only any variation in the processing of the stimuli list, but also the influence of the re-presentation of a sub-set of these items, the presentation of new items (fillers) and the accompanying variation in processing modes and presentation times. The division of recall into P, R and F-item categories is in an attempt to reduce the number of potential effects, represented by each score. The magnitude of the word category effects, in this and the previous experiment, would seem to validate such a division.

4.2. The operation of ANCOVA.

Prior to further discussion, it may be useful to consider the operation of the analysis of covariance. Such a discussion may prove useful not only for the elucidation of the present experimental results, but also for subsequent experiments in which ANCOVA is the primary method of data analysis.

Analysis of covariance is the product of the combination of two statistical techniques: (linear[7]) regression analysis and analysis of variance. As well as recording a measure on the dependent variable for each subject, one or more concomitant (control) variables must be measured. The concomitant variables represent a source of variance that is not controlled by the experimental procedures and is assumed to influence the dependent variable. The ANCOVA procedure adjusts the dependent variable such that the effect of the variation represented by the concomitant variable is removed. The benefit of such a procedure is that it allows a more accurate analysis of the effects of the experimental variables.

To remove the variation in the dependent variable scores that is associated with the concomitant variable, the weighted average (pooled) regression coefficient of the regression of the dependent variable, across each of the experimental conditions is calculated. The linear model for the completely randomised one factor, one covariate ANCOVA is therefore,

(1) $$Yij = u + Aj + b(Xij - X) + ei(j),$$

for i = 1, ..., n and j = 1, ..., p, where Yij is the i'th subject's score in the j'th treatment, u is the general mean of all Y scores, A is the effect of the j'th treatment level, b is the regression coefficient of

---

7 ANCOVA can be carried out using non-linear regression. However, even in situations where non-linearity is present transformations to achieve linearity are usually carried out in preference to the application of non-linear regression. Some of the reasons for this are the greater complexity of non-linear regression calculations and the difficulty of interpreting non-linearly adjusted effects (Draper & Smith, 1981).

the dependent variable (Y) on the predictor variable (X), Xij is the covariate (predictor variable) score corresponding to the Yij, X is the mean of all covariate scores, and ei(j) reflects random variation due to any uncontrolled source. The model for the completely randomised ANOVA is,

(2) $$Yij = u + Aj + ei(j).$$

Applying a little bit of algebra to the linear models above, the model for the values free of the effects of the concomitant variable can be obtained,

(3) $$Yaij = Yij - b(Xij - X) = u + Aj + ei(j).$$

The models for analyses employing more than one factor are simple generalisations of the one factor case, with additional terms to represent the extra factors and their interactions. For example, the two factor model for ANCOVA is,

(4) $$Yijk = u + Aj + Bk + (AB)jk + b(Xijk - Xjk) + ei(jk),$$

with the term B representing the effect of the k'th treatment level and the term AB representing the interaction between the treatment levels. The corresponding ANOVA model is,

(5) $$Yijk = u + Aj + Bk + (AB)jk + ei(jk).$$

Consequently, the model for the adjusted scores is,

(6) Yaijk = Yijk - b(Xijk - Xjk) = u + Aj + Bk + (AB)jk + ei(jk).

Apart from providing an estimate of the normal ANOVA terms as in models (2) and (5), the ei(j) in models (3) and (6) are likely to be smaller than those in (2) and (5) due to the extraction of variance associated with the concomitant variable. The reduction in the error term is despite a decrease in the error term degrees of freedom: the denominator producing the error means square, which is a consequence of the regression adjustment. Usually, the variance removed by the regression (unless systematic) can be accommodated only by the error parameter in the ANOVA model.

Useful illustrations of the effect of such ANCOVA operations are provided by Kirk (1982). Figure 7.3. graphically presents two hypothetical experimental conditions. The relationship between the dependent variable (Y) and the concomitant variable (X) forms a scatterplot for each condition represented by the two ellipses.

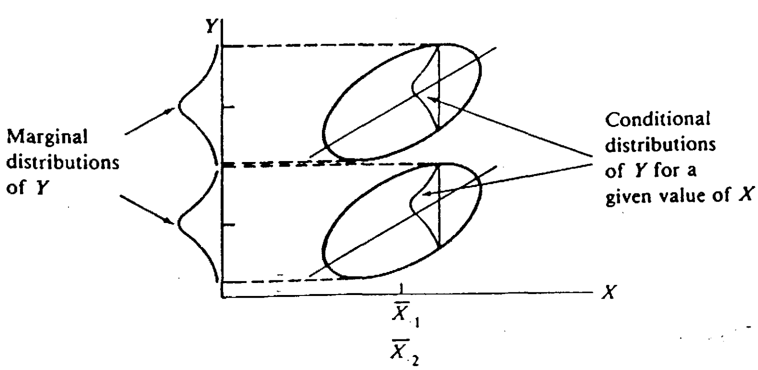

Figure 7.3. Scatterplots for two treatments when covariate means are equal (from Kirk, 1982).

Through each ellipsis is a line representing the regression of Y on X. In ANOVA the error variance is determined by the dispersion of the marginal distributions, while in ANCOVA it is determined by the dispersion of the conditional distributions. The greater the correlation between X and Y, the narrower are the ellipses and the greater is the reduction of the error variance in the ANCOVA.

Figure 7.3. reflects a situation in which the concomitant variable group means are equal. However, it is possible, even with random allocation and particularly without it, for situations to arise in which the concomitant variable means will not be equal across conditions. Figure 7.4. presents some examples of the types of effect that can be obtained in such circumstances.

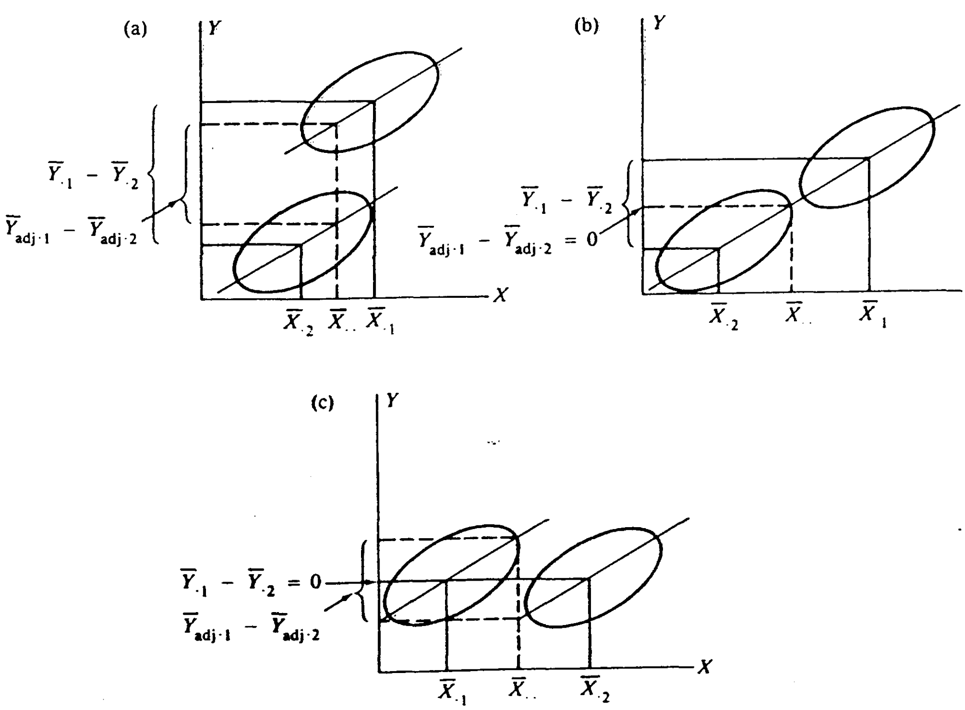

Figure 7.4. Scatterplots for two treatments when covariate means are different (from Kirk, 1982).

In example (a) and (b) the difference between adjusted dependent variable means is less than the difference between the unadjusted means, whereas in example (c) the converse is true.

There may be two direct consequences of the ANCOVA procedures. The first, which has already been discussed is that the error variance is likely to be reduced. This in itself will allow a more powerful test. The second consequence is that the adjustment carried out on the dependent variable may alter the group mean values of the dependent variable. As the previous examples demonstrate this can either increase or decrease the difference between the Y variable group means, depending upon their associated X variable group means.

Essentially, ANCOVA can be considered as consisting of two procedures. The first procedure is an adjustment of dependent variable scores on the basis of the regression of these scores on the concomitant variable scores. The second procedure is an analysis of variance on these adjusted scores.

4.3. Evidence of non-elaborative processing.

A quick comparison of the number of P-item words correctly recalled in this experiment with the number obtained from subjects in the last, gives much credence to the claim that with respect to forming descriptions of the word stimuli, a lower level of processing was in operation. Despite the other differences between the two experiments, it seems unlikely that any factor other than a change of processing mode could explain such an overall decrement in recall.

Further evidence from the subjects' recall supporting the view that low level processing was being carried out on the words, comes from the different patterns of word type recall effects reported in the two experiments. In both experiments the filler items are introduced after the major word presentation and in both experiments they serve the same purpose in the recognition task. However, differences occur in the type of processing the fillers receive as part of the recognition task in the two experiments. As has been suggested, in experiment one little attention need be given to filler items due to the easy identification of the recognition items from the word list. It would be expected that as recognition would be more difficult in experiment two, greater consideration would be given to the filler items. The cross-over of P and F-item recall rank positions between experiments conforms with this viewpoint and suggests a difference in processing mode to the detriment of P-item recall in this experiment.

All in all, the changes in word type recall imply that the form of the task presented to subjects achieved the goal of eliciting a low level (semantic) processing mode.

4.4. P-item analysis: general conclusions.

The pattern of P-item results obtained in the present experiment replicates the pattern observed in experiment one and the general pattern of effect reported by Smith (1979, expt.1). The pattern of results obtained in all of these studies supports the view that the environmental context effect is a consequence of a match/mismatch between presentation and test environment contexts.

Despite the alterations to the two rooms, the observation of an environmental context effect indicates that sufficient differences (of some sort) between rooms still remain. Indeed, the magnitude of effect is greater in this experiment than it was in experiment one. This is despite the feeling that there was a reduction in the degree of difference between rooms and that this could have precluded any effect. Given such a situation, it is tempting to attribute the increase in effect to the change of processing mode. However, apart from the changes in environments and experimental procedure, the use of ANCOVA and the increase of power associated with such an analysis also contributes to the general muddle of factors varying between experiments that confounds any direct comparison.

Under the ANCOVA form of analysis, the nature of the room factor requires some mention. As will be discussed in more detail in chapter eight, ANCOVA equates recall scores as if they were from subjects that had obtained the overall covariate mean. Consequently, in the ANCOVA the potential benefit (and effect) of any presentation environment is eliminated. The ANCOVA procedure simplifies the interpretation of the results, by restricting effects to recall. Consequently, the lack of room-specific effects is to be expected. However, the lack of influence of any combination of presentation room and recall environment, continues to suggest that some aspect(s) common to, or equal across rooms is involved in the production of the environmental context effect.

Although it is not possible to determine all the consequences of each of the changes in procedure, the general impression is that an experiment more sensitive to the influence of environmental context has been carried out. Certainly, the demonstration of an environmental context effect indicates that the changes in procedure have had no effect on the basic result of the experiment.

4.5. Assessment of the familiarity hypothesis.

Given the task of drawing the intermediate environment which is likely to make the subject more familiar with this room than any other and the overall experimental results, one would seem able to conclude that environmental familiarity has no influence on subjects' recall.

As a last attempt to resurrect the notion that environmental familiarity has influence on recall, it could be argued that the lack of a difference between DF and DU conditions was due to the similarity of rooms A and B. To maintain this position in the light of the environmental context effect observed, it would have to be said that higher levels of difference between environments are necessary to produce differences in recall on the basis of the familiarity of recall environment, than are necessary to effect differences on the basis of a match or mismatch between presentation and test environments. Such an account would require a considerable change in the model underlying the original familiarity hypothesis, and still would have to admit that the influence of recall environment familiarity cannot account for the data observed in these and presumably other environmental context studies.

4.6. ANOVA of P, R and F-item recall.

Apart from examining the main effect of word type on recall, interest in the ANOVA also was focussed on the possibility of a significant interaction between the presentation-recall environment match/recall environment familiarity, and word type factors.

The argument presented to account for the main effect of word type was that subjects processed recognition task items more elaborately than

P-items because of the task differences between the original presentation and the recognition task, while the further increment of R-item recall was attributed to these items having been presented twice. P-items should have been processed only once in a non-elaborative manner. Within the experiment therefore, there are different levels of processing of word types. As a result it might be expected that there would be some evidence pertaining to the predictions drawn from the model presented in chapter six. If the predictions of an effect of environmental context with low level, but not high level processing are supported by the data, there should be an interaction between the two factors mentioned above. The specific form of the interaction should be an increment in the recall of P-items in the SF condition in comparison to P-item recall in DF and DU conditions, with R and F-item recall being unaffected by the particular SF, DF or DU condition.

Although the pertinent interaction did not reach the 0.05 level of significance, it came close enough to warrant consideration. However, an inspection of the form of the interaction suggests that rather than it being due to the pattern of P-item recall alone, the major contribution is from the pattern of F-item recall. There would seem to be no theoretical reason as to why F-item recall should be so affected. However, tracing the greater recall of F-items in the DF condition back to the raw data reveals that this increment is caused by three (out of sixteen) subjects recalling an unusually high number of filler items. So, whatever the reason for these subjects' recall including a large number of F-items, it does not seem to have been shared by the other subjects in the same condition.

In relation to the lack of an effect caused by a greater recall of P-items in the SF condition, a lack of power would seem to be the cause. In the original P-item analysis greater power was provided by the use of RHTS as a covariate (concomitant/control variable). It is also worth noting that other than reducing the error term, the effect of the ANCOVA was to increase the difference between the recalls obtained in the SF, and DF and DU conditions. This is because; despite random allocations of subjects, the degree of elaborate processing being carried out by subjects across SF, DF and DU conditions was not equal. The degree of relationship between RHTS and recall is such that even small differences in RHTS can have a significant and potentially distortive effect on the pattern of results obtained.

It would appear that this retrospective analysis has not provided a short-cut answer to the question of whether or not mode of processing interacts with the influence of environmental context. As has been said already, this question will be returned to; with more appropriate data upon which to base an answer, in the next chapter.

4.7. Open-ended questionnaire data.

One, if not the major problem in analysing the questionnaire data is determining an appropriate null hypothesis to test. In only one occasion did it seem reasonable to define the likelihood of response. In situations where subjects had mentioned strategies other than dividing the to-be-searched sets into presentation list and recognition task list, it seemed impossible to identify what the range of alternative strategies were. Consequently, the determination of the probability of the subjects' choice of strategy on the basis of chance would seem to be impossible. Such an inability to relate subjects' performance to some estimate of its general likelihood means that it can be said only if the data conforms, or counters some intuitive notion.

Beyond the problem of drawing inferences from such data, it is worth mentioning that the hope that subjects might provide some insight into their psychological activity was not really fulfilled. Despite spending, sometimes extraordinary lengths of time filling out the questionnaire, the type of self-reports provided did not really tap into the processes that are the focus of study. Typically subjects would write "I could remember the words with more than three vowels easiest", or "I remembered words that had some personal association". As these examples demonstrate, subjects were stating what the result of their attempts to remember were, rather than how they were able to remember such things.

However, despite the difficulties encountered in interpreting and relating the open-ended questionnaire data to psychological processes, there was some indication that environmental context was being used to aid recall. Not only were some subjects consciously aware of their use of environmental context, but they also considered it important enough to write down.

4.8. Similarity ratings and recall.

It is suspected that several subjects employed a nominal type of scale, or just "agreed" with the reminder of the form of the scale, ie. 1 = totally different (see appendix D). This would explain why there was such high variation in subjects' scores. The median value of 25 would seem to indicate a general opinion that rooms A and B were quite dissimilar. Although a negative relationship; indicating greater recall with lower similarity ratings was obtained, there was no significant prediction of P-item recall in either the SF or DF conditions.

The relationship between similarity ratings and P-item recall requires some consideration in its interpretation. Previously, (chapter two, section 5.3.; this chapter, section 1.1.), the view was expressed that large differences in environmental contexts were more likely to give rise to environmental context effects than small differences. The assumption was that in similar environments, subjects in the DF condition would be able to obtain environmental information compatible with that encoded with the nominal stimuli items and would therefore be able to recall as much as subjects in the SF condition. In the SF condition, as the original presentation environment is also the test environment, the degree of similarity between the presentation environment and the other environment should be a redundant factor. Consequently, if there is a relationship between environmental similarity and subjects' ability to recall, evidence pertaining to this relationship should be found only in the DF conditions.

The first analysis demonstrated there was no significant prediction of P-item recall using only similarity ratings. However, as it had been demonstrated already that apart from the SF or DF type condition, RHTS also predicted P-item recall, it was decided to investigate the ability of similarity ratings to predict P-item recall when the variance explained by RHTS had been removed. These separate analyses revealed that only in the DF condition was there a significant prediction of P-item recall made with RHTS and similarity ratings. However, the actual contribution of SIM was not significant. When RHTS alone was used as the predictor variable; although the prediction of P-item recall was no longer below the 0.05 level, there was no significant loss in the ability of the model to predict P-item recall. The situation is somewhat confused if the role of the alpha level is not appreciated as being quite an arbitrary cut off point.

Although there are theoretical reasons for expecting the degree of difference between environments to influence the environmental context reinstatement effect, it may have been too optimistic to expect this to be reflected in such a crude measure as similarity ratings. However, an alternative is that a non-significant relationship was observed due to a lack of statistical power. This possibility is examined in experiment four (chapter nine), where a sample size of almost double that of the present DF condition contributes relevant data.

4.9. Examination of model predictions.

The observation of an environmental context effect with non-elaborative processing is one piece of evidence in accord with the model of encoding and retrieval presented in chapter six. Of course, the observation of the effect under these conditions only supports one of the predictions made. The second of the specific predictions is examined in the following chapter.

CHAPTER EIGHT

EXPERIMENT THREE

THE INFLUENCE OF HIGH LEVEL SEMANTIC ELABORATION ON THE ENVIRONMENTAL CONTEXT REINSTATEMENT EFFECT

A COMPARISON WITH LOW LEVEL SEMANTIC ELABORATION

1. Introduction

1.1. Aim.

In chapter seven; in line with the prediction made in chapter six, it was shown that low level processing was conducive to the production of an environmental context effect. The main purpose of the experiment reported in this chapter is to examine the validity of the related prediction that high level processing should prevent the manifestation of such an effect.

1.2. Implications of prediction.

The rationale behind this prediction has already been explicated in chapters six and seven, but essentially an environmental context effect is not expected with high level processing because there should be many more ways that a target memory record could be uniquely described. Of course, if the assumption that the inclusion of environmental context information is a robust aspect of processing is adhered to, the implication must be that there is something "better" about the use of non-environmental context information to distinguish between memory records.

Alternatively, if the assumption of the robust inclusion of environmental context information is dispensed with, it could be easily argued that environmental context information is ineffective in distinguishing memory records, because it is not encoded when high level processing is in operation. The question as to whether environmental context information is encoded, but not used, or not encoded at all, will be retuned to in chapters ten and eleven. However, a prerequisite to such consideration is the presentation of credible evidence supporting the hypothesis that high level processing does indeed inhibit the manifestation of an environmental context effect.

1.3. Experimental promotion of elaborative processing.

For the experiment described in chapter seven, a major problem was preventing subjects from engaging in elaborate processing. In the present experiment there is the comparatively easy task of promoting elaborate processing.

One of the main reasons for the ease with which elaborate processing can be elicited from subjects, is their tendency to search for meaning (Bartlett, 1932). In terms of the present model, the search for meaning translates to elaboration, as the cognitive (and emotive) structures that accommodate the stimuli are instantiated.

The benefit of mnemonic strategies such as constructing a story using stimuli words, or forming an integrated image of the stimuli items is well documented (eg. Bower, 1972b; Bower & Clark, 1969; Bower & Winzenz, 1970). The explanations of why such benefit should be obtained from the use of such encoding strategies have much in common with the type of consequences of elaborate processing suggested by the schema model. Baddeley (1976, p.355-357) in particular, identifies the discriminative capacity and redundancy of coding available with semantic elaborate processing.

In an attempt to ensure an elaborate representation of the word stimuli; as well as anticipate what some subjects might do anyway, it was decided to instruct subjects to employ both story construction and imagery mnemonics. Subjects were told to form a story using the words as they were presented and at the same time, imagine the events of their story in their minds eye.

2. Method.

2.1. Subjects:

48 Glasgow University students; males and females of approximately equal numbers, participated as subjects. All subjects were naive regarding the purpose of the experiment and were paid one pound for taking part.

2.2. Environmental contexts:

Rooms A and B were employed in the experiment. Descriptions of these rooms can be found in appendix A.

2.3. Stimuli, apparatus and questionnaire:

The stimuli, apparatus and questionnaire used in this experiment were identical to those used in experiment two; described in chapter seven.

2.4. Design:

A two factor (2 x 2) between subjects design was applied. The first factor was defined by the relationship between the presentation environment and the recall environment; same or different. The second factor was defined by the two learning environments, rooms A and B.

2.5. Procedure:

Again, this was a replication of the procedure employed in experiment two. The only procedural difference between the two experiments was the type of instructions given to subjects prior to the presentation of the stimuli words. Subjects were told that their task was to memorise all of the words that they would be presented with, but in a particular way. It was explained that their ability to remember the words would be aided if they incorporated the words into some sort of story format and at the same time imagined the events of the story; in their minds eye, as they constructed it. Subjects were then asked to use this mnemonic, bearing in mind that their use of the words in the story did not have to follow the order of the word presentation.

2.6. Scoring:

Each subjects' verbal recall protocol was transcribed and scored in the same manner as reported in chapter seven.

3. Results.

3.1. ANCOVA Of P-Item Recall.

3.1.1. selection of a covariate.

As in chapter seven, a stepwise linear regression was carried out to determine the most efficient predictor of P-item recall and to ensure that the overall regression of predicted and predictor variables was linear. The two predictor variables employed were; RHTS and RFAS.[1]

Again, the best prediction of P-item recall was obtained using RHTS only. Although the F-to-enter value for RFAS was 5.01, this value was not significant at the 0.05 level.[2] Table 8.1. gives those values associated with the simple regression equation.

variable	std reg coeff	R^2	adj R^2	F(1,46)	adj p	predictor corr
RHTS	0.451	0.204	0.177	7.68	0.018	0.013

Table 8.1. Summary of simple regression of P-item recall on RHTS.

---

1 For the same reasons as detailed in chapter seven (section 3.1.1.), d' scores were not employed as a predictor.

2 see chapter seven, section 3.1.1.

A chi-square test of homoscedasticity suggested the tenability of this assumption, $X^2(1) = 0.02$, $p = 0.88$. An examination of the plot types described in chapter seven (section 3.1.1.) and obtained with the present data, also suggested that the assumptions underlying the regression analysis were tenable and consequently, that linear regression was appropriate.

3.1.2. preliminary analysis for ANCOVA.

The same type and order of ANCOVA assumption checks described in chapter seven (section 3.1.2.) were followed in this analysis.

The linearity of the overall regression line was supported by the results of the regression analysis described in the previous section. The test of homogeneous within group regression lines did not provide any evidence to the contrary, $F(3,24) = 1.00$, nor did the test of the linearity of the between group means, $F(2,27) > 0.00$.

The top section of table 8.4. presents the data upon which Hartley's F-max test was carried out. This revealed no significant departures from the assumptions of normality and variance homogeneity, $F\text{-max}(4,6) = 2.17$.

The results of these tests indicated that the assumptions underlying the ANCOVA were tenable.

CONDITION	ABA	BAB	ABB	BAA
X RECALL	20.75	17.75	19.12	19.37
S.D.	8.31	8.38	8.89	7.63

CONDITION	SAME FAMILIAR(SF)	DIFFERENT FAMILIAR(DF)
$\bar{X}$ RECALL	19.25	19.24
S.D.	8.34	8.28

Table 8.2. Unadjusted mean recall and standard deviation of P-items.

CONDITION	ABA	BAB	ABB	BAA
X RECALL	8.37	7.37	7.62	8.12
S.D.	0.52	1.30	1.51	0.83

CONDITION	SAME FAMILIAR(SF)	DIFFERENT FAMILIAR(DF)
$\bar{X}$ RECALL	7.87	7.87
S.D.	0.99	1.22

Table 8.3. Covariate mean and standard deviation.

CONDITION	ABA	BAB	ABB	BAA
X RECALL	19.10	19.39	19.95	18.55
S.D.	7.43	6.68	9.18	6.23

CONDITION	SAME FAMILIAR(SF)	DIFFERENT FAMILIAR(DF)
$\bar{X}$ RECALL	19.24	19.25
S.D.	7.06	7.84

Table 8.4. Adjusted mean recall and standard deviation of P-items.

3.1.3. overall ANCOVA F-tests.

A 2 x 2 (presentation-recall environment match/mis-match x presentation environment) completely randomised design was applied. The dependent variable was subjects' P scores and the covariate was their RHTS scores.

Unlike the ANCOVA analysis reported in chapter seven (section 3.1.3.), the main effect of presentation-recall environment match did not even approach significance, $F(1,27) < 0.00$, MSe = 57.86. As before, neither presentation environment, $F(1,27) = 0.04$, nor the presentation-recall environment match by presentation environment interaction, $F(1,27) = 0.09$, was significant. The pooled within group regression coefficient of P-item recall on RHTS was 3.290.

3.1.4. Power.

Assuming that an effect of environmental context; if manifested under these conditions, would have the same magnitude as that observed in chapter seven, the power (1-B) of the relevant F-test in this analysis would be operating at approximately 0.10. In order to achieve the same power at the same level of significance as the comparable F-test reported in the previous ANCOVA analysis, the sample size would need to be increased from 32 to 288 subjects. Alternatively, the power of the F-test could be increased by reducing the level at which the F-value is regarded as significant. However, the probability that the observed difference between the two means occurred by chance, is unity (ie. $p = 1.00$). Obviously, it is not credible to set the level of significance at 1.00, irrespective of the power achieved by doing so.

3.2. Questionnaire data.

3.2.1. encoding and retrieval strategies.

All subjects reported that they had attempted to construct a story and image of the story's events from the words presented to them; as directed by the experimenter's instructions, and that they had reconstructed the story and images as part of their attempt to recall the presented words. However, it seems that most subjects were unable to incorporate all the presented words into one coherent story and instead adopted the (sub) strategy of forming several smaller stories.

In addition to using the story and image mnemonics at retrieval, nine subjects (4 SF, 5 DF) distinguished between the presentation and list words, while one subject in the DF condition reported using "the room itself to jog...memory". Another subject in the SF condition mentioned an attempt to "recall moods and atmospheres that ..(she)..had related with these words" to help her remember.

Another interesting aspect of the recall reports was that only one subject mentioned any difficulty in identifying the presented words from their total remembrances. By way of explanation, the subject attributed this to using the imagery mnemonic too heavily.

3.2.2. rating of environmental similarity.

The median rating of the similarity of the two rooms; A and B, was 30 ($\bar{X}$ = 34.89, S.D. = 21.50).

3.3. Combined Analysis Of Experiments Two And Three.

3.3.1. introduction.

A comparison of the ANCOVA results obtained in experiments two and three reveals that the reinstatement effect observed in the former experiment was not observed in the latter experiment. Nevertheless, in strict statistical terms, this is not equivalent to saying that there is a significantly different reinstatement effect in one experiment as compared to the other. To assess the significance of difference between effects, it is necessary to examine the interaction between the experiments (ie. type of processing) and the environmental reinstatement/change factor. The only way to do this is to carry out a combined analysis of the two experiments. However, there are two issues to consider in relation to such an analysis: one is primarily methodological and the other is primarily statistical, but as always, the boundary between the two is hazy.

3.3.2. subject allocation.

The first issue concerns the preference of random allocation of subjects to conditions. The use of the word "preference" stems from the fact that in many studies where random allocation is impossible, conventional statistical analyses are still carried out. As this is generally acceptable; provided care is taken in the interpretation of the results, obviously it is not an absolute necessity to randomly allocate subjects to conditions.

The preference for random allocation derives from the assumptions underlying the determination of the sampling distribution of a data set and consequently, the test of the null hypothesis. The assumption of randomness is essentially that there is no systematic effect on the allocation of scores. Of course, in most scientific investigations some source of systematic effect on the allocation of scores is hypothesised, but all other effects on scores are assumed to be random. On this basis, if the estimated probability of such a partition of scores is low enough, the null hypothesis may be rejected. However, it is only if the systematic effect on the scores can be attributed to <u>only</u> that source identified as the independent variable, that the rejection of the null hypothesis and acceptance of the alternative hypothesis can be considered to support the model with which the latter hypothesis conforms. Therefore, random allocation of subjects to conditions is one way of trying to ensure that no other source of systematic bias effects the scores. If random allocation is not carried out, it must be carefully considered whether the method of subject allocation would produce any systematic effect on the scores.

In each of the experiments two and three, subjects were randomly allocated to the different conditions. In all experiments reported, subjects were obtained from the same population: Glasgow University undergraduate students. Across all the experiments reported in this thesis, the only difference between experiments in terms of subject allocation, is the time period within which each group of subjects was obtained. However, several experiments were carried out in comparatively quick succession, with each experiment extending over a period of time longer than the gap between experiments. One consequence of this was that subjects recruited towards the end of one experiment were likely to be taking part in this experiment nearer in time to subjects participating at the beginning of the next experiment, than subjects taking part in the same experiment, but who were recruited at its start. So although there is a difference in terms of the periods of time in which the subjects were allocated, these periods are not as distinct as at first they may appear.

If clearly distinct time periods are assumed, the question this raises is what kind of influence could this have? The only plausible effect of such sampling is that different types of subject would be available and so different types of subject would take part in the different experiments. This could happen possibly, if certain experiments were carried out around exam times. It might be expected that in these situations the type of subject taking part in experiments would not have exams, or would be less concerned about them than other subjects who may have participated in experiments conducted at other times. However, this argument seems rather unlikely given the attitudes of the majority of potential subjects to their exams and especially when this was combined with the prospect of earning a pound in about thirty minutes. When it is also considered that there was a subject pool of some ten thousand students, in different and changing pre, post and non-exam states, any real influence is doubtful. In addition, for this factor to exert an effect on the results of any combined analysis, it would have to be assumed that significantly different sub-groups would be selected and that these group types would <u>interact</u> with the variables manipulated <u>between</u> experiments. Such occurrences would seem to be quite unlikely.

Beyond the details of random allocation and the potential consequences of non-randomness, there would seem to be a contradiction between the apparent acceptance of the validity of comparing the <u>results</u> of completed experimental analyses (as in reviews etc.) and the assumed equivocality of a more rigorous and formal statistical comparison. Although it might be argued that the potential differences between experiments; in such terms as have been discussed, are implicit in any review comparison, it seems that they are unlikely to be made explicit unless it suites the purposes of the reviewer.

3.3.3. statistical analysis.

The second issue for consideration concerns the manner of the statistical comparison. The purpose of the ANCOVAs was to examine the experimental manipulations when that variation known to be associated with RHTS was removed. The relationship between RHTS and P-item recall is regarded as reflecting the degree of elaborative processing carried out by subjects at presentation. In ANCOVA the covariation of covariate and dependent variable is eliminated by extracting that variation determined by the deviation from the overall covariate mean. In other words, the variation in the dependent variable due to variation in the covariate is eliminated by adjusting the dependent variable score to that expected if all subjects had obtained the overall covariate mean as their covariate score. Consequently; in order that only variation unaffected by the experimental treatments is removed, the covariate must not be influenced by the experimental manipulations. It is a failure to meet this requirement that prohibits the use of RHTS as a covariate in a combined ANCOVA. Although RHTS is independent of the differences in recall environment within each experiment (as being measured before their introduction assures), this independence does not extend to the type of instructions that subjects received. As would be expected, RHTS is influenced by the level of processing induced by the task instructions. This is clearly revealed by an ANOVA comparison of the covariate scores obtained in the two experiments, $F(1,78) = 25.68$. As there are no alternative acceptable covariates, an ANCOVA cannot be performed.

Although a combined analysis cannot be carried out using ANCOVA procedures, still it is possible to extract the variation associated with the covariate, and compare the effects of the two experiments. In fact, one way of doing this is suggested by the ANCOVA linear model (see chapter seven, section 4.2.).

The model underlying the ANCOVA defines each subject's score as consisting of the usual ANOVA model components, plus a component that is a function of the relationship between the dependent variable and the covariate. By subtracting this component representing the relationship between the covariate and dependent variable, the adjusted score is obtained. The linear model therefore, provides a means of calculating a score for each subject that is free of the variation associated with the covariate.

The two separate ANCOVAs carried out in this and the previous chapter, provide the values required for the calculation of the adjusted scores. Each adjusted score would be based on the particular relationship between P-item recall and RHTS as described by the within group regression coefficient for that processing condition and the average RHTS score for that processing condition.

Adjusting scores on the basis of specific group regression lines, rather than on the basis of an average regression line has advantages and disadvantages. On the plus side is the fact that as more unique information is preserved regarding the different conditions, more accurate adjustments are likely to be made to the scores (Winer, 1971). In addition, as the adjustment to each subject's score is made with respect to the particular covariate group mean, any difference between groups will be maintained. On the minus side is the fact that a degree of freedom is lost with every regression line that is used.

In chapter seven (section 4.2.), the ANCOVA was heuristically described as consisting of two procedures: an adjustment to the scores identical to that described above, and an ordinary ANOVA of these adjusted scores. If this was a totally accurate description of the ANCOVA procedure, it may be wondered why such restrictive assumptions as homogeniety of within group regression coefficients are made, when the different groups could have their adjusted scores calculated separately using different regression coefficients. In fact, some of the ANCOVA assumptions such as the homogeniety of within group regression coefficients and the validity of the covariate overall mean are made on the basis of computational convenience, rather than theoretical necessity.[3]

Unfortunately however, the ANCOVA adjustment is not as straight-forward as previously described and actually involves the use of two regression lines: the overall regression of the dependent variable on the covariate and the pooled within group regression of the dependent variable on the covariate. The former is used to adjust the total sum of squares while the latter is used to adjust the error sum of squares. Treatment effects are obtained by subtraction.

This roundabout method of adjustment is in an attempt to maintain the independence of the F-ratio numerator and denominator, which is one of the F-test assumptions. The adjustment on the basis of the two

---

3 For example, Searle (1971) has provided alternative computational procedures for ANCOVA with heterogeneous within group regression coefficients.

regression lines described above, maintains this independence, whereas an adjustment on the basis of the single within groups regression line could produce correlated scores. However, Lindquist (1953) has shown that if a dependent variable is normally distributed, the numerator (MSBG) and denominator (MSWG) are statistically independent. This is because MSBG depends on the means of the samples, whereas MSWG depends on the variance of the samples and with normally distributed data, the mean and variance are independent. Consequently, if it can be assumed that the adjusted scores are normally distributed, the assumptions of the F-ratio would be satisfied and it would be valid to perform an ANOVA on the adjusted scores.

Although an F-max test has already been carried out on the adjusted scores and this test is sensitive to departures from normality as well as variance homogeniety, it would seem valid to examine the data further for two reasons. First, the F-max test is not very powerful and though this power was satisfactory in relation to the previous analyses, it seems appropriate in this rather novel situation that more stringent criteria be applied. Secondly, a glance at the within group variances of the two experiments suggests that a transformation to achieve homogeniety of variance may be in order. An assessment of the tenability of the normality and homogeniety of variance assumptions will be required after the data transformation.

3.3.3. preliminary analysis.

One method of assessing the assumption of normally distributed data is to carry out a chi-squared test, as suggested by Hays (1973). In order to carry out this test, some set of arbitrary class intervals must be arranged. The number of scores falling into each class interval is compared to that expected under the null hypothesis of a normal distribution. Hays states that the expected frequency in each interval must be five or more. Unfortunately, as there are only sixteen subjects per group (collapsing across the non-significant factor of presentation environment), the maximum number of class intervals possible, with five expected scores, is only three. As the calculation also involves the use of the mean and standard deviation, a total of three degrees of freedom are lost. This leaves a chi-squared value with zero degrees of freedom. Therefore, this test cannot be applied sensibly to this data.

An alternative to the chi-square assessment is that suggested by Engelman (1983) and previously applied to the data reported in chapters three and seven. Engelman employs the ratio of the sample skew and kurtosis values to the sample standard error. This value is read as a Z-score and so can provide a probability of the skew and kurtosis values deviating from the normal by chance. In other words it operates as a significance test.

An examination of the adjusted scores of experiments two and three in terms of the skew and kurtosis values just described (skew/S.E. [SK], kurtosis/S.E. [KU]) suggested that the assumption of normally distributed data was tenable. However, an F-max test indicated a significant departure from the assumption of variance homogeniety, F-max(4,14) = 6.00.

Logarithmic and square root data transformations were investigated, with least variance heterogeniety being observed after the log transformation Y' = log(Y + 2), F-max(4,14) = 3.26. Subsequent examination of SK and KU values suggested that the assumption of normally distributed data was tenable also.[4] Table 8.5. presents the mean, standard deviation, SK and KU values.

processing	LOW		HIGH	
recall cf. presentation.	SAME	DIFF	SAME	DIFF
$\bar{X}$	1.00	0.79	1.18	1.18
S.D.	0.14	0.25	0.19	0.21
SK.	-0.42	-1.49	-0.49	-0.16
KU.	-0.42	-0.12	0.64	-1.14

Table 8.5. Mean, standard deviation, SK and KU values per condition.

3.3.5. overall ANOVA F-tests.

A 2 x 2 (level of processing x presentation-recall environment match/mismatch) completely randomised design was applied. The DU conditions of experiment three were excluded from the analysis to achieve a completely crossed design. The dependent variable was transformed adjusted P-item recall.

---

4 This assessment of normal distribution is actually more stringent than the nominal (two tailed, alpha = 0.05) level would suggest. As the number of SK or KU values increases, there is a greater chance that one or more of these would be above the critical value of 1.96. The same rational leads to the use of the specially constructed F-max tables; rather than the normal F tables, and the alterations to alpha levels when multiple comparison tests are carried out.

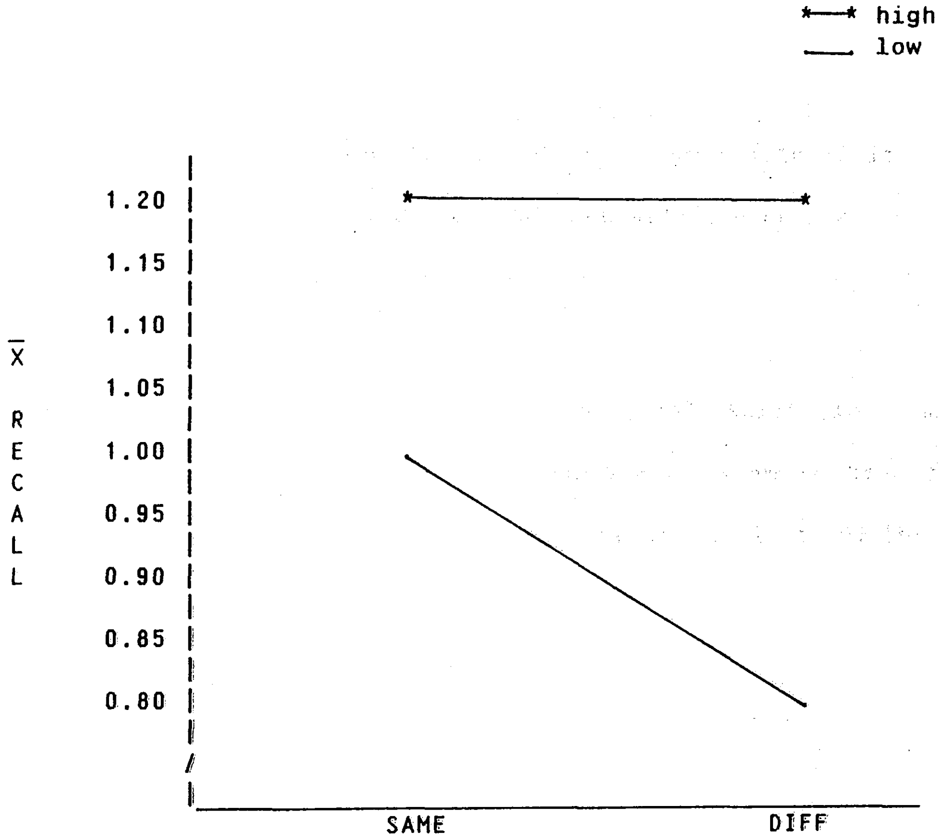

Figure 8.1. Mean transformed adjusted recall by group.

Three significant effects were obtained: level of processing, F(1,58) = 32.26, MSe = 0.04, presentation-recall environment match/mis-match, F(1,58) = 5.10 and the interaction between these two factors, F(1,58) = 4.34. Figure 8.1. illustrates these effects.

4. Discussion.

4.1. Questionnaire data.

One of the most important contributions of this data is its additional support for the presumed view of subjects' encoding and retrieval strategies.

In comparison to the variety of retrieval strategies employed in experiment two, subjects in this experiment seem to have been quite conservative. This probably reflects the productivity of the story and imagery retrieval mnemonics. Subjects are much more likely to change their retrieval strategy when it is not producing any to-be-remembered words, than when it is operating quite successfully.

It is also interesting to note the difference between the responses provided in this experiment compared to the last. Here subjects give much clearer basic accounts of what they did to obtain recall, rather than describing the category of remembered item as they had in experiment two. This is probably a consequence of the conscious mediation of retrieval through the use of the mnemonics. As a larger part of the retrieval process is occurring consciously, there should be a greater ability to introspect.

Over the course of the experiment, at no time were subjects informed that reconstructing/remembering the story would aid recall, although they were informed that the use of such mnemonics at presentation should help their memory for the words. However, if subjects were not aware, or had not assumed the method of employing the mnemonics at retrieval, then the appropriate use the mnemonics at retrieval by all subjects indicated that they learned very quickly.

Although the problem of establishing a null hypothesis to test remains: even on an anecdotal level, the accounts provided by subjects suggest some interesting manners of psychological process operation. Despite the apparent uniformity of the subjects' retrieval strategies in this experiment in comparison with the last (see section 4.2.), the number observed serves to indicate the potential variety of retrieval strategies that subjects can employ.

The comparable ratings of environmental similarity across this and the previous experiment excludes any account of the "non-reinstatement effect" in terms of differing perceptions (descriptions) of the two environments. In turn this emphasises the role of the encoding and retrieval strategies employed in eliminating the environmental reinstatement effect.

4.2. Levels of processing and the relationship between RHTS and P-item recall.

The vastly superior recognition (RHTS) and (P-item) recall performance in experiment three lends strong support to the view that the different tasks engaged in by subjects produced two different levels of semantic processing. One feature of the different levels of semantic processing, is the greater variation of P-item recall in the high level processing condition, before and after the covariate adjustment. The greater variation in high level processing condition scores suggests that a greater variety of successful and unsuccessful retrieval (sub) strategies were employed by subjects in this condition than in the low level processing condition. This result contrasts with the apparent conservativism of the strategies reported by high level processing condition subjects when compared with the reports of low level processing condition subjects. The discord probably reflects a difference in the criteria applied to what merits being reported. It seems likely that a much higher criterion is applied by the high level processing condition subjects, as they have a greater repertoire available. In addition, the subjects view of the purpose of the experiment will influence their criteria of what to report. Consequently, basic reports of the use of the story and imagery mnemonics predominate.

The effect of the greater variation in the high level processing condition pre-adjustment recall scores can be seen in the relationship between the standardised and unstandardised regression coefficients across processing conditions. Although the unstandardised regression coefficient in the high level processing condition is nearly three times that of the low level condition, the standardised regression coefficients (equivalent to the correlation coefficient in simple regression) are quite comparable.

Essentially, what the two kinds of regression coefficient reveal is that while the type of relationship between RHTS and P-item recall is different between conditions; with almost three times as many words being recalled per RHTS item in the high level condition, the extent of the association between the two variables across conditions is similar. Presumably, the ability to predict subjects' P-item recalls from their RHTS scores derives from the successful use of common processes of remembrance in both tasks.

Mandler (1980) has suggested that recognition performance is dependent upon two separate processes: retrieval (of context) and familiarity.

Familiarity is considered by Mandler to be the phenomenonological experience of a match with information in memory that has been integratively processed (see chapter seven, section 1.3). What Mandler describes as the retrieval of context corresponds with that which has been described here as constructing an appropriate retrieval specification and as has been discussed, obtains a large degree of benefit from previous elaborate processing.

It should be expected therefore, that the processes common to recognition and recall would be those that contributed to the production of elaborative memory structures. Certainly, the superiority of RHTS scores in experiment three cf. experiment two is in accord with this view. However, given the degree and extent of elaborative processing carried out in experiment three, there seems to be relatively little association between RHTS and P-item recall. Indeed, it is little more than that observed in experiment two, when elaboration was at a minimum. This, in conjunction with the high P-item variation in experiment three already attributed to the different uses of the story and imagery mnemonics, suggests that the manner in which these mnemonics are employed at recall, does not match the way in which they benefit recognition.

4.3. Level of processing and environmental context effects.

Figure 8.1. clearly reveals the nature of the interaction effect that had been suggested by the separate ANCOVAs carried out on the data collected in experiments two and three. The observation of an environmental context effect when subjects engage in low level semantic encoding, but not when they engage in high level semantic encoding, conforms with the predictions made in chapter six, and so provides support for the schema model underlying the account of encoding and retrieval.

The demonstration that the manifestation of an environmental context effect is influenced by the type of encoding of the nominal stimuli must encourage a reappraisal of the view that environmental context effects are unreliable phenomena.[5] As the majority of studies examining environmental effects have employed intentional learning paradigms, it is perhaps surprising that so many effects have been obtained and reported. However, it is extremely unlikely that the observation of environmental context effects only depends upon the level of semantic processing engaged in by subjects at encoding. Indeed, all of the factors identified in chapter two could exert some influence on the manifestation of environmental context effects.

4.4. Encoding vs retrieval.

All of the factors listed in chapter two are encoding factors. The view of their operation is that in some way they may bias the formation of the nominal item descriptions such that use is not, or cannot be made of environmental context information at retrieval. The obvious consequence of environmental context information incorporated in retrieval specifications being unable to provide any retrieval advantage, is that no environmental context effects are observed.

Focusing attention on the process of memory representation formation seems to lead naturally to consideration of whether or not environmental context information is being encoded. Of course the only means of examining this is to test memory, and this depends on retrieval

---

5 Of course, it is not the phenomena that are unreliable, but rather our knowledge of the conditions that produce them.

processes. So, before any attempt is made to try and determine if the lack of an effect indicates that environmental information has not been incorporated into the nominal items' descriptions, or that it has, but this information is redundant when so many other forms of retrieval specification are available, it would be useful to have more information on the ways in which subjects use environmental context information at retrieval.

In the two experiments reported here, subjects' ability to recall was influenced by manipulating the level of semantic processing engaged in at encoding. Although this was expected to have an effect on retrieval, no specific attempt was made to direct the type of retrieval engaged in by subjects. However, in the next chapter an experiment is reported that attempts not only to influence the level of semantic encoding engaged in by subjects, but also to direct the type of retrieval strategy that they employ at recall. It may be that the manifestation of environmental context effects are affected as much by conditions at retrieval, as at encoding.

CHAPTER NINE

EXPERIMENT FOUR

AN ATTEMPT TO MANIPULATE THE STRATEGIC USE OF ENVIRONMENTAL CONTEXT INFORMATION AT RETRIEVAL

1. Introduction.

1.1. Strategic use of environmental context information.

The work of Smith (1979, expts.2 & 3; 1984) has demonstrated that the use of the retrieval strategy of self-generating the encoding environment when in a different environment can enable subjects to recall as much as subjects recalling in the presentation environment. The success of this strategy would seem to depend on how well the subject can remember the encoding environment. Such a conclusion is suggested by Smith's (1979, expt.3; 1984) findings that the benefit of self-generating the presentation environment is eliminated if subjects experience several environments over the course of the experiment. Presumably, as Smith (1979) has suggested, memory for environmental context is similarly influenced by the same factors as influence other types of memory.

1.2. Object and design of experiment.

In chapter seven, many subjects' self-reports of the retrieval strategies they employed mentioned the use of some environmental context information as a means of identifying a search set. The fact that subjects in the DF condition were employing self-generation of encoding context retrieval strategies raises the possibility that the recalls in the DF condition may have been lower had no such strategies been used.

To properly examine this possibility; using the same basic paradigm as described in chapter seven, two new conditions must be compared. In one condition, subjects should receive instructions that would prevent them from employing a self-generation of encoding environment retrieval strategy, while in the other condition subjects should be instructed to use just such a strategy.

Instructing subjects to think of some alternative environment rather than the encoding environment has the merit of not only being easy to carry out, but also of closely matching the self-generation of encoding environment strategy. Presumably there would be little difference in the psychological processing requirements of thinking of two simple places. In fact, for the purposes of the experimental investigation, different processing requirements are tolerable, provided they are biased in favour of the non-encoding environment condition. Smith (1979, expt.2) had subjects think of a room at home in an attempt to prohibit the use of the self generation of encoding environment strategy, but rather than risk any excessive overlap between the subjects' rooms and the encoding environment, it was decided to instruct subjects to think of a totally different type of environment; namely a beach at sunset.

If the use of a self-generation of encoding environment strategy is effective, then the recalls of those subjects in this condition should be greater than the recalls of those subjects prevented from employing such a strategy. In relation to the experiment reported in chapter seven; if the self-generation strategy is effective, subjects employing it should recall as much as subjects in the SF condition. If subjects in the DF condition of experiment two had benefited from the use of the self-generation of encoding environment strategy and the attempt to prevent the use of this strategy is effective, then these subjects should recall less than those subjects in the DF condition of experiment two.

However, given the experimental paradigm of the two conditions just described, there is still one other potential source of influence on the results. As mentioned before, Smith has demonstrated that subjects who experience several environments before recall in the DF condition obtain no benefit from the self-generation strategy. In the paradigm employed here, subjects in the DF condition spend ten minutes in the recall environment, prior to returning there to recall. It just may be that this is enough to cause a decrement in their ability to remember the encoding environment. A simple modification to the room order for the DF condition enables any such influence to be eliminated. Having subjects return to draw the encoding environment, rather than the recall environment before recall, should if anything, aid their ability to remember the encoding environment at the time of recall.

1.3. Questionnaire alterations.

In the two previous experiments an open-ended questionnaire was administered to subjects at the end of the experimental session. However, in terms of the purpose of the questionnaire, a problem encountered was that subjects' responses to the questions were often at a tangent to the type of response that the questions were intended to elicit. In an attempt to overcome this problem the questionnaire was modified. This involved a slight rewording of some questions and the inclusion of some new ones, which were intended to force the subjects to be more specific about certain aspects of their psychological activities. A specimen copy of the questionnaire is presented in appendix D.

1.4. Paradigm modifications.

In experiments one, two and three, no effect of presentation environment was observed. As a consequence of this and the extra effort involved in changing the presentation environment equipment control, it was decided to abandon the complementary sub-conditions (eg. for ABB, BAA) and present all subjects with the stimuli words in environment A only.

In the previous two experiments, several subjects had come very close to obtaining a recognition hit score (RHTS) of ten out of ten, but fortunately no subject actually achieved a perfect score. If some aspect of recognition performance is to be used as a covariate in the ANCOVA, it is important that this performance is free to covary with recall performance. If the indice of measurement of recognition has a ceiling occurring prior to the recall performance ceiling, then any covariation in recognition and recall performance above this cannot be represented. To reduce the likelihood of this happening, the number of RHTS (and filler) items was increased to fifteen. As the five new RHTS items were randomly selected from the P-item list, the proportions of recognition, filler and P-items were altered (see section 2.3. for details). Increasing the number of RHTS and filler items will obviously increase the time required to complete the recognition task. However, it had been noticed that the vast majority of subjects had completed the recognition test by the first one and a half minutes. Given the approximate time schedule of twenty words in one and a half minutes, it was thought that subjects could quite easily cope with thirty, in two and a half minutes. If comparisons are to be made with previous experiments with the RHTS score used as a covariate in the P-item analysis, any change in the relationship between RHTS and P-item recall would be a potentially disruptive factor. Fortunately, if this was the case then it would be identified by the checks on the assumptions underlying the ANCOVA. Any change in the relationship between RHTS and P-item recall would be indicated by heterogeniety of the within-group regression lines.

1.5. Learning.

So much for the intended format of the experiment. As the results and discussion will reveal, the subjects' behaviour in the experiment would not quite conform to the restrictions imposed by the different conditions.

One of the problems in the present experiment that had been encountered previously in experiment two was the tendency of subjects to try and learn some of the presented words. In experiment two, a stiff criterion was set and any subject reporting an attempt to learn any of the stimuli words had not even had their verbal protocols transcribed. However, probably as a consequence of asking subjects directly if they had made any attempt to learn any of the presented words; in the new version of the questionnaire, a greater proportion of those doing so were identified. As the number of subjects identified as having made an attempt to learn some of the words approached fifty-percent, economy of subjects and effort suggested that it would be prudent to examine the effect of the type of learning reported by these subjects. Generally this learning took the form of an attempt to remember particular words, rather than anything as elaborate as subjects were instructed to carry out in experiment three.

2. Method.

2.1. Subjects:

34 Glasgow University students; males and females of approximately equal numbers, participated as subjects. All subjects were naive regarding the purpose of the experiment and were paid one pound for taking part.

2.2. Environmental contexts:

As before, rooms A and B were employed in the experiment. Descriptions of these environments can be found in appendix A.

2.3. Stimuli:

For the reasons described in section 1.5., five words were randomly selected from the P-item list and added to the list of recognition items. Five more words of a similar type to the whole stimuli set were randomly selected from the Kucera and Francis (1967) word count, to serve as filler items. There were therefore, fifteen recognition and fifteen filler items typed in upper case and randomly ordered in one column on slips of paper, which were presented to subjects to carry out the recognition task.[1]

---

1 A full listing of all stimuli and filler words is presented in appendix B.

2.4. Apparatus:

This was identical to that used in experiments two and three, and described in chapter seven.

2.5. Questionnaire:

A simple questionnaire; based on that presented to subjects in experiments two and three, was administered to subjects after they had completed their experimental session. A specimen copy of the questionnaire can be found in appendix D.

2.6. Design:

A one factor, between subjects design was applied. The three levels of the factor were defined by the order of rooms: AAB or ABB, and the type of retrieval strategy the subject was instructed to use: self-generation of the encoding environment, or the self-generation of an inappropriate environment (ie. a beach at sunset).

2.7. Procedure:

The procedure employed in this experiment contained only slight modifications to that employed in experiment two. Again the experiment is primarily distinguished by the type of instructions given to the subjects. As in experiment two, all subjects were instructed that their task was to keep a count of the number of words presented to them that had three or more vowels. At recall, after being told that they had to try and recall all the words presented to them during the first experimental period, subjects were told that their ability to remember the words they had been presented with would be aided if they imagined a particular scene as they attempted to recall. The particular scene they were asked to imagine depended on their allocated condition. Those subjects in conditions 1 and 2; where they were asked to imagine the encoding environment, were told the "real" reason for doing this, ie. it should help them to remember the stimuli words, while those subjects in condition 3; who were asked to imagine a beach at sunset, were told that the relaxation caused by such a thought should help them to remember. The room order for subjects in condition one was AAB, while for subjects in conditions 2 and 3 it was ABB.

2.8. Scoring:

Each subject's verbal protocol was transcribed and scored in the same way as described in chapter seven. The only difference, due to the change in proportions of recognition, filler and P-items, was the total number of each type of item that could be scored. As well as the subjects' verbal recall, each subject's questionnaire data was scored on four counts. Three of these were nominal designations and identified if the subject had or had not i) attempted to learn any of the words at presentation, ii) consistently maintained the instructed retrieval strategy and iii) utilised the self-generation of encoding environment as a retrieval strategy. The fourth score obtained from the subjects completed questionnaire was their rating of the two environments similarity.

3. Results.

3.1. ANCOVA Of P-item Recall.

3.1.1. identification of P-item predictors.

A stepwise linear regression analysis was carried out to determine the best predictor of P-item recall. Previously, the variables included in such an analysis had been RHTS and RFAS, but here whether or not a subject had made an attempt to learn any of the words (LR) was included as a (potential) predictor variable.

The best prediction of P-item recall was obtained using RHTS only. The F-to-enter values of RFAS and LR; after the entry of RHTS, were $F(1,32) = 0.89$ and 0.68 respectively. Table 9.1. gives those values associated with the simple regression equation of P-item recall on RHTS.

variable	std reg coeff	R^2	adj R^2	F(1,32)	adj p	predictor corr
RHTS	0.479	0.229	0.205	9.51	0.010	0.510

Table 9.1. Summary of simple regression of P-item recall on RHTS.

A chi-square test of homoscedasticity suggested the tenability of this assumption, $X^2(1) = 0.18$, $p = 0.67$. An examination of the previously described plot types (see chapter seven, section 3.1.1.) obtained with the present data, also suggested that the assumptions underlying the regression analysis were tenable.

3.1.2. preliminary analysis for ANCOVA.

As before, the ANCOVA assumptions were examined as part of the complete data analysis. The linearity of the overall regression line was supported by the results of the regression analysis described in the previous section. The tests of homogeniety of within group regression lines, $F(2,28) = 0.20$ and linearity of the between group regression line, $F(1,30) = 0.00$, suggested that the homogeniety and linearity assumptions underlying the ANCOVA were tenable.

Hartley's F-max test, carried out on the data presented in table 9.4. revealed no significant departures from the variance homogeniety and normality assumptions, F-max $(3,10) = 1.66$.[2]

3.1.3. overall ANCOVA F-tests.

A one factor completely randomised design was applied. The three levels of the factor represented the three conditions described in the introduction and method sections. The dependent variable was subjects' P scores and the covariate was their RHTS scores.

The ANCOVA revealed no significant differences between conditions, $F(2,30) = 0.23$, MSe = 7.06. The pooled within group regression coefficient of P-item recall on RHTS was 0.563.

---

2 As there were unequal numbers of subjects across groups, the df of the F-max denominator was calculated from the number of subjects in the largest group (ie. 12). The result of such a procedure is an increase in the stringency of the test (Kirk, 1982).

CONDITION	AAB	ABB	ABB
$\bar{X}$ RECALL	4.70	4.00	5.42
S.D.	2.41	3.16	3.09

Table 9.2. Unadjusted mean recall and standard deviation of P-items per group.

CONDITION	AAB	ABB	ABB
$\bar{X}$ RECALL	9.50	8.92	10.08
S.D.	1.35	3.23	1.97

Table 9.3. Covariate mean and standard deviation per group.

CONDITION	AAB	ABB	ABB
$\bar{X}$ RECALL	4.70	4.33	5.09
S.D.	2.16	2.79	2.76

Table 9.4. Adjusted mean recall and standard deviation of P-items per group.

3.1.4. power.

A post-hoc analysis of the power of the analysis; assuming a mean square effect that would have provided a significant F-value when divided by the obtained MSe value, suggested that the F-test would operate with a power (1-B) of approximately 0.45.

3.2. Questionnaire Data.

3.2.1. subjects making an attempt to learn some words.

Out of the 34 subjects participating in the experiment, 15 responded that they had made an attempt to learn some words. No subject reported that they had specifically tried to link words in any way, but subjects did state that they had "noticed" some links between words.

3.2.2. subjects maintaining the instructed strategy.

Only 10 out of the 34 subjects maintained the instructed strategy. Most subjects employed several strategies to try and recall words. These strategies ranged from randomly generating candidate words to see if any could be identified as having been presented, to using some more sophisticated means of generation such as those items with three or more vowels, or trying to remember the words that had made them think of some particular topic.

3.2.3. subjects in the non-encoding environment generation condition.

Despite being instructed to think of a beach at sunset to "aid" recall, 9 out of the 12 subjects in this condition reported using a self-generation of encoding environment strategy. (These subjects are included in the 24 not maintaining the instructed strategy, reported in the previous section). Apparently, only 3 subjects in this condition did not use the strategy that the condition was supposed to prevent and all of these subjects reported employing some other method of retrieval such as mentioned in section 3.2.2..

3.2.3. similarity of environment ratings.

The median rating of environmental similarity (SIM) was 19 ($\bar{X}$ = 23.81, S.D. = 20.26).

3.3. Regression Model Of P-Item Recall.

As in chapter seven (section 3.4.), a stepwise linear regression was carried out to determine the model that most efficiently predicted P-item recall. The predictor variables employed were; RHTS, RFAS, LR, constant retrieval strategy or not (CS) and SIM. As described in section 3.1.1., only RHTS significantly predicted P-item recall. All other variables failed to make any significant improvement to the RHTS regression model. Table 9.5. lists the F-to-enter values of all the potential predictor variables after the entry of RHTS.

variable	$F(1,30)$[3]
RFAS	1.46
CS	1.15
LR	0.70
SIM	0.14

Table 9.5. Non-significant predictor variables' F-to-enter.

3 Two subjects failed to provide SIM ratings, so for this analysis N = 32.

3.4. Combined Experiment 2 And Experiment 4 ANCOVA.

3.4.1. selection of a covariate.

To equate the RHTS, RFAS and P-item recalls across experiments, the proportion of words correctly recognised or recalled was calculated for each subject. An examination of the proportion distributions revealed RFAS to be the only variable significantly deviating from the assumptions of normality, skew/standard error = 6.31, kurtosis/standard error = 9.47. A square root transformation (Y' = sqrt[Y] + sqrt[Y+1]) was found to restore normality of distribution most effectively.

A stepwise linear regression using the proportion of RHTS (PRHTS) and the transformed proportion of RFAS (RPRFAS) as predictor variables again identified PRHTS (ie. RHTS) as the only significant predictor of proportion P-item recall (PP). The F-to-enter for RPRFAS; after the entry of PRHTS, was 2.53. Table 9.6. gives the values associated with the simple regression of PP on PRHTS.

variable	std reg coeff	R^2	adj R^2	F(1,80)	adj p	predictor corr
PRHTS	0.437	0.191	0.181	18.86	<0.001	0.033

Table 9.6. Summary of simple regression of PP on PRHTS.

As before, a chi-square test of homoscedasticity, $X^2(1) = 1.05$, $p = 0.31$, and the plotted data (see chapter seven, section 3.1.1.) suggested that the assumptions underlying the regression analysis were tenable.

3.4.2. preliminary analysis for ANCOVA.

Prior to the ANCOVA analysis all P-item and RHTS proportion scores were multiplied by 100. This was to reveal differences between scores in terms of more comprehensible percentages rather than between the previous decimal places. The linearity of the overall regression line was supported by the results of the previous regression analysis. The tests of homogeniety of within group regression lines, $F(2,76) = 0.76$ and the linearity of the between group regression line, $F(2,78) = 0.93$, suggested that these assumptions were tenable.

Hartley's F-max test; carried out on the data presented in table 9.9., revealed no significant departures from the homogeniety of variance and normality assumptions, F-max $(3,32) = 1.42$.

CONDITION	S	D1	D2
$\bar{X}$ RECALL	11.15	7.16	8.56
S.D.	7.04	5.48	5.29

Table 9.7. Unadjusted mean recall and standard deviation of percentage P-items per group.

CONDITION	S	D1	D2
$\bar{X}$ RECALL	54.37	65.31	63.33
S.D.	16.72	15.65	15.68

Table 9.8. Covariate mean and standard deviation per group.

CONDITION	S	D1	D2
$\bar{X}$ RECALL	12.64	7.21	8.37
S.D.	5.36	4.81	4.66

Table 9.9. Adjusted mean recall and standard deviation of percentage P-items per group.

3.4.3. overall ANCOVA F-tests.

A one factor completely randomised design was applied. The three levels of the factor represented; the same retrieval context as at presentation (condition SF in experiment 2) and designated here as S, a different retrieval context cf. presentation (conditions DF and DU in experiment 2) and designated here as D1, and a different retrieval context cf. presentation, but with subjects employing the self-generation of encoding/ presentation environment as a retrieval strategy (conditions 1, 2 and 3 in the present experiment) and designated here as D2. The dependent variable was subjects' percentage P-item recall and the covariate was their PRHTS scores.

Figure 9.1. illustrates the significant effect obtained, F(2,78) = 6.22, MSe = 24.34. The pooled within group regression coefficient of percentage P-item recall on PRHTS was 18.72.

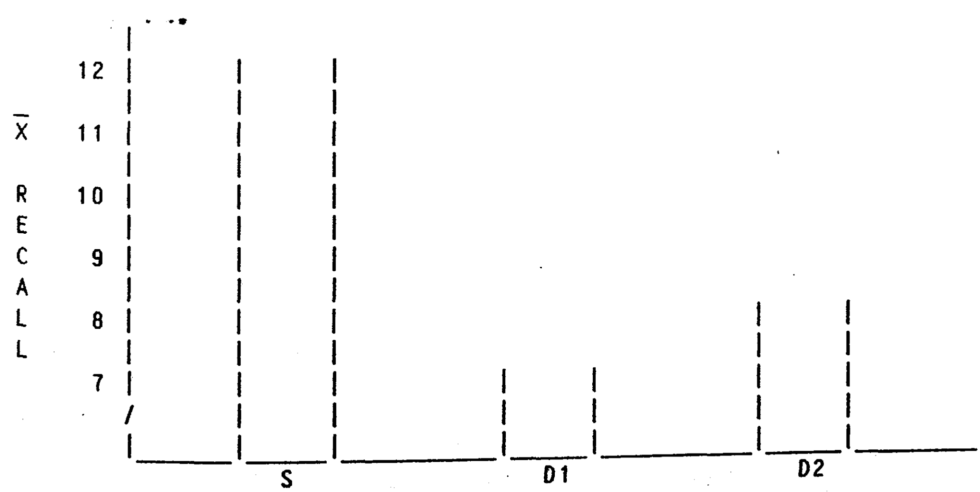

Figure. 9.1. Adjusted mean recall of RPP per group.

3.4.4. specific comparisons.

As the pairwise comparisons between means that were of interest were not orthogonal, the Dunn-Sidak procedure was employed (Kirk, 1982). The (one-tailed) comparison of conditions S and D1 again revealed a significantly greater recall in condition S, $t_{DS/3,78} = 3.538$. The (two-tailed) comparison of conditions S and D2 also found significantly greater recall in condition S, $t_{DS/3,78} = 2.806$, but no significant difference was obtained in the (one-tailed) comparison of conditions D1 and D2, $t_{DS/3,78} = 0.944$.[4]

3.4.5. power and degree of association.

A post-hoc analysis of the power (1-B) of the F-test identifying the significant effect revealed it to be operating at approximately 0.82. The degree of association (W^2) between the effect and subjects' RPP scores was calculated to be 0.114.

---

4 As there is still much debate about the most appropriate method of making multiple comparisons, Fisher's LSD was also calculated for the non-significant contrast. Despite this more powerful test, there was still no significant difference observed, $t(78) = 0.959$.

Critical $t_{DS/3,78}$(alpha = 0.05, one tail) = 2.151; (two tail) = 2.440.

4. Discussion.

4.1. Experiment four.

4.1.1. assessment of design.

Although the power of the ANCOVA in the analysis of experiment four does not match that of the previous analyses identifying significant effects, there are good reasons to suspect that an increase in power still would not reveal any significant effects. First, a comparison of the adjusted and unadjusted means of the three conditions does not suggest that their "differences" are being masked by a large variance component. Of course this argument really only applies to the two conditions in which subjects were instructed to use a self-generation of encoding environment retrieval strategy. As nearly all subjects instructed to use a dummy strategy in fact used the same strategy as all the other subjects, a difference on this basis is hardly expected. As this contrast between the types of self-generation strategy was originally of primary interest, the failure to achieve different retrieval strategies was the reason for concluding the study with such a comparatively small number of subjects in each condition. However, the nature of this failure may reveal as much about the types of processes underlying memory retrieval as would a significant difference between two successfully manifested conditions.

4.1.2. influence of learning.

The result of the regression analysis reported in section 3.1.1., did not find the addition of LR to significantly benefit the prediction of recall. After the variation that could be accounted for by RHTS was removed, LR did not account for a significant amount of variation. RHTS was entered into the regression model first because, when the amount of variation that each variable alone could account for was calculated, RHTS could account for most. However at this point, the F-to-enter value for LR was $(F1,32) = 4.08$. The reduction in the amount of variation accounted for by LR after the entry of RHTS is due to their correlation (r) of 0.46. In other words, RHTS and LR are measuring much the same thing. This suggests that under these conditions, the effect of the degree of elaborative encoding engaged in by subjects, when they report making an attempt to learn some words is being accommodated by their RHTS scores.

Of course there is not a perfect correlation between RHTS and LR. One reason for this is that the RHTS score does not reflect just elaborative encoding. Recognition performance also is influenced by integrative organisation (Mandler, 1979; 1980; Craik, 1981, see chapter seven, section 1.3.. This topic will be returned to in chapters ten and eleven).

4.1.3. self-generation of encoding environment.

Considering the lack of effect between the two self-generation of encoding environment conditions distinguished by room order, it would seem that the ten minute period spent drawing the recall environment did not hamper subjects' ability to recall, nor did the ten minute drawing of the encoding environment increment their ability to recall. To obtain the type of interference effects reported by Smith, more extreme circumstances probably are required. It is most likely that subjects would have to be presented with more than just two rooms over the course of the thirty-four minute experiment, before their ability to remember one of them would be affected.

4.1.4. questionnaire data.

Some of the most interesting findings in the experiment were revealed by the questionnaire. The type of learning engaged in by about half of the subjects has already been discussed. This in itself again demonstrates that subjects in an experiment are not necessarily operating only in the manner that their instructions may require of them.

However, most revealing was the finding that nearly all subjects who were told to think of the dummy environment to "help them remember", actually thought of the same alternative: the encoding environment. Presumably the reason subjects began to think of the encoding environment was because they were having little success with the dummy environment and so adopted a more familiar strategy that had provided them with some success in the past. The number of subjects adopting this single strategy, despite specific instructions to use another, supports the view that indeed this is a commonly employed method of aiding memory, as was suggested in chapter one.

That most subjects did not stick to one (ie. the instructed strategy) method of trying to remember words is not very surprising. A little introspection is enough to suggest that retrieval attempts use many different methods to attain their goal. It seems to be the way of retrieval; and particularly so when difficult, to attempt to employ several routes to the target item. Unfortunately, this aspect of retrieval makes it very difficult to control precisely, especially in free recall type situations, where subjects are provided with relatively long periods of time in which to produce their remembrances. Even if it was possible for subjects to exert total conscious control over their retrieval(s), in a situation where the conditions are purposefully made difficult for them to remember, the naive subject is bound to be tempted to try other potentially more successful strategies to achieve a "better" performance, even if it is "just" to avoid looking quite so stupid.

4.1.5. similarity ratings and overall regression model.

In common with the previous experiments, the ratings of environmental similarity indicate that subjects regarded the two rooms as quite dissimilar. In fact, the mean rating obtained in the present experiment is "just" significantly less than that obtained in experiment three, $t(57) = 2.000$, but is not significantly less than that obtained in experiment two, $t(58) = 1.265$.

As with the regression model for the DF condition in experiment two, SIM did not have any significant influence on the prediction of P-item recall (see section 3.3.). Despite the increase in power through the increase in sample size (error terms comparable), no significant relationship between SIM and P-item recall was observed. In fact, over the two regression analyses, the nature of relationship changed: in the DF condition of experiment two a negative relationship was obtained, whereas in the present conditions a positive relationship was observed.

Whether the lack of a stable and significant relationship between similarity ratings and P-item recall is a consequence of an inappropriate theoretical model or an inappropriate method of investigating the relationship; through the use of a rough and variable rating scale, will have to be determined by further research.

4.2. Combined ANCOVA of experiments two and four.

4.2.1. issues raised by the combined analysis.

The results of the ANCOVA and the specific comparisons raise two major questions. First, why did the self-generation of encoding environment fail to increment the recalls of subjects in the different recall environment conditions to a level comparable with that of subjects in the same recall environment condition; as has been observed by Smith on several occasions. And secondly, why were subjects in the different recall environment conditions that had not been instructed to use the self-generation of encoding environment strategy able to recall as much as those subjects who had been so instructed?

One simple answer to both of these questions is that the self-generation of encoding environment is an ineffectual strategy. However, Smith has produced considerable evidence supporting the utility of such a strategy, and indeed much of the evidence presented in chapter one and the information obtained here from the questionnaire would suggest that there must be some value in employing this strategy, or it would not be so commonly reported and applied.

4.2.2. lack of effect of retrieval strategy; D1 vs D2.

It may be remembered that in experiment two, as well as in this experiment, a number of subjects in the DF condition reported employing a self-generation of encoding environment strategy. Although fewer subjects in the DF condition of experiment two than in the dummy retrieval strategy condition of experiment four, reported employing this strategy, it may only appear that this strategy was less employed. Such an effect could be a consequence of the differences in the questionnaire. In the version given to those subjects in the condition reporting a widespread use of such a strategy, there was a direct question asking if just such a strategy was used, while in the other condition, subjects were only asked what they had done at recall.

In other words it is possible that all the subjects in the different recall conditions may have been employing similar self-generation of encoding environment strategies. This would certainly explain the lack of difference observed between conditions D1 and D2, as in effect there is no difference to be detected.

If the account of the equality of D1 and D2 conditions is valid, the second major question that is raised by the results of the combined analysis of experiments two and four is substantially changed. Rather than asking why condition D1 and D2 recalls are similar, it has to be asked why Smith was able to obtain conditions in which subjects apparently did not use the strategy of self-generating the encoding environment at recall, when such conditions could not be attained here.

It has already been suggested that subjects in the dummy retrieval strategy condition adopted the self-generation of encoding environment as a strategy because of their lack of success with their instructed strategy. Subjects in Smith's experiments were asked to memorise the words as they were presented, so presumably they would find much less difficulty recalling the words than subjects recalling under the present experimental conditions. It may be therefore, that it is the difficulty of recall and or the lack of an alternative that causes subjects to consciously use this strategy. This is in accord with the anecdotal evidence presented in chapter one (section 1.1.), which suggests that the awareness of use of environmental context information tends to occur when retrieval becomes more difficult.

However, presumably as subjects in Smith's experiments were not filling their whole recall period with the constant outpourings of their successful memory retrievals, one would expect them also to have some difficulty in recalling the stimuli words. In such a case it would be expected that at some point they too would employ a self-generation of encoding environment strategy. Therefore, the difference in the use of such a strategy would be one of degree, rather than of use or non-use. The more easily a subject can remember information; which should be related to the amount of elaborative processing engaged in at encoding, the less likely they are to employ a self-generation of encoding environment strategy.

This account of the influence of the degree of elaborative encoding upon the extent of the self-generation of encoding environment retrieval strategy use also accords with the explanation of the failure to increment the recalls obtained in different recall environment conditions, when this retrieval strategy was employed in the present experiment. The premise underlying the explanation is that a memory of something will never be as detailed as the actual entity that it is a record of.

The influence of a high degree of elaborative encoding and retrieval has been described in detail in chapter seven and to a lesser extent in chapter eight. However, to reiterate briefly, a high degree of elaborative processing at encoding provides the means for multiple access to the memory records. In such circumstances the dependence upon any one type of information for successful retrieval is reduced.

Consequently, environmental context information, being one type of information that can be employed to describe an entity, becomes less crucial in the specification of that entity.

If subjects have employed only a certain degree of semantic elaborative processing in their encoding of the stimuli items, there may be benefit to be had from the additional use of environmental context in their retrieval specifications. In such cases the provision of environmental information in its original form will produce recall superior to that obtained if this information is not provided, ie. an environmental context reinstatement effect. With a fairly elaborate representation there will be several ways that the target item could be unambiguously specified, so the benefit to the retrieval process will be limited. In other words, there will only be a slight environmental context effect.

Although the self-generation of encoding environment only produces a memory record that partially represents the environmental information present at encoding, where little further information is required to uniquely specify the target items, this strategy will have the effect of making up the difference in the number of items remembered. The information provided by the self-generation of the encoding environment is sufficient to "top-up" the retrieval specifications such that target items can be uniquely identified and so properly retrieved.

In the situation where virtually no semantic elaborative processing has occurred at encoding, there will be a great dependence on retrieval specifications using alternative sources of information. In such circumstances the retrieval specifications formed will be more ambiguous and generally will require much more information to achieve unique identification of target items. Consequently, the effect of that environmental information provided by the self-generation strategy will not be so great, as the discrepancy between the amount of information required for unique specification cannot be provided by the partial representation of the encoding environment obtained through the use of the self-generation strategy. However, the re-presentation of the encoding environment and the full detail of information it contains will provide sufficient information to allow a greater number of target items to be uniquely specified and successfully retrieved.

4.2.3. when is environmental context information used?

It may seem from the present account of the use of environmental context in retrieval that it is employed only as a last conscious resort, when all other types of strategy have failed. However, it has also been argued that the psychological processes underlying most behaviours only reveal themselves by entering into consciousness, when they encounter difficulty (Bobrow & Norman, 1975; Williams & Hollan, 1981). Later, recognition tasks in which many subjects reported an obliviousness to the colour contexts at test, will be seen to demonstrate environmental context phenomena. However prior to this, the present paradigm, involving free-recall and room environment, will be employed in the examination of another factor indicated as having effect on environmental context phenomena. The following chapter reports an investigation into the influence of increased presentation-test gaps on the environmental context reinstatement effect.

CHAPTER TEN

EXPERIMENT FIVE

THE INFLUENCE OF RECALL DELAY ON THE ENVIRONMENTAL CONTEXT REINSTATEMENT EFFECT.

1. Environmental Context Effects Over Time.

1.1. Evidence and current account.

Although extremely few in number, those studies which have examined the influence of varying presentation-test gaps on environmental context phenomena have found greater effects with larger delays (see chapter two, section 5.5.). However, the two studies manipulating this variable have employed very different levels. Smith (1982) examined presentation-test gaps of thirty seconds and five minutes, while Mayes, Meudell and Som (1981, expt.2) looked at the difference between immediate cued-recall and cued-recall after one week. The common finding of larger effects with longer presentation-test gaps suggests that as time passes, the utility of environmental information in nominal item descriptions for retrieval increases, in relation to the alternative forms of specification. This interpretation is consistent with the previous accounts of environmental context effects based on the model presented in chapter six, but it is not the only account of the effect of time that has been presented.

1.2. Differential accessibility.

Mayes, Meudell and Som (1981) suggest that environmental "features" become less accessible more rapidly than other types of features. As contextual features are important in retrieval (in some unspecified manner), the "cuing" provided by the restoration of the encoding environment aids retrieval and so increments recall.

Although the Mayes et al. account may seem quite simple, there are some difficulties in determining its consequences. This is one result of providing a generally descriptive, rather than process orientated account of the data (see chapter four).

There are actually four separate assumptions involved in the Mayes et al. account. The first is that for retrieval, environmental context information is important. The second is that environmental context information becomes less accessible quicker than other types of information. Thirdly, that restoration of the presentation environment "cues" the environmental context information and finally, that this cued information can be utilised effectively by the system.

This account is handicapped by the vaguery of its assumptions. That environmental context information is important is already demonstrated by the observation of environmental context effects. What is required is an account that details how environmental context information is important. Is it important because it activates environmental information in the memory record? If so, how does it find the relevant memory record? Is it important because without its presence it is unlikely to be incorporated in some sort of retrieval specification? This latter possibility seems unlikely given the style of the Mayes at al. account.

As all other assumptions, but for that regarding the relative accessibility of environmental context information depend directly, or indirectly upon the unspecified manner in which environmental context information is important for retrieval, this would seem to be the only part of the account that could contribute further insight into the retrieval processes that produce the environmental context effects observed.

The hypothesis of differential loss of accessibility to environmental context information raises some interesting theoretical questions. How would the requirement of differential accessibility influence the view of the model underlying the accounts of the environmental context effects? Can an item description be considered as being comprised of autonomous components?

An examination of the structure of the model presented in chapter six would suggest that a memory record can be broken down into smaller component parts, although the extent to which this can be done will depend on the nature of the particular memory record/description. Indeed, it is this ability that allows relevant information to contribute to the new retrieval specification on the subsequent retrieval cycle. Therefore, there is no major obstacle to prevent accommodation of the hypothesis that environmental context information becomes less (or more) accessible than other types of information. But would there be any advantage to such an inclusion? Could inclusion of the relative accessibility of environmental context information improve the account of the environmental context effects reported in chapters seven and eight, and the similar increase in effect with the multiple presentation paradigm?

1.3. Level of semantic processing effect.

Consideration of the potential benefit of the differential accessibility hypothesis in relation to the effects observed in experiments two and three, can be summarised quite succinctly: there is none. As the only difference between the two processing level conditions was the instructions received by the subjects, the different effects of environmental context observed in these two conditions cannot be attributed to any variation in the presentation-test gap affecting the accessibility of environmental context information.

Although the hypothesis of differential accessibility cannot account for the level of semantic processing effects in isolation, it may be worth considering how the hypothesis could interact with the present model to provide an alternative account of these effects.

One possibility is that environmental context information used in the specification of nominal stimuli items becomes less accessible as a consequence of the type of high level processing employed at retrieval.

However, it would seem unlikely that environmental context information would become more inaccessible when high level retrieval processes were operating. Apart from the majority of studies observing environmental context effects using intentional learning paradigms, and so it would seem fair to assume, having some degree of high level processing employed by subjects at retrieval, both theoretical and introspective accounts suggest autonomous retrieval cycles that each utilise many different and varied types of information; especially as retrieval difficulty increases. Given such retrieval processes, it is incredibly difficult to imagine any reason why the use of one type of information on one cycle should inhibit the use of another type of information on another cycle. In the previous experiment it was exactly this type of efficient use of information that eliminated one of the planned differences between conditions.

1.4. Multiple presentation environment effect.

It seems probable that at least one difference between multiple presentation and reinstatement paradigms will be the degree of environmental context effect that they manifest. In chapter nine it was observed that although subjects employed self-generation of presentation environment strategies at recall, this did not increment their recalls to equal that of subjects in the reinstated presentation environment condition. It was argued that the reinstated condition advantage could be attributed to the difference between the ability to remember something and actually experiencing it again. The observation of multiple presentation environment effects indicates that retrieval can be aided by that environmental information obtained from memory, while the results of experiment four would suggest that reinstatement effects should be larger than multiple presentation environment effects due to the superior availability of environmental information. Consequently, reinstatement paradigms are likely to be more sensitive to any benefit provided to retrieval by environmental context information.

Despite the paradigms and presentation-test gaps employed by Mayes et al. and Smith being considerably different, still it would be expected that similar processes would underly the general increase in environmental context effects over time. Although the schema model easily achieves such a parsimonious account, the same cannot be said of the Mayes et al. account.

The Mayes et al. account would require memory information to be accessed differently in the multiple environment situation, in comparison to the reinstated environment situation. The reason for this is that in their account the reinstatement condition obtains benefit through the cueing action of the reinstated environment. Without such cueing, environmental features cannot be accessed to aid retrieval. However, the observation of an increased multiple presentation recall environment effect over time, when such effects can be obtained with recall in neutral (non-cueing) environments, demonstrates that cueing is an unnecessary requisite for the effect.

Of course as an alternative and perhaps last resort, advocates of the Mayes et al. account could argue that the the five minute delay is not long enough to allow environmental context information to become inaccessible and attribute the increase in effect to chance, citing as evidence the lack of an interaction in Smith's analysis.

1.5. Comparison of accounts.

The unique aspect of the Mayes et al. account is the hypothesis that over time environmental context information becomes less accessible than other information. If it is assumed that this does not happen, then in order to account for the increase in environmental context effects with time, it must be assumed that at least the use, if not the presence or accessibility, of other forms of retrieval information decline with time.

Without any qualification; in terms of how the subjects' style of processing could affect the benefit obtained by subjects recalling in the presentation environment, the view taken by Mayes et al. makes the simple prediction that with time, the benefit obtained from a reinstatement of the presentation environment always will increase. This means that given sufficient time between presentation and test, an environmental context effect always will be observed. The problem is identifying what sufficient time is.

With respect to the conditions described in chapters seven and eight, it would be expected that with longer presentation-test gaps, the environmental context effect observed with low level processing would increase, while an effect of environmental context should become manifest in the high level processing conditions. However, if the environmental context reinstatement effect is attributed to a loss of accessibility to environmental context information over time (unless cued by the reinstatement of the presentation environment), retrieval must become more dependent upon alternative forms of nominal item description, especially semantic elaboration. In such circumstances, subjects in a low level processing condition, attempting to recall after some time in a different environment to presentation, would be at a particular disadvantage. Not only would the environmental information have diminished in accessibility, but also they would have very little semantic information from which to construct retrieval specifications. Although subjects in the comparable high level processing condition should experience a similar lack of environmental information accessibility, they still should have access to semantic information upon which to base retrieval.

However, if it is assumed that the emergence of larger environmental context effects over time is a consequence of the decline in the use of other forms of retrieval information, similar predictions are derived. According to this account, low level processing, non-reinstated presentation environment subjects would experience most difficulty due to the lack of appropriate environmental context information and from the further decrement in semantic information available to be used in retrieval as a consequence of the passage of time.

It should be expected that there would be differences in the decrements of recall as a consequence of these different factors. Unfortunately, it is as yet impossible to estimate with any degree of accuracy the size of such decrements and so provide any prediction of effects.

1.6. Redundancy of differential accessibility.

Essentially, the Mayes et al. account depends on the hypothesis that environmental context information becomes more accessible than other forms of information over time. But in fact, if it is assumed that all forms of information become equally inaccessible over time, an advantage for reinstated presentation environment conditions still is predicted. This is because still there would be more environmental context information pertinent to the formation of appropriate retrieval specifications (or even to activate environmental features in a target memory representation) in this condition. This is the account of the reinstatement effect presented in chapter six. Therefore, in terms of the schema model, the hypothesis of greater inaccessibility of environmental context information is unnecessary.

1.7. Some theoretical clarification.

In the preceding discussion there has been much use of the terms environmental context information or "features", and semantic information. It is worth emphasising that such terms should not be taken to refer to any objective types of "real" world information defined in an a priori manner.

If the schema model of processing is examined, it can be seen that the terms environmental and semantic information actually refer to the products of psychological processing. It is the form of the schemata that succeed in accommodating the information received from the sensory apparatus that determines the persons unconscious and conscious perception and interpretation of the "real" world.

Using the terms environmental and semantic information too freely can suggest objective types of information, with some pre-ordained definition in the "real" world, which people do or do not encode. All information enters the system via the sensory apparatus. Consequently, the person's consideration of whether or not this information has some deep meaning depends on the type of psychological processing the original physical information receives.

One consequence of failing to appreciate that there is no objectively defined a priori information types (only common processes) is the application of inappropriate terms of conception in the attempt to comprehend the influence of these factors in psychological processing. The inappropriate abstraction of such information creates a false dichotomy that restricts the manner in which it seems valid for information to be of use.

1.8. Experimental design.

Having discussed the different accounts of the apparent phenomenon of greater environmental context effects with longer presentation test gaps at some length, it would seem appropriate to put the phenomenon and accounts to test. Subjects recall was tested after twenty-four hours using the basic reinstatement paradigm in conjunction with a level of semantic processing manipulation. The reinstatement paradigm was employed because of its presumed sensitivity to environmental context effects, while the level of processing manipulation was included to allow comparisons to be drawn with the previous experiments.

The recall delay of twenty four hours was chosen for both theoretical and practical reasons. It was thought that despite the degree of semantic elaboration in high level processing conditions, a free recall paradigm, rather than a cued-recall paradigm, should manifest any benefit of environmental context. In addition, as a low level processing condition comparison would be possible, any change in effect over time should be able to be detected. Finally, pilot studies employing forty eight hour and one week presentation test gaps, had demonstrated a similar pattern of results across all recall delays and a great difficulty in getting subjects to return after more than twenty four hours.

With the lack of specification of what constitutes sufficient time for such effects to be observed, the lack of any significant increment in environmental context effect can always be accounted for by claiming insufficient time between presentation and test. It is a simple enough exercise to test this hypothesis by increasing the recall delay, but subsequent failures to obtain any change in effects can always be accommodated by recourse to the same argument. Obviously in these circumstances, other criteria would have to be applied to assess the validity of the account.

1.9. Experimental hypotheses.

As detailed previously, it is expected that with increasing presentation-test delays, environmental context effects will increase. In addition, there should be an interaction between the presentation-recall environmental context match/mis-match factor and level of processing, with greater presentation delays.

2. Method.

2.1. Subjects:

64 Glasgow University students; males and females of approximately equal numbers, participated as subjects. All subjects were naive regarding the purpose of the experiment and were paid one pound for taking part.

2.2. Environmental contexts:

Rooms A and B were employed in the experiment. Descriptions of these rooms can be found in appendix A.

2.3. Stimuli, apparatus and questionnaire:

The stimuli used in this experiment were identical to those used in experiment four; described in chapter nine, while the apparatus employed was the same as had been used since experiment two; described in chapter seven.

2.4. Design:

A two factor (2 x 2) between subjects design was applied. The first factor was defined by the type of instructions given to subjects at presentation, ie. the level of processing induced. The second factor was defined by the relationship between the presentation environment and the recall environment; same or different.

2.5. Procedure:

Subjects were required to attend two 10 minute sessions, 24 hours apart. At the first session subjects were randomly allocated to one of the processing conditions. This determined if they received task instructions as described for experiment two, or as described for experiment three. After the presentation of the stimuli words, subjects were given the recognition task to complete. As before, subjects were given 2.5 minutes to mark the words they recognised. At the end of the recognition test after the ten minute session was over, the subjects were taken to the waiting area where they were told that the first part of the experiment was complete and that the next day they would have to do the same thing again, but with a different set of words. The reason for the experiment was given as an investigation into the development of cognitive strategies. It was hoped that this would prevent subjects from being tempted to rehearse any of the presented items prior to their next visit. When subjects returned the following day, they were again randomly allocated to either the same, or a different environment (cf. presentation) to recall. All subjects received the same instructions to try and recall all the words that they had been presented with the day before, including any from the recognition test, that they could remember. As before, subjects were told to say each word out loud; as it would be recorded, to start as soon as the experimenter left the room and to keep trying to remember more words until he returned. After ten minutes had passed, the experimenter returned to stop the subjects' free recall. All subjects then returned to the waiting area where they were paid and debriefed. For the same reasons as described in chapter nine there were no complimentary room order sub-conditions. Therefore, the room order for subjects in the same (S) conditions was AA, and for the different (D) conditions it was AB.

2.6. Scoring:

Each subjects' verbal recall protocol was transcribed and scored in the same manner as reported in chapter nine.

3. Results.

3.1. ANCOVA Of P-Item Recall.

3.1.1. selection of covariates.

As before, stepwise linear regressions were carried out to determine the most efficient predictors of P-item recall in the two processing conditions and to ensure that the overall regression of predicted and predictor variables was linear. For the same reasons as detailed earlier, the two predictor variables employed were; RHTS and RFAS.

In the low level processing condition, again RHTS proved to be the variable most highly correlated with P-item recall with no significant contribution being made by RFAS (F-to-enter = 0.05), but in the high level processing condition, this situation was almost reversed. Here, the variable most highly correlated with P-item recall was RFAS. However, in this condition the prediction of P-item recall made by RFAS was not significant and obviously, nor was the contribution of RHTS (before entry of RFAS, F-to-enter = 1.10, after entry of RFAS, F-to-enter = 0.69). Table 10.1. gives those values associated with the simple regression equations.

Chi-square tests of homoscedasticity (RHTS $X^2[1] = 0.81$, $p = 0.37$; RFAS $X^2[1] = 1.16$, $p = 0.28$) and an examination of the plot types described in chapter seven (section 3.1.1.) and obtained with the present data, suggested that the assumptions underlying the regression analyses were tenable and consequently, that linear regression was appropriate.

cndt	varbl	std reg coeff	R^2	adj R^2	F(1,30)	adj p	predictor corr
LOW	RHTS	0.611	0.374	0.353	17.89	<0.001	0.301
HIGH	RFAS	0.230	0.053	0.021	1.68	0.368	-0.124

Table 10.1. Summary of simple regression of P-item recall on RHTS.

Although the prediction of P-item recall by RFAS was not significant, it was decided to use RFAS as a covariate in the ANCOVA as still it increased the power of the analysis by reducing the error variance.

3.1.2. preliminary analysis for ANCOVA.

The linearity of the overall regression line in each processing condition was supported by the results of the regression analyses described in the previous section. The test of homogeneous within group regression lines did not indicate significant heterogeniety in either condition; low level $F(1,28) = 2.18$, high level, $F(1,28) = 0.52$.

As there are only two groups in each ANCOVA, the usual method of assessing the linearity of the within group regressions, ie. by testing the linearity of the between group regression line, cannot be applied. With only two groups, there is no way of telling if the straight line between the two points identified by the covariate and dependent variable means is, or is not appropriate. Expressed more formally, it means that there are insufficient degrees of freedom on the numerator to allow an F-test to be performed. As an alternative to such an analysis, the graphical procedures suggested by Draper and Smith (1981) and referred to by Huitema (1980) were employed. These procedures are identical to those used to assess the tenability of the assumptions underlying the overall regression analyses, but applied to each group, rather than to the whole data set.

For both low level processing conditions, tests of homoscedasticity did not indicate any significant departure from the assumptions, $X^2(1) = 0.41$, $p = 0.52$; $X^2(1) = 0.04$, $p = 0.85$, nor did examination of any of the plot types described in chapter seven (section 3.1.1.). A similar conformity to the regression analysis assumptions were observed with both high level processing conditions, $X^2(1) = 0.28$, $p = 0.60$, $X^2(1) = 1.32$, $p = 0.25$.

Hartley's F-max test; carried out on the data presented in table 10.4., indicated that there was a significant departure from the normal distribution and/or heterogeniety of variance in the high level processing condition, $F\text{-max}(2,15) = 4.48$, but not in the low level processing condition, $F\text{-max}(2,15) = 2.03$. An examination of the distributions of the significantly heterogeneous conditions revealed that two unusually large scores (unadjusted; 41 and 28) were the primary cause of the failure to meet the ANCOVA assumptions. However, as these scores were also the main cause of the difference between the two means, it was decided to rely on the robust nature of the ANCOVA,[1] rather than attempt a rectifying transformation.

---

1 Huitema (1980) provides a discussion of the consequences of assumption failures in the ANCOVA. For a more detailed account see Glass, Peckham and Sanders (1972).

	LOW LEVEL PROCESSING		HIGH LEVEL PROCESSING	
CONDITION	AA	AB	AA	AB
$\bar{X}$ RECALL	3.56	2.62	15.69	11.56
S.D.	1.55	1.75	9.05	4.21

Table 10.2. Unadjusted mean recall and standard deviation of P-items per group.

	LOW LEVEL PROCESSING (RHTS)		HIGH LEVEL PROCESSING (RFAS)	
CONDITION	AA	AB	AA	AB
$\bar{X}$ RECALL	9.31	8.37	0.81	0.81
S.D.	2.63	2.58	1.22	0.83

Table 10.3. Covariate mean and standard deviation per group.

	LOW LEVEL PROCESSING		HIGH LEVEL PROCESSING	
CONDITION	AA	AB	AA	AB
$\bar{X}$ RECALL	3.39	2.80	15.69	11.56
S.D.	1.55	1.75	8.76	4.14

Table 10.4. Adjusted mean recall and standard deviation of P-items per group.

3.1.3. overall ANCOVA F-tests.

Two separate ANCOVAs were carried out with low and high processing condition data. In both analyses the two groups were distinguished by whether subjects' recall; after 24 hours, took place in the same, or in a different environment to presentation. In the low level processing condition the covariate was RHTS and in the high level processing condition the covariate was RFAS. In both ANCOVAs the dependent variable was P-item recall.

In neither of the ANCOVAS was there a significant effect of presentation-recall environment match/mismatch: low level processing condition, $F(1,29) = 1.45$, MSe = 1.82, regression coefficient = 0.376; high level processing condition, $F(1,29) = 2.81$, MSe = 48.52, regression coefficient = -1.620.

3.1.4. power.

Assuming the same percentage difference between groups (relative to the overall mean) as obtained in experiment three, the power (1-B) of the F-test on the low level processing condition analysis would have been operating at > 0.99.

3.2. Low and high level processing conditions combined analysis.

3.2.1. introduction.

As in chapter eight, a combined analysis of adjusted P-item recall was undertaken.

3.2.2. preliminary analysis.

Hartley's F-max test carried out on the adjusted scores of the low and high level processing conditions revealed that there was a significant and extreme departure from the assumptions of normality of distribution and/or the assumption of homogeniety of variance, $F\text{-max}(4,14) = 65.55$.

Logarithmic and square root data transformations were investigated, with greatest conformity to the assumptions being observed after the log transformation, $Y' = \log(Y + 1)$, $F\text{-max}(4,14) = 3.33$. Subsequent examination of SK and KU values suggested the tenability of the assumption of normally distributed data. Table 10.5. presents the mean, standard deviation, SK and KU values.

processing	LOW		HIGH	
recall cf. presentation.	SAME	DIFF	SAME	DIFF
$\overline{X}$	0.62	0.56	1.16	1.07
S.D.	0.14	0.13	0.24	0.15
SK.	0.82	-1.06	-1.19	-0.57
KU.	-0.66	-0.01	0.78	-0.69

Table 10.5. Mean, standard deviation, SK and KU values of transformed adjusted data per condition.

3.2.3. overall ANOVA F-tests.

A 2 x 2 (level of processing x presentation-recall environment match/mis-match) completely randomised design was applied. The dependent variable was transformed adjusted P-item recall.

In contrast to the combined analysis reported in chapter eight, in this analysis only the level of processing factor was observed to exert a significant effect, $F(1,58) = 49.41$, $MSe = 0.03$. Neither presentation-recall environment match/mis-match, $F(1,58) = 1.37$, nor the interaction between this and level of processing, $F(1,58) = 1.14$, was significant.

3.3. Combined analysis of 14 minute and 24 hour environmental context reinstatement experiments.

3.3.1. introduction.

To formally examine the relationship between recall delay, level of processing and presentation-recall environment match/mis-match, an analysis of the adjusted data from experiments two, three and five was undertaken.

3.3.2. preliminary analysis.

As the total number of P-items that could be recalled by subjects in experiment five was less than that possible in experiments two and three (50 cf. 60), the proportion of total recall achieved by each subject was calculated. As Hartley's F-max test indicated extreme heterogeniety of variance, F-max(8,14) = 65.54, arcsine, logarithmic, square root and inverse transformations were investigated. The most successful transformation in reducing variance heterogeniety was $Y' = 1/y + 12$, where y represents the proportion recall, F-max(8,14) = 9.17. Fortunately, this F-max value is just under the value at which it is suggested by Keppel (1982) that the validity of the ANOVA should be of some concern.

The major disadvantage with the transformation described above is that the order of highest to lowest score is reversed. This can cause some confusion in the interpretation of the means. Table 10.6. presents the means, standard deviations, SK and KU values.

delay	14 MINUTES			
processing	LOW		HIGH	
recall cf. presentation.	SAME	DIFF	SAME	DIFF
$\bar{X}$	3.14	3.83	2.00	2.03
S.D.	0.52	0.73	0.43	0.47
SK.	0.43	1.62	0.51	0.17
KU.	-0.46	-0.16	-0.68	-1.19

delay	24 HOURS			
processing	LOW		HIGH	
recall cf. presentation.	SAME	DIFF	SAME	DIFF
$\bar{X}$	3.60	3.73	2.16	2.40
S.D.	0.34	0.28	0.64	0.43
SK.	-1.14	0.43	1.29	0.56
KU.	-0.46	-0.58	0.63	-0.73

Table 10.6. Mean, standard deviation, SK and KU values of transformed adjusted data by condition.

3.3.3. overall ANOVA F-tests.

A 2 x 2 x 2 (recall delay x level of processing x presentation-recall environment match/mis-match) completely randomised design was applied. The dependent variable was transformed adjusted P-item recall.

All main effects were significant; recall delay, $F(1,112) = 5.02$, $MSe = 0.27$, level of processing, $F(1,112) = 248.96$ and presentation-recall environment match/mis-match, $F(1,112) = 9.95$.

Only one interaction effect was significant: recall delay x level of processing x presentation-recall environment match/mis-match, $F(1,112) = 5.45$. Figure 10.1. illustrates the effects reported. To counter the reversal of mean order caused by the transformation and bring the form of the figure more in line with previous plots of such effects, the Y-axis has been inverted.

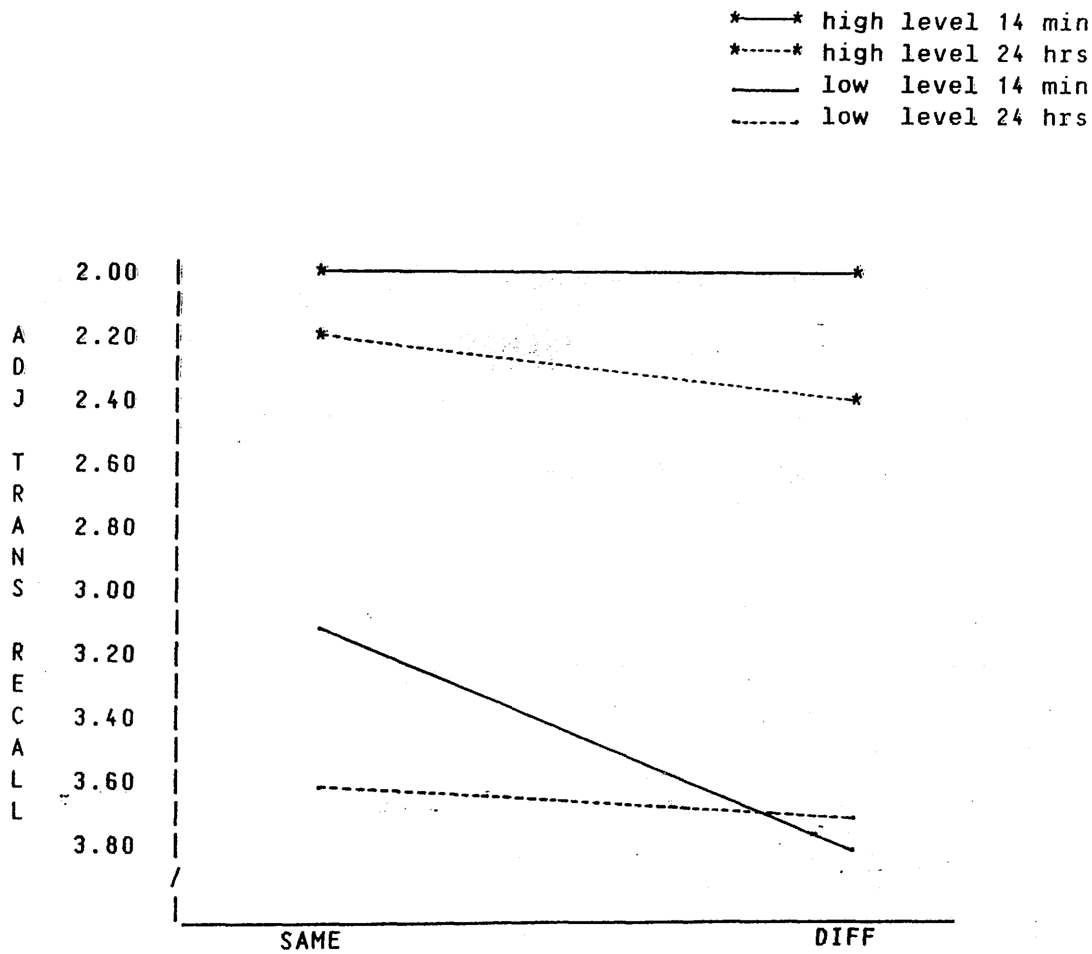

Figure 10.1. Mean transformed adjusted P-item recall by group.

3.3.4. specific comparisons.

The (simple main effect) comparisons of interest are between the two levels of the presentation-recall environment match/mis-match factor under the different combinations of the other two factors. Essentially, it is to check that; under the same power, only the low level processing, 14 minute delay conditions exhibit an environmental context reinstatement effect.

Planned orthogonal contrasts were carried out on the two means representing the environmental reinstatement and non-reinstatement conditions under each combination of recall delay and level of processing. A significant difference was observed between the means of the presentation-recall environment match and mis-match conditions in the 14 minute recall delay and low level processing combination, $t(112) = 4.161$, but not between the similar means in any of the other recall delay and level of processing combinations: 14 minute recall delay and high level processing, $t(112) = 0.114$; 24 hour delay and low level processing, $t(112) = 0.708$; 24 hour recall delay and high level processing, $t(112) = 1.329$.

The comparisons above make use of the robustness of the t-test with equal sample sizes, to cope with the variance heterogeniety observed. However, the validity of the results of these comparisons is supported by the fact that the only other comparison to approach significance (24 hr delay and high level processing), if calculated separately, would provide an MSe greater than that employed in the previous comparisons, $MSe = (0.64^2 + 0.43^2)/2 = 0.30$.

4. Discussion.

4.1. Low level processing RHTS and P-item recall relationship over twenty-four hours.

The finding that over twenty-four hours the unstandardised regression coefficient reduces from 1.246 to 0.376, while the standardised regression coefficient increases from 0.417 to 0.611 requires some explanation. The reason for these changes is the variance associated with the covariate and (particularly) the dependent variable. As recall decreases over the twenty-four hour period, this is reflected in the number of P-items recalled per RHTS item: the unstandardised regression coefficient. However, it is also the case that as recall delay increases, so the variance of P-item recall decreases. This leads to the increase in the standardised regression coefficient, as this indice takes variance into account. Therefore, the pertinent question is why does the P-item recall variance decrease over the twenty-four hours. This variance decrement is impressive and significant as F-tests carried out on the error variances from the 14 minute and 24 hour unadjusted and adjusted score ANOVAs reveal; unadjusted, $F(42,30) = 5.11$, adjusted, $F(41,29) = 5.65$.

It would be expected that a greater number of items could be recalled in the fourteen minute condition in comparison to the twenty-four hour delay condition. This difference should provide a greater potential range of scores. However, a greater potential range does not explain why subjects actually produce recalls that are more or less alike. Presumably, the reduction in recall variation reflects a reduction in the number of different types of strategy being employed to retrieve the target items. As the relationship with RHTS increases over time, it suggests that the strategies employed by subjects at recall must be those that have communality with those employed at recognition.

As was mentioned in chapter eight (section 4.2.), Mandler (1980) has argued that recognition performance can operate on the basis of integrative and elaborative processing. Although Mandler's account of recognition performance is in terms of a "two-process theory" (Eysenck, 1984; Tulving & Thompson, 1973), the ideas it contains can be expressed conveniently in terms of the schema system.

As outlined in chapter seven (section 1.4.), in the schema system, integrative and elaborative processing can affect the form of both the encoding descriptions and retrieval specifications created. A retrieval specification may be constructed by employing the recognition item encoded via integrative processing and incorporating other related and potentially useful information obtained via elaborative processing.

If the information provided by the recognition probe is sufficient in itself to provide a unique specification of the relevant memory record on an early retrieval cycle, then the matching process will occur extremely quickly and may appear as some authors have described it, as "direct access" (Baddeley, 1982b; Mandler, 1980; Tulving & Thompson, 1973; Wickelgren, 1979). However, in the terms of the model presented in chapter six, "direct access" is a consequence of the greater speed with which a discriminable retrieval specification can be constructed (a full account of the rational underlying this argument will be presented in the next chapter).

Mandler (1980) claims that the utility of integrative processing for "retrieval" decreases at a greater rate than elaborative processing.

Presumably this is because the integrated representation employs perceptual schemata, which normally identify only a small number of features in relation to the number of semantic features that can be identified. As time passes, the discriminability of these "perceptual" descriptions will decrease as similar descriptions are constructed. For these reasons it would seem that the reduction in low level semantic processing P-item recall variation after twenty-four hours must be attributed to a reduction in the efficiency; if not the number, of successful retrievals based on integrated information across subjects. The nature of the successful retrieval strategies engaged by subjects becomes more uniform as the possibility of retrieving items on the basis of integrated information (alone) decreases. Therefore, the strategies common to recognition and free-recall should be those employing elaborately processed information.

The question that follows this account is how could subjects in the fourteen minute delay conditions have obtained benefit from integrative processing, when they were presented with a free recall task? The answer to this is that as reported in chapter seven (section 3.3.1.) a good number of subjects turned their free-recall task into a recognition task by self-generating and imagining candidate words on the wall (where the stimuli had been projected). It is interesting to note that in the SF condition where most subjects reported such a strategy, the prediction of P-item recall by RHTS was greatest. Although there are many problems associated with any attempt to draw conclusions on the basis of the questionnaire data, it is tempting to suggest that the highly significant prediction in this condition is in part a consequence of the particular correspondence between integrative and elaborative processing engaged by subjects at recognition and free-recall.

4.2. High level processing RHTS and RFAS relationship with P-item recall over twenty-four hours.

Unlike the quantitative changes observed in the prediction of P-item recall in low level processing over twenty four hours, high level processing prediction undergoes a qualitative change over this time period. The account of the similarity of high and low level processing condition standardised regression coefficients, presented in chapter eight (section 5.1.), provides the basis of the account of the RHTS to RFAS predictor change.

Here and in chapter eight, it has been suggested that comparable standardised regression coefficients are a consequence of similar retrieval strategies underlying recognition performance. However, it has been suggested also that the greater variance in adjusted recall in the high level processing conditions was due to the employment of other retrieval strategies; specifically those based on the story and visual imagery mnemonics. As these strategies were very unlikely to be employed in the recognition task, only that part of recall performance that was not based on those strategies could contribute to the relationship with RHTS. Consequently, the variance produced by the mnemonics was not accounted for and was taken to be the cause of the greater variance in adjusted recall.

Applying a similar logic to the change in predictor variable, the loss of RHTS predictive power is taken to be a result of dissimilar strategies being employed at recognition and recall. As the vastly superior adjusted recall must be attributed to the efficient retrieval enabled by the use of the story and visual imagery mnemonics, it would follow that the reason for the change in predictor variable was also caused by subjects concentrating on the use of strategies based on these mnemonics, rather than fully re-employing those retrieval strategies used at recognition. The insignificant, but superior prediction of RFAS is probably a consequence of this variable representing the extent of subjects' guessing because they have not uniquely specified the word stimuli. Without unique specification, subsequent retrieval cannot be selective. The consequence is that unless the retrieval specification is generalised, recall will diminish. However, it should be remembered that the prediction of P-item recall made by RFAS was not significant and was very similar to that of RHTS (1.10 cf. RFAS, 1.68).

4.3. Recall delay, level of processing and environmental context effects.

The failure to observe any effect of environmental context with either high or low level processing after twenty four hours is contrary to the findings of the other studies examining environmental context effects over time. As mentioned in the introduction, it could have been argued that the reason for the lack of an environmental context effect, is insufficient time between presentation and test, either to allow the loss of accessibility to environmental features without cueing, or to allow the discriminability of nominal items on the basis of semantic specification to decrease, such that retrieval on the basis of environmental context information provides advantage. However, as the environmental context effect with low level processing is eliminated by the twenty four hours recall delay, it is extremely unlikely that insufficient time is the cause of the lack of effect. Presumably an account of the reduction in the environmental context effect observed in the low level processing conditions over time will provide some insight into why no environmental context effect emerged in the high level processing conditions over the same time period.

The observation of almost exactly the opposite effects after a recall delay of twenty four hours to that predicted indicates not only the incorrect nature of the accounts presented, but perhaps also a slight naivety of the accounts. It would seem reasonable to expect the components of environmental context phenomena to conform to the style of normal memory operation, yet the accounts of environmental context effects after recall delay make no allowance for this. This is perhaps a consequence of designing models specifically to give an account of an effect, rather than designing a model that primarily attempts to achieve the goal of memory retrieval within defined constraints (see Marr, 1982, p.347-349). Effects can indicate the constraints under which the processes operate and as such should be accommodated by the model, but they should not be the sole reason for the model's existence.

The lack of an effect of environmental context in experiment five indicates that no advantage was obtained by the subjects in the reinstated presentation environment at recall. The ability of subjects in both conditions to recall something indicates that they were able to construct some retrieval specifications that were successful in retrieving information from memory. In a situation where subjects are having difficulty in recalling the nominal information and they have ample time to explore alternative retrieval strategies, it seems unlikely that they would employ strategies that ignored available and potentially useful environmental context information. Experiment four has already demonstrated how unlikely this is. The only alternative therefore, is that the environmental context information available for inclusion in the retrieval specifications; with which successful retrieval was previously achieved, is now no longer useful. It seems most likely that for such information to have lost its utility, the form of the nominal item memory records must have changed, such that the discriminability of nominal items on the basis of unique environmental context information no longer affords significant advantage. In other words, the environmental context information incorporated in the nominal item memory records is represented with less specificity.

One of the basic characteristics of memory; revealed in a multitude of studies, is the way in which experiences become abstracted over time. Experiments employing text (eg. Bartlett, 1932; Thorndyke, 1977) shapes (eg. Carmichael, Hogan & Walters, 1932), faces (Reed, 1972), etc. have shown that while in the short term memory may maintain considerable detail, over time this is lost until eventually, only the gist of the experience remains. Although the state of knowledge with respect to the exact representation of "natural" environment makes the conception of an abstract form a rather speculative exercise, the ubiquity of the phenomenon is such that it would seem unreasonable not to expect similar changes to occur with the representation of environmental context.

The effect of time on the memory for most materials; the decline of detail and the emergence of gist, has been well accommodated by schema type models (eg. Abelson, 1981; Alba & Hasher, 1983; Graesser & Nakamura, 1982). Minsky (1975) suggests that each frame (or schema) would be organised in a hierarchical manner, with the information most important or fundamental to the particular entity represented at the top of the hierarchy, and the less important information at successively lower levels of the hierarchy. Variation in essential aspects of a representation (the top levels of the frame) would reduce the predictive utility and therefore the validity of the representation. The loss of detail and the emergence of gist in remembrance is explained by the the loss of instantiated values at the lower levels of the frame, but the retention of the fact that the frame was instantiated. Over time therefore, remembrance comes to be based more and more on the prototypical representation.

If the same interpretation is applied to the representation of the two room environments employed here, it would suggest that the same prototypes are used in the formation of the descriptions of the two different environments. The differences between the two environments are represented at relatively low levels of the representational hierarchy and are lost with the passage of time. As the detail of information required to distinguish one room environment from the other is no longer available in the memory records of the nominal stimuli items, no advantage is obtained from attempting to recall in the presentation environment. By the time recall is required, the environmental context representations of these particular rooms have become very similar through their prototypical abstraction. Subjects may have representation of the oddity of the rooms and the bare, sparse and blandly decorated building annexe in which the rooms were situated. Indeed, it is probably the similarity in oddity and location (see appendix A), that leads subjects to employ similar prototype schema in their representation of the two environments. This in turn contributes to the loss of discriminability as unique information is represented only at low levels.

In contrast, many other studies of the effect of environmental context on memory; including Mayes et al., have used a "normal" environment and an "unusual" environment. Using this combination of environments may facilitate the environmental context effect over time. It is very likely that the prototypes employed to form a description of an ordinary room or environment are different to those employed to form a description of an unusual environment. In such a situation as the Mayes et al. study, the differences in the environmental context representation would not be restricted to the lower levels of the representations. As time passed, discrimination would be possible on the basis of environmental context information, but not with some alternative forms of specification due to a parallel loss of detail. Consequently, the potential utility of environmental context information for retrieval would increase, resulting in greater benefit to reinstated presentation environment conditions and therefore, larger environmental context effects.

The general account of how environmental context information aids retrieval of nominal items also suggests why Mayes et al. observed that after the week delay, subjects tested in the unusual environment reinstatement condition could remember less than subjects tested in the ordinary environment reinstated condition. It would be expected that the remembrance of an unusual environment would be more difficult than the remembrance of an ordinary environment. The description of an ordinary environment is likely to have been formed longer and retrieved more often than an unusual environment memory record. There is good evidence indicating that prior retrieval benefits subsequent retrieval (Bjork, 1975; Bjork & Geiselman, 1978; Tulving, 1966). In addition, default values on the prototype are likely to be better approximations to the real instantiated values of the ordinary environment than the unusual environment, so reconstruction of a memory record utilising the defaults will be more accurate.

Although the environmental context representation is part of a memory record, the model presented suggests that the record is a composite. As the previous account suggested, it is possible that there would be difficulty in obtaining the exact form of the original memory record that is to match with the target description. For those reasons above, it would be expected that ordinary room environmental context descriptions would be more accurately defined and easily assembled as part of the memory record, with a consequent increase in nominal item discriminability and success of retrieval.

4.4. A model assumption and prospective tests.

Throughout this chapter there has been the implicit assumption that irrespective of the form of encoding processes, environmental context information would be incorporated in the description and consequently, the memory record formed. However, in chapter eight (section 1.2.) an alternative account of the lack of environmental context effect was mentioned. This account simply asserts that under certain conditions environmental context information is not encoded and therefore cannot provide advantage at recall.

The following chapter attempts to examine this possibility with respect to the high and low level semantic elaboration encoding conditions. The experiment reported also examines the generality of the schema model account by employing a different form of environmental context and a different test of memory.

CHAPTER ELEVEN

EXPERIMENT SIX

AN INVESTIGATION OF THE RELATIONSHIP BETWEEN RECOGNITION, LEVEL OF PROCESSING AND PRESENTATION TIME, IN TERMS OF THE ENVIRONMENTAL REINSTATEMENT EFFECT, USING BACKGROUND COLOUR AS THE ENVIRONMENTAL CONTEXT

1. Introduction.

1.1. Control of environmental context.

One of the biggest problems in the study of the influence of environmental context is obtaining and maintaining effectively different environments. As discussed in chapter three, at present there is some difficulty determining the extent to which room type environments differ. For more detailed investigations of the specific influence and effect of environmental context, the ability to determine and easily control environmental context would be extremely useful.

In the short term, one way of obtaining accurate description and control of environmental context is to employ simple univariate types of environment such as stimuli background colour. Although the study of the influence of such environments has its own validity and application, for it to be useful in the manner outlined above, similar manipulations with different types of environmental context must be shown to produce similar environmental context effects.

1.2. Design restrictions with an unintentional learning paradigm.

The use of an unintentional learning paradigm in the presentation of nominal stimuli items restricts the use of a within subjects design. Obviously, subjects given one stimuli presentation, followed by a test of their memory for these items, cannot be expected to be naive enough to believe the experimenter when again he says that they do not have to learn the items on the second presentation. An alternative procedure is to have a single presentation and to incorporate the different types of test conditions in one test session. Unfortunately, with recall this could result in an interaction between retrieval strategies and remembered items between test conditions. In turn, this can give rise to theoretical complications and ambiguity regarding the influence of the former and latter retrieval strategies, and the originally recalled items on subsequently recalled items, in relation to the particular effect under consideration.

Recognition procedures can alleviate the problems of testing memory under different conditions in the one test session without overlap from previous conditions. This is achieved by presenting <u>independent</u> sets (conditions) of nominal stimulus items and/or cues. As each subject is presented with the task of indicating whether or not the unique item and condition combination is a member of some previously seen stimulus group, there should be less likelihood of any relations between items being formed.

In room type environmental context studies, there is the extra problem of organising subject transits from one room to another, first to manipulate the independent variable, but additionally to maintain adequate experimental controls in the within subjects design.

Happily however, apart from providing a greater ability to control the (nominal) environmental context variable, the use of background colour also facilitates the use of a within subjects design. As background colour context can be manipulated independently of the room type environmental context, there is no need for numerous subject transits. Subjects can remain stationary, while the experimental variables and controls are manipulated in terms of the classical conception of the experimental materials.

The use of background colour; as the nominal environmental context, and recognition procedures allow the construction of an experimental methodology that seems able to overcome the problems normally encountered when an unintentional learning paradigm, in conjunction with a within subjects design, is considered as a means of investigating the effect of environmental context on memory. However, the suggestion that recognition should be employed as the memory test procedure raises the issue of the different abilities of various memory tests; polarised in recall and recognition, to detect environmental context effects.

1.3. Comparison of recall and recognition.

The differences between recall and recognition effects have been discussed and debated in theoretical terms for some time (eg. Brown, 1976; Broadbent & Broadbent, 1977; Mandler, Pearlstone & Koopmans, 1969; Rabinowitz, Mandler & Patterson, 1977a; 1977b). The major difference between views is whether or not the two behaviours are served by the same psychological processes.

Two-process models (eg. Anderson & Bower, 1972; Kintsch, 1970) have distinguished between a retrieval stage and a second recognition/decision stage. Recall involves both stages. However, when a copy-cue is presented for recognition it is assumed that access to the memory representation is direct. Consequently, recognition is assumed to involve only the latter of the two stages.

Other views of memory performance; such as that expressed by Tulving (1982; 1983) in his GAP system and that presented by Norman and Bobrow (1979) in their account of descriptions, regard both recall and recognition as requiring the operation of retrieval processes.

Essentially the former and latter views are in dispute regarding the manner in which access can be gained to a representation in memory. The question is whether access can be achieved directly or must involve retrieval.

At present there would seem to be some confusion amongst psychologists as to what direct access actually means. For example, Baddeley (1982b) uses the term direct access, but seems to be using it to express a phenomenon, or phenomenological experience. However, authors such as Mandler (1980), Tulving and Thompson (1973) and particularly Wickelgren (1979), seem to regard direct access as a process akin to that employed in the types of associative memory systems presented by Hinton and Anderson (1981).

Unfortunately, both of these uses of the notion of direct access are unacceptable in terms of a process account of memory performance. This is most obvious with Baddeley's use of the term. Baddeley uses direct access to describe a situation and not to describe any process underlying memory performance.

However, it is probably the use of the term by such as Mandler, Tulving and Thompson, and Wickelgren, that has the potential to cause most confusion. The reason for this is that the notion of direct access is only valid within a particular form of descriptive abstraction. However, within any of the forms of use in providing a process account of memory performance, the notion of direct access is invalid. By including the notion of direct access in their accounts of recognition, these authors seem to be prepared to allow some sort of magic to "explain" recognition, but not recall. It is the double standard of explanation that is likely to cause confusion and the continued use of an invalid notion under the guise of a component process that is likely to maintain it.

In terms of the computational theory of a process (ie. the what and why), the use of the notion of direct access is insufficient to constrain any procedure that may attempt to access information in memory. A particular elaboration of the concept of direct access would be necessary for it to have validity within an account expressed in such terms of abstraction.

As presented by the likes of Mandler, Tulving and Thompson, and Wickelgren, the notion of direct access would seem to be presented as an attempt to explain <u>how</u> the access to memory is achieved. In this sense, the notion is part of the abstraction expressed in terms of representations and algorithms. However, the manner of direct access as presented by these authors is fully detailed by the term direct access. In other words, no form of computation (formal or informal) is described that would achieve direct access.[1]

In fact, in terms of a representation and algorithm account of the processes underlying memory performance, it can be shown that the notion of direct access is a contradiction in terms. If the purpose of an account expressed in such terms of abstraction is to describe at a level sufficient to allow such processes to be carried out, then stating that access to memory occurs directly begs the question of <u>how</u> direct access is achieved.

---

1 Wickelgren does mention the relation between "direct access" and associative memory. However, despite hinting that procedures are required to achieve direct access, he continues with the use of this misleading term.

In the associative memory systems presented by Hinton and Anderson, great detail is gone into regarding how this "direct" access could be achieved. In effect there can be no direct access to memory representations. This is most easily illustrated if the requirements for comparing a stimulus with a memory representation are considered. To begin with, prior coding of this stimulus must transform its external representation to that which is compatible with the form of representation employed by the nervous system. In addition and perhaps more importantly from an information processing point of view, the neural representation also must have representation of those aspects of the stimulus that define it as distinct from others. The construction of such representations (on the basis of internally, as well as externally generated information) is a necessary preliminary to any comparison with information in memory.[2] In such a system how can any memory access ever be described as direct? At best the term "direct access" is an extremely bad simile that expresses some combination of ease and speed of one example of memory performance relative to some other.

1.4. The effect of memory tests and contexts.

In the review of context effects presented in chapters one and two, one of the main conclusions was that the major difference between semantic, (perhaps) process, physiological and environmental contexts was in terms

2 The account of target description formation in the schema model is one attempt to detail; albeit in a different form of abstraction, the processes involved in producing such a representation.

of the paradigms that were sensitive to their effect. These paradigms differ in the amount of information they provide subjects with, pertinent to the formation of retrieval specifications for the to-be-remembered items. This manner of scaling concords with the degree of environmental context effect observed. The free-recall paradigm, which provides the least retrieval information, is most sensitive to changes in environmental context, while the paradigm providing the most retrieval information: recognition, is rarely effected by such changes.

The schema model presented in chapter six provides a process oriented account of the manner in which this information could be used. According to the model, the lack of effect of environmental context with recognition is due to the superior constructability and discriminability of the retrieval specifications made possible by the provision of all the information constituting a copy-cue. However, as the schema model advocates a communality of processes for recall and recognition, it should be possible to obtain an advantage for reinstated environmental context conditions even with recognition, provided the redundancy of environmental context information for retrieval specifications can be eliminated.

Using the same method as described in previous experiments; namely to have subjects count the number of words containing three or more vowels, it should be possible to restrict the amount of semantic elaboration carried out in the production of an encoding description of each word. As before with free-recall, this should emphasise the role of environmental context information in the descriptions formed.

Consequently at retrieval, environmental context information should be a more effective information component in discriminating between memory records. The greater availability of this information in reinstated environmental context conditions at recognition should increase the constructability of appropriate discriminatory retrieval specifications and should result in an environmental context (reinstatement) effect with recognition.

The review also concluded that all the different types of context had essentially similar effect. The model presented to account for environmental context effects makes no special provision for environmental information; never mind different types of environmental information, but instead regards all types of information as (potentially) equivalent. Therefore, it would be predicted that provided the particular environmental context information enables and is employed to discriminate between target and other memory records, environmental context effects will be observed.

As the schema model makes no distinction in terms of the potential use of different sources of information, environmental context information in the form of nominal stimuli background colour should be predicted to affect memory retrieval in the same manner as room type environmental context. Similar effects should be observed under similar conditions.

However, in order to replicate the relationship between nominal stimulus items and room type environmental context with background colour context, a slightly different design to that normally employed in experiments investigating the influence of colour context should be applied. Often these studies present the nominal stimuli items on several different background colours. This is equivalent to presenting some nominal stimulus items in one room and other nominal stimulus items in other rooms. Consequently, these studies have more in common with the investigations of environmental context interference and multiple encoding environments carried out by Smith (1979; 1982; 1984). To achieve equivalence with the normal single environment reinstatement paradigm, the nominal stimulus items should be presented on only one background colour and tested with this colour, or with one other colour as the (nominal) changed environmental context.

1.5. A different account of the lack of an environmental context effect.

Throughout this thesis, any failure to obtain an effect of environmental context has been discussed primarily in terms of alternative forms of nominal item retrieval specification making discrimination on the basis of environmental context information redundant. This account is based on the assumption that, although different forms of description were created at encoding, environmental context information always was a basic part of each of these and consequently, always was incorporated in the nominal stimuli item memory records.

However, as has been mentioned (see chapter eight, section 1.2.), there is another way of looking at such results. According to this view, the lack of an environmental context effect simply indicates that environmental context information has not been incorporated in the nominal stimuli item memory records.

Differences in processing resources have generally been considered in terms of a general pool of processing resources that enable the operation of the particular processes (eg. Broadbent, 1958; 1971; Kahneman, 1973; Moray, 1967; Neisser, 1967; Welford, 1952; 1968). In these terms, the lack of an environmental context effect may be attributed to the exhaustion of the pool of processing resources; such as general processing space, through the demands made by elaborative processes, which precludes or inhibits specification in terms of environmental context information.

However, Allport (1972; 1980) has suggested that the notion of a general pool of processing resources is inappropriate. Instead, Allport suggests that processing resources, or capacity is distributed throughout the nervous system by virtue of the existence of autonomous, modular systems, which are oriented to particular processing purposes. From this perspective, processing capacity is limited by the nature of the specific processing function and so requirements, by the specific type of data input, or output from other processes and by the ability to keep different goals concurrently active (Allport, 1980).

Although elaborate description formation has been outlined only informally, it seems plausible that all three of the factors listed above could be involved in preventing the incorporation of environmental context information in the descriptions formed at encoding.

Unfortunately, an examination of the effect of varying the presentation time of nominal stimulus items in relation to the two tasks intended to manipulate the degree of elaborative processing is unlikely to provide information enabling some conclusion to be drawn with regard to the debate between the different views of processing limitations. In both accounts, the processing limitations could be overcome by adopting a sequential strategy; specification via environmental context information, then via semantic referents, or vice versa. However, varying the presentation time and so effectively the time that each subject has to carry out the required task (per item) should allow some determination to be made as to whether the lack of environmental context effect is a consequence of time restrictions; as both accounts of processing limitation suggest, or the nature of the elaborated description (representation) formed at encoding, irrespective of the ability to encode environmental context information in the time available.

The nature of memory research is such that performance is always based on memory retrieval. Consequently, when constructing models of memory it is important to bear in mind that the lack of any particular effect need not have consequence for any process underlying memory performance, other than retrieval. The initial discussion in this section illustrates this. Just because no environmental context effect is observed does not mean that environmental context information has not been incorporated in the particular memory records. It remains a possibility that the lack of effect is a consequence of the type of memory retrieval employed. Consequently, the amount and degree of information required as the basis for drawing inferences regarding the nature of the memory representation requires that specific patterns of effects; rather than any single effect, be obtained.

1.6. Model predictions, experimental hypotheses and design.

If, as the review of the literature on environmental context effects suggests, the observance of an environmental context effect depends upon the subject having enough time to encode the environmental information in the description formed at encoding (see chapter two, esp. section 5.4.), then reducing the presentation time should reduce the environmental context effect. Also, if the lack of an environmental context effect in high level processing conditions was due to the inability of those processes encoding environmental context information to operate effectively through a limit on processing ability, caused by the introduction of those processes underlying semantic elaboration, then an increase in presentation time should increase the chances of the serial application of processes being completed and therefore should increase the likelihood of observing an environmental context effect in memory performance.

To test these ideas, a recognition paradigm using background colour as the (nominal) environmental context in a mixed design was employed. Four independent groups of subjects were established. Two groups received low level semantic processing instructions, as reported in chapter seven and the other two groups received high level semantic processing instructions as reported in chapter eight.

One of the low level groups received the nominal stimulus items at a presentation rate identical to that reported in all previous experiments (excluding experiment one), while the other low level group were presented with each nominal stimulus item for only half of the normal presentation time.

Similarly, the two high level groups differed in the presentation times of nominal stimulus items. One group received items at the normal presentation rate, while the other group were presented with each nominal stimulus item for twice as long as normal.

As was mentioned earlier, to equate the use of environmental context between rooms and background colour, at presentation each subject received all nominal stimulus items on the one background colour. At recognition, subjects were presented with half of the nominal stimuli items previously presented, on the same background colour and the other half of the items on a (common) different colour.

The first experimental hypothesis; upon which the others rest, is that the results of experiments two and three will be replicated. In other words, with a nominal stimulus presentation rate of one item for 3 seconds and an inter-stimulus gap of 1.5 seconds, there should be an environmental context (reinstatement) effect with low level processing, but not with high level processing.

Beyond this, if presentation time is a determining factor in the obtainment of environmental context effects for any of the reasons previously described, then in the low level processing, short presentation time condition, the environmental context effect should be reduced, while the environmental context effect should increase in the high level processing, long presentation time condition.

2. Method.

2.1. Subjects:

64 Glasgow University students; males and females of approximately equal numbers, participated as subjects. All subjects were naive regarding the purpose of the experiment and were paid one pound for taking part.

2.2. Environmental contexts:

Two colours: red and blue, were used.

2.3. Stimuli:

2.3.1. Form of presentation:

Each stimulus word was presented in black upper case letters in the middle of a rectangle of colour on a black/blank screen (see figure 11.1.).

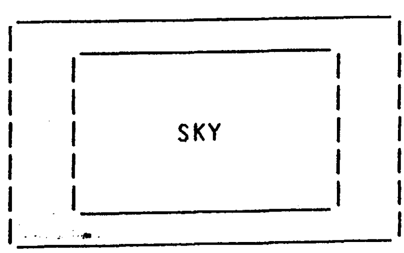

Figure 11.1. Form of presentation of word on colour.

2.3.2. Low level and high level 3 seconds presentation conditions:

40 of the 85 items used in the previous experiments were randomly selected and employed as presentation items. Another 40 items; randomly selected from the remaining set of words employed in the previous experiments, were used as filler items in the recognition test.

2.3.3. High level 6 seconds presentation condition:

To prevent a ceiling occurring with respect to the number of items that could be correctly recognised in this condition, an extra 28 items; similar to those in use, were selected from the Kucera and Francis (1967) word count. Half of these items were presented as stimuli items, while the other half were presented as filler items in the recognition test.

2.4. Apparatus:

An Apple II microcomputer was linked to a Sony portable colour television that had been modified to behave as a VDU. A footswitch and a response box also provided input to the micro.

2.5. Design:

Initially, a two factor mixed design was applied. The first factor (between subjects) was defined by (i) the level of processing induced by the instructions and (ii) the stimulus presentation times. The resulting four conditions were low level processing, 1.5 seconds presentation (L1.5), low level processing, 3 seconds presentation (L3), high level processing, 3 seconds presentation (H3) and high level processing, 6 seconds presentation (H6).

The second factor (within subjects) was defined by the relationship between the presentation background colour environment and the recognition background colour environment: same or different.

Within each first factor condition, half of the subjects were presented with the stimulus items on the red background and the other subjects received the stimuli items on the blue background colour. In conditions L1.5, L3 and H3, for each colour background presentation group, two word sets were created. One word set consisted of the 40 presentation items and the other word set consisted of the 40 filler items. From each of these word sets, 20 randomly selected items were presented at recognition on the same background colour as at presentation and the other 20 items were presented on the other background colour. In condition H6, there were 54 presentation items and 54 filler items, and consequently, 27 randomly selected items in each of the two background colour recognition conditions. The colour background pairing with words was constant across subjects, while a new order of presentation was randomly generated for each subject.

A counter-balance was also operated for preferred hand. In each of the experimental conditions, half of the subjects responded YES with their preferred hand, while half responded NO with their preferred hand.

2.6. Procedure:

The same basic procedure as used in experiments two and three was employed here. Subjects were randomly assigned to one of the first factor conditions. This assignment determined whether they received instructions as in experiment two, or three. Subjects were seated in a softly lit sound proof booth in front of the VDU. They were told that one after the other, a series of words would be presented on the screen in front of them. Each word was separated by an inter-stimulus gap of 1.5 seconds, similar to that obtained with the projector presentations in experiments one to five. No mention was made of the colour background. Subjects were given their instructions as dictated by their assigned condition and told to open the booth door when a message appeared on screen telling them that the presentation was complete. After the presentation the experimenter returned and asked the subjects to draw the booth and everything in it, in as much detail as possible. They were asked to keep drawing until the experimenter returned, turning the paper over and starting a new drawing if they felt they had finished the original. After 14 minutes the experimenter returned and explained that all the words that they had been presented with earlier were going to be shown again along with as many words that they had not seen earlier. They were told that their task was to try and identify as quickly, but as accurately as possible, all those words that they had been presented with previously. If they recognised a word from this set, they were to press the response key marked YES and if they decided that it did not belong to this set they were to press the response key marked NO. Subjects were asked which their preferred hand was; and depending upon their random allocation to preferred or non-preferred hand group, the response box was appropriately linked and labelled. Subjects were then asked to find a comfortable position in which to hold the response box before they initiated the presentation of each stimulus item. Subjects were told that a single word would be presented onto the screen every time they pressed the footswitch and would remain on screen until they made a response. To allow time for the subjects to prepare for each stimuli item, it was explained that there would be a 2 second gap between each footpress and each word appearing on screen. Subjects were asked to begin the task as soon as the experimenter left the booth and to come outside when a message appeared on screen informing them that the task had been completed. After completing the experiment, subjects were paid and debriefed.

2.7. Scoring:

The number of correct responses: recognition hits (RHTS) and correct rejections (CR), and incorrect responses: false rejections (FR) and false alarms (FA), were calculated for each subject. In addition, the times taken to make these decisions were also recorded and each subject's median time for each response type was calculated.

3. Results

3.1. Response type.

As subjects in the high level processing 6 seconds presentation encoding condition (H6) had been presented with twenty-seven, rather than twenty words of each type (ie. recognition item on presentation colour, recognition item not on presentation colour, filler item ...etc.), the proportion of correct responses made was calculated for each word type for each subject.

Although d' values could be calculated from the data available, the fact that some subjects did not produce false alarms raises problems similar to those mentioned in chapter seven (section 3.1.1.). In this particular situation it is also the case that as subjects with no false alarms are potentially high d' scorers, they would be expected to be most numerous in the condition(s) affording most advantage to performance. Consequently, the result of excluding these subjects would be to reduce the effect of any advantage provided by any such condition(s).[3]

---

3 There is a parallel here with the conditions under which a weighted means analysis should be carried out.

3.1.1. Hits.

3.1.1.2. **preliminary analysis.**

The proportion of items correctly recognised by each subject was converted to a percentage. The only reason for such a procedure was to simplify the comprehension of these values. The percentage values were then analysed as a two factor (4 x 2) mixed design. An F-max test of the between subjects factor normality and homogeniety of variance assumptions suggested that both were tenable, F-max(4,15) = 1.92.

CONTEXT			CONDITION L1.5	L3	H3	H6
	same	X	53.44	73.44	85.00	81.25
		S.D.	20.95	12.34	11.55	14.28
	diff	X	55.31	57.19	83.44	85.65
		S.D.	16.48	15.91	13.99	14.62

Table 11.1. Mean and standard deviation of percentage hits per condition.

3.1.1.2. **overall F-tests.**

The two factor mixed ANOVA (processing/presentation encoding condition x retrieval colour context) was carried out on the hit percentages. This revealed main effects of encoding condition, F(3,60) = 16.97, MSe = 399.46, and retrieval context, F(1,60) = 4.02, MSe = 66.28, and an interaction between these factors, F(3,60) = 10.30 (see table 11.1. and figure 11.2.).

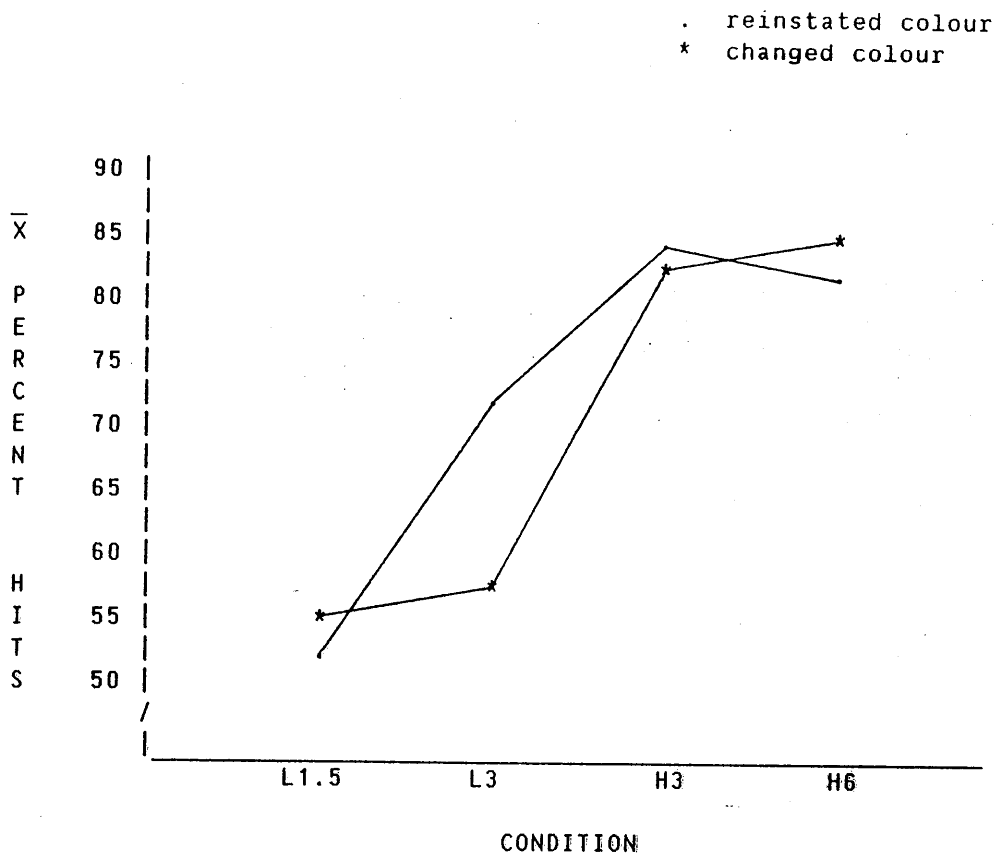

Figure 11.2. Mean percentage hits per condition.

3.1.1.3. specific comparisons.

The comparisons of interest are between the four levels of the processing/presentation time encoding factor, and between the two levels of the colour context reinstatement/change factor under the different levels of the processing-presentation time factor. These comparisons should determine which processing-presentation time conditions differ and which encoding conditions exhibit a colour context reinstatement effect.

Planned (one tailed) orthogonal contrasts were carried out on the means representing the conditions described above. It was predicted that overall performance would follow a downward monotonic trend from condition H6 through to condition L1.5. Also, it was predicted that reinstated colour conditions would exhibit superior performance in comparison with changed colour conditions.

No significant difference was observed in overall performance between conditions H6 and H3, $t(60) = -0.154$. However, a significant difference was observed between the encoding conditions H3 and L3, $t(60) = 3.784$, and L3 and L1.5, $t(60) = 2.189$.

Of the reinstated vs. changed colour comparisons, only two were in the predicted direction. These were the L3 and H3 conditions. Of these only L3 exhibited a significant advantage of colour context reinstatement, $t(60) = 5.646$. The calculated t-value of the H3 one tailed comparison was $t(60) = 0.543$. The remaining two (two-tailed) comparisons were not significant; L1.5, $t(60) = -0.651$ and H6, $t(60) = -1.528$.

3.1.2. False alarms.

3.1.2.1. preliminary analysis.

The proportion of items incorrectly identified as presentation items was analysed in the same manner as recognition hits. An F-max test of the between subjects factor normality and homogeniety of variance assumptions suggested that one or both, were not tenable, $F(4,15) = 5.35$. Subsequently, logarithmic, square root and arcsin transformations were examined in terms of their ability to reduce the F-max statistic. The transformation producing the lowest F-max value was, $Y' = \mathrm{sqrt}(Y) + \mathrm{sqrt}(Y + 1)$, $F\text{-max}(4,15) = 2.98$.

3.1.2.2. overall F-tests.

A two factor ANOVA (factors defined as in section 3.1.1.2.); carried out on the root transformed data, revealed a main effect of encoding condition, $F(3,60) = 9.09$, $MSe = 0.07$, but not of retrieval context, $F(1,60) = 1.69$, $MSe = 0.03$. However, the interaction between these two factors approached significance, $F(3,60) = 2.39$, $p = 0.078$ (see table 11.2. and figure 11.3.).

CONDITION

CONTEXT		L1.5	L3	H3	H6
same	X	26.25	27.50	8.75	9.49
	S.D.	15.76	13.17	7.42	7.65
diff	X	24.06	18.12	10.31	9.49
	S.D.	16.35	14.13	10.71	6.04

(a)

CONTEXT		L1.5	L3	H3	H6
same	X	1.60	1.63	1.29	1.30
	S.D.	0.26	0.19	0.20	0.21
diff	X	1.55	1.46	1.32	1.33
	S.D.	0.29	0.26	0.22	0.14

(b)

Table 11.2. Mean and standard deviation of non-transformed (a) and transformed (b) false alarms per condition.

3.1.2.3. specific comparisons.

A significant difference was observed between the overall means of encoding conditions H3 and L3, $t(60) = 3.583$, but not H6 and H3, $t(60) = -0.197$ and L3 and L1.5 $t(60) = 0.393$.

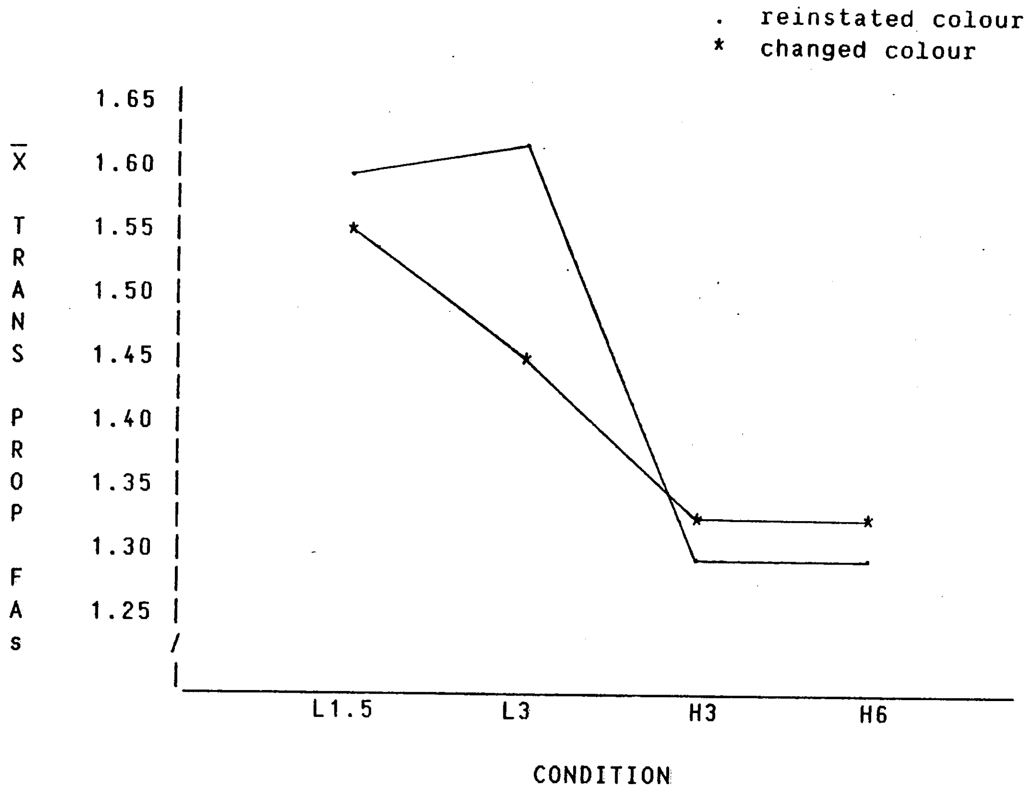

Figure 11.3. Mean transformed proportion of false alarms per condition.

Although the interaction between the encoding condition and retrieval context factors was not significant, the F-test operated is two-tailed. As a prediction is made with respect to the direction of the effect, ie. more false alarms should be made with those items presented on the reinstated colour context, it is legitimate to investigate the pairwise comparisons of interest with a <u>one</u>-tailed t-test.

Only the two low level processing encoding conditions exhibited the predicted trend. Of these only the L3 condition exhibited a significant effect, $t(15) = 2.820$, L1.5 condition, $t(15) = 0.738$. Neither of the two high level processing conditions exhibited a significant effect; H3, $t(15) = 0.486$ and H6, $t(15) = 0.456$.

3.2. Reaction times.

3.2.1. Introduction.

Although there is a perfect and inverse relationship between the number of hits and misses, and the number of false alarms and correct rejections, all the reaction times (RTs) obtained can be regarded as stochastically independent.

For each subject a median RT was calculated for each type of response (ie. hit, miss, correct rejection and false alarm) under each processing/presentation time and environmental context reinstatement/change condition.

The major purpose of the RT analysis is to determine the conditions under which an environmental context effect is manifested. Given the problem of a comparatively small number of false alarm and miss responses, and the reduction in the power of analysis this causes, it would seem appropriate to examine the RTs from each response type separately, in terms of processing/presentation time and environmental context reinstatement/change factors. This allows all data collected to be analysed, whereas an analysis in terms of a three factor (processing/presentation time, by response type, by environmental context reinstatement/change) mixed ANOVA, requires that subjects not providing RTs in all response type categories are ignored.

However, both forms of analysis should provide interesting information regarding the relationship and locus of effect of the different factors. Indeed as will be seen, the separate two factor (response type by environmental context reinstatement/change) mixed ANOVAs may be considered as subsequent interaction comparison analyses of the three factor ANOVA (Keppel, 1982; Kirk, 1982).

3.2.2. Complete RT data set ANOVA.

3.2.2.1. preliminary analysis.

A three factor (4 x 4 x 2) mixed design was applied. The factors were defined by processing/presentation time (P/PT), response type (RST) and environmental context reinstatement/change (CTX). An F-max test of the between subjects factor suggested that the normality and homogeniety of variance assumptions were tenable, $F\text{-max}(4,15) = 3.23$. However, to maintain consistency of data presentation and to facilitate comparisons between analyses, a logarithmic transformation of the form $Y' = \log(Y + 1)$ and identical to those implemented prior to all subsequent ANOVAs was carried out. The transformed data also satisfied the normality and homogeniety of variance assumptions, $F\text{-max}(4,15) = 2.82$.

3.2.2.2. overall F-tests.

The three factor mixed ANOVA was carried out on the transformed RT data. The results of this analysis are summarised in table 11.3., while tables 11.4. and 11.5. present the non-transformed and transformed means and standard deviations of the significant interaction (response type x process/presentation time). Figure 11.4. presents a plot of the transformed mean values.

SOURCE	df	mean square	F	p	Huynh-Feldt p
P/PT	3	0.03657	0.50	0.685	
error	41	0.07322			
RST	3	0.20836	19.10	0.000	0.000
RST x P/PT	9	0.03606	3.31	0.000	0.000
error	123	0.01091			
CTX	1	0.00015	0.02	0.890	
CTX x P/PT	3	0.00277	0.35	0.786	
error	41	0.00781			
RST x CTX	3	0.00854	1.19	0.316	0.315
RST x CTX x P/PT	9	0.00365	0.51	0.866	0.852
error	123	0.00718			

Table 11.3. Summary of the three factor mixed ANOVA of transformed RT.

		CONDITION			
		L1.5	L3	H3	H6
hits	X	1.731	1.638	1.431	1.339
	S.D.	0.851	0.492	0.793	0.250
	n	348	418	539	721
misses	X	1.896	2.150	2.235	2.624
	S.D.	1.144	1.147	1.333	2.007
	n	292	222	101	143
c.r.s	X	1.630	1.893	1.743	1.983
	S.D.	0.617	1.024	0.583	0.888
	n	479	494	579	782
f.a.s	X	2.178	1.795	3.012	2.972
	S.D.	1.382	0.690	2.324	1.722
	n	161	146	61	82

Table 11.4. Mean and standard deviation of non-transformed response type RT and the number of RTs contributing to each mean per condition.

		CONDITION			
		L1.5	L3	H3	H6
hits	X	0.419	0.414	0.380	0.367
	S.D.	0.115	0.074	0.066	0.035
misses	X	0.437	0.476	0.491	0.518
	S.D.	0.135	0.104	0.122	0.151
c.r.s	X	0.409	0.441	0.430	0.458
	S.D.	0.091	0.115	0.087	0.123
f.a.s	X	0.469	0.435	0.553	0.569
	S.D.	0.161	0.091	0.159	0.140
	n	13	14	8	10

Table 11.5. Mean and standard deviation of transformed response type RT and the number of subjects per P/PT condition.

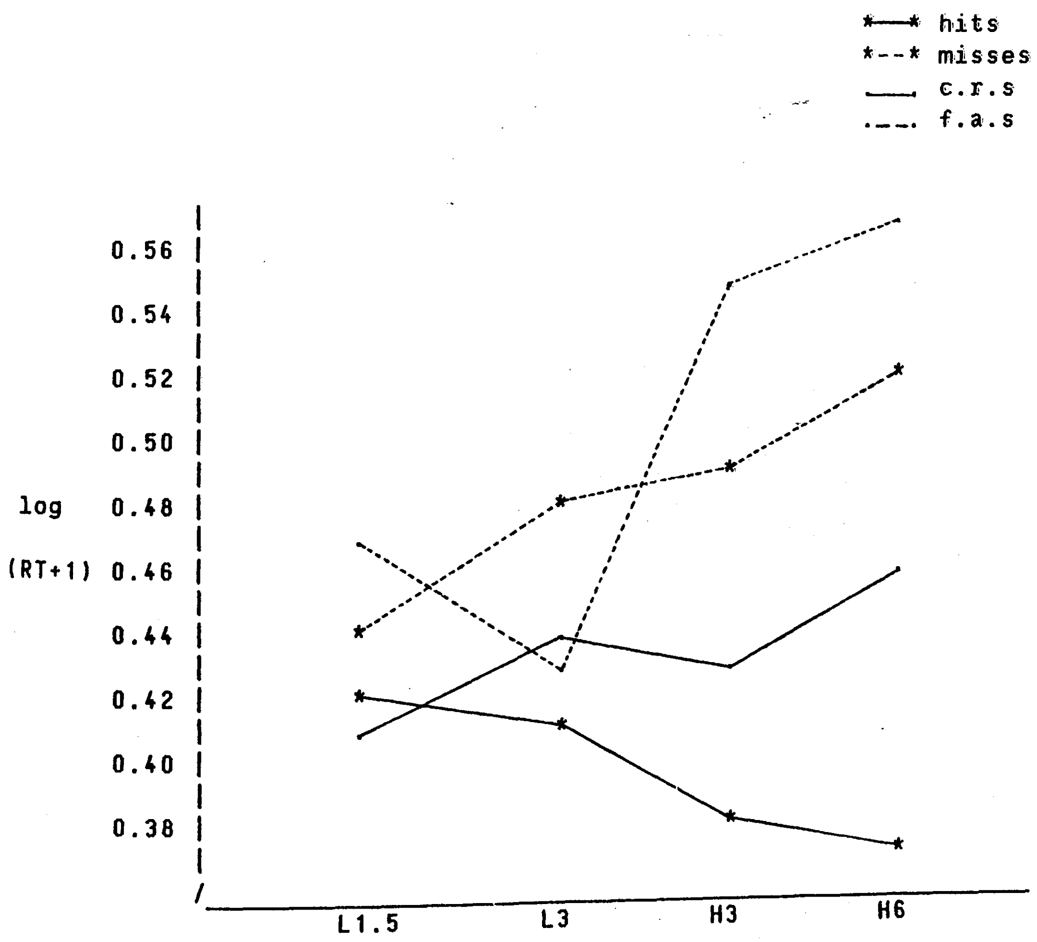

Figure 11.4. Mean transformed hits RT per condition.

3.2.2.3. specific comparisons.

As the separate response type analyses of the effect of the processing/presentation time factor to be reported can be considered as simple effect analyses, the only specific comparisons required here are those involving the main effect of response type.

The mean transformed RT of each response type: calculated using all the data available in each category, rather than just that from subjects providing RT data in all response type categories; as in the three factor ANOVA, is presented in table 11.6. below.

response type	X transformed RT	sample size
hit	0.388	64
miss	0.482	59
c.r.	0.417	64
f.a.	0.503	48

Table 11.6. Mean transformed RT by response type.

Orthogonal analyses were carried out to determine which response type RTs were significantly different. As the homogeniety of variance of differences assumption was violated, separate error terms for each comparison were required. Given such a situation, it was decided to carry out separate F-tests employing all the data available for each comparison, rather than the sub-set of data from subjects providing a RT in each response category.

The means, error terms and F-values for each comparison are presented in table 11.7. below.

comparison	respective means	df	MSe	F
hit vs cr.	0.388 vs 0.417	63	0.003	8.54
cr. vs miss	0.419 vs 0.482	58	0.006	18.94
miss vs fa.	0.477 vs 0.495	44	0.007	1.04

Table 11.7. Means, error terms and F-values of the three orthogonal comparisons.

3.2.3. Separate two factor ANOVAs of RT by response type.

3.2.3.1. Hits

3.2.3.2. preliminary analysis.

An F-max test of the between subjects factor normality and homogeniety of variance assumptions suggested that one or both were untenable, F-max(4,15) = 10.78. Logarithmic and square root transformations were examined, with the log transform providing the smallest F-max value, F-max(4,15) = 5.72.

CONTEXT		CONDITION L1.5	L3	H3	H6
same	X	1.832	1.520	1.215	1.391
	S.D.	0.958	0.391	0.265	0.291
diff	X	1.673	1.698	1.326	1.298
	S.D.	0.640	0.534	0.467	0.240

(a)

CONTEXT		L1.5	L3	H3	H6
same	X	0.431	0.397	0.342	0.376
	S.D.	0.134	0.067	0.051	0.051
diff	X	0.416	0.423	0.360	0.360
	S.D.	0.099	0.083	0.076	0.043

(b)

Table 11.8. Mean and standard deviation of non-transformed (a) and transformed (b) hit RT per condition.

3.2.3.3. overall F-tests.

The two factor mixed ANOVA (factors defined as in section 3.1.1.2.) carried out on the log transformed data revealed a main effect of encoding condition, $F(3,60) = 3.34$, $MSe = 0.011$, but not retrieval context, $F(1,60) = 0.17$, $MSe = 0.002$. However, the interaction between these two factors did approach significance, $F(3,60) = 2.27$, $p = 0.089$ (see table 11.8. and figure 11.5.).

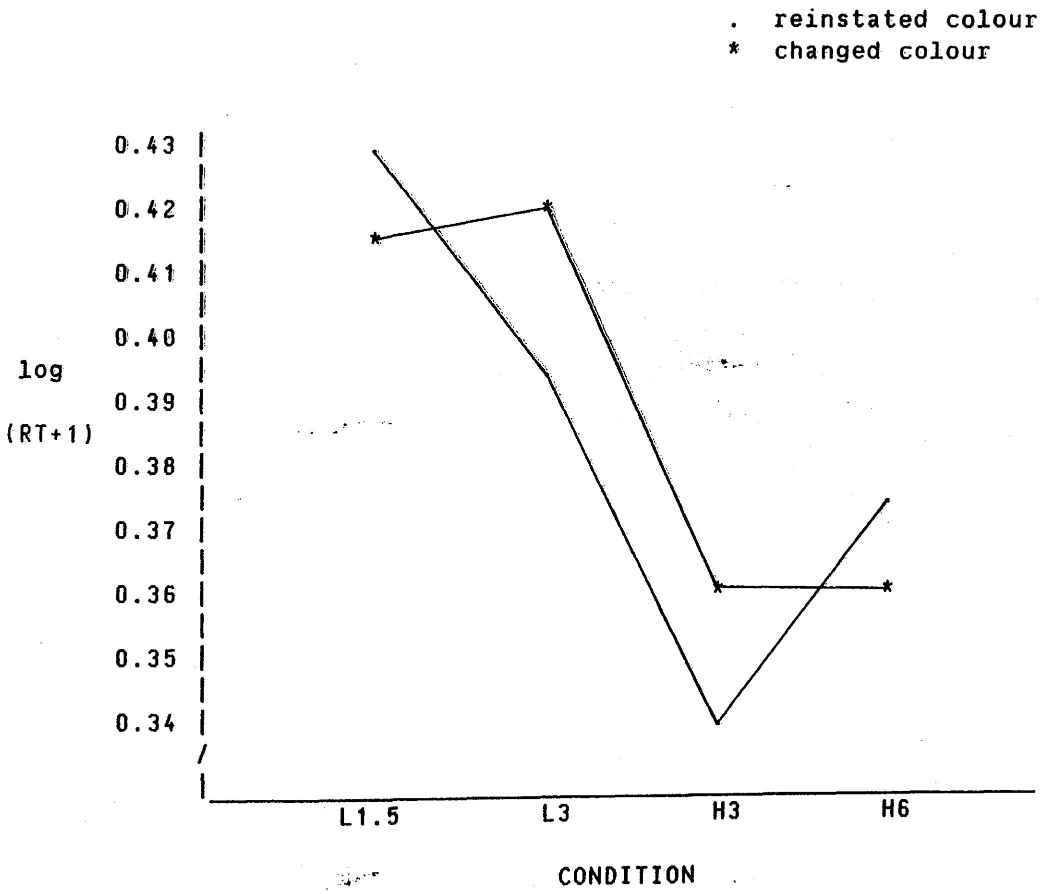

Figure 11.5. Mean transformed hits RT per condition.

3.2.3.4. specific comparisons.

No significant differences were observed between the H6 and H3 encoding conditions, t(60) = -0.615, nor the L3 and L1.5 encoding conditions, t(60) = 0.516. However, a significant difference was observed between the H3 and L3 encoding conditions, t(60) = 2.213.

Although the interaction between the encoding condition and retrieval context factors was not significant, as the superiority of reinstated colour context conditions is predicted, it is legitimate to investigate the pairwise comparisons of interest with one-tailed t-tests (see section 3.1.2.3.)

Of the four pairwise comparisons, only the L3 and H3 conditions exhibited superior speed of response with reinstated colour context and so merited further examination. Of these two comparisons only the L3 condition revealed a significant effect t(60) = 1.815; H3, t(60) = 1.172.

3.2.4. Misses.

3.2.4.1. preliminary analysis.

An F-max test of the between subjects factor normality and homogeniety of variance assumptions suggested that they were untenable, F-max(4,15) = 4.01. An examination of logarithmic and square root transformations revealed the former to provide the lowest F-max value, F-max(4,15) = 2.43.

CONTEXT			L1.5	L3	H3	H6
	same	X	1.854	2.270	2.783	2.679
		S.D.	0.870	1.461	2.192	2.261
	diff	X	2.055	1.874	2.242	2.359
		S.D.	1.201	0.539	1.241	1.260

(a)

CONTEXT			L1.5	L3	H3	H6
	same	X	0.438	0.481	0.533	0.517
		S.D.	0.123	0.170	0.189	0.189
	diff	X	0.460	0.452	0.487	0.504
		S.D.	0.145	0.076	0.143	0.135

(b)

Table 11.9. Mean and standard deviation of non-transformed (a) and transformed (b) miss RT per condition.

3.2.4.2. overall F-tests.

The two factor mixed ANOVA; carried out on the log transformed data, revealed no significant effects. Encoding condition, F(3,60) = 0.85, MSe = 0.033, retrieval context, F(1,60) = 0.72, MSe = 0.011 and condition x context, F(3,60) = 0.55 (see table 11.9.).

3.2.5. Correct rejections.

3.2.5.1. preliminary analysis.

An F-max test of the between subjects factor normality and homogeniety of variance assumptions suggested that they were tenable, F-max(4,15) = 3.06. However, for ease of comparison with the other response type RT analyses, a log transform was carried out, F-max(4,15) = 2.01.

			CONDITION			
			L1.5	L3	H3	H6
CONTEXT	same	X	1.659	1.724	1.495	1.796
		S.D.	0.574	0.670	0.478	0.791
	diff	X	1.630	1.875	1.550	1.786
		S.D.	0.573	1.233	0.535	0.706

(a)

			L1.5	L3	H3	H6
CONTEXT	same	X	0.416	0.423	0.390	0.433
		S.D.	0.092	0.104	0.080	0.109
	diff	X	0.411	0.431	0.398	0.433
		S.D.	0.090	0.147	0.084	0.100

(b)

Table 11.10. Mean and standard deviation of non-transformed (a) and transformed (b) correct rejection RT per condition.

3.2.5.2. overall F-tests.

The two factor mixed ANOVA; carried out on the transformed data, revealed no significant effects. Encoding condition, $F(3,60) = 0.67$, $MSe = 0.019$, retrieval context, $F(1,60) = 0.13$, $MSe = 0.002$ and condition x context, $F(3,60) = 0.12$ (see table 11.10.).

3.2.6. False alarms.

3.2.6.1. preliminary analysis.

An F-max test of the between subjects factor normality and homogeniety of variance assumptions suggested that one or both were untenable, F-max(4,15) = 11.78. An examination of logarithmic and square root transformations revealed the former to provide the lowest F-max value, F-max(4,15) = 3.39.

3.2.6.2. overall F-tests.

The two factor mixed ANOVA carried out on the log transformed data revealed a main effect of encoding condition, F(3,44) = 3.04, MSe = 0.039, but not of retrieval context, F(1,44) = 0.05, MSe = 0.016, nor did the interaction between the two factors approach significance, F(3,44) = 0.35 (see table 11.11.).

CONTEXT			CONDITION L1.5	L3	H3	H6
	same	X	2.259	1.711	3.836	2.773
		S.D.	1.705	0.444	3.807	1.855
	diff	X	2.096	1.879	2.875	3.092
		S.D.	0.955	0.869	1.688	1.414

(a)

CONTEXT			L1.5	L3	H3	H6
	same	X	0.467	0.427	0.594	0.546
		S.D.	0.200	0.073	0.274	0.155
	diff	X	0.471	0.442	0.556	0.587
		S.D.	0.136	0.127	0.172	0.156

(b)

Table 11.11. Mean and standard deviation of non-transformed (a) and transformed false alarm RT per condition.

3.2.6.3. specific comparisons.

No significant difference was observed in overall performance between encoding conditions H6 and H3, t(60) = -0.174. However, a significant difference was observed between encoding conditions H3 and L3, t(60) =

2.855. No significant difference was observed between encoding conditions L3 and L1.5, $t(15) = -0.495$.[4]

---

4 Superior performance with false alarms; in common with misses, should be represented by larger erroneous response RTs. The assumption is that greater time reflects greater uncertainty, manifested in more retrieval cycles to obtain more information regarding the status of the word in question.

4. Discussion.

4.1. General conclusions.

As in the previous experiments, the pattern of effects of encoding condition would seem to indicate the success of the different instructions in eliciting different modes of processing. The major differences between encoding conditions occur between the H3 and L3 conditions. This suggests that it is processing mode rather than presentation time, that exerts the most significant effect.

Beyond this, the replication of the environmental context reinstatement effects (as observed in chapters seven and eight), but detected here with a recognition paradigm (in terms of the number of hits and false alarms, and hits RT) and using background colour as the nominal environmental context supports not only the view of context suggested by the model, but also its emphasis on a communality of processes in recall and recognition.

4.2. Nature of reaction time data.

Of the analyses of reaction time data presented, most importance should be placed on the hit and correct rejection response times analyses. The reason for this; as mentioned in section 3.2.1., is that these analyses employ most data. As most subjects performed the recognition task well, comparatively few misses and false alarm responses were made. For the analyses of these reaction times the consequence was not only that some subjects did not provide any responses in one category (ie. false alarms), but that many of those median values obtained in the miss and false alarm categories were derived from a much smaller number of reaction time scores than was the case for the other two response categories. A small sample size allows a greater influence of any rogue score. In turn this can produce greater variance of scores and so a less powerful and even misleading analysis. Although the mean retains more information about each score than the median, this aspect of the mean provides greater accuracy only if there is a superior ratio of scores genuinely representing the process(es) under investigation to rogue scores. The critical ratio will depend on the magnitude of deviation of the rogue scores. The use of mean values when there is a small sample size, or extreme rogue scores can provide a summary statistic that is quite unrepresentative of the values associated with the performance of the task under investigation. Given reaction time data, where slight inattentiveness to the task can produce very large rogue values and the small sample of scores produced by some subjects in the correct rejection and false alarm categories, it was decided to employ medians and not means.

4.3. RT ANOVAs discrepancy.

As an environmental context reinstatement effect was identified in the L3 condition of the hits RT analysis, but not in any of the other response type analyses, an interaction between all three factors in the three factor ANOVA should have been expected. A lack of statistical power would seem to be the reason for the non-significance of this effect. There are several reasons for a reduction in power between the hits RT analysis and the three factor analysis. The most obvious of these is the difference in magnitude of the relevant interaction error terms. In the hit RT analysis this is 0.002, while in the three factor analysis it is 0.007. An examination of interaction error terms across the other response type analyses reveals that only the analysis of correct rejections provides an interaction term of comparable magnitude. The relevant miss and false alarm error terms are much higher. The size of the interaction error terms is inversely proportional to the number of data points contributing to its calculation. This finding supports the view expressed in section 4.2..

In addition to the effect of sample size on variation, the manner in which subjects with incomplete data records were eliminated from the three factor ANOVA would seem to have consequence for the results observed. This elimination is carried out on the same basis as was considered improper to enable d' analyses. As with the d' analyses discussed, eliminating subjects that did not provide false alarm or miss responses eliminates those subjects who performed the task best. The greatest proportion of these subjects should be found in the conditions that facilitate performance and this will include reinstated, rather than changed colour background environmental context conditions. As high performing subjects would be expected to complete hit responses quicker than other subjects, it is likely that the elimination of subjects on such grounds would lessen the differences between conditions in terms of RT as well as the number of response types.

Finally, it is noteworthy that even the violation of the homogeniety of variance of differences assumption is operating to increase the alpha level (make significance less likely), rather than as normal, decrease the alpha level. Fortunately, the slight effect this has on the alpha level is countered by the Huynh-Feldt adjustment.

4.4. Retrieval strategies across conditions suggested by RTs.

An examination of figure 11.4. suggests two major components of the three factor ANOVA, processing/presentation time by response type interaction. First, there appears to be a general decrease in hits RT with an increase in processing/presentation time, in contrast to a lack of such effect with correct rejection responses and a general increase in RT over processing/presentation time with miss and false alarm responses. The second component is the apparent reduction in false alarm RT in the L3 condition. However, an examination of the separate response type RT two factor ANOVAs reveals support for only two of these conclusions. Both miss and correct rejections are uneffected by changes in processing/presentation time, while false alarm RT over processing/presentation time differs only between low and high level processing (ie. L3 cf. H3) and is not effected by presentation time (ie. L1.5 cf. L3, H3 cf. H6).

The particular pattern of an increase in response type RT from hits to correct rejections to misses and false alarms suggests that at the end of each retrieval cycle, some assessment is made as to whether further retrieval cycles will be implemented. Presumably, hit item retrieval cycles are terminated when a valid match is considered to be made with information in memory. Such correct matches can be determined quicker as processing/presentation time increases. This gives support to the prediction made in chapter seven (section 1.4.) that there would be an exponential rate of approach to an appropriate target description.

Another interesting feature of the hit data obtained in this experiment is the way in which they appear to contradict the model of recognition presented by Atkinson and Juola (1974). One of the fundamental assumptions of this model is that as if the "familiarity" of an item cannot be employed to make a decision as to the stimulus item's previous occurrence, then a search of event knowledge (E/K) memory is initiated. The longer this memory search continues, the greater the probability that a correct response will be made. With the present hit data however, the presumed relationship between the probability of a hit and RT is not observed. In fact the opposite relationship is observed. As the probability of a hit increases (across conditions L1.5, L3, H3 and H6), the RT decreases. Proponents of the Atkinson and Juola model may attempt to account for such observation by attributing the decrease in RT to decisions made on the basis of high familiarity achieved by the elaborative processing engaged in the H3 and H6 conditions. However, this argument ignores two awkward facts. First, it seems odd that the RTs based on such "familiarity" judgements are not influenced by presentation time. Second, it is even more odd that judgements based on "familiarity" should be influenced by the form of processing engaged by subjects, when this seems to be distinguished by the manner of use of other knowledge. The form of memory representations formed at encoding in the conditions exhibiting high(er) hit probabilities should facilitate what Atkinson and Juola have termed "extended memory search" and not decisions based on familiarity information.

As correct rejection responses take longer than hit responses, it suggests that more information and so more retrieval cycles are required before it is decided that the information sought is not present. This finding is consistent with most studies of memory search carried out since Sternberg (1969).[5]

---

5 See McNicol and Stewart (1980) for a review of RT in memory research.

The RTs associated with misses suggests that in an attempt to provide support to an earlier assessment that the item was not presented, more retrieval cycles are required before a (wrong) decision as to the items prior presentation is made. This decision could be based on verification criteria, or possibly on a combination of verification criteria satisfied and the number of retrieval cycles required to achieve this. The latter measure would approximate to retrieval time. Presumably, some limit to the number of retrieval cycles engaged would be set by the subject in response to the experimenter's instructions of responding with speed, but accuracy.

The false alarm RTs also seem to be accountable in a similar manner. However, the fact that the type of decision and response made is of false recognition suggests that prior to the final retrieval cycles, some assessment has indicated that the particular item is a member of the previously presented word set, otherwise retrieval would be terminated with the same RT as correct rejections. The extra retrieval cycles, probably in comparison to hits, but certainly in comparison to correct rejections, can be regarded as attempts to confirm or contradict this prior assessment, as in the overshoot phenomenon (see chapter six, section 1.4.).

The increase in RT observed with false alarm responses as the level of semantic processing increases is more difficult to account for than the converse effect observed with hit response RT. However, one possibility is that after semantic elaboration, a vague retrieval specification may (via the processes described in chapter seven, sections 1.4. & 1.5.) produce several candidate memory records, whereas a vague retrieval specification employing mainly integrated information may have difficulty in producing any memory records at all. Consequently, more time will be required to assess the various candidates produced in the former situation than is required to terminate the latter retrieval process.

The observation of an environmental context effect with hit RT, but not false alarm RT, may seem slightly paradoxical given the presence of an environmental context effect in terms of the number of hits and false alarms. The advantage in terms of the number of hits and their RTs in the L3 condition indicates not only that environmental (colour) context information has been encoded and employed in retrieval specifications, but also that benefit has been derived from such inclusion. Likewise, the effect of environmental context in terms of the number of false alarms indicates that environmental context information is being used, but this time to the subjects' performance disadvantage. Given this finding and interpretation, an environmental context effect in terms of RT might be expected with false alarms as well as hits. However, there are two (non-mutually exclusive) reasons why such an effect might not be observed; one primarily statistical and the other psychological. The first is the limited sample size. As has been mentioned, this is manifest not just in terms of the number of subjects contributing a false alarm RT, but also in terms of the number of RTs that contribute to the subjects' median scores, which are input to the ANOVAs. The result is greater variance and fewer degrees of freedom in false alarm RT scores and so, less powerful analyses. The second reason is that although false alarm response items eventually may be regarded as hit items by the subject, the greater RTs associated with false alarms; suggesting an attempt to obtain corroborative or contradictory evidence, may mask the environmental context influence. There may be greater variation in RT on the basis of the "checking" retrieval cycles than as a consequence of the environmental context factor, although in the retrieval specification verification criteria hierarchy, the environmental context information correspondence is placed in a position of sufficient importance to affect the final decision and response.

The lack of effect of environmental context on correct rejections; despite comparable power between this and the hits analysis, strongly suggests that environmental context information is not an essential component in the retrieval specifications constructed. Likewise; drawing on an assumed similarity of processing between correct rejections and misses, neither does it play a large part in the retrieval specifications constructed for items wrongly judged as not being members of the previously presented word set. Some sort of information other than that distinguished by the experimental manipulations must be used in the retrieval specifications resulting in correct rejections and miss responses. This implies that there is more than one mode of retrieval and; as there is no reason to believe that in condition L3 environmental (colour) context information was not encoded as part of the miss items' descriptions, that at least the inclusion of environmental context information in retrieval specifications is under strategic control.

If the view that false alarm item "checking" masks the environmental context effect in terms of RT is accepted, then the pattern of results obtained would suggest that retrieval specifications incorporating environmental context information are constructed early in the retrieval process and if they do not disambiguate satisfactorily, other information is incorporated in the subsequent retrieval specifications instead. Such a procedure would comply with a general strategy of employing higher level retrieval specifications, only if lower level retrieval specifications failed to disambiguate and achieve a successful retrieval.

However, consideration of the way in which the order of application of the different retrieval specifications that can be constructed is probably misplaced here, especially with the doubt concerning the validity of the false alarm and miss RT means as robust measures of the time taken to engage and complete the processes underlying such responses. Although much of the account based on the RT data is speculative; addressing issues beyond that which the experiment was originally designed to examine, it does serve to illustrate the potential of the approach employed.

4.5. The nature of the environmental context effect obtained.

The magnitude of the environmental context effects observed; especially in terms of the number of hits (W^2 = 0.659), may seem extraordinary given the traditional failure to obtain any such effects with recognition paradigms. However, several factors may contribute to the magnitude of the effect. One important aspect of the experimental method is that the nominal stimulus words were presented on a background colour. Although this is stating the obvious, there would seem to be an important difference between the presentation of words in a room; apparently by convenience on a white wall, and the suspiciously significant fact that all presentation words were presented on the same block of colour. In the latter case, which describes the recognition-word colour experiment, the manner of presentation would seem to do more than increase the perceptual saliency of the nominal environment. The use of the colour background; although only verbally referred to in the same way as room environment, seems to connote that it is an important aspect of the experiment. If this is so, then it should be expected for subjects to employ an encoding strategy that incorporates information concerning the colour. This then could provide a basis for an advantage of reinstated colour context at retrieval.

Although the greater significance attributed to the nominal environmental context at encoding may account for some of the difference in magnitude of environmental effect between chapters seven and eleven, it cannot provide a complete account when the comparison is drawn with other environmental context studies. In such a comparison the type of processing carried out at encoding would seem to play a critical role. An interesting study would be to examine the relative contribution to the environmental context reinstatement effect of level of processing and cognitive (rather than perceptual) saliency.

It is also interesting to note that a prediction of an effect of colour environmental context effect in recognition; given low level processing is one of the consequences of Mandler's (1980) account of recognition performance. Mandler suggests that integrative organisation/processing is heavily based on perceptual information, ie. low level schemata. As there should be a strong bias towards intra-item processing in the low level conditions, the encoding descriptions formed should include a considerable amount of colour environmental context information. As recognition employs this sort of information primarily, it should be expected that a change in colour environmental context would be detrimental to such performance. To what extent room type environmental context would exert a similar effect should depend on the degree of encoding; as may be affected by the cognitive saliency of the nominal environment,[6] and the level of discrimination such information provides

---

6 This would suggest a less automatic method of integrative processing than is implied by Mandler (1980).

at retrieval. With the present RED-BLUE distinction, this should be relatively high.

4.6. Environmental context information unavailability, or non-use at retrieval.

The observation of an environmental context reinstatement effect in the L3, but not the L1.5 condition would suggest that some minimum time for the encoding of environmental context information is required, if it is to be of use in subsequent retrieval specifications. Unfortunately however, the lack of a significant effect of context reinstatement in either of the two high level processing conditions prevents an unequivocal determination of whether environmental context is encoded, but is redundant at retrieval, or simply is not encoded. Despite this, the pattern of results obtained strongly suggests that the former account is the more accurate.

In the H3 condition, in both the number of hits and hits RT, an advantage (although not significant) for reinstated colour context was observed. It would seem reasonable to assume that if the size of this effect was inhibited by limiting the ability of processing environmental context information by inducing story and image construction processes, so elaboration could occur within the allocated time, then increasing the time of presentation should facilitate the encoding of environmental context information. However, as the H6 condition provides no evidence of advantage to colour context reinstatements, nor any evidence of a general increment in performance, it implies that the H3 condition does not exhaust the processing ability of subjects. Consequently, the cause of the lack of environmental context reinstatement effect is very unlikely to be caused by any processing limitation. Given such a conclusion, it would seem that the best account of the lack of environmental context effect observed in high level processing conditions remains the redundancy of environmental information in memory records and for retrieval specifications, due to the variety of retrieval routes available when the nominal stimuli items have been elaborately encoded.

4.7. Reorientation.

Having initiated the consideration of the influence of environmental context on memory with anecdotes of attempts to remember in naturalistic situations, and having manifested and examined the reinstatement phenomenon with traditional experimental variables, it seems appropriate that the last experiment to be reported should return to consider a more commonly experienced type of remembering. In the following chapter, an attempt is made to recreate, in a more formal setting, the need to recall goal-oriented, intentional action.

CHAPTER TWELVE

EXPERIMENT SEVEN

AN INVESTIGATION OF ENVIRONMENTAL CONTEXT AS AN AID TO THE REMEMBRANCE OF INTENDED ACTION

1. Introduction.

1.1. Different aspects of memory performance.

In the previous chapters the experiments that were reported dealt with the influence of environmental context on the ability to remember discrete verbal stimuli. As with most memory research, the reason for the use of verbal stimuli has more to do with their ease of presentation and general utility, rather than for any theoretical reason or commitment. However, although such verbally based tasks provide an easy way to examine the influence of environmental context on many aspects of memory performance, their ability to reflect in an ecologically valid manner, one aspect of memory performance that seems to be distinctly tied to environmental context; namely intentional action, is much more questionable.

1.2. Environmental context and intended action.

An often cited example of environmental context influence that has also been cited to illustrate the effect of environmental context on intended action (Norman, 1981) is that anecdote reported by William James (1890),

> very absent minded persons in going to their bedroom to dress for dinner have been known to take off one garment after another and finally get into bed, merely because that was the habitual issue of the first few movements when performed at a later hour (p.115).

The influence of environmental context on the remembrance of intended action can be further illustrated by considering the difficulty there is in describing how to carry out such activities as driving a car (when not in a car), or locating a book in a library through the use of the cataloguing system (again, when not in the library).

It may be argued that the difficulty in providing such descriptions is due to the relevant information being contained only within some procedure which controls these acts. Although this is certainly part of the difficulty, also it seems that access to the relevant procedures and access at appropriate points to the relevant sub-routines is achieved much more easily in the relevant environmental context.

Of course, in these sorts of situations there is much more than an arbitrary relationship between the environment and performance. In such cases, elements of the environment have a functional significance with respect to the performance of the task. This aspect of real environments may be considered a nuisance in determining the "true" influence of environmental context on memory performance. However, the fact that such performance is inextricably linked with elements in the environment not only indicates their importance in psychological processing, but also the consequent need for some theoretical account that can accommodate the relationships between such seemingly different entities (ie. environments, objects, plans and actions) and the ability to access one form of representation from another.

1.3. Models of action comprehension and intended action.

Models of intended action have been presented by Brewer (Brewer & Dupree, 1983; Lichtenstein & Brewer, 1980) and Norman (1981). In both of these accounts, the notion of schemata is employed as the basis for cognition and action. Brewer's work focuses on subjects' comprehension and remembrance of actions, while Norman addresses the issue of initiating and accounting for the frequent errors of action.

Brewer suggests that information mapped into plan-schemata (in the same manner as described in chapter six, section 1.2.) provides the means by which an abstract underlying structure can be identified from a temporal sequence of actions or events. In other words, action intent, or goals are derived from the structural relationships provided by plan schemata. In accord with memory performance observed with other types of information, Brewer and Dupree have noted that over time, memory for (deduced) goal-directed action exhibits a loss of detail and an emergence of gist.

Norman proposes an activation-trigger-schema system (ATS) model that controls action via sensori-motor schemata. For the operation of each schema, an activation value and a triggering condition must be fulfilled. Norman views the schemata as being organised in a heterarchical control structure, with each schema being relevant (ie. containing information and directing the control of motor activity) for only a small number of actions. Therefore, any sequence of actions (as most "actions" are) require several, or many schemata.

The highest level schema is termed the parent schema and represents the abstraction of intent. All those schemata activated by the parent schema are termed child schemata. Actions are initiated by the triggering of an activated parent schema. Child schemata are activated subsequently, with their triggering conditions controlling the particular order of schemata operation. Determination of the appropriateness of schema operation is carried out by means of descriptions (Norman & Bobrow, 1979).

Norman identifies three main sources of action errors: intention formation, activation and triggering. Of the catalogue of action slips provided by Norman, a sub-category of activation errors; capture errors, are associated most frequently with environmental context. However, as descriptions play a crucial role in intention formation and the activation and triggering of the appropriate schemata, environmental context information could exert an influence at all stages of action execution.

An interesting correspondence exists between the anecdotal accounts of the sort of task influenced by environmental context, the models of intentional, goal-directed action and the suggestion from the state-dependent literature that particularly tasks involving a serial-order component are affected by state changes. Actions with intent and therefore with goal-direction are well ordered in temporal sequence. Usually, prior actions set-up the conditions that enable subsequent actions to occur. Consequently, the theoretically oriented schema based models of intentional action may provide a means of accounting for the effect of physiological context (state) on such tasks.

1.4. The operationalisation of intended action.

The similarity of the models presented to account for the environmental context effects reported in the previous chapters and intended action suggests that there is substance to the anecdotes of environmental context influencing such action. The next step in the investigation of the potential relationship between environmental context and intended action would seem to be to attempt to formulate an intentional, goal-directed task that can be formally examined under laboratory conditions. Of course, the problem faced by any such attempt is achieving a valid correspondence between the "real-world" and laboratory situations on those factors which have effect on the behaviour under study, when these factors have yet to be identified.

For compatibility with previous experiments and in order that the goal-oriented task could be carried out with equal ease in two environments, it was important that no advantage to task performance was provided by any one environment. Tasks that wholly replicated real-world activities; apart from requiring a large amount of apparatus, tended to suffer from the problem mentioned in section 1.2., of being inter-twined with environmental features. To achieve equality in both room environments with such tasks would require the two environments to become more alike, as those environmental elements involved in in the task were replicated in each environment. Consequently, a simple task was devised that was independent of the environment in which it was performed. The task involved the ordering of a set of thirteen unique objects. The subjects' goal was to place these objects in a particular order in a straight line across the table that they would be sitting at. Such a task not only maintains the goal oriented nature of the tasks described anecdotally and allows easy measurement of how well performance achieves this goal, but it also enables separation to be made between the ability of subjects to remember a particular pattern of objects and the order of object placement that leads to such a pattern.

1.5. Experimental hypotheses.

In common with other experiments reported in this thesis, this particular experiment is quite exploratory in nature. As a result, few specific a priori hypotheses can be presented. However, the general view and hypothesis is that goal-oriented actions practiced in a particular environment will be better remembered (ie. re-enacted) in that environment, as compared to any other.

2. Method.

2.1. Subjects:

24 Glasgow University students; males and females of approximately equal numbers, participated as subjects. All subjects were naive regarding the purpose of the experiment and were paid one pound for taking part.

2.2. Environmental contexts:

Rooms A and B were employed in the experiment. Descriptions of these rooms can be found in appendix A.

2.3. Stimuli and apparatus:

13 different objects were used in the experiment. These items were: a hook, a button, a screw, a (wooden) crescent, a hinge, a (wooden) cube, a key, a paper-clip, a battery, a coin, a (small piece of) string, a (wooden) triangle and a di.

A set of instructions regarding the placement of the objects was recorded and presented to subjects via a mono tape recorder; operated by a remote switch and hidden speakers (the instructions are detailed in section 2.5). Subjects verbal protocols were recorded on the same JVC stereo cassette deck via the microphones employed in experiments two to five.

2.4. Questionnaire:

A simple questionnaire; based on that presented to subjects in experiment nine, was administered to subjects after they had completed their experimental session. A specimen copy of the questionnaire can be found in appendix D.

2.5. Design:

Initially, a two factor (2 x 2) mixed design was applied. The first factor was defined by the subjects' second session environment, ie. same (S), or different (D). The second factor was defined by the session number, ie. first or second.

2.6. Procedure:

The experiment was carried out over two, approximately ten minute sessions, separated by twenty-four hours. On first arrival, subjects were led into room A and asked to sit at the table at the end of the room. The experimenter then informed the subjects that he would not be remaining in the room and that instead they would be directed in the task they had to perform by tape recorded instructions. The experimenter explained to the subjects that all they had to do was listen to the recording and follow the instructions that would be given at appropriate intervals. After ascertaining that the subjects understood what was about to happen, the experimenter left the room and switched on the tape. The script of the recording was as follows:

If you look at the table in front of you, you will see a variety of different objects. Look at the objects and make sure that you can unambiguously identify each one. There is a hinge, a screw, a hook, a piece of string, a battery, a triangle, a button, a paper-clip, a coin, a cube, a dice, a crescent and a key. Your task here is to place these objects in a horizontal line across the table. However, what you have to do is place these items along the line in a particular order. What will happen is that I will instruct you, item by item, to place each object somewhere along the line. Now just to avoid any mis-understandings, I will repeat what you have to do. You have to arrange the objects in a horizontal line across the table in front of you. To do this I will instruct you, item by item, where to place each object. OK. then, let's start.

First put the button in the middle of what will be your horizontal line of objects.

Place the hook to the right of the button.

Place the screw to the right of the hook.

Put the crescent between the hook and the screw.

Put the hinge at the right hand end of the horizontal line.

Put the cube at the left hand end of the horizontal line.

Place the key half-way between the button and the cube.

Put the paper-clip between the button and the key.

Put the battery to the right of the screw.

Put the coin between the paper-clip and key.

Place the string to the right of the cube.

Put the triangle between the battery and hinge.

Place the dice between the string and key.

Well that's you finished the first part. Now, put all the objects into a pile as they were before.

At this point the experimenter returned and explained that the subject's task was to reform the line of objects by placing each object in the line in the same order as they had just been instructed. As they did this, subjects were asked to give a detailed commentary of what they were doing, especially with respect to the order of placing the objects. Subjects were informed that their commentaries would be recorded (via the microphone in position and employed since experiment two) and from these, their ability to reform the line of objects in the correct order would be assessed. They were asked to begin as soon as the experimenter left the room (prior to entering to instruct the subjects, the experimenter had started the tape recording). Subjects were instructed to return to the waiting area outside; leaving the reconstructed line behind them, after they had completed the task. Subjects were then reminded that they had to return the next day at the same time to carry out an identical task, but with another set of objects. On returning for the second session, subjects were randomly allocated to either room A (same condition), or room B (change condition). In both rooms subjects were invited to sit at the desks in front of the same pile of objects as they had seen the day before. The experimenter then informed the subjects that rather than as had been described the day before, their task on this day was to reform the original line of objects, placing the objects in the same order as they had originally been instructed. It was also explained that the reason for leading them to believe that they would be asked to carry out a different task was to reduce the likelihood of them thinking about, or practicing the task. Subjects were then asked to reform the line in the same order, provide commentaries of their actions as they did so and return to the waiting area when they were finished, leaving the intact line on the table, just as they had done on the previous day. When the subjects returned to the waiting area after completing the experiment, they were asked to return to the room they had just been in to complete the questionnaire (presented on a clipboard). After doing this the subjects were debriefed regarding the purpose of the experiment and were paid one pound for their participation.

2.7. Scoring:

The order of object placement was transcribed from each subject's two recorded commentaries. The final position of each object was recorded also. Employing this information it was possible to calculate Spearman rank order correlation coefficients, which estimated the degree of correspondence between the subjects' performance and any particular standard.

3. Results.

3.1. Introduction.

The values for the analyses reported here were obtained by calculating (for each subject) Spearman rank order correlation coefficients between two placement orders, or between two completed line orders. Correlation coefficients are identified by the notation rs. Three correlation coefficients were calculated for placement and line orders. The first (rs1) is the correlation between the first session performance and the instructed performance. The second (rs2) is the correlation between the second session performance and the instructed performance. The third (rs3) is the correlation between the first session and second session performances.

3.2. Placement order analyses.

3.2.1. Comparison of correlations.

3.2.1.1. preliminary analysis.

As a two factor (2 x 3) mixed ANOVA design was to be applied to the data, an F-max test of the between subjects factor homogeniety of variance and normality of distribution assumptions was carried out. This suggested that one, or both of these assumptions were untenable, $F\text{-max}(2,11) = 6.10$. As each rs value is expressed as a decimal point, squaring this value achieves a result comparable to that obtained by taking the square root of a number greater than 1.

An F-max test on the transformed data ($Y' = [Y]^2$) indicated that the homogeniety of variance and normality of distribution assumptions were tenable, F-max(2,11) = 2.86.

A test of the assumption of homogeneity of variance of differences was carried out (symmetry test, Anderson, 1958, p.259). This indicated that the probability of the heterogeneity of variance of differences observed occurring by chance was 0.063.

3.2.1.2. overall F-tests.

The two factor (subject group x correlation type) mixed ANOVA was carried out on the transformed data. The main effect of subject group approached significance, F(1.22) = 3.72, MSe = 0.11 (two-tailed test, p = 0.067), suggesting superior performance by subjects in the S group. A significant main effect of correlation type was observed, F(2,44) = 4.81, MSe = 0.04 (Huynh-Feldt p = 0.016), but no significant interaction between factors was indicated, F(2,44) = 0.83 (see table 12.1.).

		(a) S	(a) D	(b) S	(b) D
rs1	$\overline{X}$	0.787	0.722	0.635	0.565
	S.D.	0.133	0.216	0.196	0.270
rs2	$\overline{X}$	0.760	0.511	0.594	0.388
	S.D.	0.135	0.372	0.207	0.319
rs3	$\overline{X}$	0.859	0.724	0.746	0.572
	S.D.	0.099	0.228	0.166	0.283

Table 12.1. Correlations(a) and transformed correlations(b) by group.

3.2.1.3. specific comparisons.

As the symmetry test approached significance, it was decided to employ specific error terms for the pair-wise comparisons. As the pair-wise contrasts were a posteriori and non-orthogonal, initially Tukey's HSD was calculated. This suggested that no two condition means were significantly different, rs3 vs rs1, $q3,(44) = 1.077$, rs3 vs rs2, $q3,(44) = 3.156$ and rs1 vs rs2, $q3,(44) = 2.674$. Subsequently, calculation of Fisher's LSD identified the rs3 vs rs2 comparison as significant, $t(44) = 2.229$, but not rs3 vs rs1, $t(44) = 0.761$, nor rs1 vs r2, $t(44) = 1.886$.

3.2.2. The effect of environmental context on performance as measured against original instructions (ie. rs1 and rs2).

3.2.2.1. preliminary analysis.

In this analysis, as a two factor (2 x 2) mixed design was to be applied, an F-max test of the between subjects factor homogeniety of variance and normality of distribution assumptions was carried out. This suggested that one, or both of these assumptions were untenable, $F\text{-max}(2,11) = 4.03$. However, as before a square transform brought the data within acceptable limits, $F\text{-max}(2,11) = 1.80$.

3.2.2.3. overall F-tests.

A two factor (subject group x session) mixed ANOVA was carried out on the transformed correlation data. A significant main effect of session was observed, $F(1,22) = 4.78$, $MSe = 0.03$, indicating superiority of performance in the first session. However, there was no significant main effect of subject group, $F(1,22) = 2.33$, $MSe = 0.10$, nor was there a significant interaction between factors, $F(1,22) = 1.84$.

3.2.3. The effect of environmental context on inter-session performance.

3.2.3.1. preliminary analysis.

As a one factor; two level, completely randomised design was to be applied, an F-max test of the homogeniety of variance and normality of distribution assumptions was carried out. The F-max test indicated that one, or both of these assumptions were untenable, $F\text{-max}(2,11) = 5.32$. However, an F-max test, subsequent to the same square transform applied previously, indicated that the tenability of these assumptions with this data set were tenable, $F\text{-max}(2,11) = 2.89$.

3.2.3.2. F-test.

A one factor completely randomised ANOVA was carried out on the transformed data. A significant effect indicating the predicted superiority of S group subjects' performance was observed, $F(1,22) = 3.52$, $MSe = 0.03$ (one-tailed test, $p = 0.037$, see table 12.1.).

3.3. Completed line order analyses.

3.3.1. Comparison of correlations.

3.3.1.1. preliminary analysis.

As a two factor (2 x 3) mixed ANOVA design was to be applied to the data, an F-max test of the between subjects factor homogeniety of variance and normality of distribution assumptions was carried out. This suggested that both of these assumptions were tenable, F-max(2,11) = 2.78.

A test of the assumption of homogeneity of variance of differences was carried out (symmetry test, Anderson, 1958, p.259). This indicated that the probability of the heterogeneity of variance of differences observed occurring by chance was 0.004.

3.3.1.2. overall F-tests.

The two factor (subject group x correlation type) mixed ANOVA was carried out on the correlation data. No significant effects were observed. Group, $F(1,22) = 0.00$, MSe = 0.05, correlation type, $F(2,44) = 2.49$, MSe = 0.02 and group x correlations, $F(2,44) = 0.07$.

		S	D
	$\overline{X}$	0.828	0.828
rs1	S.D.	0.171	0.147
	$\overline{X}$	0.738	0.754
rs2	S.D.	0.224	0.142
	$\overline{X}$	0.843	0.826
rs3	S.D.	0.244	0.156

Table 12.2. Correlations by group.

3.3.2. The effect of environmental context on performance as measured against original instructions (ie. rs1 and rs2).

3.3.2.1. Preliminary analysis.

As in this analysis, a two factor (2 x 2) mixed design was to be applied, an F-max test of the between subjects factor homogeniety of variance and normality of distribution assumptions was carried out. This suggested that both of these assumptions were tenable, F-max(2,11) = 1.48.

3.3.2.2. overall F-tests.

A two factor (subject group x session) mixed ANOVA was carried out on the correlation data. No significant effects were observed. Group, F(1,22) = 0.02, MSe = 0.04, session, F(1,22) = 4.75, MSe = 0.02 and group x session, F(1,22) = 0.05 (see table 12.2.).

3.3.3. The effect of environmental context on inter-session performance.

3.3.3.1. preliminary analysis.

As a one factor; two level, completely randomised design, was to be applied, an F-max test of the homogeniety of variance and normality of distribution assumptions was carried out. The F-max test indicated that both of these assumptions were tenable, F-max(2,11) = 2.44.

3.3.3.2. F-test.

A one factor completely randomised ANOVA was carried out on the correlation data. No significant effect was observed, $F(1,22) = 0.04$, $MSe = 0.04$ (see table 12.2.).

3.4. **Questionnaire data.**

3.4.1. **open-ended question responses.**

Only 3 subjects (out of the twenty-three participants that returned their questionnaires) replied that they had not imagined themselves doing the task the day before, as an aid to performing the task in the second session. However, only 3 subjects replied that they had imagined more than the desk and objects in front of them. These 3 subjects also had imagined the room they had performed the task in previously to help them remember. Of these 3, only 1 was in the D condition. The only other strategies to aid task remembrance mentioned by subjects were trying to imagine the tape recorded instructions and "recalling actions". 13 subjects mentioned the former strategy (7, D) and 4 subjects mentioned the latter strategy (2, D). Of course, most subjects (14; 7, D) reported employing more than one recall strategy and of these 3 subjects (2, D) reported using three strategies. No subject indicated that they had used more than three recall strategies.

3.4.2. similarity of environment ratings.

The median rating of environmental similarity was 8.75 ($\bar{X}$ = 11.67, S.D. = 11.40).

4. Discussion.

4.1. Placement order analyses.

4.1.1. differences between correlation types.

The pattern of results obtained in the comparison of the placement order correlation types suggests that although subjects' performance in the first session was based on their remembrance of their instructions, their second session performance was based on their remembrance of their first session performance, rather than the original instructions. Although no direct evidence is available on the topic, as subjects were asked to try and reform the line in the same order as they had been instructed originally, it seems likely that their second session performance is due in part to an inability to distinguish between their remembrance of the original instructions and their remembrance of their first session performance. Of course it is possible that subjects could make some discrimination between their two previous activities, but as certain gaps occurred with respect to the instructed performance, the easiest solution to this problem was to provide movements based on their more accessible memory for the session most recently performed. This explanation puts much more emphasis on the influence of time on remembrance. However, given that there was twenty-four hours between first session and second session performance, and roughly three minutes between performance carried out under instruction and first session performance, the latter account does seem less reasonable.

Another interesting finding was the effect of subject group. Although subjects were randomly allocated to reinstatement or change groups, on this occasion it so happened that this provided the reinstated group with superior performers. Fortunately, the use of a repeated measures design enabled this artifact to be identified. If a simple one factor ANOVA is calculated with rs2 data; employing S and D groups as the two levels, a significant difference in favour of the S group is obtained (two-tailed $p = 0.04$). However, as the two factor ANOVA, employing rs1 and rs2 data demonstrates, this effect is primarily a function of the lower overall performance of the D subject group.

4.1.2. the effect of environmental context on the remembrance of placement order.

The lack of a significant interaction effect in the subject group by session ANOVA indicates that environmental context is not affecting placement order performance as measured against the original instructions. However, this is not particularly surprising given the results of the comparisons between correlation types, which indicated that subjects were basing their second session placement order performance on their memory for their first session placement order performance. Consequently, the measure of performance that is likely to reveal the effect of environmental context is rs3. The ANOVA employing this variable indicated the predicted superiority of reinstated (S) environmental context conditions.

As outlined in the introduction to this chapter, there is much theoretical use of environmental context as a source of information that "activates", and/or "triggers" action schemata (see also, Norman & Shallice, 1985; Reason, 1979; 1984a; 1984b). In the Norman (1981) and Norman and Shallice (1985) accounts, one of the most interesting aspects from the perspective of environmental context involvement is how often environmental context is able to exert an effect on performance

Influence can be exerted not only in the activation of a whole set of schemata, but also in the specific triggering of each schema in (or sometimes out of) sequence. Given the theoretical importance of environmental context in willed and automatic action, perhaps it is surprising that the environmental context effect observed is only marginally significant. However, this issue is discussed in the context of the differential effect of environmental context on placement and completed line order performance, towards the end of the next section.

4.2. Completed line order analyses.

In some respects, subjects' remembrance of completed line order is nearer to the type of remembrance required in previous experiments. Although an order has to be remembered, it can be defined in more ways than is possible for placement order. For example, completed line order performance can make use of not only sequence information (ie. defining in terms of previous, or subsequent entities), but also any "pattern" information available. The "pattern" allows the line to be defined from any point, eg. left, right, middle, or from any object, including reference points not actually part of the nominal object line, such as aspects of the table geography. The physical existence of the line also avoids the difficulties and demands that maintaining and operating a mental model of the order makes on processing. In comparison, the order of object placement has few sources of information that can be employed in defining the sequence of moves. Although the sequence of moves may be defined by start and finish points, between these there is little but the previous move and the subsequent move that can be employed in definition. Fortunately, in real life situations this predicament is likely to be alleviated by structured temporal environments.

Nevertheless, despite the differences in placement and completed line ordering, the total lack of effect of either time, standard of comparison, environmental context, or subject group is as puzzling as any of the effects observed in the analyses of placement order. However, as a similar ordering of high to low correlation types is observed, there is a suggestion that similar factors are operating to influence completed line order performance.

The different effects of environmental context reinstatement/change on placement and completed line orders suggests that different processes are involved in these two aspects of the subjects' task. The lack of effect of environmental context on completed line order after twenty-four hours compares with the similar lack of effect observed with the recall of nominal stimulus words in experiment five (chapter ten). Consequently, these differences in the effect of environmental context support the view that there is greater similarity of processing between the recall of nominal stimulus words and completed line order, as opposed to placement order performance. In both cases, the recall performance was required after a presentation-test gap of twenty-four hours. Presumably, although the environmental context representation altered over time in the manner described in chapter ten, its degree of use by the processes underlying placement order performance highlights any discrimination that can be made.

4.3. Questionnaire data.

As in previous experiments, subjects ratings of the similarity of the environments suggested that they perceived them as being considerably different.

In common with all previous questionnaires, many subjects responded that they had imagined part of the environment in an attempt to aid their recall. Only three subjects responded that they had imagined the whole room, but most subjects indicated that they had envisaged the immediate area occupied by the task objects. So, although only a small part of the environment was involved, it would seem that the principle that environmental context information can be used as an aid to remembrance is adhered to. In fact, this suggestion of the local nature of the environmental information employed may contribute to the account of the comparative lack of effect of a change in environment. If subjects were confining their attention to the immediate locality of the objects, and as the objects were exactly replicated in both environments, there would be little to distinguish the two situations except the difference between table tops.

Finally, the variety of recall strategies operated by subjects and observed in previous experiments again was evident.

CHAPTER THIRTEEN

GENERAL SUMMARY AND DISCUSSION

1. The Approach To The Investigation Of Environmental Context Effects.

1.1. Origins.

As stated in chapter six, much of the work reported in this thesis is exploratory in nature. Although several experiments examined hypotheses generated by particular theoretical accounts, the general lack of specific theoretical accounts predicting environmental context effects sometimes forced a more pragmatic approach. However despite this, the conditions determining environmental context effects, and the environmental context reinstatement effect in particular, were sought as part of an attempt to understand the psychological influence of environmental context.

This general approach emphasises the role of psychological processing in the account of environmental context phenomena, with the term "conditions" including the different forms of psychological processing underlying the subjects' overt behaviour in the various experimental situations. This may seem to be a very obvious and perhaps redundant point, made as it is with regard to psychological research.

Nevertheless, it seems it is this consideration of what the person actually does in the various experimental situations that as much as any other factor distinguishes between the present research and other research into environmental context phenomena.

Previous research has tended to define situations and subjects' average responses, rather than involve itself with trying to describe the psychological processes that give rise to such environmental context effects. More recently, research has considered such processes, but only in a shallow manner. Usually, a very basic model of encoding and retrieval is referred to, as the the purpose or results of an experiment are discussed. It would seem that one of the reasons why subjects have been assumed to operate uniformly in such simplistic manners is because there were no alternative theoretical accounts that could accommodate non-verbal information (Tulving, 1979).

The fundamental problem faced by any theoretical account of environmental context phenomena is how to represent nominal and environmental context information. This is a fundamental problem as the form of representation determines to a very large degree the type of encoding and retrieval operations that need to and may be carried out. Without the theoretical development of a representational system that can provide an account of the basic environmental context effects to explore, the only direction that research on this topic can take is to demonstrate the situational variables that influence environmental context phenomena. Essentially, this is the difference between process oriented accounts and what previously were described as descriptive accounts. However, the success of an approach that attempts to account for environmental context effects in terms of psychological processing depends on the availability of good theory.

1.2. Some consequences.

The account of representation, encoding and retrieval provided by the schema model presented in chapter six allows for a great variety of different encoding and retrieval strategies. A variety of encoding and retrieval strategies can be employed not only by different subjects, but also by each subject in any experiment. Such ways of remembering are experientially and intuitively obvious, but they have not been accommodated, or explicated very well by previous theoretical accounts. Indeed, the high error variances commonly observed in psychology experiments (Cohen, 1977) are probably due to this theoretically and statistically unaccommodated aspect of behaviour, which is normally categorised under the heading of individual differences (Kirk, 1982).

For theoretical reasons, obviously it is important that the form of processing operated by subjects should not contradict that upon which predictions and hypotheses are based. However, as evidenced by several of the experiments reported here, in an experiment people do not always behave in such conforming ways as some research reports have tended to portray them (Claxton, 1980, p.13). As more sophisticated models are developed and more specific predictions and hypotheses are generated, the degree of control of subjects' processing will be required to increase. In the present work some steps were taken to ascertain and control the manner of processing engaged by subjects. Eventually, it was discovered that the use of a relatively simple multivariate methodology was sufficient to statistically control the effect of subjects' processing on the dependent variable. Multivariate procedures are not only easier to apply as access to computing resources increases, but they also provide great potential for the experimental investigation of psychological processing; such as the examination of the relationship between recall and recognition, in addition to their more commonly recognised value in non, and quasi-experimental designs.

2. Findings And Implications.

2.1. Mode of processing.

Probably the most important finding reported in this thesis is the interaction between the form of psychological processing operated and the environmental context reinstatement effect. In itself this factor is likely to account for much of the apparent unreliability of the environmental context reinstatement effect. The importance of this factor in any account of, or prediction concerning environmental context effects demonstrates the need to be able to describe psychological activity. Without an insight into the psychological consequences of different tasks, the behavioural effects would remain enigmatic, as observed in the Fernandez and Glenberg (1985) study.

2.2. Self-generation of environmental context.

The difficulty encountered in trying to prevent subjects from employing a self-generation of encoding environment strategy provides an interesting insight into the retrieval process. The common use of this strategy supports the argument presented in the opening chapter that the use of environmental context to aid remembrance is quite common. Also, such common use would suggest that the self-generation strategy usually is an effective method of retrieval. Nevertheless, the question as to whether the use of such a strategy is a specific consequence of the difficulty of the attempt to recall has been raised. As mentioned in chapter one, an awareness of employing environmental context does seem to occur most often when there is difficulty remembering. Bobrow and Norman (1975), and Williams and Hollan (1981) argue that the difficulty of a task does not particularly change the way in which processing occurs, other than bringing the processes involved into consciousness. Although it is interesting to note the similarity of retrieval strategies revealed by the questionnaires and that proposed by Williams and Hollan, again this may be a consequence of similar recall difficulty, rather than the general applicability of the account. At present however, the weight of evidence indicating the utility of environmental context information in a variety of experimental situations; particularly those in which the response times suggest a lack of conscious mediation, supports the view that environmental context information is regularly and successfully employed at retrieval in a manner that conforms with the schema model and the particular description of its operation provided by Williams and Hollan.

2.3. The influence of time.

The observation that increasing the presentation-test delay does not necessarily increase environmental context effects provoked consideration of the relationship between environmental context effects and the memory representation of the environmental context information involved. The suggestion that the representation of environmental context information "becomes" more abstract over time is consistent with several other ideas and opinions which have been expressed with respect to environmental context representation.

In chapter one, a relationship between environmental context and semantic context was observed. If semantic contexts were considered, each seemed to be associated with some environmental context. Likewise, if environmental contexts were considered, each seemed to have some meaning attached to it. Fernandez and Glenberg (1985) discuss some pertinent aspects of environmental context as they consider why they did not obtain the environmental context reinstatement effects they expected. As was discussed in chapter twelve, they suggest that natural environments; unlike those constructed for experimental purposes, are tied to events in special ways. Often environments enable, or cause certain events. The restaurant and its props in Schank and Abelson's (1977) notion of a restaurant script provides a theoretically compatible example. Another and very pertinent point that Fernandez and Glenberg make is that a critical feature of experimental environments is the fact that they are part of the experiment. This aspect of environmental context representation was raised in relation to Norman and Rumelhart's

(1975) model of perception and memory, in chapter five (section 5.4.). In other words, the subjects' context may be "taking part in an experiment," which includes information about environmental context features, but at a lower level in the representational hierarchy. In such a situation, a change in the subjects' environment would not change the subjects' context. Clearly, there is a relationship between these ideas and the notion of a type of semantic abstraction of environmental context representation.

Eich (1985) has suggested that one reason for the lack of an environmental context effect may be the over-loading of the environmental context cue(s). However, the present discussion indicates how vague the notion of an environmental context cue is, even after eighty years of use. Another suggestion from Eich is that the studies of environmental context reinstatement with the greatest magnitude of effect may have succeeded by inducing a change in subjects' mood between environments, as well as changing environmental context information. Again this raises the question, what is involved in the representation of environmental context?

The account of the effect of the increased presentation-test delay on the environmental context reinstatement effect emphasises the inter-relations between memory processes that give rise to this and probably many other effects. After consideration of basic schema accounts in conjunction with the architecture of the memory record proposed by the model, it may even be claimed that such effects should have been predicted. The account also suggests that much more attention need be given to the representation of environment, if a proper and full understanding of environmental context effects is to be achieved. The relationship between the theoretical notions of semantic and environmental context and the apparent need to account for such a relationship supports the view of the essential underlying similarity of physical (ie. environmental and state) context and semantic context effects.

2.4. Observations with recognition measures.

The observation of an environmental context reinstatement effect with a recognition paradigm given the appropriate form of processing, again demonstrates the validity and utility of the attempt to account for such effects in terms of processing which can vary in degree as well as form, rather than in terms of discrete experimental situations. The pattern of results obtained across the different conditions in the recognition experiment indicates as suggested, that there are factors at both encoding (processing time) and retrieval (memory record and retrieval specification compatibility) that may operate independently to influence the manifestation of environmental context effects.

Eich (1985) suggests that environmental context information may not be encoded with nominal items as frequently as has been presumed. Unfortunately, as described in chapter eleven (section 1.5.), there are problems associated with any simple attempt to infer that environmental context is not encoded. However, it seems that research into the nature of environmental context information processing and the use to which such processing is put, will lead to a much greater understanding of environmental context effects. At present however, the evidence from the recognition experiment would suggest that provided subjects attend to/process environmental information sufficiently, and enough processing time is available per item, environmental context will be encoded as part of the memory record. Nevertheless, the manifestation of an effect still will depend on the retrieval processes that are engaged.

The results of the recognition experiment were not compatible with the predictions derived from the Atkinson and Juola model of recognition. This is quite ironic, as Mandler (1980) cites this model as one of the major influences on his account of recognition, and it is Mandler's account which provides the basis for the detailed application of the schema model to situations requiring recognition. One of the basic problems with the Atkinson and Juola account as a model of general recognition is that it fails to allow for the fact that different types of information "carry more weight" than others. By simplifying the forms of information available they are able to mathematically model a particular form of assessment, but it is probably a form of assessment that takes place only within an information hierarchical level, if at all.

2.5. Intentional action.

The experiment carried out to determine the effect of environmental context on intentional action was primarily an attempt to replicate an aspect of behaviour that experience and intuition indicated should be influenced by environmental context. However, it also turned out that this was one of the few areas where recent theoretical and anecdotal accounts had acknowledged the specific, rather than general and vague influence of environmental context. The demonstration of the effect of environmental context on intended action not only provides further support to the schema based account, but it also provides the experimental evidence that previously was lacking.

2.6. Conclusions.

Overall the demonstration of environmental context effects illustrates a important theoretical and practical aspect of psychological processing that largely has been ignored. The reason for this aspect of psychological processing being ignored is probably because its relation to other aspects of psychological processing has not been understood.

In chapter two (section 7.), it was argued that the oddity of environmental context effects was due to their implication of simple mechanistic control of intellectual function. Hopefully, this thesis has shown that the logic leading to this conclusion is flawed. Such logic is one consequence of the inappropriate abstraction of "semantic" and "physical" information. This inappropriate abstraction creates a false dichotomy between information types that restricts the view of how such information is able to be employed in psychological processing.

The psychological system alluded to by the descriptions presented here certainly is not simple, but neither is it wholly unfathomable. While the account may be termed mechanistic, the standard connotations of this term now must be obsolete. The effect of physical context is the consequence of a sophisticated processing system that strategically utilises information in a practical and effective manner.

2.7. Postscript.

There were several purposes to the presentation in chapter five of the possible and actual accounts of environmental context phenomena. These included the illustration of the use of environmental context as a theoretical construct and the changing requirements of theory. However, the main purpose was to demonstrate that as judged by empirical and metatheoretical criteria, these accounts failed to provide adequate explanations of the environmental context effects recorded in the literature and described in chapter two.

Consequently, the assessment that these models are inadequate and/or inappropriate formulations to accommodate the environmental context phenomena recorded in the literature makes redundant the exercise of considering the ability of these models to account for the results of the experiments reported in chapters seven to twelve. The irrelevant nature purpose of

> (s)ubsequent chapters..(was to)..report studies which continue(d) what..(was)..basically an exploration of the processes through which environmental context exerts its influence. As this..(was)..carried out, constant reference..(was)..made to the model of psychological processing outlined in this chapter in an attempt to obtain a theoretical, as well as an empirical grip on the phenomena (chapter six, section 3.4.).

All of the data obtained from the experiments reported in chapters seven to twelve could be accommodated by each of the accounts presented in chapter five. If the basic description employed by an account is accepted, it is not particularly difficult to to include or presume mechanisms that will manifest the effects observed. An example of this exercise of account remoulding is provided with Norman and Rumelhart's model. How unsatisfactory or ridiculous some of these accounts subsequently appear depends upon the metatheoretical criteria that are applied in their assessment. It is not possible to evaluate an account without reference to some such set of criteria. However, it would seem that the remit of this appendage is to assess briefly the worth of the unmodified accounts presented in chapter five with respect to the data presented in chapters seven to twelve, ignoring the important criteria that some of these accounts already have failed to meet.

The experiments reported in chapters seven to twelve have revealed an effect of environmental context reinstatement/change with mode of processing or degree of semantic elaboration, self-generation of presentation context, recall delay, recognition and intended action. Of these presented, the simple cue based models, ie. Carr and McGeoch, Smith, Godden and Baddeley, and Norman and Rumelhart, make no allowances for differences in processing or the effect of time on the environmental context reinstatement effect, nor do they provide any mechanisms as would be required to account for the self generation of context or intentional action. The effect of environmental context on recall and recognition also requires a more subtle theoretical account than can be provided within the strictures of a pure and simple cue based model. This is demonstrated by Godden and Baddeley's account of the "lack" of effect of environmental context on recognition.

Without employing metatheoretical criteria, it is possible to say only that Tulving's and Glenberg's characterisations do not provide any mechanism to account for intended action and consequently, the effect of environmental context upon intended action, nor for the effect of time on environmental context effects. The other effects observed should be able to be accommodated, although this would have to be accomplished without reference to any specific representational structure or unambiguous description of the processes that give rise to these effects.

The final account to be considered and probably the one which has caused most concern, is that of Anderson and Bower (FRAN & HAM). This account apparently has renewed candidacy given the observation of an environmental context effect with recognition. However, there is no way that the tenability of this account can be evaluated without introducing metatheoretical criteria.

If forced, Anderson and Bower's model can give some account of the different effect of environmental context reinstatement/change on recall and recognition apparent in the literature. Recall items are identified by list markers on the links between nodes in the associated memory network. If it is assumed that these list markers are identified and assessed as at the recognition phase, then a benefit to reinstated condition recall might be expected as two assessments involving environmental context information are required for recall, whereas only one assessment is required for recognition. However, it seems very unlikely that such a full scale assessment of list markers would be carried out at the point of identifying potential recall items. It would seem economical to carry out this operation on the basis of some simple nominal code. Either way, the Anderson and Bower account still cannot explain the relative size of the recall and recognition effects observed here, unless, as with the schema model account, it is assumed that the experimental paradigm biases the sort of information referenced by the list marker. However, this assumption requires some mechanism other than the probability based operation described (see chapter five, section 4.2.) to control the associations to and from the list markers. In addition, the Anderson and Bower account has the theoretical drawback of assuming "direct access" (see chapter eleven, section 1.3.), while HAM has been criticised with respect to its naive parsing assumptions (Kolers & Smythe, 1984), its lack of a formal semantic specification of its representational structure, its method of long-term memory search and its ability to reflect natural language effects (Anderson, 1976; Foss & Harwood, 1975).

As demonstrated previously and also by this brief consideration of the explanative ability of alternative models, one of the main attractions of the schema system is the ease with which the description of its representational structure and encoding and retrieval operations can provide a coherent account of so many observations.

REFERENCES

Abelson, R.P. "The psychological status of the script concept." American Psychologist (1981) 36 715-729.

Abernethy, E.M. "The effect of changed environmental conditions upon the results of college examinations." Journal Of Experimental Psychology (1940) 10 293-301.

Alba, J.W. and Hasher, L. "Is memory schematic?" Psychological Bulletin (1983) 93 203-231.

Allport, D.A. "Patterns and actions: cognitive mechanisms are content specific." In G. Claxton (ed), Cognitive Psychology, New Directions. London: Routledge and Keagan Paul, 1980.

Allport, D.A. "Alternatives to the computational view of mind: the baby or the bath water." Journal Of Verbal Learning And Verbal Behaviour (1984) 23 315-324.

Allport, D.A., Antonis, B. and Reynolds, P. "On the division of attention: a disproof of the single-channel hypothesis." Quarterly Journal Of Experimental Psychology (1972) 24 225-235.

Anderson, J.R. _Language, Memory And Thought._ Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1976.

Anderson, J.R. and Bower, G.H. "Recognition and retrieval processes in free recall." _Psychological Review_ (1972) _79_ 97-123.

Anderson, J.R. and Bower, G.H. _Human Associative Memory._ New York: John Wiley and Sons, 1973.

Anderson, J.R. and Bower, G.H. "A propositional theory of recognition memory." _Memory And Cognition_ (1974) _2_ 406-412.

Anderson, T.W. _An Introduction To Multivariate Statistical Analysis._ New York: Wiley, 1955.

Atkinson, R.C. and Juola, J.F. "Search and decision processes in recognition memory." In D.H. Krantz, R.C. Atkinson, R.D. Luce and P. Suppes (eds), _Contemporary Developments In Mathematical Psychology: Vol 1; Learning, Memory And Thinking._ San Francisco: Freeman, 1974.

Baddeley, A.D. _The Psychology Of Memory._ New York: Basic Books, 1976.

Baddeley, A.D. "Domains of recollection." _Psychological Review_ (1982)a _89_ 708-729.

Baddeley, A.D. Your Memory: A User's Guide. Harmondsworth, Middx.: Penguin, 1982(b).

Baddeley, A.D. and Scott, D. "Short-term forgetting in the absence of proactive inhibition." Quarterly Journal Of Experimental Psychology (1971) 23 275-283.

Barclay, J.R., Bransford, J.D., Franks, J.J., McCarrell, N.S. and Nitsch, K. "Comprehension and semantic flexibility." Journal Of Verbal Learning And Verbal Behaviour (1974) 13 471-481.

Baroni, M.R., Job, R., Peron, E.M. and Salmaso, P. "Memory for natural settings: role of diffuse and focused attention." Perceptual And Motor Skills (1980) 51 883-889.

Barnes, J.M. "Extinction of associations in four transfer paradigms." Psychological Monographs (1964) 78 1-21.

Bartlett, F.C. Remembering. Cambridge: Cambridge University Press, 1932.

Berger, T.W., Alger, B. and Thompson, R.F. "Neuronal substrate of classical conditioning in the hippocampus." Science (1976) 192 483-485.

Berger, T.W., Clark, G.A. and Thompson, R.F. "Learning-independent neuronal responses recorded from limbic system brain structures during classical conditioning." Physiological Psychology (1980) 8 155-167.

Bilodeau, I.M. and Schlosberg, H. "Similarity in stimulating conditions as a variable in retroactive inhibition." Journal Of Experimental Psychology (1951) 41 199-204.

Birnbaum, I.M. "Context stimuli in verbal learning and the persistence of associative factors." Journal Of Experimental Psychology (1966) 71 483-487.

Bjork, R.A. "Retrieval as a memory modifier: an interpretation of negative recency and related phenomena." In R.L. Solso (ed), Information Processing And Cognition. Hillsdale, New Jersey: Erlbaum, 1975.

Bjork, R.A. and Geiselman, R.E. "Constituent processes in the differentiation of items in memory." Journal Of Experimental Psychology: Human Learning And Memory (1978) 4 347-361.

Block, R.A. "Temporal judgements and contextual change." Journal Of Experimental Psychology: Learning, Memory And Cognition (1982) 8 530-544.

Block, R.A. and Reed, M.A. "Remembered duration: evidence for a contextual change hypothesis." Journal Of Experimental Psychology: Human Learning And Memory (1978) 4 656-665.

Bobrow, D.E. "Dimensions of representation." In D.G. Bobrow and A. Collins (eds), Representation And Understanding: Studies In Cognitive Science. New York: Academic Press, 1975.

Bobrow, D.G. and Norman, D.A. "Some principles of memory schemata." In D.G. Bobrow and A. Collins (eds), Representation And Understanding: Studies In Cognitive Science. New York: Academic Press, 1975.

Bobrow, D.G. and Winograd, T. "An overview of KRL, a knowledge representation language." Cognitive Science (1977) 1 3-46.

Boden, M.A. Artificial Intelligence And Natural Man. Hassocks: Harvester Press, 1977.

Bower, G.H. "Stimulus-sampling theory of encoding variability." In E. Martin and A. Melton (eds), Coding Theory In Learning And Memory. New York: Academic Press, 1972(a).

Bower, G.H. "Mental imagery and associative learning." In L.W. Gregg (ed), Cognition In Learning And Memory. New York: Wiley, 1972(b).

Bower, G.H. "Mood and memory." American Psychologist (1981) 36 129-148.

Bower, G.H. and Clark, M.C. "Narrative stories as mediators for serial learning." Psychonomic Science (1969) 14 181-182.

Bower, G.H., Clark, M.C., Lesgold, A.M. and Winzenz, D. "Hierarchical retrieval schemes in recall of categorised word lists." Journal Of Verbal Learning And Verbal Behaviour (1969) 8 323-343.

Bower, G.H., Gilligan, S.G. and Monteiro, K.P. "Selectivity learning caused by affective states." Journal Of Experimental Psychology: General (1981) 110 451-473.

Bower, G.H., Monteiro, K.P. and Gilligan, S.G. "Emotional mood as a context of learning and recall." Journal Of Verbal Learning And Verbal Behaviour (1978) 17 573-583.

Bower, G.H. and Winzenz, D. "Comparison of associative learning strategies." Psychonomic Science (1970) 20 119-120.

Box, G.E.P. and Anderson, S.L. "Permutation theory in the determination of robust criteria and the study of departures from assumptions." Journal Of The Royal Statistical Society, Series B (1955) 17 1-34.

Bransford, J.D. and Johnson, M.K. "Contextual prerequisites for understanding: some investigations of comprehension and recall." Journal Of Verbal Learning And Verbal Behaviour (1972) 11 712-726.

Brewer, W.F. and Dupree, D.A. "Use of plan schemata in the recall and recognition of goal-directed actions." Journal Of Experimental Psychology: Learning, Memory And Cognition (1983) 9 117-129.

Brewer, W.F. and Treyens, J.C. "Role of schemata in memory for places." Cognitive Psychology (1981) 13 207-230.

Broadbent, D.E. Perception And Communication. London: Pergamon, 1958.

Broadbent, D.E. Decision And Stress. London: Academic Press, 1971.

Broadbent, D.E. and Broadbent, M.H.P. "Effects of recognition on subsequent recall: comments on 'Determinants of recognition and recall: accessibility and generation' by Rabinowitz, Mandler and Patterson." Journal Of Experimental Psychology: General (1977) 106 330-335.

Brown, J. "Some tests of the decay theory of immediate memory." Quarterly Journal Of Experimental Psychology (1958) 10 12-21.

Brown, J. "An analysis of recognition and recall and of problems in their comparison." In Brown, J. (ed), Recall And Recognition. Chichester: Wiley, 1976.

Carlson, N.R. Physiology Of Behaviour. Boston: Allyn and Bacon, Inc., 1980.

Carmichael, L., Hogan, H.P. and Walters, A.A. "An experimental study of the effect of language on the reproduction of visually perceived form." Journal Of Experimental Psychology (1932) 15 73-86.

Carr, H.A. "Maze studies with the white rat." Journal Of Animal Behaviour (1917) 7 259-306.

Carr, H.A. Psychology: A Study Of Mental Activity. New York, Longmans Green, 1925.

Carroll, R.M. and Nordholm, L.A. "Sampling characteristics of Kelley's E^2 and Hay's W^2." Educational And Psychological Measurement (1975) 35 541-554.

Claxton, G. "Cognitive psychology: a suitable case for what sort of treatment?" In G. Claxton (ed), Cognitive Psychology, New Directions. London: Routledge and Keagan Paul, 1980.

Cohen, S. Statistical Power Analysis For The Behavioural Sciences. New York: Academic Press, 1977.

Craik, F.I.M. "Encoding and retrieval effects in human memory: a partial review." In A.D. Baddeley and J. Long (eds), Attention And

Performance IX. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1981.

Craik, F.I.M. and Jacoby, L.L. "Elaboration and distinctiveness in episodic memory." In L-G. Nilsson (ed), Perspectives On Memory Research. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1979.

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

Crowder, R.G. Principles Of Learning And Memory. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1976.

Dallet, K. and Wilcox, S.G. "Contextual stimuli and proactive interference." Journal Of Experimental Psychology (1968) 78 475-480.

Devane, J.R. and Parkman, R.A. "Effects of imagery instructions when the context is changed." Perceptual And Motor Skills (1978) 46 716-718.

Devenport, L.D. and Hollaway, F.A. "The rat's resistance to superstition: role of the hippocampus." Journal Of Comparative And Physiological Psychology (1980) 94 691-705.

Deweer, B., Sara, S.J. and Hars, B. "Contextual cues and memory retrieval in rats: alleviation of forgetting by a pretest exposure to background stimuli." Animal Learning And Behaviour (1980) 8 265-272.

Dixon, W.J. (chief editor), BMDP Statistical Software Manual, 1983 Printing With Additions. Berkeley: University Of California Press, 1983.

Dolinsky, R. and Zabrucky, K. "Effects of environmental context changes on memory." Bulletin Of The Psychonomic Society (1983) 21 423-426.

Dooling, D.J. and Danks, J.H. "Going beyond tests of significance: is psychology ready?" Bulletin Of The Psychonomic Society (1975) 5 15-17.

Dooling, D.J. and Lachman, R. "Effects of comprehension on retention of prose." Journal Of Experimental Psychology (1971) 88 216-222.

Draper, N.R. and Smith, H. Applied Regression Analysis. New York: Wiley, 1981.

Dulsky, S.G. "The effect of a change of background on recall and relearning." Journal Of Experimental Psychology (1935) 18 725-740.

Ebbinghaus, H. Uber Das Gedachtnis. Leipzig: Dunker, 1885. Translation by H. Ruyer and C.E. Bussenius, Memory. New York: Teacher's College Press, 1913.

Eich, J.E. "State-dependent retrieval of information in human episodic memory." In I.M. Birnbaum and E.S. Parker (eds), Alcohol And Human Memory. Hillsdale, New Jersey: Lawrence Earlbaum Associates, 1977.

Eich, J.E. "The cue-dependent nature of state-dependent retrieval." Memory And Cognition (1980) 8 157-173.

Eich, J.E. "Context, memory and integrated item/context imagery." Journal Of Experimental Psychology: Learning, Memory and Cognition (1985) 11 764-770.

Elio, R.E. and Reutener, D.B. "Colour context as a factor in encoding and as an organisation device for retrieval of word lists." Journal Of General Psychology (1978) 99 223-232.

Eliot, G. The Mill On The Floss. Originally published 1860. Harmondsworth, Middx.: Penguin, 1985.

Engelman, L. "P2D: detailed data description, including frequencies." In Dixon, W.J. (chief editor), BMDP Statistical Software Manual, 1983 Printing With Additions. Berkeley: University Of California Press, 1983.

Estes, W.K. "Statistical theory of spontaneous recovery and regression." Psychological Review (1955) 62 145-154.

Estes, W.K. "The statistical approach to learning theory." In S. Koch (ed), Psychology: A study Of A Science. Vol. 2. New York: McGraw-Hill, 1959.

Eysenck, M.W. A Handbook Of Cognitive Psychology. London: Lawrence Erlbaum Associates, 1984.

Falkenberg, P.R. "Recall improves in short-term memory the more recall context resembles learning context." Journal Of Experimental Psychology (1972) 95 39-47.

Farnsworth, P.R. "Examinations in familiar and unfamiliar surroundings." Journal Of Social Psychology (1934) 5 128-129.

Farnsworth, P.R. "Continued training with the omission of certain nebenreize." Journal Of Genetic Psychology (1937) 50 277-282.

Fernandez, A. and Glenberg, A.M. "Changing environmental context does not reliably affect memory." Memory And Cognition (1985) 13 333-345.

Fisher, R.P. and Craik, F.I.M. "Interaction between encoding and retrieval operations in cued recall." Journal Of Experimental Psychology: Human Learning And Memory (1977) 3 701-711.

Forsyth, A.B. "It wasn't an accident (F-to-enter, F-to-remove)." In Dixon, W.J. (chief editor), BMDP Statistical Software Manual, 1983 Printing With Additions. Appendix C2. Berkeley: University Of California Press, 1983.

Forsythe, A.B., May, P.R.A. and Engelman, L. "Prediction by multiple regression. How many variables to enter?" Journal Of Psychiatric Research (1971) 8 119-126.

Foss, D.J. and Harwood, D.A. "Memory for sentences: implications for human associative memory." Journal Of Verbal Learning And Verbal Behaviour (1975) 14 1-16.

Gaffan, D. "Loss of recognition memory in rats with lesions of the fornix." Neuropsychologia (1972) 10 327-341.

Gaffan, D. "Recognition impaired and association intact in the memory of monkeys after transection of the fornix." Journal Of Comparative And Physiological Psychology (1974) 86 1100-1109.

Geis, M.F. and Winograd, E. "Semantic encoding and judgement of background and situational frequency for homographs." Journal Of Experimental Psychology: Human Learning And Memory (1975) 104 385-392.

Glass, G.V. and Hakistan, A.R. "Measurements of association in comparative experiments: their development and interpretation."

American Educational Research Journal (1969) 6 403-414.

Glass, G.V., Peckham, P.D. and Sanders, J.R. "Consequences of failure to meet assumptions underlying the fixed effects analysis of variance and covariance." Review Of Educational Research (1972) 42 237-288.

Glenberg, A.M. "Component-levels theory of the effects of spacing of repetitions on recall and recognition." Memory And Cognition (1979) 7 95-112.

Glenberg, A.M. and Adams, F. "Type 1 rehearsal and recognition." Journal Of Verbal Learning And Verbal Behaviour (1978) 17 455-463.

Godden, D.R. and Baddeley, A.D. "Context-dependent memory in two natural environments: on land and underwater." British Journal Of Psychology (1975) 66 325-331.

Godden, D.R. and Baddeley, A.D. "When does context affect recognition memory?" British Journal Of Psychology (1980) 71 99-104.

Gottlieb, W. and Lindauer, M.S. "The effect of contextual stimuli on retroactive inhibition." Pychonomic Science (1967) 9 331-332.

Graesser, A.C. and Nakamura, G.V. "The impact of a schema on comprehension and memory." The Psychology Of Learning And

Motivation (1982) 16 59-109.

Greenspoon, J. and Ranyard, R. "Stimulus conditions and retroactive inhibition." Journal Of Experimental Psychology (1957) 53 55-59.

Guthrie, E.R. The Psychology Of Learning. New York: Harper, 1935.

Hartley, H.O. "Testing the homogeniety of a set of variances." Biometrica (1940) 31 249-255.

Hartley, H.O. "The maximum F-ratio as a short-cut test for heterogeniety of variance." Biometrica (1950) 37 308-312.

Hays, W.L. Statistics. New York: Holt, Rinehart and Winston, 1973.

Hewitt, K. "Context effects in memory: a review." Unpublished manuscript, Cambridge University Psychological Laboratory, 1977.

Hill, S.Y., Schwin, R., Powell, B. and Goodwin, D.W. "State-dependent effects of marihuana on human memory." Nature (1973) 153 395-406.

Hinton, G.E. and Anderson, J.A. (eds), Parallel Models Of Associative Memory. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1981.

Hirsch, R. "The hippocampus, conditional operations and cognition." Physiological Psychology (1980) 8 175-182.

Horel, J.A. "The neuroanatomy of amnesia: a critique of the hippocampal memory hypothesis." Brain (1978) 101 403-445.

Huitema, B.E. The Analysis Of Covariance And Alternatives. New York: Wiley, 1980.

Hull, C.L. Principles Of Behaviour. Appleton-Century-Crofts, 1943.

Iversen, S. "Do hippocampal lesions produce amnesia in animals?" In C.C. Pfieffer and J.R. Smythies (eds), International Review Of Neurobiology. New York: Academic Press, 1976.

Jacoby, L.L. and Craik, F.I.M. "Effects of elaboration of processing at encoding and retrieval: trace distinctiveness and recovery of initial context." In L.S. Cermack and F.I.M. Craik (eds), Levels Of Processing In Human Memory. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1979.

Jacoby, L.L. and Dallas, M. "On the relationship between autobiographical memory and perceptual learning." Experimental Psychology: General (1981) 110 306-340.

James, W. The Principles Of Psychology. New York: Holt, 1890.

Jensen, L.C., Harris, K. and Anderson, D.C. "Retention following a change in ambient contextual stimuli for six age groups." Developmental Psychology (1971) 4 394-399.

Jones, B.T., Rana, R.V. and McGonigle, B.O. "The role of context in the perception of orientation by the rat." Perception (1980) 9 591-598.

Johnson-Laird, P.N. "The correspondence and coherence theories of cognitive truth." The Behavioural And Brain Sciences (1978) 1 108-109.

Kahneman, D. Attention And Effort. Englewood Cliffs, New Jersey: Prentice Hall, 1973.

Keppel, G. Design And Analysis: A Researcher's Handbook. 2nd edition, Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1982.

Kesner, R.P. "An attribute analysis of memory: the role of the hippocampus." Physiological Psychology (1980) 8 189-197.

Kirk, R.E. Experimental Design: Procedures For The Behavioural Sciences. Monterey, Ca.: Brooks/Cole Publishing Company, 1982.

Kintsch, W. "Models for free-recall and recognition." In D.A. Norman (ed), Models Of Human Memory. New York: Academic Press, 1970.

Kolers, P.A. and Smythe, W.E. "Symbol manipulation: alternatives to the computational view of mind." Journal Of Verbal Learning And Verbal Behaviour (1984) 23 289-314.

Kucera, H. and Francis, W.N. A Computational Analysis Of Present-Day American English. Providence R.I.: Brown University Press, 1967.

Lane, D.M. and Dunlap, W.P. "Estimating effect size: bias resulting from the significance criterion in editorial decisions." British Journal Of Mathematical And Statistical Psychology (1978) 31 107-112.

Lehr, D.J. and Duncan, C.P. "Effect on priming on spontaneous recovery of verbal lists." Journal Of Verbal Learning And Verbal Behaviour (1970) 9 106-110.

Leight, K.A. and Ellis, H.C. "Emotional mood states, strategies and state-dependency in memory." Journal Of Verbal Learning And Verbal Behaviour (1981) 20 251-266.

Ley, R. and Huba, M. "Encoding specificity and experimenter supplied semantic associates as cues in recall." Journal Of

General Psychology (1980) 102 27-31.

Lichtenstein, E.H. and Brewer, W.F. "Memory for goal-directed events." Cognitive Psychology (1980) 12 412-445.

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

Lindquist, E.F. Design And Analysis Of Experiments In Psychology And Education. Boston: Houghton Mifflin, 1953.

Lisman, S.A. "Alcoholic 'blackout': state-dependent learning?" Archives Of General Psychiatry (1974) 30 46-53.

Locke, J. "An essay concerning human understanding." Originally published 1690. Reprinted by Everymans Library, London: Dent, 1961.

Lockhart, R.S., Craik, F.I.M. and Jacoby, L.L. "Depth of processing, recognition and recall: some aspects of a general memory system." In J. Brown (ed), Recall And Recognition. Chichester: Wiley, 1976.

McGeoch, J.A. The Psychology Of Human Learning. New York: Longmans Green, 1942.

McNicol, D. and Stewart, G.W. "Reaction time and the study of memory." In A.T. Welford (ed), Reaction Times. London: Academic Press, 1980.

Madigan, S.A. "Intraserial repetition and coding processes in free-recall." Journal Of Verbal Learning And Verbal Behaviour (1969) 8 828-835.

Mandler, G. "Organisation and repetition: an extension of organisational principles with special reference to rote learning." In L-G. Nilsson (ed), Perspectives In Memory Research. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1979.

Mandler, G. "Recognising: the judgement of previous occurrence." Psychological Review (1980) 98 252-271.

Mandler, G., Pearlstone, Z. and Koopmans, H.J. "Effects of organisation and semantic similarity on recall and recognition." Journal Of Verbal Learning And Verbal Behaviour (1969) 8 410-423.

Marcel, A.J. and Steel, R.G. "Semantic cueing in recognition and recall." Quarterly Journal Of Experimental Psychology (1973) 25 368-377.

Marr, D. "Artificial intelligence - a personal view." Artificial Intelligence (1977) 9 37-48.

Marr, D. _Vision: A Computational Investigation Into The Human Representation And Processing Of Visual Information._ San Francisco: W.H. Freeman and Company, 1982.

Marr, D. and Nishihara, H.K. "Visual information processing: artificial intelligence and the sensorium of sight." _Technology Review_ (1978) _81_ (October), 1-23.

Marr, D. and Poggio, T. "From understanding computation to understanding neural circuitry." _Neurosciences Research Progress Bulletin_ (1977) _15_ 470-488.

Martin, E. "Stimulus meaningfulness and paired-associate transfer: an encoding variability hypothesis." _Psychological Review_ (1968) _75_ 421-441.

Martin, E. "Verbal learning theory and independent retrieval phenomena." _Psychological Review_ (1971) _78_ 314-332.

Martin, E. "Memory codes and negative transfer." _Memory And Cognition_ (1973) _1_ 495-498.

Mayes, A., Meudell, P. and Som, S. "Further similarities between amnesia and normal attenuated memory: effects with paired-associate learning and contextual shifts." _Neuropsychologia_ (1981) _19_ 655-664.

Melton, A.W. "The situation with respect to the spacing of repetitions and memory." Journal Of Verbal Learning And Verbal Behaviour (1970) 9 596-606.

Metzger, R.L., Boschee, P.F., Haugen, T. and Schnobrich, B.L. "The classroom as learning context: changing rooms affects performance." Journal Of Educational Psychology (1979) 71 440-442.

Milner, B. "Amnesia following operation to the temporal lobe." In C.W. Whitty and O.L. Zangwill (eds), Amnesia. London: Butterworth and Co., 1966.

Milner, B. "Preface: material specific and generalised memory loss." Neuropsychologia (1968) 6 175-179.

Minsky, M. Semantic Information Processing. London: MIT Press, 1968.

Minsky, M. "A framework for representing knowledge." In P.H. Winston (ed), The Psychology Of Computer Vision. New York: McGraw-Hill, 1975.

Moore, J. and Newell, A. "How can MERLIN understand?" In L.W. Gregg (ed), Knowledge And Cognition. Potomac, Md.: Lawrence Erlbaum Associates, 1973.

Moore, J.W. and Stickney, K.J. "Formation of attentional-associative networks in real time: role of the hippocampus and implications for conditioning." Physiological Psychology (1980) 8 207-217.

Moray, N. "Where is capacity limited? A survey and a model." Acta Psychologica (1967) 27 84-92.

Morton, J., Hammersley, R.H. and Bekerian, D.A. "Headed records: a model for memory and its failures." Cognition (1985) 20 1-23.

Nabokov, V. Speak, Memory: An Autobiography Revisited. London: Weidenfeld and Nicolson, 1967.

Nadel, L. and Willner, J. "Context and conditioning: a place for space." Physiological Psychology (1980) 8 218-228.

Nagge, J.W. "An experimental test of the theory of associative interference." Journal Of Experimental Psychology (1935) 18 663-682.

Neisser, U. Cognitive Psychology. New York: Appleton-Century-Crofts, 1967.

Neisser, U. Cognition And Reality. San Francisco: Freeman, 1976.

Newell, A. "You can't play 20 questions with nature and win." In W.G. Chase (ed), Visual Information Processing. New York: Academic Press, 1973.

Newell, A. "Physical symbol systems." In D.A. Norman (ed), Perspectives On Cognitive Science. New Jersey: Ablex/LEA, 1981.

Nixon, S. and Kanak, N.J. "The interactive effects of instructional set and environmental changes on the serial position effect." Bulletin Of The Psychonomic Society (1981) 18 237-240.

Norman, D.A. "Perception, memory and mental processes." In L-G. Nilsson (ed), Perspectives On Memory Research. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1979.

Norman, D.A. "Twelve issues for cognitive science." In D.A. Norman (ed), Perspectives On Cognitive Science. New Jersey: Ablex/LEA, 1981.

Norman, D.A. "Categorisation of action slips." Psychological Review (1981) 88 1-15.

Norman, D.A. Learning And Memory. San Francisco: Freeman, 1982.

Norman, D.A. and Bobrow, D.G. "On the role of active memory processes in perception and cognition." In C.N. Cofer (ed), The Structure Of Human Memory. San Francisco: Freeman and Co., 1976.

Norman, D.A. and Bobrow, D.G. "Descriptions: an intermediate stage in memory retrieval." Cognitive Psychology (1979) 11 107-123.

Norman, D.A. and Rumelhart, D.E. "A system for perception and memory." In D.A. Norman (ed), Models Of Human Memory. New York: Academic Press, 1970.

Norman, D.A. and Rumelhart, D.E. and LNR Explorations In Cognition. San Francisco: Freeman, 1975.

Norman, D.A. and Shallice, T. "Attention to action: willed and automatic control of behaviour." In R.J. Davidson, G.E. Schwartz and D. Shapiro (eds), Consciousness And Self-Regulation: Advances In Research. Vol. VI. New York: Plenum Press, 1985.

Norton, D.W. "An empirical investigation of some effects of non-normality and heterogeniety of the F-distribution." Unpublished doctoral dissertation, State University Of Iowa, 1952.

O'Keefe, J. and Conway, D.H. "On the trail of the hippocampal engram." Physiological Psychology (1980) 8 229-238.

O'Keefe, J. and Nadel, L. The Hippocampus As A Cognitive Map. Oxford: Clarendon Press, 1978.

Olton, D.S., Becker, J.T. and Handelmann, G.E. "Hippocampal function: working memory or cognitive mapping." Physiological

Psychology (1980) 8 239-246.

Palmer, S.E. "Fundamental aspects of cognitive representation." In E. Rosch and B.B. Lloyd (eds), Cognition And Categorisation. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1978.

Pan, S. "The influence of context upon learning and recall." Journal Of Experimental Psychology (1926) 9 468-491.

Pessin, J. "The effect of similar and dissimilar conditions upon learning and relearning." Journal Of Experimental Psychology (1932) 15 427-435.

Peterson, L.R. and Peterson, M.J. "Short-term retention of individual items." Journal Of Experimental Psychology (1959) 58 193-198.

Proust, M. Remembrance Of Things Past. Originally published 1913-1927. Translation by C.K. Scott Moncrieff and Terence Kilmartin. Harmondsworth, Middx.: Penguin, 1983.

Rabbitt, P. "Current paradigms and models in human information processing." In V. Hamilton and D.M. Warburton (eds), Human Stress And Cognition: An Information Processing Approach. Chichester: John Wiley and Sons, 1979.

Rabinowitz, J.C., Mandler, G. and Patterson, K.E. "Determinants of recognition and recall: accessibility and generation." Journal Of Experimental Psychology: General (1977)a 106 330-335.

Rabinowitz, J.C., Mandler, G. and Patterson, K.E. "Clarifications, refutations and resolutions." Journal Of Experimental Psychology: General (1977)b 106 336-340.

Rand, G. and Wapner, S. "Postural status as a factor in memory." Journal Of Verbal Learning And Verbal Behaviour (1967) 6 268-271.

Reason, J. "Actions not as planned: the price of automatisation." In G. Underwood and R. Stevens (eds), Aspects Of Consciousness. Vol. 1. London: Academic Press, 1979.

Reason, J. "Lapses of attention." In R. Parasuraman, R. Davies and J. Beatty (eds), Varieties Of Attention. New York: Academic Press, 1984(a).

Reason, J. "Absent-mindedness and cognitive control." In J. Harris and P. Morris (eds), Everyday Memory, Actions And Absent-Mindedness. New York: Academic Press, 1984(b).

Reddy, B.G. and Bellazza, F.S. "Encoding specificity in free recall." Journal Of Experimental Psychology: Learning, Memory And Cognition (1983) 9 167-174.

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

Reed, H.J. "The influence of a change of conditions upon the amount recalled." Journal Of Experimental Psychology (1931) 14 632-649.

Reed, S.K. "Pattern recognition and categorisation." Cognitive Psychology (1972) 3 382-407.

Rogan, J.C. and Keselman, H.J. "Is the ANOVA F-test robust to variance heterogeniety when the sample sizes are equal?: an investigation via a coefficient of variation." American Educational Research Journal (1977) 14 493-498.

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

Rumelhart, D.E. and Norman, D.A. "Accretion, tuning and restructuring: three modes of learning." In J.W. Cotton and R. Klatzky (eds), Semantic Factors In Cognition. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1978.

Rumelhart, D.E. and Norman, D.A. "Analogical processes in learning." In, J.R. Anderson (ed), Cognitive Skills And Their Acquisition. Hillsdale, New Jersey: Lawrence Erlbaum Assoc., 1981.

Rumelhart, D.E. and Norman, D.A. "Simulating a skilled typist: a study of skilled cognitive-motor performance." Cognitive Science (1982) 6 1-36.

Rumelhart, D.E. and Ortony, A. "The representation of knowledge in memory." In, R.C. Anderson, R.J. Spiro and W.E. Montague (eds), Schooling And The Acquisition Of Knowledge. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.

Rundus, D. "Analysis of rehearsal processes in free recall." Journal Of Experimental Psychology (1971) 89 63-77.

Rutherford, A. and Hamilton, J. "The colebrook hypothesis." Unpublished draft manuscript concerning the possible natural history of morality, 1981.

Salmaso, P., Baroni, M.R., Job, R. and Peron, E.M. "Schematic information, attention and memory for places." Journal Of Experimental Psychology (1983) 9 263-268.

Saltz, E. "Compound stimuli in verbal learning: cognitive and sensory differentiation versus stimulus selection." Journal Of Experimental Psychology (1963) 66 1-5.

Sanford, A.J. Cognition And Cognitive Psychology. London: Weidenfeld and Nicolson, 1985.

Sanford, A.J. and Garrod, S.C. Understanding Written Language: Explorations Of Comprehension Beyond The Sentence. Chichester: John Wiley and Sons, 1981.

Schank, R.C. "Language and memory." In D.A. Norman (ed), Perspectives On Cognitive Science. New Jersey: Ablex/LEA, 1981.

Schank, R.C. and Abelson, R.P. Scripts, Plans, Goals And Understanding: An Inquiry Into Human Knowledge Structures. New York: Academic Press, 1977.

Searle, S.R. "Alternative covariance models for the 2-way crossed classification." Communications In Statistics (1979) A8 799-818.

Skinner, B.F. The Behaviour Of Organisms. New York: Appleton-Century-Crofts, 1938.

Smith, S. and Guthrie, E.R. General Psychology. New York: Appleton-Century-Crofts, 1921.

Smith, S.M. "Remembering in and out of context." Journal Of Experimental Psychology: Human Learning And Memory (1979) 5 449-530.

Smith, S.M. "Enhancement of recall using multiple environmental contexts during learning." Memory And Cognition (1982) 10 405-412.

Smith, S.M. "A comparison of two techniques for reducing context-dependent forgetting." Memory And Cognition (1984) 12 477-482.

Smith, S.M., Glenberg, A.M. and Bjork, R.A. "Environmental context and human memory." Memory And Cognition (1978) 6 342-353.

Solomon, P.R. "A time and a place for everything? Temporal processing views of hippocampal function with special reference to attention." Physiological Psychology (1980) 8 254-261.

Spear, N.E., Smith, G.J., Bryan, R.G., Gordon, W.C., Timmons, R. and Chiszar, D.A. "Contextual influences on the interaction between conflicting memories in the rat." Animal Learning And Behaviour (1980) 8 273-281.

Sternberg, S. "Memory scanning: mental processes revealed by reaction time experiments." American Scientist (1969) 57 421-457.

Stillman, R.C., Weingartner, H., Wyatt, R.J., Gillan, J.C. and Eich, J.E. "State dependent (dissociative) effects of marijuana on human memory." Archives Of General Psychiatry (1974) 31 81-85.

Strand, B.Z. "Change of context and retroactive inhibition." Journal Of Verbal Learning And Verbal Behaviour (1970) 9 202-206.

Thompson, R.F., Berger, T.W., Berry, S.D., Hoehler, F.K., Kettner, R.E. and Weisz, D.J. "Hippocampal substrate of classical conditioning." Physiological Psychology (1980) 8 262-279.

Thorndyke, P.W. "Cognitive structures in comprehension and discourse." Cognitive Psychology (1977) 9 77-110.

Tolman, E.C. Purposive Behaviour In Animals And Men. New York: Century, 1932.

Tulving, E. "Theoretical issues in free-recall." In T.R. Dixon and D.L. Horton (eds), Verbal Behaviour And General Behaviour Theory. Englewood Cliffs, New Jersey: Prentice Hall, 1968.

Tulving, E. "Recall and recognition of semantically encoded words." Journal Of Experimental Psychology (1974) 102 778-787.

Tulving, E. "Memory research: what kind of progress?" In L-G. Nilsson (ed), Perspectives On Memory Research. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1979.

Tulving, E. "Synergistic ecphory in recall and recognition." Canadian Journal Of Psychology (1982) 36 130-147.

Tulving, E. Elements Of Episodic Memory. Oxford: Oxford University Press, 1983.

Tulving, E. "Subjective organisation and effects of repetition in multi-trial free recall learning." Journal Of Verbal Learning And

Verbal Behaviour (1966) 5 193-197.

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

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

Tulving, E. and Watkins, O.C. "Continuity between recall and recognition." American Journal Of Psychology (1973) 86 739-748

Tversky, B. and Hemenway, K. "Categories of environmental scenes." Cognitive Psychology (1983) 15 121-149.

Underwood, B.J. "Stimulus selection in verbal learning." In C.Cofer (ed), Verbal Behaviour And Learning: Problems And Processes. New York: McGraw-Hill, 1963.

Underwood, B.J. "Attributes of memory." Psychological Review (1969) 76 559-573.

Underwood, B.J. Temporal Codes For Memories: Issues And Problems. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977.

Velten, E. "A laboratory task for the induction of mood states." Behaviour Research And Therapy (1968) 6 473-482.

Warrington, E.K. and Weiskrantz, L. "A study of learning and retention in amnesic patients." Neuropsychologia (1968) 6 283-291.

Warrington, E.K. and Weiskrantz, L. "Amnesic syndrome: consolidation or retrieval." Nature (1970) 228 628-630.

Warrington, E.K. and Weiskrantz, L. "An analysis of short-term and long-term memory deficits in man." In D. Deutsch (ed), The Physiological Basis Of Memory. New York: Academic Press, 1973.

Waters, R.H. "A case of environmental change with an accompanying loss of language." American Journal Of Psychology (1934) 46 336-337.

Watkins, O.C. and Watkins, M.J. "Buildup of proactive inhibition: a cue overload effect." Journal Of Experimental Psychology: Human Learning And Memory (1975) 1 442-453.

Watson, J.B. "Kinesthetic and organic sensations: their role in the reactions of the white rat to the maze." Psychological Monographs (1907) 8.

Weingartner, H. and Fallice, L.A. "Alcohol state-dependent learning in man." Journal Of Nervous And Mental Diseases (1971) 153 395-406.

Weiskrantz, L. "A comparison of hippocampal pathology in man and other animals." In Functions Of The Septo-Hippocampal System.

CIBA Foundation Symposium 58. Oxford: Elsevier, 1978.

Weiss, W. and Margolius, G. "The effect of context stimuli learning and retention." Journal Of Experimental Psychology (1954) 48 318-322.

Welford, A.J. "The 'psychological refractory period' and the timing of high-speed performance: a review and a theory." British Journal Of Psychology (1952) 43 2-19

Welford, A.T. Fundamentals Of Skill. London: Methuen, 1968.

Wickelgren, W.A. "Alcoholic intoxication and memory storage dynamics." Memory & Cognition (1975) 3 385-389.

Wickelgren, W.A. Cognitive Psychology. New Jersey: Prentice-Hall, Englewood-Cliffs, 1979.

Williams, M.D. and Hollan, J.D. "The process of retrieval from very long-term memory." Cognitive Science (1981) 5 87-119.

Winer, B.J. Statistical Principles In Experimental Design. 2nd edition. New York: McGraw-Hill, 1971.

Winocur, G. "The hippocampus and cue utilisation." Physiological Psychology (1980) 8 280-288.

Winocur, G. and Bindra, D. "Effects of additional cues on passive avoidance learning and extinction in rats with hippocampal lesions." Physiology And Behaviour (1976) 17 915-920.

Winocur, G. and Kinsbourne, M. "Contextual cueing as an aid to Korsakoff patients." Neuropschologia (1978) 16 671-682.

Winocur, G. and Olds, J. "Effects of context manipulation on memory and reversal learning in rats with hippocampal lesions." Journal Of Comparative And Physiological Psychology (1978) 92 312-321.

Winograd, E. and Geis, M.F. "Semantic encoding and recognition memory: a test of encoding variability theory." Journal Of Experimental Psychology (1975) 102 1061-1068.

Winograd, T. "On some contested suppositions of generative linguistics about the scientific study of language." Cognition (1977) 5 151-179.

Woodall, W.G. and Folger, J.P. "Encoding specificity and non-verbal cue context. An expansion of episodic memory research." Communication Monographs (1981) 48 39-53.

Zentall, T.R. "Effects of context change on forgetting in rats." Journal Of Experimental Psychology (1970) 80 440-448.

APPENDICES

Appendix A: Environmental Contexts used in experiments.

Appendix B: Listings of stimuli, recognition and filler words.

Appendix C: Computer controlled slide projection system.

Appendix D: Questionnaires.

APPENDIX A

Environmental Contexts Used In Experiments: Descriptions

Room A (37).

This was an oblong room 480cm in length, 210cm in breadth and with a ceiling height of 347.5cm. It was entered through a plain olive green door at the opposite end of the room to that at which the subjects were seated during the time they were in the room. Subjects sat on a black metal framed, black plastic cushioned office chair, with their backs to a window, facing the wall adjacent to the door. Just in front of them was a small black wooden table (55.5cm x 55.5cm x 70cm) upon which stood an Eagle International PRO.M30 microphone and beneath which was a foot-switch (both microphone and foot-switch were absent in experiment 1).

A brown storage radiator (72cm x 25cm x 65cm) occupied the corner 7.5cm from the back wall and 30cm from the wall to the subjects' right. On the radiator stood a Kodak Carousel projector. Halfway along the wall to the subjects' right was a large wooden fireplace which was blocked off. At the far end of this wall was an alcove of shelves which extended from the floor to a height of 240cm. The slides were projected onto the right half of the wall facing the subject. All walls and ceiling were painted white and the skirting, shelves and fireplace were painted dark brown, while the floor was covered in smooth grey carpeting.

An 11.5cm diameter ventilator pipe ran along the length of the room from the top corner of the wall facing the seated subject on their left hand side 46cm from the ceiling and 3cm from the left wall, connecting to the window. As explained in chapter seven, after the completion of experiment 1 the window was boarded over with dark hard wood fixed in a light soft wood frame.

The room was illuminated by two 80 watt white strip lights house in a double fitting, but without a casing cover and positioned in the middle of the ceiling, running along the length of the room. No significant change in light level was recorded; from the position where subjects sat, before and after the window was blocked off. The light level recorded from this position was between 16-18 foot candles. All wires to and from equipment (ie. projector, microphone and foot-switch) were hidden.

Overall the room gave a long, barren and empty impression.

Room B (38).

This measured 350cm in length, 160cm in breadth and had a ceiling height of 347.5cm. A green curtain hanging from a point 22cm below the ceiling divided the length of the room at a point 222cm from the entrance: a plain olive green door, on the top wall.

Subjects sat at a large white formica covered table (137cm x 76cm x 72cm) which was placed with its breadth against the curtain and its length along the wall nearest to the door. Subjects sat on a red plastic chair placed at the middle of the table's length, facing the wall. Under the table at this point was placed a foot switch. Directly in front of the subjects, with its back against the wall was an Electrohome video monitor.

Joining with the green curtain dividing the room was a "square" of the same material (170cm x 132cm) fixed to the same wall beginning at a height of 64cm above that of the table. Behind this curtain hung a microphone. Behind and to the subjects' right in the corner stood a black metal framed, wooden topped stand (72cm x 46cm x 120cm). Another "square" of green curtain material (also 170cm x 132cm) followed the shape of the corner in which the stand was placed. The "square" was hung so that at the bottom it skirted the top of the stand along the corner walls, beginning at the corner of the stand nearest the door and extending 48cm past the length of the stand along the wall behind the subject. On top of the stand stood a Kodak Carousel projector.

During the experiment, illumination was provided by three desk lamps each containing one 60 watt Osram fireglow bulb. Two lamps were placed at each corner of the table against the wall and one lamp was placed on top of the video monitor. Slides were projected to the subjects' right above the lamp on the right corner and below the "square" of curtain. By using white on black slides a low illumination level could be maintained. From where the subject sat the light level recording was 6-8 foot candles.

All walls were painted white, the ceiling olive green and all skirting was painted dark brown, while the floor was covered in smooth grey carpeting. All wires to and from equipment were hidden. The overall impression was of a small, crowded and cramped, red and dark room.

The Apple II computer, interfaces between computer and peripherals, JVC stereo cassette deck and various other pieces of apparatus, were located on the opposite side of the dividing curtain from the subjects' environment.

Room C1 (41).

This measured 235cm in length, 250cm in breadth, had a ceiling height of 347.5cm and was entered from another small room. On the same wall as the entrance: a plain olive green door, a white formica topped table (137cm x 76cm x 72cm), was placed. The table's length extended out toward the parallel wall. Upon the table stood a 3 field tachistoscope, an electronic timer and a variety of electronic switches. On the left of the wall opposite the entrance were three shelves on brackets, cluttered with various pieces of equipment; such as morse switches, electronic timers and large quantities of different coloured wire.

The room was illuminated from the ceiling by two 80 watt strip lights in a double fitting and casing. All walls and ceiling were painted white and all skirting was painted dark brown. On the same wall as the three shelves, but to the right and flush with it, was a white wooden cupboard door.

Subjects were taken into the room round the table and sat on a black plastic chair facing the table with the exit to their right.

Overall the room appeared small, bright and technically cluttered.

Room C2 (36).

This room was entered from a short corridor, again through a plain olive green door. It measured 425cm in length, 250cm in breadth and had a ceiling height of 352cm. To the left of a door on entering, at a distance of 45cm were 3 bracket shelves. The shelves extended from this point along the full length of this wall. All of these shelves were bare except for one bicycle tyre (originally found in the room) which hung from the corner of the top shelf next to the door.

On the left hand wall was a white formic topped table (137cm x 76cm x 72cm). The length of this table was placed against the wall, beginning at a point 140cm from the corner where the shelves met this wall. Above the middle of this table on the left wall at a height of 284cm was a ventilator. A brown storage radiator (72cm x 25cm x 65cm) occupied a space against the skirting board on the wall facing the door and 40cm to the right of the left hand wall. Sitting on top of this storage radiator was an orange traffic cone (also found in the room) 155cm along the right wall from the entrance, was an olive green wooden panelled door. This door opened into another room. Immediately after the doorway (still in the first room) was another formica topped table of similar dimensions to that described above. The length of this table was placed along the right wall.

Subjects sat upon a black plastic seat at the table which was against the left wall, with their backs towards the shelves and entrance. Illumination was from two sets of ceiling lights, each of two 80 watt strip lights housed in a double fitting without a casing cover and running along the breadth of the room. Each fitting was placed at a point 100cm in from its respective parallel wall.

From where the subjects sat a large part of the adjacent room could be seen. This room measured 255cm x 270cm and had a ceiling height of 347.5cm The room contained one white formica topped table of dimensions similar to those described above and two black plastic chairs placed at the mid points of each length of the table. The table was positioned with its length running across the breadth of the room with its far end placed against the left wall at a point beginning 92cm from the wall containing the entrance. One the back wall facing the entrance, at a point 52cm from the right wall and at the same height as the former, was another ventilator. Illumination was from two 80 watt strip lights, housed in a double fitting without a casing cover, placed in the middle of the ceiling and running along the breadth of the room. In both rooms all walls were painted white, ceiling olive green, skirting brown and the floor was covered in grey carpeting. Overall the impression was of a large bright partitioned room.

APPENDIX B

Listings of stimuli, recognition and filler words.

Experiments 1, 2 and 3.

Words on initial presentation.

KING	FILM	SHIP	STONE	CLAY	HOLE
EDGE	CAR	BEAR	CITY	COFFEE	KNIFE
SKY	AIR	CHAIN	MOON	DOOR	KEY
RIVER	WATCH	BALL	GUN	MONEY	BABY
GAS	ROAD	FIELD	HOTEL	STATION	WHEEL
DOCTOR	MOTOR	COURT	RAIN	BANK	ROSE
FLOOR	SQUARE	BAR	TREE	SAW	BOOK
FIRE	ISLAND	NIGHT	KITCHEN	WELL	BOX
SAFE	PICK	HAT	WALL	EYE	BOTTLE
CASE	WORD	WAR	GLASS	BED	BRITAIN
POLICE	BLOOD	FARM	CHAIR	OFFICE	
RECORD	HOUSE	ROOM	MARKET	PHONE	

R-items

TREE	COFFEE
CASE	POLICE
MOON	ROOM
WHEEL	WAR
BALL	EDGE

F-items

BRIDGE	FOOD
SIGNAL	HELL
VALLEY	DUST
RADIO	TELEVISION
MAN	GOD

Experiments 4 and 5.

Words on initial presentation.

KING	FILM	SHIP	STONE	CLAY	HOLE
EDGE	CAR	BEAR	CITY	COFFEE	KNIFE
SKY	AIR	CHAIN	MOON	DOOR	KEY
RIVER	WATCH	BALL	GUN	MONEY	BABY
GAS	ROAD	FIELD	HOTEL	STATION	WHEEL
DOCTOR	MOTOR	COURT	RAIN	BANK	ROSE
FLOOR	SQUARE	BAR	TREE	SAW	BOOK
FIRE	ISLAND	NIGHT	KITCHEN	WELL	BOX
SAFE	PICK	HAT	WALL	EYE	BOTTLE
CASE	WORD	WAR	GLASS	BED	BRITAIN
POLICE	BLOOD	FARM	CHAIR	OFFICE	
RECORD	HOUSE	ROOM	MARKET	PHONE	

R-items

TREE	COFFEE	BOOK
CASE	POLICE	CAR
MOON	ROOM	EYE
WHEEL	WAR	HOUSE
BALL	EDGE	ISLAND

F-items

BRIDGE	FOOD	FORM
SIGNAL	HELL	BACK
VALLEY	DUST	LIGHT
RADIO	TELEVISION	POINT
MAN	GOD	WORLD

Experiment 6, conditions L1.5, L3 and H3.

Words on initial presentation.

KING	FILM	ISLAND	WAR	CHAIR	WHEEL
EDGE	CAR	WORD	FARM	MARKET	BABY
MAN	AIR	VALLEY	GOD	BED	NIGHT
RIVER	WATCH	BALL	GUN	KNIFE	KITCHEN
GAS	SAFE	BRIDGE	HOTEL	STATION	
DOCTOR	MOTOR	COURT	SQUARE	KEY	
FOOD	CASE	BAR	BRITAIN	SAW	
WELL					

Filler items on recognition test.

PICK	PHONE	POLICE	BOTTLE	STONE	FIRE
TELEVISION	EYE	MONEY	DOOR	BOOK	OFFICE
SKY	ROOM	SIGNAL	RADIO	HOLE	DUST
TREE	ROSE	HOUSE	FLOOR	CLAY	SHIP
FIELD	RAIN	MOON	GLASS	BLOOD	
BOX	NIGHT	HAT	RECORD	HELL	
WALL	CHAIN	ROAD	BEAR	COFFEE	
COLLEGE					

Experiment 6, condition H6.

Words on initial presentation.

KING	CAR	BALL	SQUARE	WHEEL	STATE
EDGE	AIR	BRIDGE	BRITAIN	BABY	HALL
MAN	WATCH	COURT	CHAIR	KITCHEN	LEAD
RIVER	SAFE	BAR	MARKET	KNIGHT	SHOW
GAS	MOTOR	WAR	BED	WINE	BOTTOM
DOCTOR	CASE	FARM	KNIFE	LIST	FAT
FOOD	ISLAND	GOD	STATION	CLUB	LINE
WELL	WORD	GUN	KEY	METAL	ANGLE
FILM	VALLEY	HOTEL	SAW	MUSIC	DATE

Filler items on recognition test.

PICK	EYE	HOUSE	RECORD	FIRE	HILL
TELEVISION	ROOM	MOON	BEAR	OFFICE	COUPLE
SKY	ROSE	HAT	STONE	DUST	PARTY
TREE	RAIN	ROAD	BOOK	SHIP	SHAPE
FIELD	NIGHT	BOTTLE	HOLE	DINNER	FIGURE
BOX	CHAIN	DOOR	CLAY	POOL	LIGHT
WALL	POLICE	RADIO	BLOOD	SEASON	BOARD
COLLEGE	MONEY	FLOOR	HELL	POINT	DANCE
PHONE	SIGNAL	GLASS	COFFEE	LAND	HANDLE

APPENDIX C

Computer Controlled Slide Projection System

Each Kodak Carousel projector was operated individually by the one Apple II computer (ie. at no time were the two projectors linked and operated simultaneously).

After receiving an input from a foot switch, the Apple initiated the slide presentation by sending a signal via a six pole socket to the forward slide change. Reverse slide change was not used.

A photo-cell placed in the Carousel projector lamphouse, between the light shutter and slide gate, signalled the onset of slide presentation.

The Carousel projector slide change procedure, after the initiation signal, involves the light shutter descending (by gravity feed) to block the light to the side, the slide being lifted into the magazine, the magazine turning round, the next slide dropping into the slide gate (again by gravity feed), and the shutter lifting allowing the light to pass through and project the slide. The whole operation to change from presenting one slide to another took approximately 1 to 1.5 seconds.

Due to the gravity feed method of light shutter operation, a variable delay of 100 to 170 milliseconds (average 150ms.),existed between the signal to change a slide and the completion of the light's shutter descent (ie. slide offset). To reduce this variation the signal from the Apple was programmed to occur 150 milliseconds prior to the end of the chosen slide presentation time. This had the effect of reducing the variation from the chosen slide presentation time to between -50 and +20 milliseconds.

APPENDIX D

QUESTIONNAIRES

Questionnaire employed in experiment two.

Condition _________

Name _______________________

As the words were presented to you, what did you do?

When you were asked to recall the words, what did you do to remember them?

On a scale from 1 - 100, how similar do you think that the two rooms are?

N.B. 1 = totally different

100 = identical

YOUR RATING ______

Questionnaire employed in experiment three.

Condition ________

Name ____________________

As the words were presented to you, what did you do to learn them?

When you were asked to recall the words, what did you do to remember them?

On a scale from 1 - 100, how similar do you think that the two rooms are?

N.B. 1 = totally different

100 = identical

YOUR RATING ______

Questionnaire employed in experiment four.

Condition _________ Name _______________________

1/ As the words were presented to you, what did you do?

2/ Did you make any attempt to learn any of the words as they were presented? (tick the appropriate answer) YES NO

3/ Over the course of the recall time did you : (tick answer)
a) maintain the suggested retrieval strategy
b) have a break at any time or try to recall words in a different way

4/ What did you do if you stopped or changed? and did this help you to remember any words?

5/ Did you remember any of the words due to their association with any others? (tick answer) YES NO

6/ At any time did you think of, or imagine the situation where you were being presented with the words? (tick answer) YES NO

7/ If you answered yes to question (6), please explain exactly what you thought, or imagined.

8/ On a scale from 1 - 100, how similar do you think that the two rooms are? (If you have only been in one room ignore this question)

N.B. 1 = totally different and 100 = identical YOUR RATING _______

Questionnaire employed in experiment seven.

Condition _________ Name _______________________

1/ Did you imagine yourself doing the task yesterday in order to help you do it today? YES NO (please tick appropriate answer)

2/ If you did imagine doing the task yesterday, could you explain exactly how you did so (if you answered NO to the previous question, please move on to the next),

3/ If you did not imagine yourself doing the task yesterday, how did you help yourself to remember the way to place the items today?

4/ Although you may have already stated in a previous answer, please say if you ever imagined, or tried to imagine the room you were in yesterday.

5/ If you did imagine being in the room yesterday, did you feel that it helped you remember at all?

6/ On a scale from 1 - 100, how similar do you think that the two rooms are? (If you have only been in one room ignore this question)

N.B. 1 = totally different and 100 = identical YOUR RATING _______