STUDIES OF RECALL AND STORAGE

IN SHORT-TERM MEMORY

by

Michael John Anthony Howe

Thesis submitted at the
University of Sheffield
for the degree of Doctor
of Philosophy. 1966.

ABSTRACT

Experiments were designed to measure the effects upon short-term retention of verbal material of certain variables relating to (i) serial organization and (ii) the activity of recalling.

In Section 1, some effects of sequential redundancy upon short-term verbal memory are described. It is shown that retention is positively related to the similarity in structure of the material to language. A series of experiments was carried out, using sequences of both word and letter units, in order to provide information about the stage or stages within a memory task at which sequential redundancy is directly influential. It has been suggested that sequential organization has its main influence at the time of recall. The present results indicate that this is incorrect, and show that memory is already affected by redundancy in a sequence at a stage prior to the recall of verbal items.

Section 2 is concerned more directly with effects of the activity of recalling verbal material. An experiment is described which shows that accuracy of reproduction may be related to order of recall, the first items to be recalled in a short-term memory task being more accurately reproduced, on the whole, than those recalled later. Consolidation in the storage of items is found to be related to the order of presentation. Some experiments are described which aimed to explain these results, and it is concluded that both rehearsal and storage time contribute to consolidation in short-term memory. The result of a further experiment confirms the observation that verbal items are sometimes most accurately reproduced when recalled in an order different from that in which they were presented.

ACKNOWLEDGEMENTS

The writer is indebted to the many people whose help has made possible the research described in this thesis. Thanks are due to numerous members of the Department of Psychology, University of Sheffield, for generous clerical and technical assistance, and for useful and stimulating advice, discussion and encouragement, without which the work described here could not have been produced.

Special mention should be made of the large number of subjects who freely gave up their time for experimental sessions, which were often arduous and dreary.

Mr. J. K. Clarkson kindly and patiently gave invaluable advice relating to the statistical treatment of the data.

The writer is particularly grateful to his supervisor, Professor H. Kay, who provided much help, encouragement and constructive criticism, from which the present research has benefitted considerably.

The work was carried out while the writer held a Research Studentship, granted initially by the Department of Scientific and Industrial Research, and later supported by the Scientific Research Council.

INDEX

LIST OF TABLES

LIST OF FIGURES

SECTION 1

AN INVESTIGATION INTO SOME EFFECTS OF REDUNDANCY IN HUMAN LANGUAGE UPON RECALL

CHAPTER 1.

WORD SEQUENCES IN LANGUAGE

1.1. Introduction.

Human activity does not usually occur in isolated units, but in organized, temporally integrated patterns of behaviour. In attempting to explain complex behaviour such as "the coordination of leg movements in insects, the song of birds, the control of trotting and pacing in a gaited horse, the rat running a maze, the architect designing a house, and the carpenter sawing a board" (Lashley, 1951), it is necessary to understand something of how the behaviour is sequentially organized. Among human skills, sequential organization takes a striking and complex form in the use of language.

Much of the scope of human language is due to the fact that a finite number of units can be variously combined to produce an inexhaustible variety of meanings. At the same time, certain rules or constraints which can be labelled "grammar" dictate which of the possible sequences of words are acceptable in a language, and language sequences are thereby limited to those which can be dealt with by mechanisms available to the human user. In other words, certain rules ensure that language as emitted is meaningful to those who receive it. If the rules of sequential organization are followed, the language pattern is immediately perceived as acceptable. As James (1890, p.263) remarked:

> "If words (do) belong to the same vocabulary, and if the grammatical structure is correct, sentences with absolutely no meaning may be uttered in good faith and pass unchallenged. Discourses at prayer-meetings, reshuffling the same collection of cant phrases, and the whole genus of penny-a-line-isms and newspaper-reporter's flourishes give illustrations of this. 'The birds filled the tree-tops with their morning song, making the air moist, cool, and pleasant,' is a sentence I remember reading once in a report of some athletic exercises in Jerome Park. It was probably written unconsciously by the hurried reporter, and read uncritically by many readers......Nonsense in grammatical form sounds half rational; sense with grammatical form upset sounds nonsensical."

Sequential organization in language is the subject of Chapters 1-6, which form Section 1 of the present study. An experimental approach will be followed, taking as a starting-point the demonstrations by Miller and Selfridge (1950) and by subsequent workers that there is a clear relationship between the degree of sequential constraint in a list of words and the ease with which they can be processed in a human language task. Explanation of organization effects in language is clearly desirable, and has not yet seriously been attempted. To understand sequential organization in language is important in its own right, and it is probably pertinent to a more general understanding of activity in the brain. To quote Lashley (1951) again: "the problems raised by the organization of language seem to me to be characteristic of almost all other cerebral activity". Attempts to explain phenomena often take the form of providing more adequate description than has previously been available; the aim underlying the present investigation is to provide finer description by determining the location within verbal tasks of sequential redundancy effects which partly determine global task performance.

The remainder of Chapter 1 summarizes available experimental evidence concerning effects on verbal task performance of sequential organization in word lists. Chapters 2 and 3 attempt preliminary explanation, in the form of closer description of these effects. Chapters 4, 5 and 6 are concerned with the same problem, but take letters rather than words as a basic unit of language analysis. Chapter 7 concludes the present section. The subsequent chapters, which form Section 2, describe some effects of the activity of recalling verbal sequences, and will be more fully introduced at a later stage.

The mere fact that we are able to use language indicates a degree of knowledge of its structure, but, until very recently, little effort has been made towards understanding the mechanisms required for organized verbal communication. Psychologists in the future may well think it surprising that in the first half of the twentieth century a vast amount of experimental effort was directed at understanding verbal behaviour, but that in most of the work the fact that such behaviour typically occurs in sequences longer than one word was hardly considered.

Clearly, experiments in which measurable attributes of single words are related to performance ratings can often be valuable, but to examine words only in situations which bear little resemblance to the contexts in which they are normally used may be a little short-sighted.

Prior to World War Two, psychologists investigating verbal learning and memory placed relatively little emphasis on serial organization. Experimental approaches have only recently been made to an understanding of sequential organization in behaviour in general, and in language in particular. Certainly Bartlett (1932) brought the phenomena of continuous verbal discourse to the attention of psychologists, but, although many of his insights are invaluable to anyone interested in sequential organization in language, the serial organization process itself did not primarily engage his interest.

1.2. Sequential redundancy in language.

The study of sequential organization has a short history in experimental psychology. The sheer difficulty and complexity of the problems involved may have been a reason for psychologists' delay in paying attention to this field. Diebold (1965), for instance, writes that "the problems relating to sequential organization are among the knottiest of all those awaiting resolution". An important initial difficulty is to find an adequate quantitative index of degree of sequential organization. Scientific progress requires adequate measurement of the phenomena under investigation. In studying verbal learning it is easy to measure single-word attributes, for example frequency and pronunciability, but it is more difficult to obtain measures of words in organized sequences, as they actually occur in language. How can sequential organization be measured? One answer to this question was provided by C.E.Shannon. Shannon was interested in the question of how much knowledge about a sequence is provided by presenting part of the sequence. In his work on estimation of redundancy in the English language Shannon suggested a number of techniques which could be used to obtain measures of sequential organization in language. One of these (Shannon, 1951) requires subjects to guess at the identity of letters in a prose sequence. He found that accuracy of guessing improves with the amount of previous context available. Subjects use their knowledge of language structure, and the accuracy of their predictions is related to the amount of context provided. A similar method was used by Miller and Selfridge (1950), taking words rather than letters as basic units, to obtain sequences of varying orders of "Approximation to English".

Using the approximation to English method, it is fairly easy to construct word lists whose structure is of a given degree of similarity to English. For instance, to obtain a second order sequence Miller and Selfridge would present a common word to a person who is instructed to use the word in a sentence. A note is made of the word the person provides directly after the one given him, and the recorded word is presented to a second person, who is asked to use that word in a sentence. The procedure is repeated until a list of the required length has been obtained.

To construct lists of higher orders of approximation, sequences of more than one word are shown, and a person is asked to complete the sentence. Each successive volunteeris shown a sequence containing an equal number of words, but with each new person the first word shown to the previous person is dropped out, and the length of the sequence kept constant be including the first word provided. Thus a word in a sequence of nth order approximation is constrained by the requirement of having to make sense in the context supplied by the previous n-1 words. Miller and Selfridge obtained first order approximations to English by scrambling the words used in the higher order sequences. In this way they produced randomly ordered lists of words whose relative frequencies reflect their frequency of occurrence in written English. The zero order words were drawn at random from the 30,000 English words in the Thorndike and Lorge (1944) count.

The approximation to English measure has been extensively used in investigations of structured language. It enables the production of lists with varying degrees of similarity to English in sequential structure. The obvious problem in attempting to measure sequential organisation in word passages is that the technical difficulties attached to measuring the frequency of sequences longer than one word, in a sample large enough for the obtained values to be reliable guides to frequency of the sequences in the language as a whole, would be immense. The approximation to English method avoids this problem by enabling the user to construct lists differing to known extents from English in sequential organization. Another problem, that of providing a base-line for the index of sequential organization, is solved by using random words and written English as the lower and upper limits. In short, there exists a simple way of producing sequences of ordered similarity to the sequential structure of language. An invaluable prerequisite for the study of language structure has been provided.

1.3. Effects of sequential redundancy in word lists.

To demonstrate the importance in a verbal performance task of sequential structure, produced by the approximation to English method, Miller and Selfridge (1950) carried out a simple recall experiment. Word lists were made at all orders of approximations to English between zero and seventh. At each order the authors made four lists of different lengths; ten, twenty, thirty, and fifty words, and they added four prose extracts, one at each of the above lengths, making a set of 32 separate lists. Two such sets of 32 lists were constructed. Each list was read aloud to subjects, who were instructed to listen until it was finished and then immediately write down what they could remember, in as near as possible to the correct word order. In scoring, the number of correct words were counted, irrespective of order. Their results are shown in figure 1.

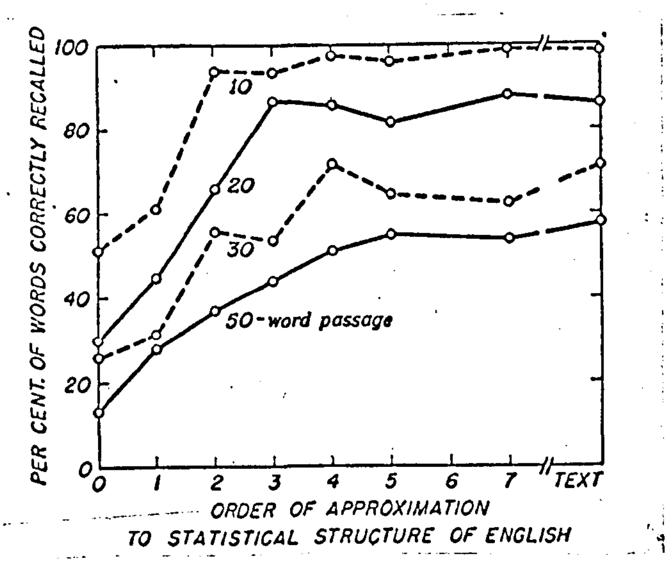

Figure 1. Results obtained by Miller and Selfridge (1950). Percentage of words from lists of different lengths that were correctly recalled at various orders of approximation to English. (From Miller and Selfridge, 1950.)

Clearly recall is strongly related to order of approximation to English, at least at the lower levels of approximation.

The demonstration by Miller and Selfridge (1950) of a correlation between recall and sequential constraint has been substantiated in a number of studies. There follows a brief survey of the experimental findings. Deese and Kaufmann (1957) and Richardson and Voss (1960) found that immediate recall of words increased with an increase in order of approximation. Marks and Jack (1952) obtained results similar to those to Miller and Selfridge, but observed more improvement than did Miller and Selfridge at the higher orders of approximation, in an immediate recall task. They suggested that the discrepancy was due to the method of scoring, since Marks and Jack scored subjects' longest correct sequences, whereas Miller and Selfridge scored words recalled, regardless of order. Strong support for this suggestion was provided by the results obtained by Coleman (1963), who measured correct recall in sequences of from one to 17 words. Advantages for the higher order approximations were greater when recall was scored in longer sequences. Degree of approximation to English was found to affect recall in a sequential memory task devised by Lloyd and Johnston (1963), in which subjects continually received information, and at unpredictable moments were required to recall some of it. In a learning task, Sharp (1958) presented passages five times, with immediate recall following each presentation. His results were in accord with those of Miller and Selfridge, as were those obtained in a six-trial learning task given by Tulving and Patkau (1962). Simpson (1965) observed that when subjects learned lists of 30 words by the method of serial anticipation, increasing the approximation of the list to sentence word-order led to faster learning. Postman and Adams (1960) found that for incidental as well as intentional learners retention increased as a function of degree of approximation. All the above authors followed Miller and Selfridge in using approximation to English as the measure of sequential redundancy, but a supporting result was obtained when another index of redundancy was used by Rubenstein and Aborn (1954). They used lists of nonsense syllables in which various rules governed the sequential order. The number of items recalled was found to increase with the degree of organization in the material.

Approximation to English has been found to affect performance in a number of other tasks requiring verbal material to be retained. For instance, Sumby and Pollack (1954) instructed subjects to reproduce material by writing with the least possible number of glances. They found that a glance was necessary for every three words in zero order word sequences, but only one for every ten words in English prose. Hogan (1961) found that when subjects were copying messages from a nine-word language with varying levels of redundancy (3.5%–89.6%) the number of times subjects looked at the messages declined as redundancy increased. A result similar to this was obtained by Lawson (1961), whose subjects read aloud word sequences. at intervals, a light was switched off so that the words could not be seen and a subject had to report as many words as he had seen but not already pronounced. Measured in this way, the average "Eye-voice Span" was found to vary between 3.52 words for second order approximations to English and at 4.67 words for twelfth order sequences. Related to the eye-voice span is reading rate, which has been investigated by Pierce and Karlin (1957). They found that English prose sequences were read at an average rate of 4.5 words per second, compared with 3.2 words per second for "Scrambled Prose" (first order approximation to English). Intelligibility of speech is related to sequential constraint. Miller, Heise & Lichten (1951) compared detection of key words in sentences and in isolation under varying levels of white noise. When the signal-to-noise ratio was such that about 80% of the words could be detected in isolation, over 95% of the same words were correctly identified if presented as part of a sentence.

O'Neill (1957) and Rubenstein & Pollack (1953) have confirmed and further quantified this finding. An interesting result obtained by Treisman (1964) is that when identical messages are played to each ear but with a time lag between them, and the subject has to attend to one message but ignore the other, the duration of lag at which the messages to the separate ears are noticed as being identical is directly related to their approximation to English. Redundancy of word sequences also affects the efficiency with which they can be reported back or "shadowed" by subjects (Moray and Taylor, 1958) and the ease of translation between French and English (Treisman, 1965). Relationships between hesitations in speech and encoder uncertainty have been examined by Dr. Freda Goldman-Eisler. A typical finding (Goldman-Eisler, 1958) is that the first lexical word subsequent to an unfilled pause is less predictable than surrounding words. Tannenbaum, Williams & Hillier (1965) have provided supporting results, but their findings did not support Goldman-Eisler's additional observation that the word preceding an unfilled pause tends to be even more predictable than when it occurs in other fluent contexts.

Clearly, performance in some verbal tasks is affected by sequential redundancy within the task material. The approximation to English method has been extremely useful, although it does not measure <u>all</u> aspects of similarity to written English (see, e.g. Epstein, 1961). Coleman (1963) has noted that the technique used to generate approximations to English produces higher order approximations which deviate from English grammatically, mainly because their is no provision for punctuation. Because of this, one would expect the recall curve for higher orders to flatten out at a level substantially below that for English prose. However, it does not, provided that recall is scored in long sequences. Coleman suggests that this is probably because higher order constraints package the elements into more familiar order than typically occur in English prose.

SUMMARY OF CHAPTER 1.

Sequential organization is required in many sorts of complex behaviour, and is essential in human language. Some questions of measurement are discussed. It is proposed to attempt close description of the effects of sequential constraints in language, as a step towards understanding the nature of the mechanisms by which redundancy influences performance. A brief survey of experimental findings about word-sequence redundancy effects is provided.

CHAPTER 2.

EXPLANATION OF REDUNDANCY EFFECTS IN WORD SEQUENCES

2.1. Explaining redundancy effects.

None of the authors of the fairly numerous studies mentioned in the previous chapter has made a very rigorous attempt to explain how sequential redundancy affects performance. It is useful to be able to measure the effects on performance of varying levels of approximation to English, but there is a need for more precise description of how sequential redundancy produces these effects, if one is to understand the mechanism by which the brain can transmit a familiar sequence more easily than the scrambled elements of that sequence. This is the sort of problem to which Lashley (1951) drew attention, although in a wider context than that of verbal language, when he discussed the inadequacy of associative explanations of serial behaviour. Certainly the problem of serial order, which Lashley defined (1951, p.122) as "the existance of generalised schemata of action which determine the existence of specific acts, which in themselves or in their associations seem to have no temporal valence" is now of undisputed importance in the attempt to understand how the brain works. The demonstrations of a relationship between sequential redundancy in verbal material and the efficiency with which humans can deal with that material indicate that quantitative measures may be used to open the general problem to experimental investigation. Further progress requires description of sequential order effects beyond that which can be provided by relating redundancy to performance in rather broadly defined tasks, such as tests on immediate memory, requiring a number of operations to be performed. The present chapter aims to provide certain parts of the necessary close description of effects of redundancy in word sequences, and the subsequent chapters in Section 1 constitute a similar attempt but with reference to redundancy in sequences of letters rather than words.

In view of the achievement of Miller and Selfridge (1950) in providing a clear demonstration of a relationship between measures of sequential constraint and performance in verbal tasks it is perhaps churlish to criticise their attempts at explanation of the results, especially since later workers using the approximation to English measure have had very little to say of additional explanatory value. However, because of the latter fact the remarks of Miller and Selfridge serve as a base-line for attempts to explain the observed effects of sequential redundancy in word lists. Miller and Selfridge suggest that the significant distinction between random words and approximations to English,

> "is not between meaning and nonsense, but between materials that utilise previous learning and permit positive transfer and materials that do not. If the nonsense preserves the short range associations of the English language that are so familiar to us, the nonsense is easy to learn.... nonsense materials that retain these short range associations are also easy to learn."

These remarks raise two separate problems. First, are Miller and Selfridge correct? The statement that only short range associations are important is supported by the fact that in their experiment very little increase in efficiency of recall occurred at orders of approximation to English beyond the third. In fact, as we have already mentioned, this result occurs solely when recall is scored in terms of single words, irrespective of order. When measures of retention are used which take order into account (Sharp, 1952; Coleman, 1963), longer range constraints are found to have significant effects. Thereford we can reasonably say that Miller and Selfridge are wrong in assuming only short range associations to be important for immediate recall. Even so, the fact that curves describing recall of word lists as a function of approximation to English are usually negatively accelerated suggests that short range constaints have greater influence than more distant ones. Secondly, Miller and Selfridge do not state precisely what they mean by "short range associations" in the context of their remarks. The phrase could refer to specific sequential associations between words and word sequences, as when "a stitch in ...." evokes the word "time". On the other hand the phrase could be taken to refer to certain <u>rules</u>. For instance, a person who understands English knows that the word immediately succeeding the sequence "the boy kicks the ...." is likely to be a noun or adjective, and unlikely to be a pronoun or verb. Here, a sequence of words does not evoke a particular succeeding word, but provides information by putting constraints on the nature of that word. This sort of rule provision is a common function of a structured word sequence; the function of suggesting specifically associated words is less usual, and its efficiency depends on the particular receiver as well as the sender of the sequence. For instance, most English speakers would think that the word immediately after the sequence "and miles to go before I....." is probably a verb, and to a minority the passage would also suggest a particular succeeding verb.

At this stage it is useful to consider what sort of explanation of the observed effects is required. The remarks made by Miller and Selfridge, irrespective of any questions concerning their correctness and possible ambiguity, do not seem to be intended as more than a fairly superficial kind of explanation. For instance, no account is attempted of how familiarity with a word sequence, irrespective of familiarity with the individual words, facilitates human performance. How could a machine be designed to profit from the sort of sequential redundancy that occurs in language? Lashley (1951) referring to problems of sequential organization, remarked that an important step would be to define more precisely what we are trying to explain. He said,

> "It is possible to designate, that is to point to specific examples of, the phenomena of the syntax of movement that require explanation, although those phenomena cannot be clearly defined. A real definition would be a long step toward solution of the problem. There are at least three sorts of events to be accounted for. First, the activation of the expressive elements (the individual words or adaptive acts) which do not contain the temporal relations. Second, the determining tendency, the set, or idea. This masquerades under many names in contemporary Psychology, but is, in every case, an inference from the restriction of behaviour within

> definite limits. Third, the syntax of the act which can be described as an habitual order or mode of relating the expressive elements; a generalised pattern or schema of integration which may be imposed upon a wide range and a wide variety of specific acts. This is the essential problem of serial order; the existence of generalised schemata of action which determine the sequence of specific acts, acts which in themselves or in their associations seem to have no temporal valence."

Lashley's own approach is largely physiological; his paper contains speculations about neural mechanisms underlying temporal integrations. A present aim of psychologists is to use their skills to describe as precisely as possible the behavioural phenomena under consideration. The present author considers that, in one important way, description of the effects of sequential organisation has so far been very vague. The vagueness lies in that the tasks in which sequential organisation has been seen to be an important variable have generally contained a number of ill-defined stages, and it is not possible to state at what point or points in the total performance this variable is effective. For instance, in the Miller and Selfridge experiment, it is found that the number of words correctly recalled is related to the order of approximation of the lists presented. If this evidence is used to answer the question "what are the effects of sequential organisation?", it is not possible to give a more precise answer than the bald statement that it influences recall scores. The subject's task in this experiment is to carry out more than one operation. He has to perceive the material, code it in a form suitable for storing, retain it over a period of time, and finally retrieve and reproduce it. Of course, these verbs may not coincide with the stages that the nervous system of the human operator uses in this task. For example, perceiving the material and coding it into a form in which it can be stored may be one and the same process at the physiological level. This difficult was appreciated by Joos (1950) who wrote that:

> "We feel that our descriptive statements fit actual speech behaviour, but we have no right to claim that they are 'correct' in the sense that they fit the neural events in the brains of the speaker and listener."

Nevertheless, there ~~are~~ is more than one stage~~s~~ in the task, irrespective of the degree of correspondence between the labels we may provide for them and the actual functions of the human brain. The implication to be drawn for results relating recall to degree of sequential constaint is that there is no evidence to show at which point or points within a given task the redundancy of the sequence is a limiting factor. It is not possible to say which of the operations comprising the total task are directly affected by this variable. This type of problem is common in investigations of serial behaviour, but for understanding the effects of sequential redundancy on verbal behaviour, precise specification of the operations or processes which redundancy directly influences is an especially important prerequisite. The achievement of such specification is made difficult by the impossibility of ensuring that the stages which are put forward in description of a task correspond exactly with what the brain does. However, it may be possible to divide the task into stages which would seem roughly to correspond to operations in its performance, and this sort of division is customary in experimental psychology.

The central aim of the experiments described in this section is simply to determine whether the effects on performance of sequential redundancy in verbal memory tasks are due to factors operating before or after recall is attempted, or both. This may appear to be a very modest step towards specifying the stages within a total task at which sequential organisation is directly important. Also, the answer may seem obvious, since one naturally tends to think that any difference in recall between different sorts of material in tests of memory demonstrates a difference in the way that the materials are stored. However, Deese (1961) has suggested, with considerable experimental evidence, that this is not necessarily the case, and so the issue is a live one.

2.2. <u>Deese's hypothesis</u>.

At the present point in the attempt to understand the mechanisms of sequential organisation it is necessary to define the stages within a total task at which the variable has direct influence. This assumption is implicit in the question for which Deese (1961) provides an answer. Deese's position is that increases in the order of approximation influence not so much the amount retained, as the subject's ability to reconstruct the material, essentially by guessing, on the basis of his knowledge of the characteristics of the language. He states (1961):

> "The recall of sequentially contingent material is in part at least a constructive process. Something is manufactured by the individual during the recalling, and the raw material for his manufacturing comes from his appreciation of the language after which the sequential pattern was modelled."

He further says:

> "It seems likely that the number of items in free recall of such sequentially structured material goes on increasing only as long as the ability to reconstruct the passage from a few elements goes on increasing. Note that I am implying that individuals actually do construct material (guess) during recall."

Both extracts express the belief that construction at recall is a major determinant of the amount recalled from sequentially structured material. By including the phrase "in part at least" in the first passage Deese avoids precluding storage differences as a possible additional influence on recall of the material. However, in the second extract Deese is clearly suggesting (with the restraint "it seems likely") that construction is all important. If Deese is correct, recall is related to sequential redundancy in language not because humans make use of past experience when coding information for retention, but because they can make use of knowledge of the language at a later stage in the task, to "reconstruct" the original passage from the elements of it which have been successfully stored. This is, in a narrow sense, analagous to Bartlett's (1932) "Construction" which is said to take place when meaningful material is being remembered. Gomulicki (1956) has suggested that when meaningful verbal passages have to be remembered, a process of abstraction occurs whereby the items which are of least importance to the meaning of the passage are eliminated, to bring the content within the limits of memory. The abstraction occurs at the time of perception, concurrently with the understanding of the passage. A similar view has been expressed by Kay (1955) who suggests that subjects make analyses of word passages at the perceptual stage, resulting in "a comprehensive digest of the material". In one sense, Deese's hypothesis is consistent with this position, since if Gomulicki is correct, a major task for subjects at the time of recall would be to construct the verbatim passage out of the retained elements. If subjects make use of their knowledge of language structure at this stage, passages of high orders of approximation would clearly benefit more than less constrained sequences.

Deese does not state very precisely what he means by the reconstruction process which, he claims, occurs at recall. Certainly, if a subject is using his knowledge of the language to guess at a word which has not been retained, the probability of the guess being correct will be directly related to the similarity to English of the structure of the word sequence. Speculation can suggest a number of possible recall strategies and combinations of strategies which might lead to organized sequences being better reproduced than random lists. For instance a subject might have an incomplete record of a word which is sufficient for recognition, but not for free recall of that item. If the subject uses his knowledge of language structure to make guesses at the time of recall, and checks those guesses against recognition memory, his degree of success will be related to the similarity to English of the sequence, since this would reduce the number of guesses required to reach the correct, and therefore recognised, word item. However, since it has not yet been established at what stage or stages in the verbal task sequential organization is directly effective, priority is due to this more elementary question. Are redundant sequences better recalled than random sequences because of reconstruction which occurs at the time of recall, as Deese believes, or are they better retained because of factors which are effective prior to recall? The remainder of this chapter will describe and evaluate the evidence which Deese (1961) produces to support his hypothesis, and other evidence relevant to the hypothesis.

2.3. Evidence relating to Deese's hypothesis.

The evidence which Deese puts forward for construction at recall comes from two studies in which he measured subjects' performance at a task requiring them to guess words deleted from passages at different orders of approximation to English. In one investigation Deese measured the agreement between individuals ~~and~~ the words supplied for given deletions at nine orders of approximation. At each order of approximation there were ten 50-word lists each containing five deletions. An index of agreement between subjects was used, by which a score of 100 was given if all subjects supplied the same word for a particular blank, and a score of 0 indicated that each subject had guessed a different word. Deese found that the average agreement score rose from 0 at zero order approximation to English, to 4.1 at first order approximation, and up to 18.4 by the fifth order of approximation. There was no further rise in the index of agreement for higher orders of approximation to English. Thuse increasing sequential constraint did lead to more agreement in guessing words, but even in the most favourable condition agreement was far from perfect. Deese also measured short-term recall of the lists as a function of order of approximation. He measured the number of single words correctly recalled, irrespective of their order, and found that scores on this index also increased with order of approximation, but again only up to about the fifth order. Table 1 summarises the results.

	Order of Approximation							
	0	1	2	3	4	5	6	8
Index of Agreement	-	4.1	5.8	10.6	13.0	18.4	18.5	17.8
Mean Recall	7.9	12.7	17.8	20.2	21.8	26.7	21.1	23.8

Table 1. Results obtained by Deese (1961). Inter-subject agreement in supplying deleted words from 50-word lists at different orders of approximation to English. A value of 0 represents no agreement, 100 denotes perfect agreement. Recall scores for the same lists (number of words correct out of 50) are also provided. Adapted from Deese (1961, p.25).

In the second investigation which Deese presents as evidence for his hypothesis, he examined subjects' ability to reconstruct passages at various approximations to English, when only a few critical words had been left. As might be expected, the results were influenced less by the order of approximation of a passage than by which particular words were left, but it was found that (Deese 1961 p.26):

> "when the remaining words preserve a good deal of the associative structure of the material (determined, unfortunately, thus far by content analysis), there is a regression between free recall and number of items correctly reconstructed by this technique....... the regression is linear. This is also to say, the number of items reconstructed is curvilinear against orders of approximation."

Deese's evidence can be summarised by saying that both the level of agreement between subjects' guesses at deleted words and the accuracy of subjects' attempts to reconstruct mutilated passages are related to order of approximation to English of the word sequences being used, as are free recall scores for words in the passages. On the basis of this correlation between guessing and recall measures he suggests that the number of items recalled goes on increasing only as long as the ability to reconstruct the passage goes on increasing, and he concludes that subjects actually do construct material during recall. The implied assumption that an observed correlation indicates a simple causal relationship may not be entirely justified, and there are points arising from Deese's results which are not too clear. Regarding the first investigation, it would be useful to know how much a particular value on the agreement index might be expected to facilitate correct construction. For instance we know that a rise from 4.1 to 5.8 (where the upper ceiling is 100) on the agreement index between first and second order approximations is accompanied by a rise from 12.7 to 17.8 in mean recall. It seems unlikely that the cause of so small an increase in guessing agreement would alone lead to the observed recall difference. Unfortunately, it is not possible to test whether this impression is correct since the operations used to calculate the measure of agreement are not precisely defined. The evidence from Deese's second investigation is weakened by the necessity to use a subjective method (content analysis) in determining the sample of the total lists for which a relationship was observed between free recall and number of items correctly reconstructed. Also, no numerical values are given for the results of this investigation. Deese's results show that there is a correlation between the success of guessing behaviour and recall in sequentially constrained lists. However it is not possible from his data to obtain an exact value for this correlation, and it cannot be determined whether or not the correlation indicates a direct causal relationship. The results which Deese puts forward are consistent with his hypothesis, but do not provide strong support for it.

Although there has been little experimental work designed to investigate why sequential organisation affects performance in verbal tasks, some empirical evidence relevant to Deese' hypothesis is available. Apart from Deese's own studies results are available from some other experiments in which subjects have been required to guess at deleted items in word sequences. For instance, in an experiment by Morrison and Black (1957) 1-6 words were deleted from sentences of 11-13 words, and subjects tried to fill in the missing items. When three words (25% of the sentence) were omitted, their position being indicated, 39% of the missing words were correctly predicted. Aborn, Rubenstein and Sterling (1959) deleted just one word from sentences of 6, 11, and 25 words, but did not tell subjects the position of the word deleted. Position and grammatical class of the deleted word were significant factors, as was the length of the sentence. Overall, the average probability of a word being guessed correctly was 47%. The average list length was 11 words. Both these experimental results suggest that word-guessing can be fairly successful, at least in prose passages. However, it is not possible to say to what extent these results can be generalised to situations in which the sequential structure of words only approximates to that of English.

The results of an experiment by Coleman (1962) show that subjects do make use of their knowledge of a language structure when they have forgotten the correct order of words in a memory task. His experimental method was to present in scrambled order the words constituting an English sentence. A subject was allowed to study the list and was then given a stack of cards, which contained the original words, printed one per card, in a different random order. The subject was told to re-arrange the words into the order in which he had originally seen them. The subject's ordering of the words was typed and presented to a second subject, who also studied them and then attempted to re-arrange a pack containing the words into the order in which he had first seen them. This second attempt was also typed, and shown to a third subject, and so on, until sixteen orderings had been produced. It was found that with each new ordering of the words the arrangement came closer to the structure of English. In fact, this tended to make the task easier for later subjects, so that in order to keep the level of correct recall approximately constant as the experiment progressed it was necessary to reduce the time allowed for successive subjects to study the lists. Judges who (subjectively) ranked the lists for similarity to English estimated that no further approach towards English occurred after about the twelfth of the sixteen orderings.

The above result indicates that when the words in a list are known, but their order is not completely retained, subjects do tend to fall back on learned language habits. However, this provides only indirect evidence for Deese's hypothesis, since the provision of the words to be reproduced in Coleman's (1962) experiment alters the nature of the task. Yet if it is true that one effect of sequential redundancy is to facilitate the reproduction of order, apart from that of single items, the effect which Coleman observes may be a determinant of the influence of sequential redundancy on performance in typical verbal tasks.

The observation by Marks and Jack (1952) that when words reproduced in their correct sequence are measured, the recall score goes on increasing at higher orders of approximation than when words are scored irrespective of order, supports the suggestion that sequential redundancy influences memory for order. Thus Coleman's evidence does suggest an effect apparent at recall which would favour a reproduction of redundant sequences. However, this evidence cannot be linked to Deese's hypothesis in a quantitative way; for instance, one cannot say how much influence the order effect would have in a typical short-term retention situation. Further, in the Coleman experiment, although the nature of the task ensured that active re-ordering occurred at the time of recall, the new order may have been based partly on how a subject had retained the order, and not only on knowledge of the language which exerted an influence when order was being reconstructed. A subject's retention of the order may itself be influenced by knowledge of the language such knowledge being effective at a pre-reproduction stage in the task.

A second way in which the construction at recall hypothesis can be experimentally examined is by comparing the serial position error curves in lists varying in sequential constraint. If construction at recall is a determinant of total recall, it would be predicted that when the xth item is correct the probability of item x + 1 being correct will be affected by sequential redundancy. The correct recall of early items should facilitate recall of subsequent items more in redundant than in nonredundant sequences. If the differences in recall between lists of high and low orders of approximation to English were entirely due to construction at recall, the serial position error curves for lists of different levels of approximation would be something like those shown in figure 2, in which recall of item 1 is independent of order of approximation, but recall of later items is increasingly affected by this

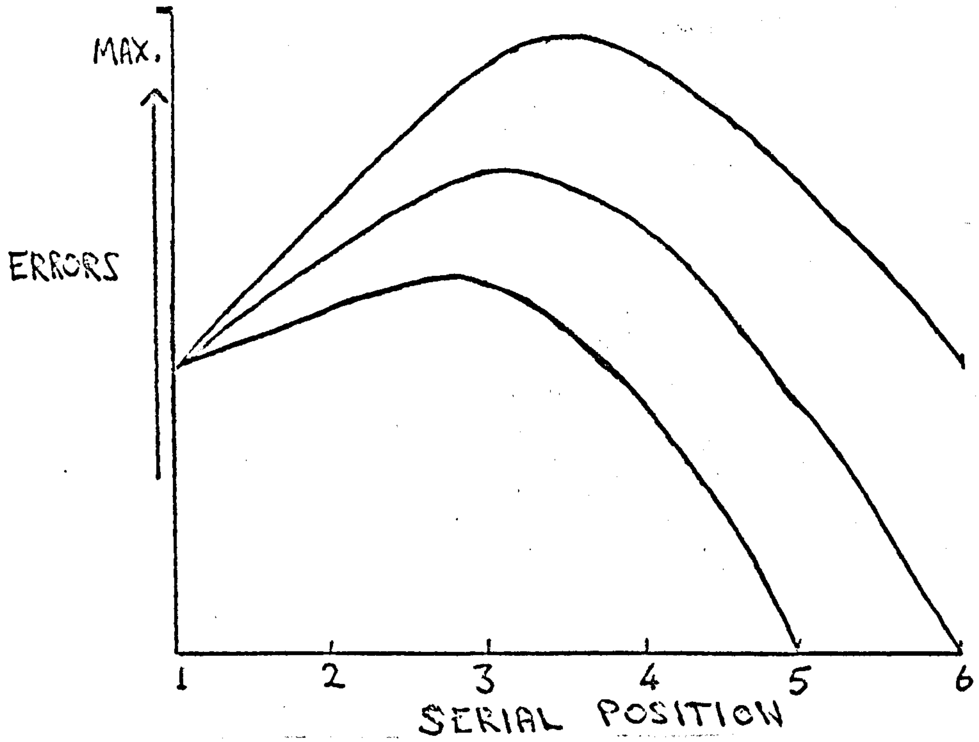

Figure 2. Hypothetical (simplified) serial position curves for errors, based on the assumption that the effects of sequential constraint on word recall are entirely due to construction which occurs at recall.

variable. Clearly, Figure 2 is a vastly over-simplified representation, but reconstruction occurring at recall would definitely influence the distribution of errors between serial positions. Two studies have supplied some empirical evidence relevant to this point. Simpson (1965) presented 30-word lists which varied in sequential restraint from sentences to randomly ordered words. He plotted the distribution of errors, transforming the raw data into percentages of total errors for each order of approximation at each serial position. From this it should be possible to determine whether the pattern of the error distribution varies between levels of sequential structure in the lists. The considerable length of the lists makes interpretation of the serial error curves difficult, but Simpson found no definite evidence of systematic differences in the general trends of the curves. Deese and Kaufmann (1957) examined the serial order curves errors for made by subjects in 10-word free recall trials, the lists being of varying orders of approximation to English. They found that for the higher order approximations a larger percentage of the total items recalled came from the first half of the list than was the case in the less constrained sequences. This result would appear to oppose Deese's hypothesis, but interpretation of the evidence is complicated by the fact that subjects were not required to recall the items in the order presented, and could recall the words in whatever order they wished. Order was not considered in scoring. Deese and Kaufmann also found that order of items in free recall of random word lists correlated positively with the ranked frequency with which the items were recalled, whereas order in free recall of words presented in textual material correlated with order of presentation. As the authors remark, since increasing redundancy in the sequentially organised lists is associated with reorganization of emission of items in recall, the effect of sequential structuring is rather complex. Deese and Kaufmann's results cannot be cited as firm evidence against construction at recall, because word emission order differs between levels of approximation to English in word lists, even when no instructions are given about order of recall. However, if an experiment were carried out with lists of similar length to those used by these authors, but with the instructions to recall words in the order of presentation, the shapes of the resulting serial position error curves could be used for evidence for or against Deese's construction at recall hypothesis.

All the studies discussed above contain evidence relevant to the question of whether successful reconstruction occurs at the time of recall, but none of the evidence can be used for quantitative evaluation of Deese's hypothesis. Deese's own (1961) results, and those of Morrison and Black (1957) and Aborn, Rubenstein and Sterling (1959) demonstrate that subjects are able to guess at words in sequentially constrained lists with better than chance probability of success. However, most of these investigations used prose passages, and it is not possible to determine from the evidence obtained how successful guessing would be in passages of lower order approximation to English. Nor is it simply possible to use the knowledge of what percentage of single words can be correctly guessed when deleted from a prose passage in order to predict numerically how much benefit successful guessing would provide in a recall situation.

Guessing words at recall also differs from guessing single deleted words in that the effects of success or failure of a guess could be cumulative. For instance, whether or not a particular word was correctly guessed in a sequence to be recalled might affect the probability of subsequent words being guessed correctly. The attempt to assess the validity of a reconstruction hypothesis from evidence provided by single-word guessing experiments is further complicated by the fact that guessing would only affect recall of words which were not remembered. The proportion of such words and their distribution are additional factors to be considered in estimating the effect on recall of a guessing strategy. Another difficulty in attempting to make predictions from guessing experiments is that, as has been mentioned, subjects may guess the correct order of remembered words (Coleman 1962) and this could influence recall scores. Single-word guessing studies cannot take into account this sort of guessing, and its effects on recall may not be simply additive to those of word construction. Finally, it may be the case in situations requiring recall of word sequences that the activity of guessing early items may interfere with the retention of items to be recalled later in the sequence.

This factor would not be present in simple guessing experiments. In short, there are so many differences between the typical recall situation and the guessing experiments which have been described that results from the latter do not provide very clear evidence on the question of construction at recall.

Turning to the evidence from serial position curves for errors, it is unfortunate, for the present problem, that the subjects of Deese & Kaufmann (1957) were not told to recall items in the order of presentation. It is difficult to interpret the results obtained by Simpson (1965), but they certainly provide no unambiguous evidence in favour of construction at recall. However, there is no reason why comparison of serial position curves in experiments which require ordered recall of lists differing with respect to sequential organisation, should not be used as evidence in the current problem. Curves obtained from suitable experiments have yet to be described, but the necessary data can be drawn from the results of an experiment to be described below.

Generally speaking, the evidence presently available which is relevant to construction at recall does not provide more than the general statement that subjects can make use of their knowledge of language structure to make better than chance guesses, which might aid recall. But whether such guessing actually does benefit recall in the typical experimental situation, and if so by how much, are questions which cannot be answered from presently available results. A more direct experimental approach is required. The results of such an approach will be described in the next chapter.

SUMMARY OF CHAPTER 2.

Problems arising from the need to measure sequential redundancy in language are discussed. A useful step towards explanation would be provided by specifying precisely the points within a task at which redundancy is directly effective. Deese has suggested that sequential redundancy in verbal material is important largely at the time of recall, at which stage reconstruction is said to occur, based on a person's knowledge of language. The evidence put forward by Deese in support of his hypothesis is surveyed, and other experimental evidence is described which is relevant to Deese's position.

CHAPTER 3.

AN INVESTIGATION OF REDUNDANCY EFFECTS IN WORD SEQUENCES

3.1. Exact description of redundancy effects.

The experiment whose description constitutes the major part of this chapter was designed to test Deese's hypothesis that differential accuracy of attempts to reconstruct material at recall underlies the observed recall differences between word passages high and low in sequential constraint. One approach has been described in a recently published paper by Lachman and Tuttle (1965, Experiments 1 and 3). Their procedure was to present 104-word sequences of high or low order approximation to English, followed by a recognition task which consisted of sorting 208 cards, on each of which a word was typed. The words on half the cards had been in the original list, and the others had not. Subjects had to sort the cards, which were randomly ordered, into two boxes marked "Yes" or "No" according to whether or not the word had been seen as part of the original sequence. Presentation of the list was auditory, at the rate of 1.5 seconds per word, and the cards containing the recognition lists were sorted by subjects at a paced rate of one card per 2.5 seconds. It was argued that the paced two-choice recognition test for each word of the recognition sequence would eliminate all opportunity for reconstruction, so that if words which had been presented in lists of high approximation to English were better recognised that words in low approximation lists, this would demonstrate a storage difference between high and low sequences. The results showed that words in lists of high approximations to English were more frequently recognised, and therefore the authors claim that material in highly organized sequences is stored more efficiently. Lachman and Tuttle point out that this result in no way precludes output construction as a possible additional factor.

The results obtained by Lachmann & Tuttle show quite clearly that an advantage in recall of highly organized, as against less organized sequences is evident when the possibility of construction at recall is reduced and therefore it seems that at least part of the advantage to recall of highly organized sequences has been acquired by a point in the recall task prior to the stage of actual recall. Further experiments would be necessary to provide more precise information concerning the stage within verbal tasks at which sequential organization first becomes important. The question of whether reconstruction at the time of recall does occur at all in word sequences can be fairly satisfactorily answered from the results of the experiment to be described.

3.2. Experiment 1.

The present experiment was carried out before Lachman & Tuttle's (1965) results were published. The general aim was similar, in that the experiment was designed to determine whether there was any evidence for storage differences between the recall of word lists differing with respect to degree of sequential organisation.

The reasoning behind Experiment I is that if there is a difference in the way in which word lists of high and low approximation to English are stored, the effect of adding an experimental variable which can be expected to influence material in store, is likely to differ between list types. Word sequences of two levels of approximation to English were presented visually, followed by a variable period during which rehearsal was permitted. Then followed a distracting task after which came written recall. It was argued that the effects of rehearsal would be to stiffen resistance to interference, the strength of the effect depending on the length of rehearsal period. If the two types of sequence are stored differently it seems likely that the consolidating effect of providing a rehearsal period of a certain length will differ between the types of sequence, leading to an interaction between the effects of the two independent experimental variables, type of list and length of rehearsal period. A weakness in this design compared with that used by Lachman & Tuttle, is that whereas in their results an equivalence in recognition between lists high and low in sequential organization would strongly indicate that no difference in retention between sequence types existed prior to recall, the same reasoning cannot be applied to the present experiment. It is possible, for instance, that types of sequence differ in the way items are stored, but that this difference does not happen to be sensitive to the particular independent variable used, at least within the values adopted for the present experiment. However, although the absence of a significant interaction between rehearsal time and list type would not be certain evidence for the absence of a difference in storage, the presence of such an interaction would indicate that an underlying storage difference does exist.

The dependent variable in this experiment is the resistance to interference from a distracting task possessed by word sequences which subjects are remembering. Such resistence, or consolidation, is measured by scoring the recall of lists by subjects after they have performed an interfering task. Rehearsal succeeded the presentation of six-word lists (i.e., the longest lists for which recall by English undergraduates after presentation or during the rehearsal period is rarely imperfect). Then following the interfering task, after which recall was tested.

Method

Subjects and Materials. Twenty-four paid undergraduates participated in the experiment. Twenty-four six-word lists were taken from the second-order approximations to English used by Taylor and Moray (1961). These will be called "B lists". The words constituting these lists were chosen at random to make a further 24 lists of six words each ("A lists").

Interference Task. Subjects had to subtract by sevens from a three-digit number. They read aloud the original number, and the subsequent three residual numbers, which were simultaneously written down. Each subject received a new number for each trial, for which all subjects in a group were given different numbers. Each commencing number for an interfering task was printed on the sheet of paper used by a subject for subsequent recall of the word sequence. A number was followed by three dashes above which the subject wrote the residual numbers, and each line on a sheet had six further dashes, for written recall.

Experimental Design. The subjects were randomly assigned to three approximately equal groups. They were all shown 48 lists, in six batches of eight. There were three values of the rehearsal time variable, zero, three and six seconds, and within batches the same value was used with all lists, subjects each seeing two batches of eight lists coupled with each rehearsal-time. Subjects rested for three minutes after the 24th list, and there were pauses of about a minute between batches. Order of lists within batches was identical for all groups, between which order of batches was varied. Within each batch there were four A lists and four B lists, ordered randomly except that no runs of over 3 were included. Before each batch subjects were told the rehearsal time to be allowed after each of the eight lists, but they were not given advance knowledge of list types.

Procedure. Lists were projected for three seconds on to a large screen. Subjects were told to watch each list until it disappeared, and then rehearse silently until a red lamp, placed just below the screen, was illuminated. In the no-rehearsal condition this coincided with the disappearance of the list. The lamp was the cue for subjects to start the interference task, on completion of which written recall was attempted. Subjects had to write from left to right, using the six printed dashes on the recall sheets. About 30 seconds elapsed between the start of the interference task after one list and the projection of the next. For practice, subjects had four trials of the interference task alone, followed by nine trials of the full experimental procedure, which included at least one combination of each list type and rehearsal-time.

Scoring. Two measures of the number of words correct in a list were used.

Ordered Recall	To be correct a word had to be recalled in the right position in a list.
Item Recall	Recall of any word which had appeared in the original list was scored as correct.

Item recall was scored in order to determine whether the magnitude of the difference between ordered and item recall varies as a function of order of approximation to English. However, ordered recall is more extensively used in the presentation and discussion of results, since it is the more sensitive measure (see Marks and Jack, 1952; Coleman, 1963) of ability to carry out the task as instructed.

Results

Figure 3 shows the relation between number of words correctly recalled from both types of six-word lists and at each rehearsal period, using both scoring methods.

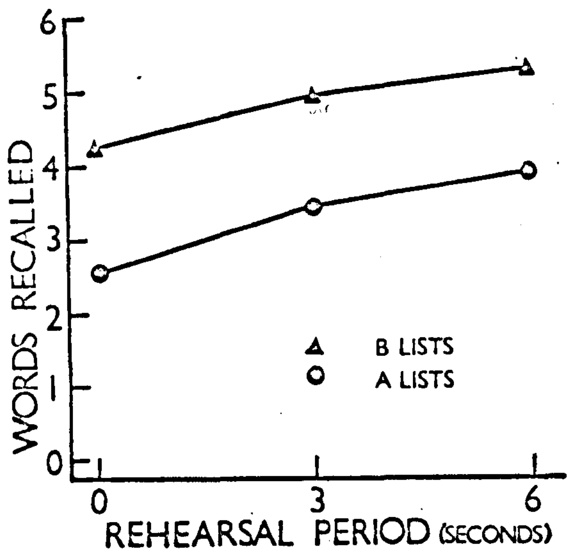

Figure 3. Mean number of words correctly recalled (out of six) in Experiment 1, as a function of sequential constraint and time available for rehearsal. (A lists are first order word approximations, B lists are second order approximations.)

Table 2 provides the mean recall scores in numerical form and gives the p values for the significance of differences between the means a) for different rehearsal times, within each list type, and b) for different list types within each rehearsal time, by both methods of scoring.

	Rehearsal Time (Secs.)	Ordered Recall: A lists	Ordered Recall: B lists	Item Recall: A lists	Item Recall: B lists
Recall Scores	0	2.55	4.27	3.53	4.94
	3	3.42	4.90	4.28	5.33
	6	3.84	5.23	4.51	5.52

		Ordered Recall: Pair		t	p<	Item Recall: Pair		t	p<
Differences between recall scores.	Rehearsal-Time Constant, List-Type Varied.	A0	B0	6.65	.001	A0	B0	6.75	.001
		A3	B3	5.45	.001	A3	B3	9.56	.001
		A6	B6	8.30	.001	A6	B6	9.48	.001
	List-Type Constant, Rehearsal-Time Varied	B0	B3	4.27	.001	B0	B3	5.55	.001
		B3	B6	1.60	.1	B3	B6	1.73	.1
		B0	B6	5.00	.001	B0	B6	6.31	.001
		A0	A3	2.02	.1	A0	A3	2.60	.02
		A3	A6	0.67	.6	A3	A6	1.53	.2
		A0	A6	3.65	.01	A0	A6	5.20	.001

Table 2. Mean numbers of words correctly recalled (out of six), and significance levels of differences between pairs of mean recall scores, in Experiment 1, for list-types (A, B.) and rehearsal-times (0, 3 and 6 seconds).

Clearly, B lists are better recalled than A lists at all values of rehearsal time. With both types of list more words are recalled when three seconds rehearsal precede the interference task than when no time is allowed for rehearsal. Between means in the three- and six-second rehearsal conditions, only one of the four comparisons is significantly different, although the other differences are all in the same direction. The absence of a significant rise in the recall of the B lists when rehearsal time increases from three to six seconds may be a ceiling effect, since five of the 24 subjects recalled all the B lists correctly when rehearsal time was six seconds (ordered recall). It is clear that there is no evidence of interaction between the effects of type of list and length of time for rehearsal. Effects of length of rehearsal period on subsequent mean recall did not differ between A and B lists. Comparing differences between recall in the zero and six seconds rehearsal conditions for A and B lists, $t=0.70$, $p=.2$, by ordered recall; $t=0.29$, $p=.5$, by item recall.)

Figure 4 shows serial position effects for errors at all rehearsal time values for both types of list, scored by ordered recall. In all conditions the least errors are made among words presented in the first position in a list, and the greatest number of errors occurs in the fourth position. The difference in shape of serial error curves between A and B lists is an artefact of graphical representation of distributions differing in magnitude. Table 3 lists errors at each serial position as a

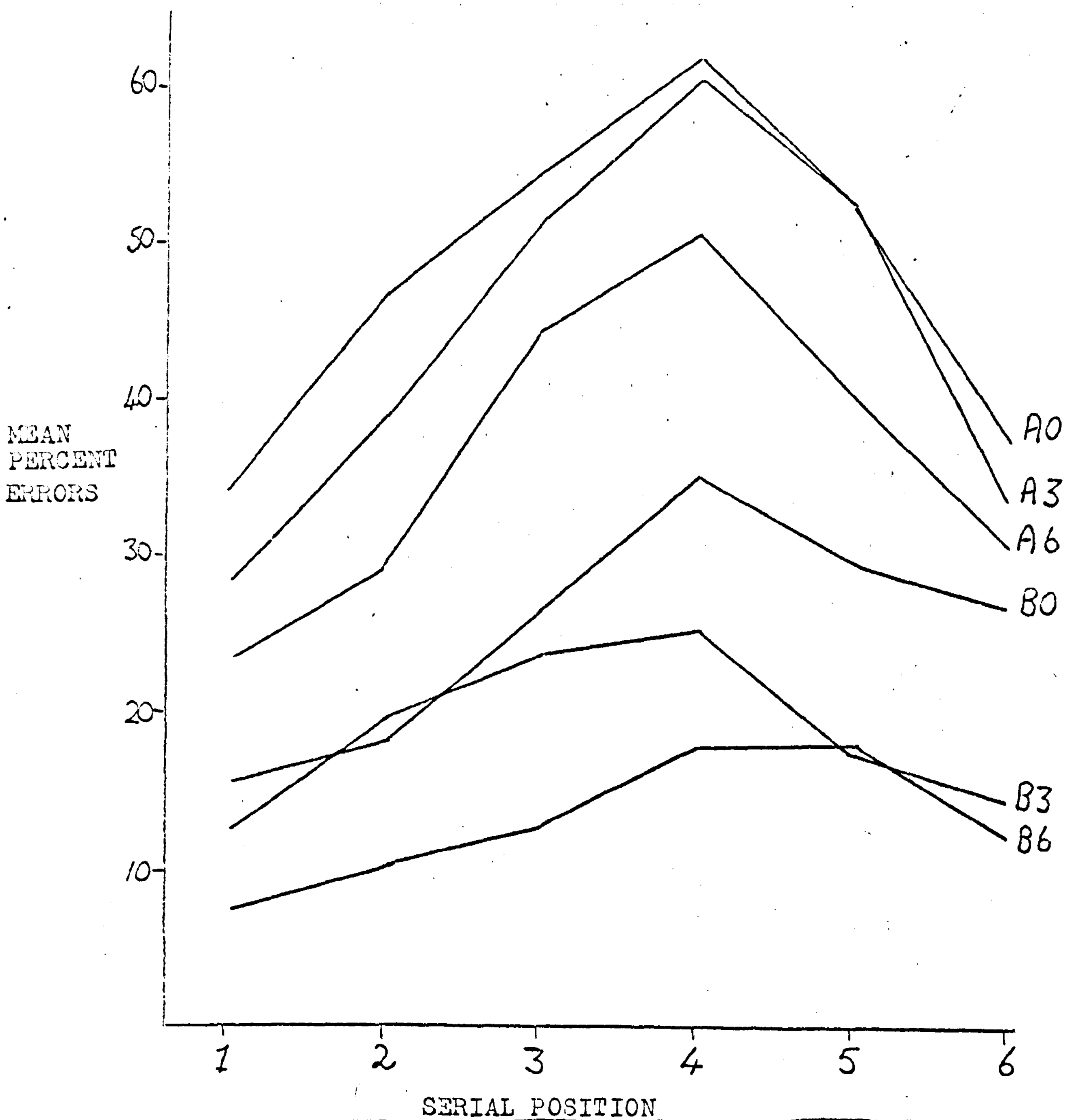

Figure 4. Serial position curves for errors in Experiment 1. A lists were first order word sequences, B lists were second order sequences. The digits refer to the length of time available for rehearsal, in seconds.

percentage of total errors in a condition. Differences between conditions in inter-position error distribution are small. Comparing the

	Rehearsal Time (secs.)	Percentage of total errors per condition Serial position.					
		1	2	3	4	5	6
A lists	0	11.9	16.2	19.0	21.6	18.3	13.0
	3	10.7	14.6	19.4	22.9	19.8	12.6
	6	10.6	13.5	20.2	23.2	18.6	13.9
B lists	0	10.6	12.2	17.6	21.6	20.0	18.0
	3	11.2	17.3	20.9	22.4	15.4	12.8
	6	9.3	12.9	16.4	22.9	22.9	15.6

Table 3. Distribution of total errors in each condition between serial positions, in Experiment 1.

distribution of errors between serial positions in A lists with the distributions we would expect if the errors in A and B lists were distributed identically among serial positions within conditions, we find no significant difference in shapes of distributions between A and B lists at any rehearsal time. ($\chi^2 = 8.61$ (5) for the difference between A0 and B0; 3.34 (5) for A3 and B3; + .96 (5) for A6 and B6. For all of these $p > .1$.)

Discussion

Both order of approximation and time available for rehearsal influence the resistance of items being retained to a subsequent distracting task. The potency of sequential constraint as a variable is illustrated in the present experiment by the fact that a period of six seconds for rehearsal does not compensate for the difference between A and B lists in recall after a distracting task (see Figure 3).

Since there is no interaction between the effects of degree of approximation to English and length of time available for rehearsal, positive evidence for a difference in storage is lacking. The absence of an increase in the difference in recall between A and B lists as time available for rehearsal increases does not necessarily preclude the possibility of a difference in the way in which words in A and B sequences are stored. This is so for the reason already given; a genuine difference may be undetected in a study using the present experimental variables. Also, although the recall difference between A and B sequences is constant over all values of rehearsal time, to maintain such a constant differential, more items in B lists than in A lists must profit by a rehearsal-time increment of a given length. That is to say, more B list words than A list words must profit by rehearsal if the recall difference between types of sequence is to be as large when say, six seconds are allowed for rehearsal as when no opportunity is provided. It may be the case that the magnitude of the effect of the distracting task as well as the post-task recall score depends on prior level of retention. However, this possibility does not complicate the present analysis, since only if there is a pre-recall difference in retention could such an effect influence the results, as the directly affected stage in the task would be prior to that of recall. Thus such a factor could not create an interaction in the experimental results which was not due to events in pre-recall retention, although it might conceivably distort the magnitude of a genuine interaction which was primarily due to such events.

It is apparent, then, that the results shown in Figure 3 and Table 2 provide no clear evidence that degree of sequential constraint in word lists affects pre-recall retention of them. However, strong support for the assumption that retention is affected by sequential constraint, albeit of a negative kind, is given by the serial error curves shown in Figure 4, and by the data presented in Table 3, which is based on them. For reasons given before, any effect of successful guessing at the time of recall will influence the distribution of errors between serial positions. Therefore, if any of the difference in recall between lists differing with respect to sequential restraint is due to reconstruction at recall, the error curves will reflect this. However, the results show quite clearly that in the present experiment inter-position distribution of errors does not differ significantly between the different levels of approximation to English. It follows that successful construction at recall is not occurring to any large extent. But total recall does vary with order of approximation, and if this is not due to factors operating when recall is being attempted, the observed difference must reflect factors operating previous to recall, that is to say, during and/or before the period of retention. This conclusion complements that reached by Lachman and Tuttle (1965), the combined evidence strongly suggesting that the observed recall differences are not due to reconstruction occurring at recall, but reflect differences which are first apparent at a prior stage in the verbal tasks used.

It is possible that the serial position curves for errors reflect factors additionlto those which have been considered, which might contribute differentially to the error distributions of verbal materials varying in sequential constraint. For instance, it is possible that some successful guessing occurs at recall, but that while this contributes positively to the recall scores for some items the activity has the additional effect of interfering with the retention of other items presented, and hence to be recalled, relatively late in a sequence. This might modify an effect on serial position distribution of construction at recall, giving rise to an erroneous impression that successful reconstruction does not occur. However, although it is conceivable that effects of this sort could be present, there is no real evidence for them, and the hypothetical nature of the factors and values involved complicate appraisal of such a possibility. In any case, if the reason for the equivalence of inter-position distributions between orders of approximation was that effects of successful guessing on recall of late items were cancelled by simultaneous interference from interpolated recall, total recall would remain constant. It is the very fact that it does not which we are trying to explain. On the whole it is probably safe to assume that correct construction at recall would be evident from the shapes of serial position curves, provided that subjects obey the instruction to attempt recall of items in the order of their initial presentation.

The conclusions reached apply to construction at recall in the way suggested by Deese, that is to say, straightforward guessing by a subject on the basis of his knowledge of the structure of a language. As was suggested earlier, recall might benefit by more complex processes containing guessing, such as a combination of guessing and recognition, and conclusions regarding the importance of construction at recall in Deese's sense to not necessarily apply to the possibility of recall being effected by relatively complex strategies in which guessing may play a part. Another hypothetical account of what might happen at the time of recall is that subjects, while not retaining a verbatim representation of the passage, have a condensed account which retains the general meaning and structure (Bartlett, 1932). When recall is being attempted a subject may be able to use his knowledge of the language to construct a verbatim representation from the condensed version he has retained.

SUMMARY OF CHAPTER 3

An experiment is described (Experiment 1) which was designed to test Deese's hypothesis that sequentially constrained word lists are more accurately reproduced after retention than random lists because recall of them profits by reconstruction which occurs at recall. The results, and those obtained by Lachman & Tuttle (1965), indicate that reconstruction is not an important determinant of the difference in recall between random and constrained word lists.

CHAPTER 4.

MEASURES OF REDUNDANCY IN LETTER SEQUENCES

4.1. Letter units in sequential analysis.

The aim so far has been to provide a more precise description than has hitherto been available of some effects of sequential constraint in language. It was argued that one step towards understanding how the brain codes material which is sequentially redundant for the human user, would be to specify the stage or stages within a verbal task at which degree of sequential redundancy is a directly limiting factor. The present chapter and the succeeding two chapters continue this attempt.

The analysis until now, and the conclusion reached, have been concerned with language organization at the level of words; effects of sequential constraints between words have been considered. The approximation to English index was used by Miller and Selfridge and by subsequent workers, as a measure of the similarity in organization of a sequence of words to English. However, analysis of language can be made at any practicable level, e.g., the sentence, word, syllable, letter or phoneme. For specific purposes there may be advantages in adapting a particular unit of analysis, but there is no "best" unit for all analyses of language. The word unit is clearly useful. Newman and Gerstman (1952) say that,

> "In many respects analysis in the level of words would be the most useful. Words come closest, at least, to the layman's idea of what are the natural units of speech and communication."

However, if the analysis is to rely heavily upon the use of experimental methods, there are disadvantages to the use of words, which may best be overcome by adopting a unit of analysis which can more suitably be used in experimental investigations.

If experimental methods are to be used in attempts to understand mechanisms by which the brain codes sequentially constrained verbal material, an important requirement is the ability to make precise measurements of the important variable, the degree of sequential constraint. We need to be able to measure with some accuracy the redundancy of the language segments which are being used. Related to this, a measure is needed which can be applied both easily and widely to materials such as those which have typically been used in published experiments in verbal learning. It would be desirable, for instance, to be able to measure sequential restraint in material which has been used in the past to show correlations between aspects of the material and ratings of verbal performance. The experimental work described in the previous chapters, in which the word was the unit of analysis, made use of the approximation to English index as a measure of sequential organisation. It is fairly easy to manufacture a passage whose structure has a given degree of ordinally ranked similarity to English. However, there are disadvantages in the experimental use of approximations to English, and since no better measure of word sequence redundancy can be obtained, these disadvantages apply generally to the use of words as units in measuring sequential redundancy in language.

The limitation of approximations to English as indices of sequential redundancy is that while it is not difficult to construct a sequence of a given order or approximation, the measure cannot be used to measure the sequential constraint in a given word sequence. It is not practicable (although it is possible, in principle,) to obtain a direct measure of sequential redundancy in word lists. Even if a very restricted vocabulary of, say, 1,000 words is used there are $1{,}000^2$ possible two-word sequences, $1{,}000^3$ three-word sequences, and so on. Some combinations of words would never occur together, but even so the task of making a count of word sequences in an adequately long sample would be enormous, and has not yet been attempted.

It is for this reason that a method such as that of making approximations to English is required to construct word lists of known degrees of sequential organization. In constructing such lists use is made of the fact that literate adults have an intuitive knowledge of the structure of their language, so that the various approximations will reflect the sequential structure, given the appropriate amounts of context. No attempt will be made to evaluate the approximation to English index; its weaknesses are inherent in any attempt to use an index of sequential structure which is not based on direct measurement. However, an important point is that some of the limitations on the assessment of word sequence structure are such that the word may not be the most suitable of analysis for all experimental investigations of sequential redundancy in language. Put simply, the first of these limitations is that, as has been mentioned above although it is possible to construct a sequence with a given degree of structural similarity to English, it is not possible to measure directly the degree of sequential organisation in a given passage. To be able to do so is clearly an advantage in experimental investigation. A second disadvantage of the approximation to English index is that even if the passages constructed are reliable samples, the experimental use of the measure is limited by its merely ordinal nature. It would be desirable to have a unit of verbal analysis for which sequential redundancy within a list of items can be measured in a way which is direct and precise. Such a unit is the letter.

Simply because the number of letters which occur in a language is much smaller than the number of words, it is much easier to measure their frequency of occurrence in sequences. Even if spaces and characters are allowed to count as one letter each, there are only 28^2 possible diagrams (i.e. combinations of two letters), 28^3 trigrams, 28^4 tetragrams, and so on. Therefore, given a large sample of English, direct measurement of sequences of letters is possible at least up to the trigram level. A number of frequency counts for letter sequences have been made; these are summarised in Table 4.

Count	Sample Length	Nature of Sample	Symbols Counted	Data Provided
Pratt, (1939)	20,000 words	Miscellaneous 500-word written modern English sequences	Letters (26)	Trigram frequency & frequency of digrams within the trigrams
Newman and Gerstman, (1952)	10,000 letters	King James Bible. Isiah xxiv-xxxi	Letters & space & punctuation (28)	Digram frequency (They also counted frequencies of any two letters when separated by 1, 2, 3...9 symbols.)
Underwood and Schulz, (1960)				
(i) "T-L Count"	2,080 words	Sample of Thorndike-Lorge's 20,000 most frequent words	Letters (26)	Trigram frequency Digram frequency
(ii) "U Count"	15,000 words	Miscellaneous 150-word written modern English sequences	"	"
(iii) (i) and (ii) combined	18,080 words	(i) and (ii) combined	"	"
Baddeley, Conrad and Thomson, (1960)				
(i)	76,150 digrams	Editorial columns of "The Times"	Letters & space (27)	Digrams. (Trigrams available from authors)
(ii)	80,766 digrams	"Mrs. Dale's Diary" (Spoken English)	"	"
Mayzner and Tresselt, (1962a)	20,000 words	Miscellaneous 200-word written Modern English passages	Letters (26)	Digrams & single letter frequency counts for various word lengths (4-7 letters) and letter position combinations.

Table 4. Summary of characteristics of letter-sequence frequency counts.

As Miller and Selfridge (1950) point out, it is possible to use the approximation to English method for making sequences with letters rather than words. A very similar technique described by Shannon (1951) took letters as the basic unit. Miller, Bruner and Postman (1954) found that reproduction of the lists by subjects after tachistoscopic presentation was successfully predicted by order of approximation. Baddeley (1964a) correctly interpreted this result as being at least partly limited by recall rather than by perception. In other words, more letters were remembered in lists at the higher orders or approximation[1]. However, measures of letter sequence approximations to English have the same disadvantages in experimental analysis as word approximations. Hence, many investigations of sequential redundancy in letter lists have made use of direct frequency measures of letter sequences.

Much of what has already been said about precise description of word sequence redundancy effects, as a step towards explanation, is applicable to the experimental investigation of effects of sequential constraints between letters. It would be useful to be able to answer with

1 There is a complication attached to the construction of long letter sequences at higher orders of approximation. Kay (1966, personal communication) has pointed out that in the English language all-letter sequences are typically much shorter than word sequences, since there are less letters in the average word than words in the average sentence. A result of this is that long letter lists at higher orders of approximation necessitate more elaborate sequences than usually occur in words, just as the structure of very long word sequences at the higher orders of approximation may be more complex than that of typical sentences (Coleman, 1963). However, it would probably be possible to devise a means of making adequate provision for spaces and punctuation marks in constructing letter sequences at various approximations to English.

certainty the question which was asked in the case of word sequences, namely, at what stage or stages within the performance of a verbal task does sequential redundancy exert a direct influence? The answer explicit in Deese's (1961) hypothesis was intended to apply equally whether words or letters are taken as the unit of analysis. Thus it would be useful to investigate letter sequences in much the same way as word sequences have been examined. In a sense this work duplicates that of the previous chapter, but although the questions asked may be identical, it is quite possible that experimental investigation may provide very different answers. For instance, everyday evidence for construction at recall is much stronger in the case of letter sequences than in word sequences. The reason for this is identical to the reason why it is possible to obtain quantitative measures of sequential organisation in letters which are more satisfactory than available measures of word sequences: there are fewer different letters than there are words. A consequence of this is that it would seem to be much easier to predict missing letters from a structured sequence of letters than words from word sequences. Likewise, if only part of a letter sequence has been remembered by a human subject, his attempts to guess at the other letters on the basis of his knowledge of language structure may be much more successful than in parallel situations with word sequences. The experimental evidence supports this reasoning. The findings will be discussed in detail later, but at this point it can be simply noted that guessing at items in letter sequences is more successful than guessing at word items.

In this and the next two chapters investigation will be directed towards specifying the operations within a verbal task at which sequential redundancy is a directly effective variable. Evidence will be presented which will be relevant to the consideration of Deese's reconstruction at recall hypothesis. The reasons for taking this general approach and for testing this particular hypothesis are identical to those put forward in Chapter 2, in connection with word sequences. As with words, most of the experimental evidence presently available about constrained letter sequences relates redundancy to global performance at verbal tasks, and since the tasks are typically fairly complex it is not possible to determine the effects of redundancy with much precision. Experimental results relating performance to sequential organisation in word lists will be summarised, and closer attention will be paid to evidence relevant to more detailed description of these effects. One important difference between word and letter sequences is that while the approximation to English index is the only available measure of sequential organisation in the former, redundancy in lists of letters has been measured in a number of ways. Some results of this difference merit discussion.

4.2. Problems of measurement.

It is valuable to be able to measure directly the frequency with which given letter sequences occur in written English. However, in doing so certain problems concerning experimental procedure are encountered. For instance, it is much easier to measure digram frequency (i.e., the frequency with which two given letters occur in adjacent positions) than it is to measure trigram frequency, but the latter is an index of closer structural similarity to English (as third order word approximations are more similar to English than second order approximations). Measurement of the frequency in English of letter groupings longer than trigrams is not practicable. Some experimenters have used letter lists which contain only consonants. They would presumably argue that the disadvantage of not using a representative sample of English letters is balanced by the greater experimental control which can be achieved. One digram count (see Table 4) takes spaces and punctuation marks into account; another count of digrams and trigrams counts spaces but ignores punctuations, while a third ignores both. Mayzner and Tresselt (1962a) have made a digram count which includes data on the positions within words of the diagrams and the lengths of the words in which they occur, factors which are additional determinants of redundancy, as Garner and Carson (1960) imply. Garner and Carson also point out that letters can be predicted backwards as well as forwards, and Shannon (1951) has provided experimental evidence which supports this statement. It is clear that anyone who intends to use redundancy in letter sequences as an experimental variable, and needs a measure of it, has to make a number of procedural choices concerning the materials and measures to be used, for which two or more alternatives may seem equally correct.

Baddeley (1963) has devised a measure of letter-sequence redundancy which he calls "Predictability".[1] This is related to the letter digram frequencies in a sample of modern English prose (Baddeley, Conrad and Thomson, 1960). The method by which values of the measure are calculated is described by Baddeley thus:

> "Treating a space (*) as a letter, a syllable (e.g. BOF) can be considered as a sequence of five letters, *-B - O - F - *. This can be split into four digrams or pairs of letters, *B, BO, OF and F*. One can then compare these digrams with the digram structure of English by taking them one at a time, removing one letter and estimating the difficulty of predicting that letter on the basis of the structure of English.
>
> The predictability of the second letter of a pair, given the first, is calculated by assuming a statistical "subject" whose "guessess" are entirely determined by a table of the digram structure of English. Given, for example, the letter T, then according to the digram table, the most probable next letter is H, which would therefore be the first prediction. The next most frequent item following T is a space, and this would be the second "guess". On the other hand, N, for instance is much less likely to follow T, and as there are 15 letters that are more likely to occur after T, the statistical subject would require 16 guesses to predict N. The digrams TH, T*, and TN would therefore have predictability scores of 1, 2 and 16 guesses respectively when the second letter must be predicted on the basis of the first. However, a guess could also be made in the other direction by supplying the second letter and seeing how many attempts were necessary to guess the first, and this would give a similar score. Thus, the four digrams comprising a nonsense syllable would give a total of eight guessing scores,each

1 Baddeley uses the term "Predictability" in a narrower and different sense to the meaning of the term in everyday English. In the present chapter the term will be used without inverted commas to refer only to Baddeley's measure, as defined here. Elsewhere, unless the word is enclosed by inverted commas, its common English meaning can be assumed.

> ranging from 1, where the adjacent letter is the most probable next letter in English, to 27 where it is the least probable. The mean of these scores can be regarded as giving a rough index of the "predictability" of the material, which represents the similarity between its digram structure and that of English."

Two reasons (Baddeley, 1963) are given for preferring predictability to a simple count of digram frequency in English. The first is that since there is a wide range of digram frequency the occurrence of a single very probable digram, e.g., TH, in a short sequence "would be sufficient to swamp any other effect giving a huge total". Secondly, "a score based on a simple frequency total has little generality, being completely dependent on the size of sample used to estimate the letter structure of English, and on the length of letter sequence to be measured." Baddeley mentions three other drawbacks to previously available methods of obtaining measures of sequential letter similarity to English. These drawbacks are due to the inadequacies of early digram counts and have been removed by the publication of entirely satisfactory digram-frequency table based on a large representative sample of modern English (Baddeley, Conrad and Thomson, 1960).

A relationship between predictability and performance in cognitive tasks has been observed in several experiments. These will be discussed in Chapter 5. Very briefly, it was found (Baddeley, 1963 and 1964) that both inter- and intra-syllable predictability affect the ease of learning nonsense syllables. Short-term memory for eight-letter sequences (Baddeley, 1964a) and for six-letter consonant lists (Baddeley, Conrad and Hull, 1965) is related to their predictability. The correlation between recall and predictability for the eight-letter sequences used was + 0.8, and between predictability and recall of the six-letter consonant lists there was a correlation of + 0.3.

Unfortunately, the predictability index does not reach one important criterion of a good measuring-device. It is necessary that a reasonably straightforward answer can be provided to the question "What is being measured?" To justify the use of a new measure, concurrent with any decrease in the ease of answering this question, compared with existing measures, there must be some other clear advantage. In the present situation for instance, in choosing between two measures of sequential similarity to English, the one most simple and clearly related to the actual structure of English would be preferred unless the other was demonstrably superior in some other respect, such as predicting performance in a verbal task, or being easier to use. It seems to be the case that the measure of predictability does not bear a straightforward relationship to the English language, and that in no other respect has it been shown to be superior to any more direct measure of sequential organization between letters.

The following clauses have been used to define predictability: "...a measure of approximation to English... predictability of the material which represents the similarity between its digram structure and that of English." (Baddeley, 1963) "...a simple-to-use approximation to second-order letter sequence redundancy." (Conrad, Freeman and Hull, 1965.) "... how closely the digram structure of any sequence approximates to that of English." (Baddeley, Conrad and Hull, 1965.) Baddeley (1964b) describes predictability of syllables as "how closely their structure resembles that of English", and "S-R Compatibility", for which the operations required in measuring are described in exactly the way more frequently used for predictability, is said to be "similar to the language in sequential structure".

Clearly, its users have considered predictability to be a simple measure of second-order letter approximation to English. It is not this at all. Any measure of structural similarity to English must take into account the fact that some letters occur very much more frequently than others. For instance, in the 75,000 letter sample of modern English prose produced by Baddeley, Conrad and Thomson (1960) the letter E occurs 8181 times, while the numbers of occurrences of Z, J, and Q are only 21, 67 and 75 respectively. For every time Z is used there are 232 A's, 236 I's, 227 O's and 305 T's. This important form of redundancy has been called "distributional constraint" (Newman and Gerstman, 1952). The measure of predictability takes little account of this sort of constraint. To calculate predictability one ranks the frequency of occurrences of any letter k immediately after any letter m, in relation to the frequencies of other letters occurring immediately after m. For example, since E is the letter most likely to occur immediately after R, a value of one is given.[1] A direct measure of second-order letter similarity to English is obtained by measuring the frequencies of occurrence of digrams in the English language, and therefore any index which is claimed to be a measure of similarity to second-order English letter structure similarity to second-order English letter structure must be simply related to digram frequency. Values of the predictability measure for particular digrams can be very different to those of digram frequency. For instance, the forward predictability value of KL (calculated from Baddeley, Conrad and Thomson, 1960) is six, since L is the sixth most likely letter to succeed K. Similarly, since A is the sixth most likely letter to follow T, the digram TA also receives a value of six. However, in the sample from which their predictability was calculated, the digram TA occurs 240 times, while KL occurs on a mere six occasions. In other words, the predictability measure takes no account of the fact that, in the sample of English from which predictability is calculated, one of the digrams occurs 240/6 times as often as the other. Conversely, both KA and NJ occur four times in the above sample, but since A is the sixth most probable letter to succeed K, a predictability value of six is given, whereas NJ receives a value of 23, because J is only the twenty-third most probable letter to succeed N. In fairness, these are extreme examples. There is a correlation between predictability and digram frequency in English, since frequent letters tend to occur more often than infrequent letters as the second letter in digrams,

1 In calculating a value of the predictability measure, Baddeley takes into account the frequency of digrams in their reverse, as well as forward order. For example, the value for AR would take into account the ranked frequency of both R, given A, and A, given R. Whether or not to take frequency of the

reversed digram into account is an interesting question, but it will not be discussed here since it is not an issue on which to compare the merits of different indices of sequential organization. Frequency of reversed-order digrams can be included in values given by an index.

---

and consequently as the first letter in the next digram in a sequence. But since the measure of predictability is independent of the frequency in English of the first letter in a digram, and because only the ranked, and not the absolute frequency of the second letter is taken into account, the measure of predictability comes to have no simple relationship to the English language. Hence predictability cannot be regarded as an undistorted measure of similarity to English, as it is claimed to be. It is possible to calculate a correlation between predictability and digram frequency for any sequence. For example, the rank correlation between the two measures for the nine consonant lists used in Experiment 1 (Page 37, 3.2. in the present work is + 0.63 ($p < .01$). Clearly such a low correlation by itself indicates the inexact relationship between predictability and the letter structure of the language.

The question "what does predictability measure?" cannot be answered more simply than by going through the procedure by which a value is obtained. It is not quite a measure of what an approximation to English would be like were every letter equiprobable, since its values are not unaffected by letter frequency, nor is it an index of what English letter structure would be if absolute letter frequency was simply related to ranked frequency, since digrams are given predictability values not in comparison to all other digrams, but only in relation to other digrams with the same initial letter.

Although predictability is not an adequate measure of similarity to English, it may still be valuable if it is found to be more highly predictive of performance in verbal tasks than other measures. If this were so it might well be useful to look for similarities between the way in which a value is calculated and how information is coded by the human operator. For instance, a subject might adopt a guessing strategy which makes use of knowledge of the ranked frequency with which letters succeed a given letter. However, in the absence of empirical evidence, there is no reason why strategy adopted by subjects should happen to coincide with the way in which predictability is measured rather than with any other of the many possible strategies which speculation can produce, even assuming that it is realistic to accept the constraint that subjects do go through some sort of guessing behaviour which is based on past experience with a language.

As it is, the onus remains on the users of predictability to show that it has value beyond that possessed by any simpler and more direct measure of English letter structure, such as digram frequency. A comparison of these two measures as predictors of human performance has yet to be made. Log. single-letter frequency is not very important in immediate memory (Baddeley, Conrad and Hull, 1965), but no direct comparisons between correlations with performance of predictability and digram frequency can be made directly from published accounts, since users of the two measures have used different task materials. However, it is possible to measure both the predictability and the average digram frequency for a given set of sequences, and then calculate which of the two correlates most highly with performance. This has been done on the data used in Experiment 1 (partial repetition), with the following results.

Rank Correlations: Predictability and Recall + 0.26(n.s.)
Digram Frequency and Recall +0.47($p < .05$)

The evidence is not so damning to predictability as these figures suggest. Their very limited reliability is indicated by the fact that if one of the twenty lists was removed the rank correlation between predictability and recall would jump to a significant ($p < .05$) + 0.47. More weight should be placed on the correlations of + .80[1] between immediate memory for letters and predictability, found by Baddeley (1964a) and + .30 between recall of consonants and predictability (Baddeley, Conrad and Hull, 1965). However, there is no evidence to indicate that these correlations are higher than those which exist between performance and direct scores of digram frequency.

Baddeley's predictability measure, then, is not an adequate index of the similarity in structure of a sequence to English, nor has it been shown to have greater predictive power than a measure which is. Nor has it the advantage of greater simplicity in use than digram frequency.

1 It is possible that this figure may be inflated by differences in letter frequency distributions between the zero-order and other orders of approximation to English in the sequences used (which are from Miller, Bruner & Postman, 1954). For instance, the lists of higher order approximation more frequently contain letter repetitions than random lists.

There remain Baddeley's objections to the use of digram frequency scores. Ways exist by which the large frequency range, which Baddeley suggests might lead to "swamping" effects, might be reduced, such as by using a measure of log. frequency or ranked frequency. If this was done so that a value was given to each digram in relation to the frequency of all other digrams, and not just to other digrams beginning with the same letter, the resulting measure would be a considerably less distorted index of similarity to English than is predictability. Ranking all digrams would also lessen the dependence of digram frequency values on sample size, to which Baddeley rightly objects. However, to justify its use, any measure other than one based directly on a simple frequency count would need to produce evidence of its superiority over the latter.

A straightforward, unambiguous relationship exists between the sequential organisation of a language and measures direct of frequency values of letter combinations in a large sample of written English, This is so irrespective of the size of sequence (n-gram) whose frequency is counted. However, whereas a count of digram frequencies can be conveniently used for measuring second order sequential structure in letter lists, trigram and higher-order measures are less convenient for experimental use. This is partly because the proportion of possible trigrams and longer letter sequences which do actually occur in the English language is much lower than is the case with digrams. The approximation to English method can be used to make sequences in which long-range constraints produce a high degree of similarity to actual language structure. However, the fact that it is not possible to measure sequential structure in given lists by this method is a pronounced disadvantage. There is no certain criteria on which to decide whether a measure of second order similarity to English, such as digram frequency, is adequate for the task of investigating effects of sequential structure. There are probably effects of sequential redundancy in language which cannot be detected by a second order measure, but it is unlikely that second order effects will be qualitatively different from those occurring in letter sequences which are more highly organised. A number of existing experimental studies of the effects of letter sequence redundancy have used digram frequency as the measure of redundancy. Others have used predictability, which is calculated from a measure of digram frequency. In the present investigation digram frequency will be used as the measure of sequential organization.

SUMMARY OF CHAPTER 4.

The letter can be used as a unit for measuring sequential redundancy in language, and has some advantages over word units. These are described, the most important being that one can obtain direct counts of the frequency with which short sequences of letters occur in a language. This is not possible for sequences of words. There follows a survey of various measures of letter-sequence organization which have been used by investigators working on language analysis, and a detailed appraisal is made of one such measure.

CHAPTER 5

EFFECTS OF LETTER SEQUENCE REDUNDANCY UPON RECALL

5.1. Survey of findings

Sequential constraint in lists of letters has been found to affect accuracy of reproduction in a variety of recall tasks. Miller, Bruner and Postman (1954) constructed eight-letter lists of zero-, first-, second- and fourth-order approximations to English. They presented the lists tachistopically for durations of 10, 20, 40 or 500 milliseconds, and asked subjects to write down the letters they had seen, in the correct order. The number of words correctly reproduced varied directly with exposure duration and with order of approximation, and the authors conluded that the number of letters perceived depends on sequential redundancy. Baddeley (1964a) carried out a similar experiment, but which allowed adequate time for accurate perception of the letter sequences. The resulting relationship between redundancy and performance was almost identical to that described by Miller, Bruner and Postman. Baddeley pointed out that since both experiments required the subject not only to perceive, but also to remember each letter sequence, the effect which Miller, Bruner and Postman attributed to perceptual limitations is probably due to limits on memory. Although perception may not have been a limiting factor in the experiment by Miller, Bruner and Postman, the results of another study (I. Miller, 1957) show an effect of redundancy in word sequences on auditory perception. Lists differing in sequential organization (defined in terms of constraints imposed by certain rules) were presented against a background of auditory noise. Degree of correct letter recognition was positively related to sequential constraint.

McNulty (1965) constructed eight-letter sequences of zero-, third- and text-order approximation to English. The task was to learn lists of 16 sequences at each order of approximation, which were presented on a memory drum for three seconds per sequence. Four learning trials were given, and learning was tested by recall or recognition. The percentage of items correctly recalled was found to vary positively with order of approximation, irrespective of which measure of learning was used. This is an interesting result, since the use of a recognition test minimizes the possibility of reconstruction occurring at recall. However, since all the sequences in a list were presented before recognition of any of them was tested, the average retention time (around 48 seconds) was too large for close experimental control of subjects' operations, and reconstruction could have occurred at any of a number of stages within the trials. All the same, it is clearly helpful to use a recognition test to prevent construction at recall in attacking the problems presently under consideration. Miller (1958) constructed four- to seven-item strings of the letters G, N, S, and X. He used a finite state generator to select a subset of strings in which items had an average redundancy of 0.5 bits per letter, compared with two bits per letter for randomly selected sequences. Subjects studied nine lists for five seconds each, and then wrote down what they could recall. This procedure was repeated over ten successive trials. Redundant lists were learned much more quickly than random lists; after ten trials subjects could reproduce the nine redundant sequences perfectly, whereas an average of about 3.5 of the random sequences were correctly recalled.

In the experiments so far described in this chapter sequential redundancy has been provided either by using lists which were generated by human language users under varying amounts of contextual constraint, or by using rules to generate sequences having certain order contraints. Other studies have used either direct measures of letter-digram frequency in English as indices of second-order sequential constraint, or an index calculated from digram frequency. Such an index is Baddeley's (1963) measure of "Predictability" (calculated from the digram frequency count produced by Baddeley, Conrad and Thomson, 1960) which has been used in a number of experiments to demonstrate a correlation between sequential organization and performance in verbal tasks. Baddeley (1964b) presented subjects with pairs of CVC syllables, chosen so that the last letter in the first syllable and the first letter in the second syllable formed a digram which occurs frequently in the English language, but so that if the syllables were reversed, the corresponding letters formed an infrequent digram. An example given by Baddeley is QEM POG, in which MP is a frequent digram in English, whereas GQ never occurs. Seven of the eight pairs were learned more easily in the forward order, showing that learning of nonsense syllables is affected by the frequency in English of the inter-syllable digram. Baddeley suggests that nonsense-syllable learning should be regarded as "a motor-skill" involving transfer from S's vastly overlearned verbal habits". Baddeley, Conrad and Hull (1965) measured the immediate recall of six-letter consonant sequences which were presented visually, one letter at a time. They used Baddeley's "Predictability" index as their measure of sequential redundancy, and observed a significant correlation of +.30 between values of it and correct recall of the sequences.

Conrad, Freeman and Hull (1965) presented six-consonant lists visually, letter by letter. They found sequential constraint correlated significantly with recall, although recall was more strongly influenced by another variable, acoustic confusibility.

Direct measure of digram frequency in language samples (taken from the count produced by Underwood and Schulz, 1960) have constituted the index of sequential redundancy in a number of studies by M.S.Mayzner and his associates. Mayzner and Schoenberg (1964) used sequences of the nine most frequent English consonants, which were visually presented to subjects for six seconds per list. Written recall of six lists composed of eight high-frequency digrams was compared with recall of lists comprising the infrequent digrams. The average number of consonants recalled in their correct position was 7.08 for the high-frequency digram sequences and 5.41 for the low-frequency digram sequences, the diffeence being highly significant. An interesting finding from the detailed analysis of results was that four times as many of the mistakes in sequences of low-frequency digrams were inversions than was the case in lists containing high-frequency digrams. Mayzner and Gabriel (1964) presented subjects with a 4 x 4 display matrix, in which each of the 16 cells contained a letter digram. There were four presentation-times, ranging from 15 to 120 seconds. Both vowels and consonants were used in the digrams. One matrix contained digrams which occur frequently in English, and the other contained infrequent digrams. Average recall of a digram was significantly affected by its frequency in the English language, by position within the matrix and by the length of time for which the matrix was presented. At each presentation time more of the frequent than the infrequent digrams were correctly recalled. There was not a significant interaction between the effects of presentation time and digram frequency, but, as presentation time increased, the average difference in number of items recalled between high and low frequency digram conditions decreased from 2.00 to 0.60. This demonstrates, the authors suggest, that the effects of digram frequency are strongest for short exposure times. To determine the effects of between-cell digram frequency, Mayzner and Gabriel constructed two matrices which each employed the same digrams, ordered to yield high and low between-cell digram frequency matrices. Each matrix was presented for 120 seconds. The results showed that recall was positively related to between-cell digram frequency. The joint effects of digram frequency and organization in visual presentation were examined in a study by Mayzner and Adler (1965). They presented 12 consonant sequences in patterns of one, two and three letters per line. Both variables, organizational pattern and digram frequency, significantly affected recall. Sequences consisting of frequent digrams were better recalled than sequences consisting of infrequent digrams.

It is evident from the results described in the present chapter that sequential redundancy assessed by any of a number of methods, is a significant variable in tasks which require perception and subsequent recall of letter lists. As was the case with word sequences, the present concern is to attempt more precise description of these redundancy effects. The majority of the above findings are very recent, and at the time that the experiment to be described was performed, it seemed desirable to carry out an experiment investigating the effects of digram frequency on recall of sequences of letters.

5.2. Experiment 2.

The experiment by Mayzner and Schoenberg (1964), described above, demonstrated a large and significant difference between the recall of consonant lists high and low in digram frequency. Their design used the same nine consonants for all sequences; this has the advantage that factors such as confusibility and single letter frequency are automatically controlled. A possible weakness in the study is that only six lists were used in each condition, and it is possible, if unlikely, that the significant difference between conditions could have been due to very extreme degrees of recall in some particular lists. It therefore seemed advisable to determine whether or not the obtained difference between conditions was evenly distributed between all the lists within conditions. Another reason for carrying out an experiment ~~similar~~ on the effects of sequential constraints in retained letter lists was to determine whether effects of digram frequency differ between visual and auditory presentation, or between fast and slow presentation rates. It was decided to control confusibility between immediately adjacent items, i.e., the extent to which confusible letters such as CBPD occur together, since there were reasons for suspecting that this factor might have affected previous results. In fact a later experiment designed specifically to determine whether adjacent-item confusibility affects immediate recall showed no effects of this variable, which will not be discussed further. Generally speaking, the present experiment was designed as an exploratory one, to repeat that of Mayzner and Schoenberg with added controls, and to determine whether certain additional variables had any effects which might facilitate understanding of the mechanisms underlying human ability to use sequential redundancy in language.

Method

Subjects and Materials. There were 53 unpaid undergraduate subjects, who performed the experiment in four approximately equal groups. Twenty-four sequences were constructed, each using the nine consonants C, D, F, H, L, N, R, S, T. In twelve of the sequences ("L Lists") the letters were arranged to comprise digrams of low average frequency (M = 27, Underwood and Schulz, 1960[1]). The mean digram frequency in the other 12 sequences ("H Lists") was high (M = 180). As far as possible, letters were distributed equally between positions, and particular digrams were not used more than three times within a condition, nor were they used more than once in a particular position.

Design and Procedure. Subjects all attended two sessions at each of which they attempted to memorize 12 lists. Presentation was visual (letters projected simultaneously) in one session, and auditory, using a tape-recorder, in the other. Two rates of presentation were used. For the slow

1 Digram frequencies were read from the Underwood and Schulz (1960) count, following Mayzner and Schoenberg (1964). The count produced by Baddeley, Conrad and Thomson (1960) is probably a more accurate index of sequential redundancy, since it takes spaces between words into account. However, even using the Underwood and Schulz count, it is extremely difficult to construct lists within which all digrams are uniformally high or low in frequency, and it is usually necessary to include one or two frequent digrams in the low digram frequency lists and vice versa, ensuring that the discordant digrams are spread equally between positions within conditions. When it is additionally necessary to take into account the frequency of the digrams comprising the space and the first letter, and the last letter and space, it becomes even more difficult to make sequences of consistent digram structure (Dr. Ann Davies, personal communication, 1966) and the resulting necessity to include more discordant digrams in a list may nullify the advantage of taking into account space-letter digrams in the frequency calculations.

rate, auditory lists were played at the rate of 1.33 seconds per consonant and duration of visual presentation was 12 seconds per sequence. These figures were halved for the fast rate of presentation. Sessions were separated by four weeks. Two groups had visual presentation in the first session and auditory presentation in the second session, and this order was reversed for the other groups. Within sessions one rate of presentation was used for lists 1-6 and the other for lists 7-12, the orders of these rates being balanced among groups, and within each six-list batch three lists were H and three were L, ordered randomly. All four groups received the same 24 lists, but the use of a particular list among the auditory and visual modes of presentation and the two presentation speeds was balanced.between groups, so that each list was used for one group with each of the four combinations of presentation rate and mode.

Subjects wrote down what they could remember of a list immediately presentation ceased. They were told to recall letters as far as possible in their correct order and positions, on prepared sheets. Ten seconds were allowed for recall of a list, after which the next sequence was presented. Only letters recalled in their original positions were scored correct.

Results

The means for all subjects of number of letters correctly recalled per list in all conditions, together with their standard deviations, are shown in Table 5.

	Total Presentation Time (Secs.)	High Digram-Frequency		Low Digram-Frequency	
		Mean	S.D.	Mean	S.D.
Visual	6	6.37	1.35	5.94	1.65
Presentation	12	7.14	1.54	6.85	1.67
Auditory	6	5.39	1.39	5.33	1.49
Presentation	12	5.75	1.89	5.50	1.56

Table 5. Mean numbers and standard deviations of words correctly recalled per nine-consonant list in Experiment 2, as a function of digram frequency, presentation time and mode of presentation.

Using a t-test for correlated scores it was found that among lists presented visually more letters were correctly recalled from lists containing high-frequency digrams (H) than from low digram-frequency (L) lists in both the 6 sec. presentation condition ($t=2.69$, $p<.01$), and there was a non-significant difference in the same direction when presentation time was 12 seconds. ($t=1.94$, $p<.10$). More items were recalled from both H and L lists after 12 seconds than after 6 seconds presentation-time ($t=2.98$. $p<.01$ for H lists, $t=2.41$, $p<.02$ for L lists). With auditory presentation, H lists were significantly better recalled than L lists at the slow rate of presentation ($t=2.05$, $p<.05$), but not at the fast rate ($t=.07$). After auditory presentation, recall of H lists was better for the slow than for the fast presentation-rate ($t=2.23$, $p<.05$), but recall of L lists presented by tape-recorder was not significantly affected by rate of presentation ($t=.75$, $p\simeq.45$). To determine whether the superior recall of H lists was due to extreme recall scores in a small number of lists, scores for each list across subjects were analysed. The difference between H and L lists was significant ($p < .05$, Mann-Whitney U test) indicating that the main results were not due to extreme scores in particular lists.

Discussion

The general pattern of results is similar to that found by other workers, and varying time and mode of presentation does not produce any striking results. Compared with the findings of Mayzner and Schoenberg (1964), the effect of digram frequency in the present experiment is rather small. For instance, in the six-seconds visual presentation condition, which is similar to that used by Mayzner and Schoenberg, subjects in the present experiment differed in average recall between the two digram frequency conditions by 0.43 letters, compared with 1.63 in Mayzner and Schoenberg's experiment. Standard deviations do not differ markedly between the two studies, being 1.35 for H lists and 1.65 for L lists in the present experiment, compared with 1.43 and 2.09 reported by Mayzner and Schoenberg. A number of possible causes of the discrepancy were considered. For instance, it is possible that the difference in average frequency between the H and L lists was higher in the Mayzner and Schoenberg study than in the present one. Mayzner and Schoenberg do not supply the relevant figures on which to make a direct comparison, but the likelihood of such a difference being responsible for the discrepancy in results is contra-indicated by the fact that checking revealed no correlation between recall and digram frequency <u>within</u> the H and L conditions of the present experiment. Another possible explanation of the discrepancy is that there was no attempt in the present experiment to ensure that subjects actually wrote down the letters they remembered in the order of presentation. A subject could commence his written recall by writing down a letter presented at the end of a sequence and then proceed to the recall of earlier letters. However, the results of an informal class experiment in which subjects were told to write from left to right in the order of presentation provided no support for the hypothesis that order of writing letters in recall had contributed to the observed discrepancy.

In the light of later experiments, it seemed possible that a cause of the difference between the results of the present experiment and that carried out by Mayzner and Schoenberg was a non-equivalence in the inter-list interval between the two experiments. The length of this interval in Mayzner and Schoenberg's experiment is not given in their paper. An experiment was carried out in which ten each of the H and L lists from Experiment 2 were presented by tape at the fast rate, but the procedure differed from that of the main experiment in that the interval between presentation of lists was 25 seconds, instead of 10 seconds. The average number of items recalled in their correct position was 7.14 for the H lists, and 6.19 for the L lists. The difference is highly significant. Across lists, $p < .001$ (Binomial Test), and across subjects $p < .001$ (Mann-

Whitney U Test). Clearly, the observed average recall difference between H and L lists is nearer to that observed by Mayzner and Schoenberg, and this strongly suggests that length of inter-list interval is the variable leading to the discrepancy. However, Dr. M.S.Mayzner (personal communication, 1966) states that subjects in the Mayzner and Schoenberg experiment "were allowed ten seconds for recall, and then were immediately presented the next card", i.e., just as in Experiment 2. Thus the discrepancy cannot be explained in terms of the inter-list interval period. However, it is certainly interesting that this variable does influence the difference between H and L lists, although to provide a definitive account of the sources of interference which produce this particular result would require experimental investigation beyond that which can be attempted here.

To conclude the discussion of Experiment 2, the results provide evidence that digram frequency in letter sequences affects recall, but a discrepancy exists between the present results and those obtained in a similar experiment by Mayzner and Schoenberg (1964). It has not been possible to provide an adequate explanation for this discrepancy. This is an unsatisfactory situation, but further investigation along the lines of Experiment 2 will not be attempted, since such an appraisal does not seem as likely to lead to better understanding of redundancy effects as do experiments designed to examine verbal sequence redundancy effects in more detail.

5.3. Letter sequence redundancy in guessing and anagram tasks.

It was suggested earlier that evidence relating degree of organization in word sequences to recall provides rather limited assistance in the attempt to explain sequential redundancy effects, because the tasks in which the effects have been observed require the performance of a number of operations, in any of which sequential redundancy could be a limiting factor. Naturally, the same limitation applies to results demonstrating a correlation between letter sequence redundancy and recall. It is not possible with the available evidence, to detect any relationship between the similarity of experimental findings in different verbal tasks and the extent and nature of overlaps in the processes used in performance of the tasks. This sort of analysis might be expected to provide information about the specific location of redundancy effects within a recall trial, but, as was the case with the data presented in Chapter 2, either the overlaps between tasks are too great for this sort of analysis to be useful, or the tasks for which results can be compared are different in too many respects for useful information to emerge. Nevertheless, there is a range of experimental situations in which sequential redundancy has been found to effect results, and these will be briefly described, since they shed some light on problems of sequential redundancy in language.

In a number of studies it has been shown that sequential organization in letter sequences significantly affects performance at anagram tasks. For instance, Mayzner and Tresselt (1959) found that anagrams whose summed digram frequency totals were low yielded significantly faster solution times than anagrams whose summed digram frequency totals were high. In another experiment (Mayzner and Tresselt, 1965b) the authors selected anagrams whose solutions were words in which the average digram frequency was either low or high. The median solution time was 12 seconds when the solutions contained digrams of high average frequency, and 16 seconds for the anagrams with low digram frequency solutions. A finding by Hunter (1961) supports this result. He observed that when subjects were required to produce words from the letters R, T, A, they tended to provide the words in the same order (RAT, TAR, ART). Hunter suggests that digram frequency was one of several influences determining the order of word emission. Mayzner, Tresselt and Helbock (1964) required subjects to solve difficult six-letter anagrams. Subjects were asked to "think out aloud" in their attempts towards solution, and the letters were presented on wooden blocks, which subjects were encouraged to rearrange, this providing information about their strategies. The authors found significant correlations between the frequencies of digrams produced by subjects in their attempts to solve the anagrams, and the frequencies with which the digrams occur in English. Mayzner and Tresselt (1963) observed that solution times for anagrams were affected by the extent to which digrams tend to occur in the same positions in the correct word answers as they commonly occur in English.

The evidence from anagram experiments shows that human subjects have an awareness of the relative frequencies of digrams, and that they can use this knowledge in experimental situations. Subjects' ability to rank digrams for frequency in written English is also apparent in the results of another experiment (Mayzner and Tresselt, 1962c). They told subjects the first letter in a digram, and gave information about its position within a word, and the length of that word. Subjects were provided with an additional three letters, and were told to rank those letters in terms of the frequency with which they believed the three letters would follow the first letter in an English word. On about 30% of occasions the three letters were correctly ranked, which was significantly above the performance in a control condition where subjects were allowed to use their knowledge of single letter frequency.

There have been a number of experiments investigating the extent to which subjects can correctly guess letters deleted from English prose passages. Shannon (1951) observed subjects' performance at guessing a 129-word prose passage. His procedure was to ask a subject to guess the first letter in the passage. If the guess was correct, the subject was so informed, and told to guess the second letter. If the first guess was wrong, the subject was told the first letter and the proceeded to guess the next letter. This was continued until the end of the passage. As the experiment progressed, the subject was instructed to write down the correct text up to the current point, for use in predicting future letters. Shannon found that 69% of the letters were correctly guessed when this procedure was followed. Miller and Friedman (1957) used 300-word passages in which varying numbers of letter deletions had been made. Subjects could see the locations of deleted letters. When 20% of the letters had been deleted (in random positions) 95% of subjects' guesses were correct, but the level of correct replacement dropped to 55% when 50% of the letters in the passage were deleted. Chapanis (1954) carried out a similar experiment, but in which neither the number or the locations of the deleted words were provided. In this situation subjects typically made more wrong guesses than in Miller and Friedman's experiment. For instance, Chapanis found that when 10% of the letters were deleted from the text, only about 80% of them were correctly guessed. With 25% deletion, about 70% of the letters were guessed correctly, and only 10% were supplied in a 50% deletion condition. The most striking point about the results of all these studies is that guessing at deleted letters in English sequences is much more successful than guessing words. For instance, in a situation fairly similar to Chapanis' (1954) 10% letter deletion condition, the subjects of Aborn, Rubenstein and Sterling (1959), who attempted to guess one word deleted from 11-word prose passages, were only correct on an average of 47% of the occasions, compared with the 80% correct guessing of letters observed by Chapanis. The main importance of this difference is clearly that any attempts at reconstruction at recall is likely to be more successful in the case of letter sequences than in the case of word sequences, and hence attempts to reconstruct material at the time of recall would appear more likely to influence recall scores for letter sequences than for word sequences. It is not possible to predict precisely the effects on a recall task on the basis of a given degree of success in guessing experiments. This is as true for letter sequences as it was for word sequences (see Chapter 2), and for the same reasons. But the general statement that greater success in a guessing task would indicate greater probability of success at reconstruction in a recall task, is valid.

From the experimental evidence which has been presented in the present chapter it is clear that sequential redundancy affects verbal task performance in letter as well as in word sequences. The indications are that attempts to reconstruct material at recall are more likely to be successful when letter sequences, as against word sequences, are being reconstructed. There is some evidence, from the letter recognition experiment by McNulty (1965) described above (p.72), that sequential redundancy in letter sequences does affect the storage of sequences, but further experiments are necessary to provide more precise information relevant to the question of whether Deese's hypothesis holds good for letter sequences. Accordingly, experiments explicitly designed to differentiate between possible effects of storage and construction at recall, were performed, and these are described in the next chapter.

SUMMARY OF CHAPTER 5

A number of studies are surveyed, in which sequential constraint has been shown to be a determinant of performance in tasks requiring subjects to perceive, retain and subsequently recall letter sequences. Experiment 2 provides further evidence of such a relationship between sequential redundancy and recall. Performance at anagram tasks is affected by sequential constraint of the letters both in the anagram order and in the solution, and subjects' knowledge of language structure is also demonstrated by their ability to rank digrams in order of occurrence in English, and to supply letters deleted from redundant sequences.

CHAPTER 6

EXPERIMENTAL INVESTIGATIONS OF LETTER SEQUENCE REDUNDANCY EFFECTS

Four experiments are described in this chapter. All are concerned with the attempt to differentiate between possible effects of storage and of construction at recall in tasks requiring retention of sequentially redundant letter lists. The rationale behind two of these experiments is similar to that underlying Experiment 1. In the remaining experiments, retention is tested by recognition rather than by recall, the aim being to preclude the possibility of construction at recall, and thus to measure the effects upon storage, if there are any, of differences in sequential organization.

6.1. Experiment 3

As in Experiment 1, the aim is to differentiate between influences of sequential constraint effective (1) at recall, and (2) prior to recall, by introducing a variable to which it would seem likely that (1) but not (2) is sensitive. Subjects heard word sequences which differed with respect to digram frequency. Recall of these sequences was sometimes preceded by recall of other items. It was reasoned that the interference to retention caused by such additional recall might well depend on how items were stored, but that the influence of construction at recall would be practically unrelated to the effects of such previous interference. Thus it might be informative to compare the differences in recall between sequences high and low in digram frequency when these sequences are recalled immediately after presentation, with the differences in recall which occur when recall is preceded by the activity of reproducing other items. Is retention of sequences high and low in sequential redundancy differentially affected by interpolated recall? If the effects of such activity differ between levels of digram frequency in lists, differences in storage are likely. If recall of consonant sequences of different average digram frequency is equally affected by an additional task, there is no evidence from the experiment for digram frequency influencing the way consonant sequences are stored. As was the case in Experiment 1, where similar reasoning was applied to the attempt to provide precise description of sequential redundancy effects in word sequences, it is possible that letter lists differing in sequential redundancy are stored differently, but that the difference cannot be detected by the present experimental variables, at least within the values used in this experiment. For instance, it is conceivable that organized and random sequences are stored differently, and yet the effects of a distracting task upon retention of the two sorts of sequence are nevertheless equivalent. Consequently, although the observation of a difference between types of sequence being stored in the effect upon recall produced by a distracting task would positively indicate a storage difference, the absence of such a difference in the results between lists types would not necessarily indicate the absence of a difference in storage.

Method

<u>Subjects and materials</u>. There were 18 paid undergraduate subjects, tested individually. Lists containing the nine consonants C, D, F, H, L, N, R, S, T, were presented by tape recorder, one item per second. There were two types of list ("A" and "B"). In the A lists average digram frequency (Underwood & Schulz, 1960) within the trigrams formed by the first three and last three letters ("First and Third Trigrams") was low (M=6.25). In the B lists the first and third trigrams contained high-frequency digrams (M=207.76). The average digram frequency within the letters presented in positions 4, 5 and 6, (i.e., P2, R2, "Buffer Trigrams") was 31.15 in A lists and 30.00 in B lists. Sequences were constructed so that the average digram frequencies for letters in positions 3, 4 and 6, 7 were similar between list types, and any letter appeared equally often in each position.

<u>Design and Procedure</u>

Subjects listened to 20 A lists and 20 B lists. Recall of ten of each type of list was required in the order of presentation ("Forward Recall"). After the other lists the required procedure ("Reverse Recall") was to recall first the third trigram, then the buffer trigram, and finally the three consonants presented first. Each combination of list type and recall order occurred five times in both successive twenty lists. Six practice lists preceded the experimental sequences, which were presented in four orders, balanced among subjects.

After presentation of a list, one of two lights was illuminated to indicate recall order. Subjects dictated their attempts to the experimenter, who scored the number of letters which were recalled in their correct position. Presentation of a list began five seconds after a subject had finished his attempt to recall the previous one.

Results

The symbols P1 and P3 are used to represent the three first and the three last-presented letters (first and third trigrams), and R1 and R3 stand for recall positions. Thus BP3R1 represents a B trigram whose letters were presented in positions 7, 8 and 9 in a list, and recalled in positions 1, 2 and 3. Table 6 shows recall of all trigrams as a function of presentation and recall position, and type of list.

A graphic replication of the recall scores for P1 and P3 is provided by Figure 5, which shows clearly the overall effects of the experimental variables. At R1, P3 trigrams are better recalled than P1 trigrams, both in A lists (AP3R1 > AP1R1, $t=12.22$, $p<.001$) and in B lists (BP3R1 > BP1R1, $t=7.53$, $p<.001$), and at R3 the same is true for A lists (AP3R3 > AP1R3, $t=2.08$, $p<.04$), but the difference is not significant (although in the same direction) for B lists (BP3R3 > BP1R3, $t=0.45$).

Recall Position	Type of List	Presentation Position					
		P1		P2		P3	
		M.	S.D.	M.	S.D.	M.	S.D.
R1	A	1.50	0.50			2.79	0.28
	B	1.88	0.55			2.79	0.27
R2	A	Forward Recall		1.08	0.47		
	A	Reverse Recall		1.28	0.47		
	B	Forward Recall		1.29	0.50		
	B	Reverse Recall		1.54	0.49		
R3	A	1.10	0.36			0.80	0.46
	B	1.28	0.41			1.32	0.56

Table 6. Mean number of letters correctly recalled per trigram in Experiment 3. P1, P2 and P3 represent presentation positions of trigrams, and R1, R2, R3 denote positions in recall. In the A lists average digram frequency within trigrams P1 and P3 is high, in B lists it is low.

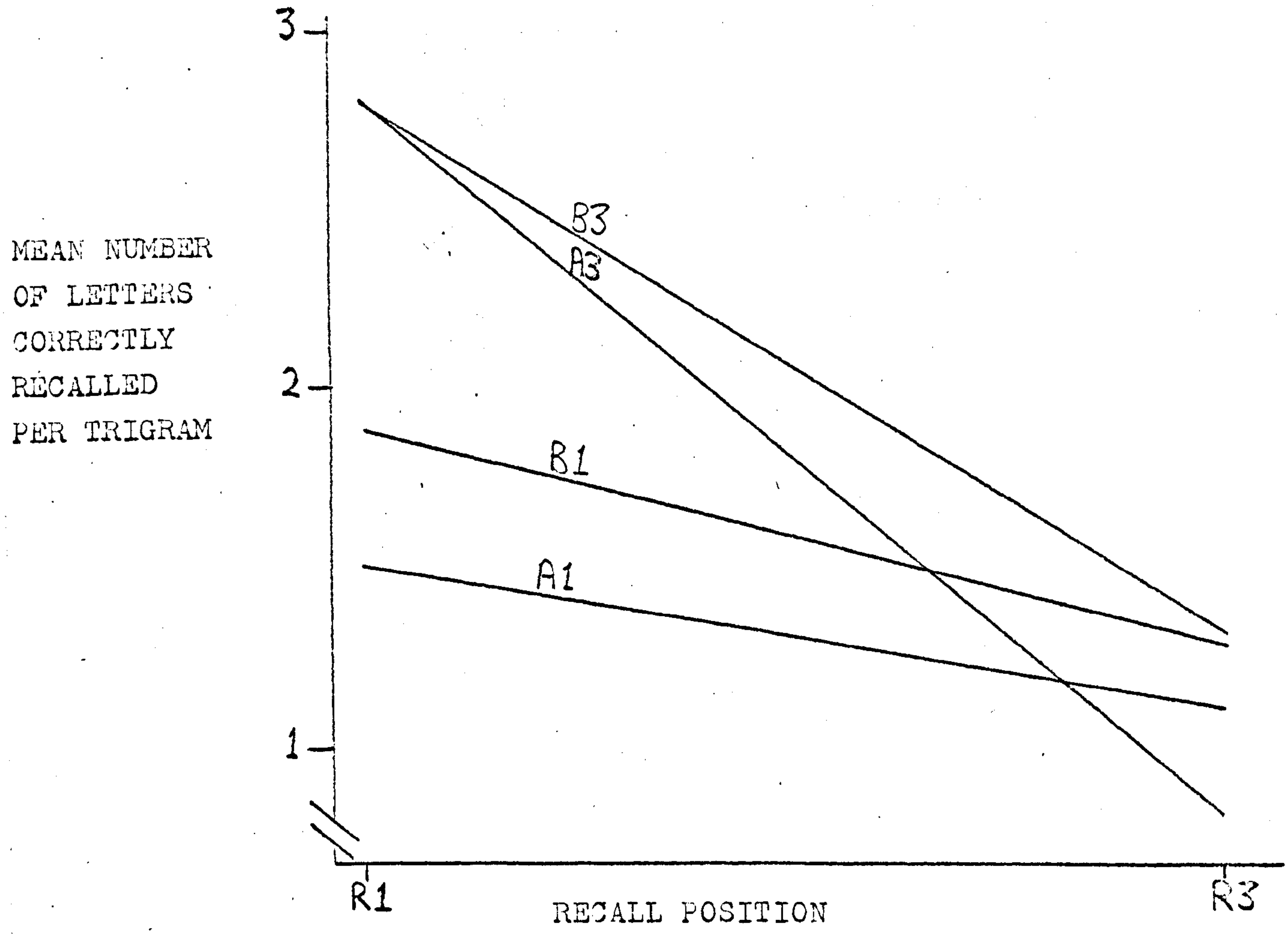

Figure 5. Average number of letters correctly recalled per trigram in Experiment 3, as a function of list type (A or B), presentation position (1 or 3) and recall position (R1 and R3).

Comparing recall of A with B trigrams, BP1 sequences are better recalled than AP1 sequences at R1 (BP1R1 > AP1R1, $t=3.96$, $p<.01$), but not significantly so at R3 (BP1R3 > AP1R3, $t=1.19$). Recall of AP3 and BP3 trigrams does not differ at R1, although BP3 trigrams are better recalled than AP3 trigrams at R3 BP3R3 AP3R3, t=5.88, $p < .001$).

The R1-R3 decrement is significantly influenced both by list type and by position in presentation of trigrams. For P3 sequences a greater decrement occurs in A than in B lists (AP3R1 - AP3R3 > BP3R1 - BP3R3, t=4.55, $p < .001$). However, for P1 trigrams the reverse is the case (BP1R1 - BP1R3 > AP1R1 - AP1R3, t=2.86, $p < .02$). Larger R1-R3 decrements occur with P3 than with P1 trigrams. This is true both for both lists (AP3R1 - AP3R3 > AP1R1 - AP1R3, t=11.80, $p < .001$) and for B lists (BP3R1 - BP3R3 > BP1R1 - BP1R3, t=7.53, $p < .001$).

Recall of buffer trigrams P2R2 is affected by both experimental variables. With reverse recall, buffer trigrams are better remembered when they occur in B lists (M=1.54) than in A lists (M=1.28, t=2.72, $p < .02$), and there is a non-significant difference in the same direction with forward recall (M=1.29 in B lists, 1.08 in A lists, t=1.40). In B lists, recall of buffer trigrams is significantly better with reverse than with forward recall (t=2.78, $p < .02$), and in A lists buffer trigrams are non-significantly better recalled when recall is forward (t=1.56).

		Forward Recall (i.e., P1,P2,P3).		Reverse Recall (i.e., P3,P2,P1).	
		M	S.D.	M	S.D.
Mean number of items recalled.	A Lists.	3.69	1.11	4.32	0.96.
	B Lists.	4.53	0.72	5.85	1.23

Table 7 Mean number of items correctly recalled per list (out of 9) in Experiment 3.

Average recall of the whole nine-consonant sequences (Table 7) is more accurate for B than for A lists, both with forward recall ($t=3.17$, $p<.01$) and with reverse recall ($t=2.71$, $p<.02$). Recall is better in reverse than in the forward order for A lists ($t=4.72$, $p<.001$) and for B lists ($t=5.09$, $p<.001$).

Discussion

Previous findings about the effects of digram frequency on immediate memory are confirmed, and effects of recall order on overall recall (see Chapter 10) can be seen to occur in the types of verbal sequence used in this experiment, items recalled first being recalled best.

There is significantly larger R1-R3 decrement in A than in B trigrams, and the previous reasoning would imply that this may indicate the existence of a difference in storage between A and B lists. However, there are three reasons why such a conclusion would not be entirely justified. First, the equivalence and near-perfection in recall at R1 of AP3 and BP3 trigrams suggests a ceiling effect, which would make difficult the interpretation of recall score decrements. Examination of the data revealed no definite evidence for a ceiling effect[1], but the

1 Evidence that the results were influenced by a ceiling effect would have been provided if it was found that those subjects whose recall in the P3BR1 situation was below average (i.e., subjects whose scores were well below 100% in the "easiest" condition) performed significantly worse in the corresponding a list situation (P3AR1). In fact, examination of the data showed that this is not the case, and therefore there is no definite evidence that a ceiling effect influenced recall scores in the present experiment.

possibility cannot be ruled out. Secondly, the decrement 3P1R1 - BP1R3 is significantly greater than AP1R1 - AP1R3, not less. Interpreting this finding is complicated by the fact that the presentation-recall intervals of the pertinent trigrams are filled not only by recall of two other trigrams, but also by presentation of them. A third objection is evident from examination of the recall scores of buffer trigrams. These scores are affected by the digram frequencies of the lists containing them. It is implicit from this that the distraction provided by presenting and recalling other trigrams cannot be regarded as equivalent between A and B lists. Similarly, proactive effects of presentation of P1 trigrams on recall of P3 trigrams may differ between list types.

It is interesting to note that recall of buffer trigrams is affected by the nature of the remainder of the lists in which they occur. It may be that the amount of capacity available for storing these trigrams is influenced by the information content of the lists. However, the present results do not show at which part of the total task storage of buffer items is differentially affected. Presentation, storage and recall of P1 and P3 can each interfere with retention of P2 trigrams, and one cannot say at which of these three stages sequential redundancy partly determines the distracting effect.

To sum up, the results of Experiment 3 suggest that there probably is a difference in the way B and A trigrams are stored. However, since a number of complications prevent unequivocal interpretation of the results, a further experiment was designed, in which the difficulty level was such that average recall in the easiest condition was well below 100%, and in which the distracting events occurring between presentation and recall were equivalent in all conditions.

6.2. Experiment 4

Method

Subjects and Materials. There were 53 unpaid undergraduate subjects, who were tested in three approximately equal groups. Lists containing the nine consonants used in the previous experiment were presented by ear at the rate of two items per second. There were two types of list. In the first ("H Lists") the digrams between items 5-9 (i.e., 5-6, 6-7, 7-8, 8-9) were of high average frequency (M=236), while in the equivalent part of the other ("L") lists average digram frequency was low (M=13.4). Average digram frequency within letters 1-4 was equivalent between types of list (M=35.3 for H lists, 35.4 for L lists) as was that of the digrams formed by items 4-5 (M=9.6 for H lists, 12.6 for L lists). No three letters occurred more than once. Letters were equally distributed between positions in lists.

Design and Procedure. Each subject heard eight H and eight L lists. Four of each were recalled in a "Forward" order, paralleling that of presentation. With the other ("Reverse") order subjects had to recall first letters 5-9, and then letters 1-4. Immediately after presentation of lists, one of two lights showed subjects which recall order was required. Both the first and second eight lists contained two of all the H-L/Forward-Reverse combinations, arranged randomly. Subjects were allowed 15 seconds to attempt written recall on sheets which contained separated spaces for the first four and last five letters in each list. Scoring was as in the previous experiment.

Results

Table 8 shows average number of letters correctly recalled as a function of list type and recall order. "H Sequences" (i.e. items 5-9 in H lists) were better recalled than "L Sequences", both with forward recall

Recall Order	Letters 1-4 (out of 4)		Letters 5-9 (out of 5)	
	H Lists	L Lists	H Lists	L Lists
Forward	2.63	2.34	3.18	2.20
Reverse	2.19	2.09	3.68	3.27

Table 8. Mean number of consonants correctly recalled in Experiment 4 as a function of type of list (H, L) and of recall order. Average digram frequency within items 5-9 is high in H lists, and low in L lists. Digrams within letters 1-4 are of medium mean frequency, equivalent between list types.

(i.e., pertinent letters recalled after items 1-4) and when the last five letters were recalled first (reverse recall) ($t=2.51$, $p<.02$ for reverse recall, $t=7.36$, $p<.001$ for forward recall). Both H and L Sequences were better recalled with reverse than with forward recall ($t=3.85$, $p<.001$ for

H Sequences, $t=4.99$, $p<.001$ for L Sequences). There is a greater recall difference between H and L sequences when they are recalled before than after letters 1-4, but the increase is not significant ($t=1.50$).

With both orders of recall, letters 1-4 were better remembered when they were followed in presentation by H than by L sequences ($t=3.77$, $p<.001$ with forward recall, $t=2.47$, $p<.02$ with reverse recall). Recall order significantly affected retention of letters 1-4 when the sequence 5-9 was H ($t=2.39$, $p<.05$) but not when consonants 5-9 made an L sequence ($t=0.82$). The effects of digram frequency within letters 5-9 on recall of letters 1-4 did not differ significantly between recall orders ($t=1.21$).

Discussion

The non-significant effect of recall order on the difference in recall scores between H and L sequences may point to a genuine population difference, since the t value of 1.50 approaches that required for significance at the .05 level, by a one-tailed test. However, even if a real difference exists, its smallness suggests that the cause is unlikely to be an important determinant of the difference in recall between digrams of high and low frequency in the present experimental situation. This does not imply that there are no differences in the way in which consonant sequences are stored which are important as determinants of recall, but merely that this experimental task provides no evidence for the contrary, positive hypothesis.

Thus the results of both the present experiment and Experiment 3 give some support to the proposition that there is a difference between consonant sequences differing with respect to sequential constraint in the way they are stored in an immediate memory task. A possible explanation for the lack of more definite evidence for storage differences is that the observed recall differences are caused largely or entirely by differential success of attempts to reconstruct at recall. However, as has been said previously, an equally valid explanation is that there are storage differences which are not clearly detected in the present experiments. Thus further investigation is required, combined with a new approach to the problem.

Before proceeding to describe further work on the main question it is interesting to consider a particular result obtained in Experiment 4. This is that letters 1-4 (which did not differ between conditions in average digram frequency) were better recalled when they were followed in presentation by H than by L sequences, with both recall orders. It will be recalled that in Experiment 3, retention of buffer trigrams (roughly equivalent to letters 1-4 in the present experiment) was affected by the average digram frequency within the remainder of the lists of which they formed a part. It was not possible to provide a precise interpretation of this result in Experiment 3, since the presentation-recall interval contained presentation retention and recall of other items. However, the situation in Experiment 4 is somewhat simpler. Considering only the forward recall conditions, it is seen that recall of consonants 1-4 is affected by the sequential redundancy of the items presented later. Since items 1-4 are the first to be recalled in the forward recall condition, any differential effects of recalling H and L sequences cannot influence recall of items 1-4. Therefore the difference between H and L sequences in their effects on recall ~~either~~ of the first four items can only be due to factors operating either at the time of presentation, or in the retention period immediately succeeding presentation. This finding is in accord with a suggestion made by Moray (1965, personal communication) that there is a central processing pool of fixed capacity but whose functions are interchangeable. Posner and Rossman (1965) have shown that the distracting effects of a numerical task upon the short-term retention of digits depends on the capacity (defined in informational terms) required by the distracting task, and the present finding provides supporting evidence that rather subtle differences between distracting materials can influence immediate memory. In Experiment 4, using more of the total capacity available for perception (and/or initial storage) of items 5-9 leaves less available for storing letters 1-4. Since a sequence containing frequent digrams is less redundant than one made up of infrequent digrams, one might predict that coding H sequences in the period immediately succeeding presentation would require less capacity than coding L sequences, and hence interfere less with retention of the first four letters, leading to the present result. It is not clear at exactly what stage in the coding process the presence of some letters interferes with retention of others. It is possible that the activity of perceiving items has an interfering effect, related to their sequential redundancy, upon the storage of other items. Alternatively, it could be that the differential interference does not occur until the period during which the distracting items are themselves being stored. An experiment was carried out to investigate these alternatives. This will be described in summary, since the problem is tangential to the main line of enquiry. Seventeen paid undergraduate subjects, who performed the experiment in three groups, each listened to forty 12-item lists. They were told that the first six items in the lists had to be recalled later. The remaining items were to be written down as they were heard, but were not required again. Rate of presentation was 40 items per minute, this being the fastest speed at which all subjects could write down letter or digit items as they heard them. In the lists, items 1-6 were either consonants or digits. The remaining items were always consonants, which were taken from a different vocabulary to that used for the first six items. Thus a letter which appeared in the first half of a list could not appear later in the same list, or in the second half of any other list. The procedure for subjects was to listen to the first six items in each list, which were later to be recalled, then to write down the subsequent letters. The letters in positions 6-12 made up digrams of either high or low average digram frequency in English, and the question which the experiment was designed to answer was whether the sequential redundancy of these items, which subjects merely had to write down, was a significant determinant of recall of the other items. As soon as presentation of a list was completed, and the necessary letters had been recorded, subjects had to write what they could remember of items 1-6. Digits were better recalled than letters, as might be expected from previous experimental results. The degree of sequential constraint within the second half of a list did not affect recall of the early items. Thus the influence of sequential organization upon the distracting effects of items presented, noted in Experiments 3 and 4 was not observed in a situation where the distracting sequences did not themselves have to be recalled. It would seem, therefore, that the stage at which sequential redundancy in material being processed first affects the capacity for storing other verbal materials is subsequent to that of initial perception. Coding of items proceeds far enough for them to be identified and written down, before their inter-item redundancy begins to use capacity at a level utilized by short-term retention. It is possible that sequential redundancy in the distracting material is important only when the distracting material has itself to be stored. Knowledge of the precise effects of material which interferes with recall, and of how these effects are partly determined by the nature of the interfering material and by what the human processer has to do with that material, is clearly pertinent to an understanding of the mechanisms of short-term human memory. This is an interesting field for further research, but not one which can be considered in any greater detail here.

Returning to the main problem, there remains a good deal of uncertainty about the possible roles of storage and reconstruction as contributors to the observed effects of sequential redundancy on recall of letter sequences. The reasoning behind Experiments 3 and 4 was similar to that underlying Experiment 1, which was designed to investigate letter-sequence effects. It will be recalled that Lachman and Tuttle (1965) used recognition tests to examine the effects of word-sequence redundancy on memory when recognition is precluded. The experiment to be described next, and the one immediately succeeding it, use recognition tasks to test memory for letter sequences.

6.3. Experiment 5

If subjects are presented with letter sequences, which differ with respect to sequential organization, and are then given a recognition test instead of one of free recall, there is hardly any opportunity for reconstruction to occur at the time retention is tested. Hence, if it is found that when a recognition test is used items in sequentially constrained lists are correct more often than items in random lists, it would seem that a greater number of the items in the constrained than in the random lists were adequately stored. The necessary assumption that the mechanisms underlying recognition memory are not markedly different from those underlying retention when tested by recall, is probably justified. (See, e.g., Wickelgren, 1966.)

Method

Subjects and Materials. Twelve unpaid subjects, all undergraduates or post-graduate students were tested individually. Each subject attempted

36 lists. The items to be retained consisted of three consonants which were typed in red capitals on a 3" by 5" card. The consonants were taken from the vocabulary C, D, H, L, N, R, S, T. In presentation of each list, the first card was followed by four others. On each of the second, third and fourth cards there were three consonants, one each of the letters B, F, G, J, K, M, P, V, Z. The fifth card contained a four-choice recognition test, which consisted of four of the letters from the vocabulary used on card one, typed in red. These four letters included one which appeared on the first card, and the subject had to identify that letter. The experimentally manipulated variable was digram frequency within the consonants on the first card. In 18 of the lists digram frequency within these sequences was high ("H"). (M=417, Underwood and Schulz, 1960.) In the remaining 18 lists average digram frequency within the corresponding ("L") sequences was low (M=13.72). The frequency and position of particular consonants within the vocabulary used for the first cards was controlled between H and L conditions, as was the frequency of particular letters on the cards containing the four-choice recognition tests. The number of times that a particular letter appeared on both of cards one and five (i.e., was the "correct" letter in the recognition test) was equivalent for H and L conditions, and the positions of such letters on cards one and five were also balanced between conditions. The orders of use of particular letters and positions were random. The letters on cards two, three and four were used in a different random order for each list.

Procedure Briefly, the task for subjects was to turn over the five cards which were used for each list, one at a time, and to say which one of the four letters printed on card 5 had also appeared on card 1. On seeing the first card in a list, subjects had to speak aloud the sequence of three letters, and repeat the sequence twice more, so that each of the three letters had been recited three times. Subjects then turned over the card, exposing card 2, the contents of which also had to be spoken aloud three times. This procedure was repeated with cards 3 and 4, after which turning the fourth card revealed the recognition test. Subjects then had to tell the experimenter which one of the four letters on that card had also appeared on card 1, and they were told to guess if no item was recognized. Presentation was self-paced, the subjects being told to proceed through a list as quickly as was possible without stumbling over words. Nine practice lists were given to enable subjects to become familiar with the procedure of speaking letters aloud quickly, and turning over the cards. It was believed that the nature of the task would preclude rehearsal during presentation, and that self-pacing would enable each subject to proceed at a rate which was slow enough to prevent stumbling and confusion, while being sufficiently fast to prevent surreptitious rehearsal. Answers by subjects to informal questions suggested that this aim was successfully achieved.

The 18 lists in each condition were ordered randomly, except that there were at least four H lists and four L lists in each successive groups of nine lists. Subjects were presented with packs containing the cards required for nine lists, and they were told to work through each pack at their self-determined maximum pace. A blank card separated each list. There was an interval of about one minute between each group of nine lists.

Results

Table 9 shows the average number of items correctly recognized,

1. Mean number of items (out of 18) correctly recognized.		
	H sequences	L sequences
M	14.16	13.75
S.D.	2.56	4.07

2. Mean number of subjects (out of 12) correctly recognizing an item		
	H sequences	L sequences
M	9.44	9.16
S.D.	0.97	1.60

Table 9. Mean correct recognition of consonants in Experiment 5 as a function of average digram frequency (H=high, L=low) of the three-item sequences in which the consonants were presented.

both across subjects and across conditions. Since four-choice recognition tests were used, the chance level of correct response is 25%. Thus chance factors alone would lead to subjects guessing correctly an average of 4.5 items, and an average item would be correctly recognized by three subjects.

It is clear from Table 9 that there are no significant differences in recognition scores between conditions, either for items across subjects, or for subjects across items. However, the non-significant differences are in the direction consistent with better recognition of the H sequences.

Discussion

Since the present results provide no definite evidence for a storage difference corresponding with the difference in sequential redundancy, they are not inconsistent with the reconstruction at recall hypothesis which Deese (1961) put forward. It is rather surprising that consonants from H sequences were not more frequently recognized than items presented in L sequences, in view of the previously mentioned result obtained by McNulty (1965). He found that when a number of letter sequences were shown to subjects, and were later presented simultaneously in a recognition test, the probability of a sequence being correctly recognized was positively related to its sequential redundancy. The present experimental situation, in which sequences are presented and subsequently tested for recognition, one at a time, differs in several respects from that devised by McNulty, but none of the differences would immediately lead one to predict the present experimental result. An important difference between McNulty's study and the present one is that the former used all letters in the English alphabet, whereas only consonants were presented for the present experiment. It seems likely that the difference in amount retained between sequentially constrained and random lists is greater when all letters are available for use in experimental lists than when only consonants can be used. However, when memory is tested by recall (see Experiment 4, 6.2.) retention of items from three-consonant lists does differ significantly between constrained and random sequences, and it remains to be explained why a similar difference is not apparent when memory is tested by recognition. A suggestion made by a subject is that it is easy to imagine two-syllable words from three consonant sequences. The ease with which this could be done would not necessarily be related to the degree of sequential constraint within the three consonants. It has also been suggested that many three-consonant sequences suggest associations, irrespective of sequential redundancy within the sequences. These factors might have the effect of modifying, or even nullifying the effect of sequential redundancy on memory for consonant lists as short as three letters. However, recall of three-consonant lists _is_ affected by sequential redundancy, and there is no reason why the factors just mentioned should discriminate between recall and recognition memory in their effects.

Thus Experiment 5, by not providing definite evidence of a storage difference, gives indirect support to the hypothesis that constrained sequences are better recalled than random sequences because of reconstruction which occurs at the time of recall. Interpretation of the results is made difficult, however, by the fact that the actual recall difference which the present experiment was designed to explain is a very small one, and hence effects of experimental variables may easily be obscured by uncontrolled sources of variance. It could be argued, for instance, that the present result is in a direction consistent with a storage explanation, but because of the weight of variance from uncontrolled sources relative to the weight of the experiment independent variable, a significant difference between conditions would occur only if a sample much larger than the present one were used. An experiment was designed to remedy this situation.

6.4. Experiment 6.

The present experiment makes use of the assumption implicit in Experiment 5 that the possibility of reconstruction of verbal material is minimized when retention is tested by recognition, rather than by recall. However, whereas the method used in Experiment 5 required single letters from the original sequences to be tested for recognition the present experiment allows the whole sequence to be presented in the recognition test, so long as the "incorrect" sequences in the recognition test have equivalent sequential organization to the original sequence. A novel feature is that instead of presenting a sequence and subsequently testing how well it has been retained, the difficulty level of the task is set so that the subject can usually recognize the material correctly, What is measured in the present experiment is not the degree of correct recognition, but the time subjects take to perceive, code and retrieve letter sequences. When this measure is adopted it is possible to test retention very shortly after presentation, without having to make use of any distracting tasks in order to produce forgetting, and without having to use sequences sufficiently long to ensure that recognition attempts are frequently incorrect. An advantage of using relatively short sequences combined with a short presentation-retrieval interval is that opportunities for subjects to adopt idiosyncratic strategies are minimized, the experimenter having tight control over the experimental task. The use of latency measure for coding and retrieval combined with recognition tests, gives two advantages over conventional scoring methods. First, the difficulty of deciding how to score recall is eliminated. It is not necessary to decide whether to measure, for instance, the number of letters correct irrespective of order, or only items recalled in the correct position. Questions which bother the researcher on information processing, such as whether an inversion should count as one error or two, or how to score a group of items recalled in their correct sequence, but displaced in position, no longer exist. Secondly, when a conventional test of immediate memory is used there is a possibility of results being distorted by certain threshold effects. A subject whose performance just reaches a certain criterion may obtain disproportionately higher recall scores than a subject whose performance level is just below that criterion. The possibility of distorted retention scores occurring in this way does not exist when time for correct retrieval is measured, rather than the amount of the material which is correctly remembered. It must be noted that the use of a response latency measurement for retention should be restricted to situations in which the percentage of items correct for each of the different conditions remains near 100%, or in which it is relatively equal between conditions (Keppel, 1965). Otherwise, as Keppel and Underwood (1962) have pointed out, measures may be biased by subject selection, owing to the fact that subjects who get least items correct are given less weight in determining mean scores.

In this experiment a subject looked at letter sequences which were typed on cards. The subject then turned over the card containing the sequences, to reveal a second card on which were typed two sequences, one of which was identical to the sequence on the first card. The task was to sort the second card into one of two piles according to the position of the sequence recognized as being the one on the first card.

Method

Subjects and Materials Twelve postgraduate students and research workers volunteered as subjects. The materials used in the recognition tests consisted of eight-letter sequences, which were typed on 3" by 5" white cards. These will be called "Item Cards". Sixteen sequences constructed by McNulty (1965) were prepared at each of the first and third orders of approximation to English. There were a further 15 sequences (constructed by Miller, Bruner and Postman, 1954) at the second order of approximation to English. The cards were put together in three packs, each of which contained all the sequences at one order of approximation. Within a pack, between each of the cards containing the sequences to be recognized, there was another ("Test") card, containing a recognition test. The packs thus contained an equal number of item cards and test cards, ordered item card, test card, item card, test card, and so on. On each test card were typed two sequences, one above the other. One of the sequences (the "Correct" sequence) was identical to that on the immediately preceding item card. The second ("Incorrect") sequence on the test card was identical to the other sequence except for one letter. For instance, the sequence on the item card might be S, T, A, N, U, G, O, P and the test card might contain S, T, A, N, U, G, O, P (correct) and S, T, A, N, U, G, I, P (incorrect). The incorrect sequences were constructed so that the letter that was different to the one used in the preceding item sequence did not change the sequential structure of the list. Thus the letters changed in first order sequences were replaced by letters of approximately the same frequency of occurrence in English (Underwood and Schulz, 1960). For the second order lists, a letter appearing in an item card but not in incorrect sequence on the succeeding test card was replaced by a letter which would retain the digram structure of the test sequence. Likewise, in constructing the incorrect sequences to succeed third order items cards it was ensured that the new trigrams in the sequences were of approximately the same frequency as those they replaced. For the first and third order lists the incorrect test sequences were those constructed by McNulty (1965). For the second order sequences, incorrect test sequences were made by the present author. The location of the changed letter varied equally among the eight possible positions, each being used twice, in random order, in a pack containing 16 item sequences. Within each pack the original sequence appeared on the text card eight times in each of the top and bottom positions, also ordered randomly.

The three orders of approximation to English made three experimental conditions. Since the experimental task required subjects to sort the cards as well as to perceive and recognize them, another condition was used in which subjects performance was measured at the sorting task alone. The materials for this sorting task consisted of a pack containing 16 pairs of cards. Typed on the first card of a pair was a single X, and on the second card there were two X's, one above the other. There was a red circle round one of these two X's, and the task was to sort the cards into piles according to the positions of the encircled X's.

Procedure A subject was given a pack containing all the item cards at one order of approximation to English, interleaved with their test cards. He was told to turn over the cards, one by one, putting each item card onto one pile after inspecting them, and then placing each succeeding test card onto one of two piles according to the position of the correct sequence on that test card. The subject was told to work through each pack as fast as was consistent with his subjective feeling that "around 90%" of the test cards were being sorted correctly. Results of both a pilot study and the full experiment showed that subjects were able to comply with this instruction fairly accurately. The experimenter measured the time taken by a subject to sort each pack, and made a note of the number of cards which were incorrectly sorted. Four packs were used in the main part of the experiment, one for each order of approximation, plus one pack (described above) which subjects were required only to sort, no recognition being required. The subjects each sorted the four packs in a different order, each pack being used in a particular position within the sequence of four by at least two, but not more than three subjects. Within packs, the pairs of cards were presented in a different order for each subject.

In order to familiarize subjects with the procedure, and to reduce practice effects, each subject sorted two practice packs. The first of these consisted of the sorting task alone. The second pack also required recognition, and was identical in form to those used in the three order of approximation conditions. It contained eight item-test card pairs consisting of sequences at the first order of approximation, and eight fourth order pairs, all arranged randomly. The item lists used in this practice pack were those constructed by Miller, Bruner and Postman (1954).

Results

Figure 6 shows the average time per pair of cards for the experimental task, as a function of sequential organization. Included on the figure is the time for the simple sorting task. The results of an analysis of variance on the data used for Figure 1 are given in Table 10. Both subjects and approximation to English are significant sources of variance.

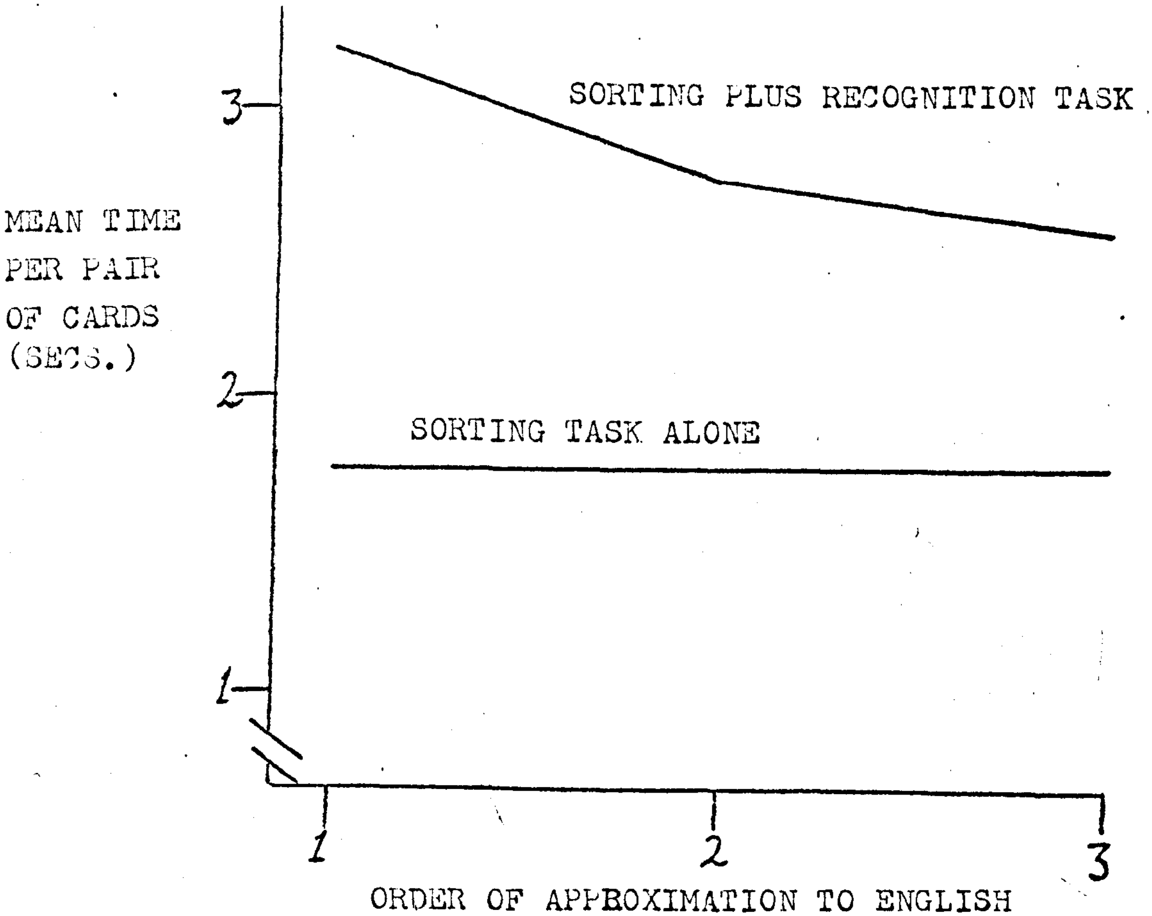

Figure 6. Mean times for the recognition task in Experiment 6. The letter sequences were first, second and third order approximations to English. The horizontal line shows the average time per pair of cards for the sorting task alone.

It was found that the average time per pair of cards in both third and second order sequences were significantly shorter than the average time for the first order sequences ($t=2.97$, $p < .01$ for the difference between

Source	Sum of Squares	d.f.	Variance	F	$p<$
Subjects (10)	4777.1	9	530.8	18.9	.001
Orders of Approximation (3)	526.9	2	263.5	9.4	.01
Residual	503.4	18	27.9		
Total	5807.4	29			

Table 10. Analysis of Variance for the recognition times in Experiment 6.

first order and third order; t=+.12, $p<.001$ for first order and second order), The difference between average times per pair of cards between sequences at the second and third orders of approximation was not significant (t=1.40, $p>.1$).

Table 11 shows the results of a second analysis of variance. The data for this were average times at each order of approximation, minus

Source	Sum of Squares	d.f.	Variance	F	$p<$
Subjects	1151	9	127.9	3.75	.01
Conditions	502	2	251.1	7.36	.01
Residual	613	18	34.1		
Total	2266	29			

Table 11. Analysis of Variance for treated recognition times in Experiment 6. Each subject's sorting time has been subtracted from his recognition time in each experimental condition.

average time for the sorting task. Thus, for each subject, the time for sorting alone is subtracted from the times for each of the whole recognition-and-sorting tasks. It is not really legitimate to refer to the results of these subtractions as times for the cognitive aspects of the task, but comparing such results for the different order of approximation conditions gives a useful rough estimate of the relative difficulties of the perceptual, coding and retrieval operations for the different levels of approximation to English. This analysis, like the previous one, shows subjects and orders of approximation to be significant sources of variance. An expected effect of the subtracting procedure is to cut down the amount of between-subjects variance (F=18.9 in Table 10, and 3.75 in Table 11), although this is still significant.

The average level of correct response was close to the 90% criterion which subjects were told to aim at. Subjects made an average 5.8 errors (σ=2.7) per 48 pairs of cards. Expressed as percentages, errors per subjects varied between 4.2% and 22.9%, the average being 12.1%. There was no significant relationship between number of errors and order of approximation, the average number of errors per pack being 2.1 at the first order, 1.5 at the second order, and 2.1 at the third order of approximation. However, in the present experiment there was a tendency for subjects who performed at a fast average rate to make more overall errors, than slower subjects, but it cannot be said whether or not this is a population effect since the observed rank correlation of -0.41 between average time per pack and total number of errors was not significant.

Discussion

The present results show very convincingly that sequential constraint in lists of letters affects the time necessary for coding and retrieval, even in situations where there is little scope for reconstruction at recall to improve performance. Thus, as was found with word sequences, redundancy does affect the retention of verbal material in the absence of active recall, and it is to coding processes which commence at perception and end at retrieval that further efforts towards more precise description of sequential redundancy effects may usefully be directed. It is possible that construction at recall may influence memory for letter sequences, since effects at recall might occur which would be additional to those on storage. However, what has been definitely established is that it would be useful to direct attempts to understand the mechanisms underlying sequential constraint effects towards the pre-recall stages in verbal memory tasks. One point worth noting is that establishing that differences occur in the absence of active recall does not rule out the possibility that retrieval is a critical stage. However memory is measured, the material has to be retrieved, and this is no less likely than any other process in the task to be directly influenced by sequential redundancy.

The sequences used in the present experiment have been made from a universe of all the letters in the English alphabet. It is difficult to determine the extent to which the result can be generalized to consonant sequences. However, it is difficult to imagine a mechanism by which sequential redundancy could influence the storage of lists which used all letters in the alphabet, but have no effect when consonants only are to be remembered. Reconstruction effects additional to the effects of redundancy on storage may contribute to the observed influence of sequential redundancy in recall tests, and if so, the relative potency of the two sources of constraint effects may differ between all-letter and consonant sequences. In principle it would be possible to use the present design for an experiment using consonant sequences only, but in practice it would be extremely difficult, if not impossible, to construct the experimental lists and the appropriate incorrect recognition items.

Comparison of inter-position error distributions between levels of constraint showed that reconstruction at recall is unlikely to be an important determinant of sequential redundancy effects on word recall (Experiment 1). However, it cannot be assumed that this conclusion is valid for letter sequences, since it has been found (5.3.) that it is easier to correctly guess letter units than words in redundant sequences. It seems more likely, therefore, that guessing at the time of recall may contribute to total reproduction of letter sequences more than is the case with word sequences. In principle, it would be possible to compare error distributions for letter sequences at different levels of approximation. Describing serial position curves for letters is less straightforward than for words, since the more restricted letter vocabulary results in a greater part of the letter errors than word errors being mistakes of position, which are less simple to indicate on a serial position curve than omission errors. Jensen and Roden (1963) and Jahnke (1963) have been able to produce serial error curves for letters, and it would be feasible to make the relevant comparison between curves for letter sequences of different orders of approximation to English.

It is not at present possible to state precisely what weight should be given to storage, and what, if any, to reconstruction at the time of recall, in leading to the total effect of sequential redundancy on immediate memory for lists of letters. However, it has been firmly established that at least a major part of the effect occurs prior to recall, which is a useful finding since, as has been suggested, it indicates the pertinance of research directed towards the facilitating effects of sequential redundancy upon memory mechanisms.

By using lists constructed by the approximation to English method, it has been possible in the present experiment to construct sequences with greater structural similarity to English than those made up from frequent digrams. A result of this is to increase the magnitude of the redundancy effect which is being examined, a clear advantage for experimental investigation, especially since none of the disadvantages of the approximation to English method are felt in the present experiment. However, it is noticeable in the results that the difference in time for the task between first and second order sequences is greater than the difference between second and third order sequences. Thus the effects of sequential constraint on recognition time would have been clearly apparent in the present experimental results even if none of the sequences used had been structured beyond the second order. Sequences of similar second order letter structure to English can be produced as easily by combining frequent digrams (using digram frequency counts) as by Miller and Selfridge's method.

The experimental technique used in Experiment 6 may possibly have further applications for the study of short-term memory. A possible disadvantage is the considerable variance produced by differences between subjects in ability to manipulate the cards in simple sorting. It is interesting to compare the results for different conditions when the time for the sorting task alone is subtracted from the time for the total experimental task, which requires subjects both to sort and to recognize items. However to equate the resulting measure with a "cognitive process" time which can be used in further quantitative analyses is not really justifiable, since this implies the acceptance of certain untenable assumptions about the additivity of times for the sorting and cognitive operations. The presence of this limitation need not lessen the utility of the technique; in the present experiment, for instance, a very clear result emerges from the data as measured for the entire task. There is no reason why all experiments on retention should have to make use of experimental situations in which considerable forgetting occurs, provided that one is sure what is being measured in a particular experiment.

SUMMARY OF CHAPTER 5

Experiments 3 and 4 compared the effects of distracting tasks on consonant lists varying in sequential redundancy. The results were consistent with the existence of pre-recall differences in the retention of differently constrained sequences. In Experiment 5 a recognition task was used to minimize the possibility of performance being affected by reconstruction at the time of recall; in this situation there was no consistent difference corresponding to degree of sequential constraint in recognition of items which had been presented in consonant sequences. However, the results of Experiment 6, which measured the time taken to perform tasks requiring letter sequences to be recognized and sorted, showed a clear and significant relation between order of approximation to English in the sequences and time to complete the task. This demonstrated convincingly that sequential redundancy in letter lists does affect retention at a stage prior to active recall.

CHAPTER 7

CONCLUSIONS TO SECTION 1

7.1. General implications of the present results

The experimental findings of Section 1 can be restated very simply. Sequences which are similar to language both in their basic units and in the structure by which those units are connected, can be better retained by human operators than sequences containing the same units ordered randomly. Deese's hypothesis concerning reconstruction at recall is substantially incorrect. It might be the case that in tasks requiring recall subjects can add to the difference in amount retained between organized and random verbal sequences, by using their knowledge of language to predict the non-retained parts of structured sequences when recall is in progress, but the present findings suggest that this is not the case to any great extent when words are taken as the basic unit of analysis, although such prediction might influence recall scores in tasks which require letter sequences to be reproduced. How far does this knowledge take us? In short, not very far. What is clear is that organized, and therefore redundant, language sequences are better retained than randomly ordered lists of the same units. There is a difference in the amount of material stored. It is clear that the nature of the mechanisms responsible for processing and short-term storage of verbal materials is such that their functioning is affected by sequential redundancy in the material. However, this knowledge does not tell us very much about memory for structured language.

For instance, it is not legitimate to deduce that the present results demonstrate the existence of a store, or memory, whose capacity must be described not only in terms of numbers of verbal units, but also in terms of the sequential redundancy of the material being stored. This is not so for the very reason underlying the present experimental study, i.e., a number of mechanisms which have functions other than storage are required in the performance of, for instance, a short-term memory task of the kind described by Miller and Selfridge (1950). Precisely what the mechanisms do, is not known, but it is clear that the raw sensory data must go through a number of different processes prior to recall in a memory task. In a sense, all these processes necessitate storage of the material. For example, even to carry out the function commonly called "perception", physical signals must be transformed into nerve impulses and then coded and retained in a form at which some sort of matching or "identification" can occur.

Thus owing to the necessity for a number of operations to occur between presentation and recall in a memory task, the finding that sequential redundancy is effective at a pre-recall stage does not add a great degree of precision to present knowledge. It is possible that redundancy has a direct influence at a relatively early stage. This would effect coding of the to-be-retained material at all subsequent stages, just as a horse who is slow in jumping the first hurdle in a race tends to remain behind the field for the rest of the race, if his speed, relative to that of the other horses, remains constant. It is possible that redundancy does not have a direct influence on processing until a relatively late stage in the total task; all that can be said with certainty is that the present evidence does not enable exact specification of the immediate effects of sequential redundancy.

7.2. <u>Some Further considerations</u>.

Work has been described which may be relevant to the task of detailing the intra-task location of sequential organisation effects. It is known, for instance, that in certain situations there is a visual limiting factor to the amount of information which can be stored over extremely short periods (up to about .5 seconds after presentation). It has also been found that auditory characteristics are a determinant of the amount of verbal material which can be retained for short periods under certain conditions. This sort of knowledge may be of value in future attempts to determine which of the brain processes used in memory tasks are affected by sequential organization in verbal material. However, the finding cannot be interpreted as evidence that all verbal material must go through a visual coding stage, and then an auditory stage, and so on. It is more likely that the processing stages which a given input of verbal material goes through are dependent not only upon the mechanisms of a multistage memory system but also upon the specific material which is being coded. In other words, what happens to the input depends partly on the nature of the input itself.

There is a limit to the extent to which experimental procedures can be used to isolate the various mechanisms used in the coding and retrieval of a message by the human operator, and this in turn may limit the value of additional work along the lines of the present investigation, directed towards precise description of redundancy effects. The present approach to understanding sequential organization in language has been fruitful up to the present, but it does not seem to promise unlimited further progress.

At present, numerous approaches to understanding the nature of serial organization in behaviour are being followed, and discussion has been limited to work directly relevant to one particular present approach. It would be foolish to attempt to predict which sort of approach will be the most fruitful, and real progress will probably depend on the combined findings of several approaches. One line of attack, the developmental study of children's progress in learning to deal with the rules of language, might be singled out as one which seems to offer considerable promise.

SECTION 2

SOME EFFECTS OF RECALL

UPON SHORT-TERM MEMORY

CHAPTER 8

RECALL AS AN INDEPENDENT VARIABLE

8.1. Introduction.

When a person recalls material which he has retained, the activity of recalling may itself reduce the capacity available for retention. This effect has been observed in a number of experimental studies, some of which will be reviewed in the present chapter. Brown (1958) has observed that when the interval between presentation and recall of a verbal sequence is partly filled by an additional activity, the accuracy with which the sequence is reproduced suffers. It is therefore not surprising that when the whole of a sequence has to be reproduced the activity of recalling that part of the sequence which has first to be reproduced may affect retention of the rest.

The recall process can be varied in many ways, and method of recall has often been used as an independent experimental variable. The experiments to be described in Section 2 are mainly concerned with two related dimensions of variability in the recalling activity, namely the proportion which has to be recalled of the total material presented, and the order in which the individual items are reproduced, relative to their order of presentation. As was the case with Section 1, a major aim is to determine the extent to which measured retention depends on factors influencing storage of verbal materials, and the extent to which factors effective at the time of recall influence the results. At this stage, the problems to be considered may be indicated most clearly by giving a concrete example. Anderson (1960) presented to her subjects lists of 12 digits, one per second, by ear. Recall followed immediately after presentation of a list. In one condition subjects were requested to recall any one of the first second, or third successive four-digit sequences constituting a digit list (i.e., digits 1-4, or 5-8, or 9-12). Subjects did not know until after presentation of a list which part they would be asked to recall. All parts were requested equally often, at random. Accuracy of recall of the four-digit parts averaged about 77%. It is reasonable to deduce from this figure that immediately prior to recall, around 77% of the items in each list were retained correctly. Anderson's subjects received another condition, identical to the previous one except that recall of all 12 digits in a list had to be attempted. From the previous result it would be expected that 77% of the 12 items would be correctly reproduced; however this was not so. When subjects were required to attempt recall of all the digits in a list, recall dropped to 64%. Thus proportionately less of the items were reproduced when recall of all the digits was attempted than when only some of them was required.

It would seem that in the process of recalling some of the items, Anderson's subjects were interfering with their own retention of the remainder of a list. The greater the proportion of the total sequence which has to be recalled, the greater the disrupting influence on the remainder, with the result that accuracy of reproduction is inversely related to the number of items recalled. It is possible that the operations carried out by a subject at recall in retrieving some of the contents of memory may interfere with that part of the stored material which is yet to be retrieved. Alternatively, recall may interfere with retention in a less specific manner, either by generally raising the noise level of a part of the system which is not solely concerned with memory but which is nevertheless essential for short-term retention, or by causing a shift in attention which results in information processing capacity which is required for storage becoming unavailable. Another possibility is that active recall may lead to loss of retention because of the time recall requires; by preventing rehearsal, recall might leave the system open to a process of spontaneous decay.

Whatever the nature of the mechanisms underlying the above recall effects, the effects are themselves of definite importance. Recall has been used as the method of testing what has been retained in the majority of experimental investigations of short-term memory, and if the accuracy of reproduction is considerably influenced by certain recall variables, then it is clearly desirable to know exactly what the effects of these variables are. Satisfactory explanation of why recall variables influence retention is also to be desired, but at this stage the more pressing need is to know precisely what the effects are.

The importance of knowledge about the influence of recall variables is also evident from the point of view of measurement in short-term memory. Decisions about the most appropriate way of measuring a quantity in psychology can influence the observed distribution of experimental results. For example, in Chapter 1, above, it was mentioned that the measured effects of sequential redundancy in word-order approximations upon recall can differ according to the unit of measurement chosen. Such a finding would normally be regarded as an irksome added difficulty to the experimenter, although in one study (Moray and Barnett, 1965) it was found that information provided by comparing the results obtained on the same performance by different indices of retention could be used to clarify the understanding of human performance in a memory task. It is well known that estimates of retention based on recall are generally lower than those based on recognition. Explanation of this particular finding has been the object of a good deal of experimental investigation (See, e.g., Davis, Sutherland and Judd, 1961; Postman and Rau, 1957.) and will not be discussed here. Again, what is especially important is that the choice of method of assessing what is retained ~~may~~ in an experimental study may effect not only the level of observations, but also the distribution of results between conditions. For instance, Lachman and Feld (1965) found that at early stages in a word learning task recognition estimates of retention were much higher than estimates based on free recall, whereas at later stages the reverse was the case. Thus the observed relationship between retention and number of previous trials differed according to whether recall or recognition was used as the index of retention; what is observed can depend on how it is measured.

The relevance to questions of measurement of the experiments to be described in the succeeding chapters may be seen by considering the following example, taken from physics.

> "Let us attempt to measure the momentum of a free electron by means of some instrument. The detector of this instrument uses the electron to trigger some macroscopic event, and in doing so, reduces the energy of the electron to the point where it is captured and becomes indistinguishable from the electrons already in the material of the detector. Thus the free electron has disappeared in the act of measurement." (McKnight, 1959).

The gist of this account can be summarised by saying that the activity of measuring a phenomenon has affected the phenomenon being recorded, producing an error in the observation. Now reconsider the previously described experiment by Anderson (1960). What occurs in her investigation is broadly similar to the effect described in the above example. In recall, that is in the process of measuring what they have retained, Anderson's subjects influence the quantity which they are attempting to record, namely the storage capacity. The activity of recalling is producing a sort of measurement error. This has certain implications for the study of short-term memory. For instance, if, as Anderson's results indicate, what is recalled in a memory task is not always synonomous with what has been retained, one needs to know how can the capacity of the memory system for retaining information be most effectively measured. The existence of problems such as this indicates that a study of recall variables is relevant to the measurement of memory, as well as constituting a field of empirical enquiry in its own right.

The subsequent chapters contain descriptions of a number of experiments which investigate effects of certain aspects of recall on retention. Section 2 clearly differs in emphasis from Section 1, although some of the problems encountered are common to both. Postman (1964) has pointed out that retention is a theoretical construct, and not a simply observable and therefore measurable quantity. It is not possible to measure exactly how much material is being retained at a particular time, by a process of simple observation. Retrieval is necessary, and as has been shown, the process of retrieving some of the material, and/or processes associated with doing so, have the additional effect of reducing the capacity to store the remaining material. Questions can be raised about the use of constructs such as "memory" and "retention". Should they be defined in terms of what is stored, in which case they may not correspond with what can be measured, or should they be defined in terms of what is observed by a measuring device to have been stored? In the latter case the definition must include description of the way measurement is made, since a value is not independent of this variable. At this stage, however, an experimental study of certain effects of recall on retention is required. This will be preceded by a brief review of the evidence on the effects of varying a) the proportion of the stored material to be recalled, and b) the order in which recall is required, in studies of short-term memory.

8.2. Partial versus total recall.

There have been a number of experiments in which it has been found that correct recall of part of a sequence of items which a subject is attempting to remember is proportionately more accurate than recall of all the items presented in the sequence. For instance, Brown (1954, Experiment 1) used a memory-drum to present sequences of four displays which each consisted of two digits and an arrow, at a rate of .78 seconds per set. Subjects had to read the numbers aloud as they appeared, at the same time signifying that they had seen the arrows by drawing appropriate lines. Each arrow could be one of four types, which varied in size and direction. Among the several conditions in which directions for recall succeeded presentation of the material, there were some where both arrows and digits had to be recalled, and others in which the items in only one of these categories were required. Recall was written and Brown found that when only the digits were required (partial recall), 51% of them were recalling correctly, whereas only 37.5% of the digits were reproduced when recall of both arrows and digits was required (total recall). In another experiment Brown (1954, Experiment 2) subjects were presented with sequences made up of two pairs of digits followed by a single digit, succeeded by two pairs of letters, which were followed by a single letter. Recall of both letters and digits was more accurate in partial than in total recall conditions. Recall of digits dropped from 48.9% with partial recall to 40.6% with total recall, and the corresponding values for letters were 74.2% and 51.5% A similar increase in efficiency with partial recall, compared with total recall, was observed by Moray, Bates and Barnett (1965). They presented lists of letters over two, three or four spatially separate auditory channels. Their technique, for example in the four channel condition, required four loudspeakers to be placed at the corners of an imaginery circle drawn in a horizontal plane around a subject's head. Each loudspeaker transmitted all the signals from one channel. After presentation, subjects were asked to recall either all the items used in a list (total recall), or just those transmitted over one particular channel (partial recall). The authors found that with partial recall subjects' performance at reproducing the letters was proportionately much better than when the items presented over all channels were required. In the most difficult condition, which was when messages were presented over all four channels, the proportion of required items correctly recalled with partial recall was 25% higher than when total recall was attempted although the particular channel from which the letters were to be recalled was not specified until all the items had been presented. Harrison (1964) presented twelve-word lists by tape-recorder at a rate of one word per second. In every list the items making each succeeding group of four words (i.e. words 1-4, 5-8, 9-12) were the entire members of a category, the categories being (a) Gospels, (b) playing card suits and (c) the four main compass-points. For total recall all twelve words in a list had to be dictated in order of presentation, and in the partial recall condition, all the four words in just one category were required, in order. Harrison found that about 75% of the required words were correctly recalled in order when partial recall was requested, compared with only about 50% when all items were required.

Thus there is evidence from a number of sources to show that the proportion of correct items recalled in verbal short-term memory tests may depend on how many of the presented items are required. Klemmer (1961) compared partial with total recall in a non-verbal memory task. He used a visual display which took the form of four simultaneously presented seven-line matrices, each in the form of a figure eight. Patterns were produced by randomly selecting any combination of lines from each matrix. Total recall took the form of subjects attempting to duplicate all four stimulus patterns on a specially prepared answer sheet, which contained dotted outlines of the figure eights used as stimuli. In the partial recall condition, a cue provided .2 seconds after presentation indicated which one of the four matrices was to be reproduced. Accuracy was then about 89%, compared with 75% when total recall was attempted. Thus an improvement in accuracy of reproduction with partial over total recall seems to occur when non-verbal materials are being retained, as well as with verbal items. However, Klemmer's finding may only apply to retention of matrices over extremely brief periods. His stimuli were presented for .04 seconds only, followed immediately by the cue for recall, so it is possible that perception was a limiting factor, and that after-images subserved by a relatively peripheral neural mechanisms were responsible to the facilitation found with partial recall. Other investigators (See, for example Averbach and Coriell, 1961.) have shown facilitating effects on the proportion of the required stimuli correctly reproduced of partial over total recall after storage for periods of less than .5 second, and in the situations described the information is probably not processed to any extent by the subjects' central cortical mechanisms. Posner (1963) suggests that experiments such as those reported by Averbach and Coriell (1961) and by Sperling (1963) indicate the existence of "a rapidly decaying immediate memory system peripheral to a limited capacity perceptual system" and it is possible that Klemmer's results are also due to events operating at stages of coding earlier than those which are necessary for storage over the periods common to short-term memory experiments.

It is clear that, at least with verbal materials, the probability of correct recall of an item which has been stored over a short period is dependent on the proportion of the total stored items which are required for recall. An additional finding of the previously mentioned experiment by Anderson (1960) was that when eight digits were to be recalled out of the twelve presented, accuracy of recall was lower than when only four digits were required, but higher than in the total recall condition. Thus the difference between partial and total recall is not independent of the proportion of the total number of presented items which is required for partial recall.

8.3. <u>Order of recall</u>.

Of equal importance to the finding that partial recall is more accurate than total recall is the fact that the number and location of correctly recalled items from a sequence is affected by the order in which recall is attempted. This was observed by Kay and Poulton (1951) whose subjects had to remember sequences of eight displays, which each consisted of two illuminated arrows. Each arrow could point in one of four directions. Two four-way joystick keys, one for each hand, were used for recall, so that subjects could indicate the directions of the arrows which had been presented. Rate of presentation was two seconds per display, and subjects had to move the keys at an equivalent rate during recall. Kay and Poulton found that when sequences were recalled in order of presentation, the first four items were more accurately recalled than the remainder (36% and 31% of displays 1-4 and 5-8, respectively, being recalled. However, when the last four items in presentation were the first to be recalled, accuracy of recall did not differ between the two halves of a sequence, being 30% for displays 1-4, and 30% for 5-8. Thus, on average, 33.5% of the items forming an entire sequence were correctly reproduced when order of recall was identical to order of presentation, against 30% when the other recall order was used. Lawrence and Laberge (1956) drew attention to the fact that accuracy of recall of particular items in a sequence is affected by their position in the recall order. Other things being equal, items recalled first tend to be recalled best, a finding which the results of partial-total recall comparisons might lead one to predict. Lawrence and Laberge presented pairs of what are known as "Wisconsin Sorting Cards". On each of these cards there are between one and four shapes in any one of four colours. There are four different shapes, all the figures on one card being identical in shape. Thus the displays have three dimensions, number, shape and colour, each of which can take one of four values. In Lawrence and Laberge's study, subjects were shown two cards, projected onto a screen for .1 second. In some conditions subjects were then asked to report what they had seen, one dimension at a time. For instance, they might be told to recall first the forms of the shapes on each card, then the colours, and finally the numbers of shapes. In this sort of situation, used in conjunction with an earlier instruction to pay equal attention to all dimensions during presentation, accuracy of recall averaged 84.8% for the first dimension to be recalled, 76.3 for the second, and 73.7 for the third. These results are supported by some obtained by Haber (1964) in a very similar experiment. Haber found that the mean accuracy of recall was 85.9% for the dimension recalled first, 83.1% for the second dimension to be recalled, and 79.6 for the third dimension.

Broadbent (1957, Experiment 1) examined recall order effects in an experiment which used verbal materials of a sort frequently found in experiments on immediate memory. Subjects heard three pairs of digits, each pair consisting of the simultabeous presentation of one item to each ear. In recall, subjects had to reproduce all the items heard in one ear, and then the items presented to the other ear. The order of ears for recall was specified after presentation. The result was that in the early stages of the experiment, subjects reproduced significantly better the material presented to the ear whose contents were recalled first. However, when the experiment was continued on a second day, this effect disappeared. Broadbent attempted an explanation in terms of his model of how subjects were able to adapt their strategies, and thus obtain this practice effect, but it is not entirely clear how it came about. Insufficient data is provided by Broadbent to enable precise comparison between the amount of practice subjects received in his experiment and in experiments reported by other workers, in which no decrease was observed in recall order effects with practiced subjects. However, Broadbent's observation of an effect of practice on the distribution of results serves as a reminder that a particular task may not be performed in an identical manner by all subjects or even by the same subject on different occasions, a fact which adds to the difficulty of attempting to understand the mechanisms underlying human memory.

Another experiment in which order of recall was found to be important has been described by Rabbitt (1962). He presented sequences of five cards, one at a time, for .75 seconds each. Every card contained a single letter, printed on one of two colours, red or black. After presentation, subjects were requested either to recall first the letters in order, and then the colours, or vice versa. Rabbitt found that both letters and colours were best recalled when recalled first. Accuracy of recall for letters was 70% in the first recall position and 62% when they were recalled second, and the corresponding figures for colours were 79% and 76%. Posner (1964) presented eight-item digit lists by ear at one of two speeds, 30 or 90 items per minute. Subjects had either to recall the lists with items 1-8 in order of presentation ("Forward" order), or to recall first digits 5-8, and then 1-4 ("Reverse" order), the numbers 1-8 standing for positions of digits in the presentation of a sequence. With both recall orders, and at each presentation rate, less errors were made in the four digits recalled first than in the remainder of a list. At the slow presentation rate accuracy of recall of a list as a whole was higher with the reverse than with the forward order of recall, but at the slow rate there was no significant difference between recall orders in the total number of items correctly recalled.

The effects of varying recall order when words are used as the items to be remembered have been investigated by Murdoch (1963). Lists containing six pairs of words were presented visually, at a rate of two seconds per pair. After presentation, one member of each of three of the pairs were shown, one at a time, and subjects had to recall the other items in those pairs. Since recall order could be manipulated by the experimenter, it was possible for Murdoch to determine probability of correct recall as a function of both presentation position and position in the recall sequence. On average, it was found that the percentage of words correctly recalled was 41% for items recalled first, and 31% and 20% for items recalled second and third respectively.

The results surveyed up to now in the present chapter can be summarized by the statement that partial recall tends to be more accurate than total recall, and recall position is a determinant of the relative accuracy of recall of the various parts within a sequence which a subject is trying to remember. The first-recalled parts are reproduced most accurately. Anderson (1960) has found that accuracy of partial recall depends not only on the number of items required, but also upon the position within the sequence in which the items were originally presented. For instance, when four out of twelve digits were requested for recall immediately after presentation, the average level of recall was around 85% for items in presentation positions 9-12, 66% for items 5-8, and 63% for the digits presented in the first four positions in a sequence. It is not surprising that accuracy of partial recall is not equivalent between the different parts of a list, since in a number of investigations of short-term memory it has been found that accuracy of recall varies within sequences, and that not all the variability can be explained in terms of interference due to recall. For sequences just above the memory span the typical serial error distribution takes the form of a curve which is highest for the middle items in a list, and lowest at the beginning and end. For instance, Jensen and Roden (1963) investigated the effects of several variables upon the skewness of serial position curves for errors in short-term recall of nine-item lists of colour forms. The items were visually presented one at a time at a rate of one stimulus every two seconds. Overt errors were about 8% for items 1-3, 10% for items 7-9 and about 15% for the middle three items. This is a fairly typical finding, and although many factors contribute to influence the precise shape of error curves in immediate recall trials, no exception can be taken to the generalization that there are intralist differences in the accuracy with which the various items are recalled. A number of factors may combine to produce forgetting between presentation and recall of a particular item. Among the most important are probably (a) interference due to the activity of recalling other items (and/or the time taken to do so) and (b) interference due to the activity of presenting other items (and/or the time taken to do so). Tulving and Arbuckle (1964) carried out an experiment which was explicitly designed to compare the effects of presentation position and recall position upon recall of items. Sequences of nine to 16 paired-associate items were presented visually on cards at a rate of two seconds per item. Presentation and recall orders were manipulated so that it was possible to score recall as a function of presentation position and recall position. The authors were interested in the relative potency of the effects of n "Inputs" (meaning the number of items presented between the presentation and recall of a particular item) and n "Outputs" (recalled items). The time factor was controlled by pacing recall at the rate used for presentation. It was found that the probability of correct recall of a single paired-associate item presented in a particular position in the input sequence was higher following two interpolated outputs than following two interpolated inputs. In other words, the presentation of subsequent items interfered with retention of a particular item more than did the recalling of an equivalent number of items.

Discussion of recall order effects is continued in Chapter 9, where further experimental evidence is provided.

SUMMARY OF CHAPTER 8

The activity of recalling material in short-term memory can affect retention of other stored items. This must be considered when memory is being measured. Studies are reviewed in which order of recall is manipulated, and in which the proportion to be recalled of the stored material is an independent experimental variable. Recall tends to be most accurate when not all the presented material has to be reproduced, and items tend to be recalled best when they are recalled first.

CHAPTER 9

AN INVESTIGATION INTO THE EFFECTS OF VARYING RECALL ORDER

9.1. Some effects of varying recall order.

An interesting result which occurred in some of the studies mentioned above (Kay and Poulton, 1951; Brown, 1954; Murdoch, 1963; Posner, 1964) is that when all the items presented in a sequence have to be recalled, and order is varied, total recall is affected by order of recall. None of the findings previously described would lead one to predict this. Murdoch (1963) has attempted an explanation. He observed that with six-item paired-associate lists recall of other items had the greatest interfering effect on the retention of the pair presented in serial position six, progressively less effect on retention of pairs in serial positions five and four, and for all practical purposes, no effect at all on retention of the remaining pairs. His explanation was that the effect of interpolated recall on retention of an item is a function of the probability of recall of that item. Thus when the probability is low, one interpolation will have little or no effect on retention. As a formal explanation, this account leaves something to be desired, since the question "Why?" remains unanswered, albeit at a slightly changed level of generality. However, Murdoch's statement is useful if it correctly indicates a general characteristic of retention. Tulving and Arbuckle (1963) have produced an explanation which is both different in form and contradictory in content to that of Murdoch, Tulving and Arbuckle visually presented lists of ten paired-associates in which the stimulus items were always single digits, and the response items were common nouns. Their experiment was designed so that a pair presented in a particular position could be required for recall in any one of the ten recall positions, and thus it was possible to determine the probability of correct recall of an item as a function of any combination of presentation and recall positions. Items presented early tended to be correctly recalled, as were items which were early in the recall sequence. The most interesting finding was that recall of items presented early in a sequence was less affected by their position in the recall sequence than was recall of items presented at positions late in the sequence. This was so to a larger extent than in Murdoch's (1963) study, with the result that whereas in early recall positions items in the final positions in the sequences were more often correctly recalled than items presented at the beginning of sequences, in late recall positions items presented early were better recalled than those presented late in a sequence. This finding is contradictory to what would have been predicted from Murdoch's statement that the effect of interpolated recall on retention of an item is a function of the probability of recall of that item. By no stretch of the imagination can the finding that the items presented in some positions in the presentation sequence are more accurately recalled than other items in the same sequence if required early, but less accurately recalled if required late, be regarded as consistent with

Murdoch's generalization. Tulving and Arbuckle interpret their results differently. They say that,

> "Items learned early in the input sequence are relatively impervious to output interference. Whether such items are recalled early or late in the output sequence seems to make little difference to their availability. Items in the middle positions in the input sequence.....are also relatively little affected by output interference, although there is a trend toward slower recall with increasing position in the output sequence. The greater the number of attempted recalls that interfere between the input and output of such items, the smaller is their availability at the time of recall".

The importance of Tulving and Arbuckle's results lies in the demonstration that as soon as a list has been presented there may be differences between items in the extent to which storage of them can resist interference by distracting events. There are differences between retained items in the degree of consolidation in memory. Immediately presentation of a list is completed, additional to the differences between items in the probability with which they will be correctly recalled, there are differences in resistance to additional cerebral activity such as is necessitated by having to recall other items. If this interpretation of Tulving and Arbuckle's results is correct there are important implications for the interpretation of results of memory experiments in general. For instance it seems likely that within a sequence of items presented and recalled in a test of short-term memory, either (a) not all the material retained is in the same "store" or (b) if all the items are in the same store, some of them are held there more efficiently than others, in which case the usefulness of a position by which memory mechanisms seen as including a number of discrete stores is decreased.

There is a difficulty in determining the relevance of Tulving and Arbuckle's findings, and the interpretation of them, to experiments more typical of those carried out in the field of short-term memory. The time between presentation and recall of items in the experiment which Tulving and Arbuckle describe varied between three and sixty seconds. This range is much wider than in most experiments designed to investigate the mechanisms of immediate memory, and the applicability of their findings to memory over shorter periods is therefore questionable. It would be useful to determine whether intra-list differences in consolidation are present within experimental trials of shorter duration. An experiment was designed which allowed systematic variations in order of recall within the framework of a situation which in duration and in the nature of the materials to be remembered was fairly characteristic of experiments which have been carried out in the field of immediate memory. It was considered that the information which it might provide on the effects of varying recall order would be useful; in particular knowledge about intra-sequence differences in consolidation, if they become apparent, would be valuable. Another reason for carrying out the experiment to be described was to determine which order of recall would be the most efficient from the point of view of accuracy of total reproduction. It would be useful to know how recall order strategies influence the efficiency of reproduction, for practical as well as for purely scientific reasons.

9.2. Experiment 7

The design of the present experiment was fairly similar to that of Anderson (1960). As in her study, presentation of a list was followed by the request to recall separately each third of the total number of items presented in a sequence. However, in the present experiment recall was usually required of all three thirds of a list, in any of the six possible orders.

Method

Subjects and Materials Twelve unpaid subjects, all of whom were either University undergraduates or research workers, were tested separately. Lists containing nine different consonants were presented by tape-recorder at the rate of one letter per second. C, N, P, S and W were not used, W because of the length of time required to pronounce it, and the others because of their high degree of confusibility (Conrad, 1964) with other consonants. Sequences of consonants occurring immediately adjacent in the English language, such as G, H, were not allowed, nor were sequences which made English words, for instance T, R, Y. Otherwise the consonants were ordered randomly.

Design and Procedure A battery of four lights was placed in front of each subject. Not more than one light was illuminated at any time. Three white lights in a line represented the different thirds of three letters which made up each list. For instance, if the list presented was H, D, X, M, K, Y, Q, F, Y the left light was a cue for recall of H, D, X, the right light requested Q, F, Y and the middle light, M, K, T. The fourth light, which was yellow, was placed below the line made by the others, and was used to indicate to a subject that no further recall of the current list was required. It was not used after lists for which all three of the upper lights were shown.

There were six conditions, representing the possible orders of recall of the three groups of three letters making each list. Thus, if order of presentation can be represented as 1, 2, 3, the numbers standing for the first, second and third parts of three letters, recall order could be any of the six combinations of the numbers 1, 2 and 3.

Subjects listened to each list of consonants. Immediately after the ninth letter had been spoken, one of the three upper lights was illuminated. This indicated to the subject which part of the list he was required to recall first. The experimenter recorded a subject's spoken report. If uncertain of a part, a subject was advised to make an attempt, unless he had no idea whatsoever of any of the required letters. If he knew some but not all of a part, the subject was asked to indicate the position of the letter being recalled, for example by saying "Blank, G, Blank."

Immediately after a subject had finished attempting to recall the first requested group of three letters, the first light was extinguished, and a second light turned on, indicating which part of the list was next to be recalled. If the subject had attempted no response at all to the first light, the lights were changed after three seconds.

After illumination of the second light, the third light might, in a sense, be considered redundant, since there was only one remaining part of the list to be required. In order to dissuade subjects from any tendency to make use of this knowledge, for example by quickly rehearsing to themselves the third to-be-recalled part of the list before dictating the second group of letters, the experimenter introduced an extra condition which was used randomly at four intervals in each trial of thirty lists. In this condition, instead of the third aligned light being illuminated, the yellow light was used, indicating that no further response was required for the current list. Subjects' responses in this condition were not used in scoring performance.

Each subject was presented with 60 lists in all, in two batches of 30 lists, which were separated by an interval of five minutes. There were additional pauses of approximately one minute after the fifteenth and forty-fifth lists. Each trial contained 24 test lists, four in each of the six conditions, arranged randomly. The remaining six lists in each trial were assigned to the pseudo-condition described above. Intervals of 15 seconds occurred between the recall of each list and the presentation of the next. Before the test trials there was a practice session which consisted of ten lists. The results indicated no significant differences in scores between the two test trials. Ten practice lists preceded the experiment.

<u>Scoring</u> Each group of three consonants was scored between 0 and 3, depending on the number of letters that were correct. To be scored as correct a letter had to be in the right position. If no attempt was made to recall any third of a list, that part was scored 0, and any subsequent parts of that list were not scored. For example, if the correct list in the order of recall was GMT LHR BZD, and a subject's attempt was G-H----BZD, the scores for the successive parts would 1/3, 0/3, 0/0. Thus attempts to recall the second- and third-requested parts of a list were only scored when they could legitimately be said to be subject to any effects of the recall of other part(s) of the list. In fact, when the above method of scoring was compared with one in which all attempts to recall any parts were scored, regardless of whether an attempt had been made to recall previously requested parts, differences in scores were small and not significant.

Results

Figure 7 shows the average number of consonants recalled out of a three-item group, as a function of presentation position. Irrespective of the position of a group in the recall sequence, items presented in the middle of a list are the worst recalled. This agrees with the observation by Jensen and Roden (1963), mentioned previously. Figure 8, which shows the average number of consonants recalled as a function of recall position, with position of presentation as parameter, presents the same data as that shown in Figure 7, but in a way which shows clearly the effect of recall position on accuracy of reproduction.

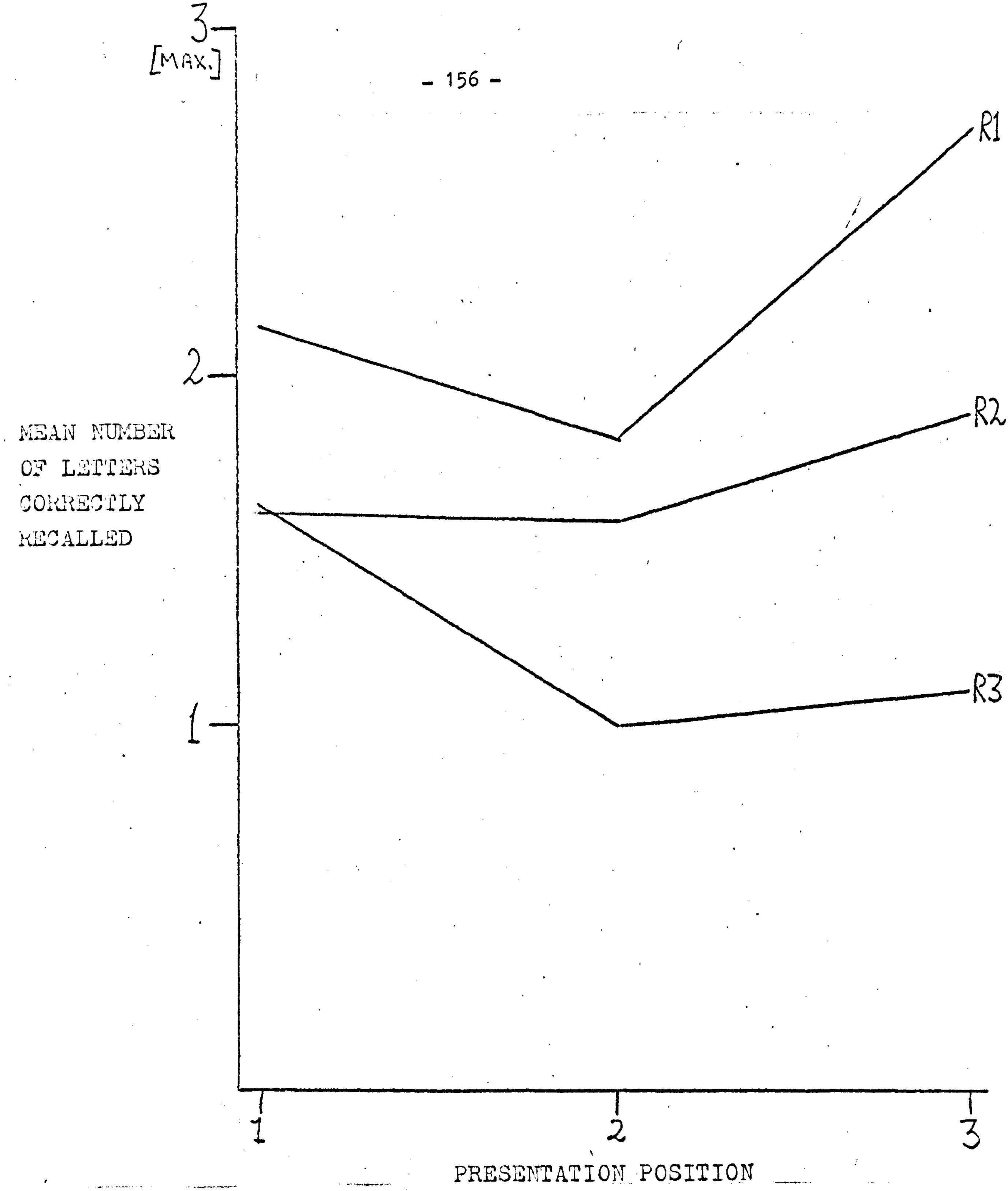

Figure 7. Mean number of letters correctly recalled per three-item group in Experiment 7, as a function of presentation position, with recall position as parameter.

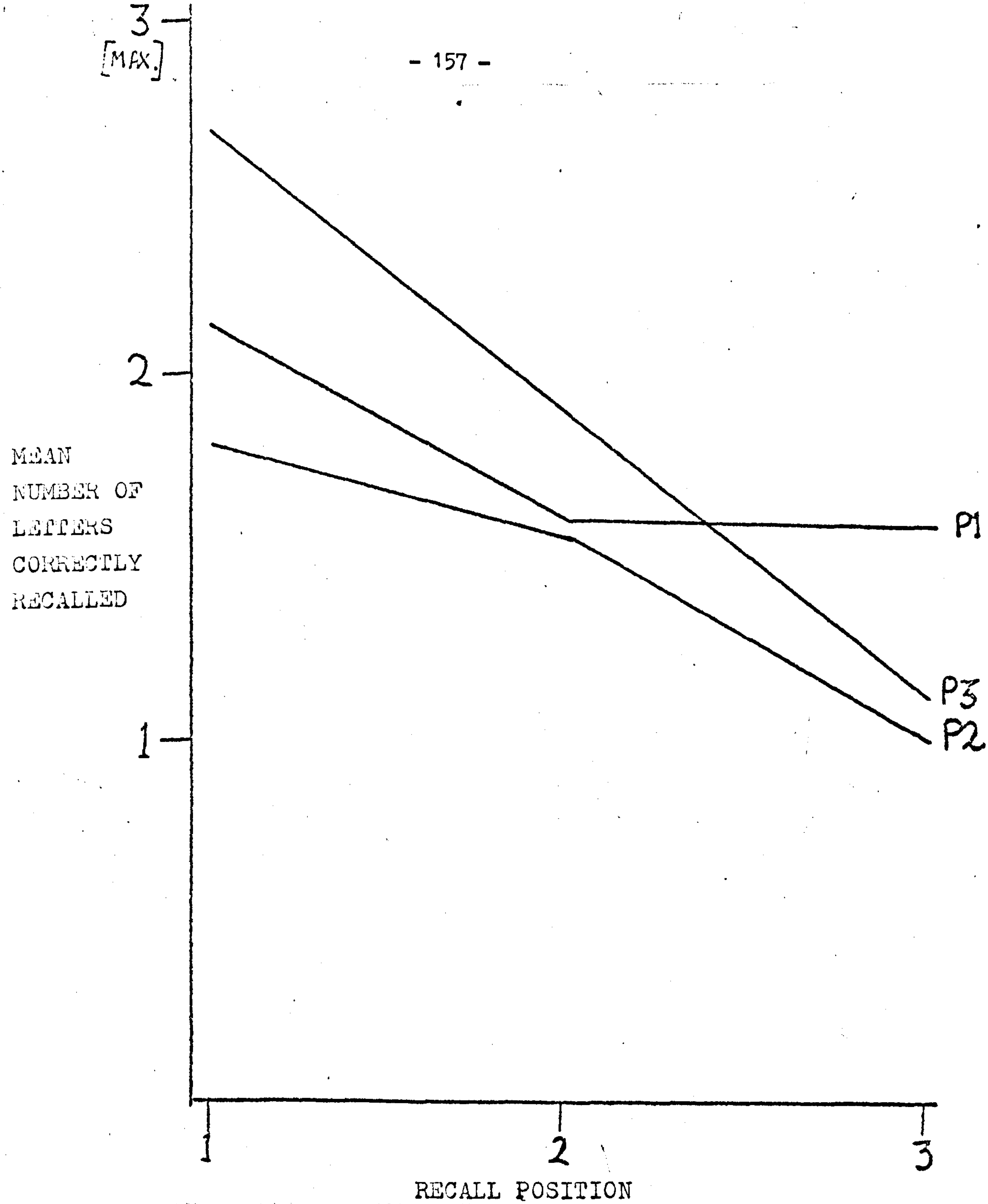

Figure 8. Mean number of letters correctly recalled per three-item group in Experiment 7, as a function of recall position, with presentation position as parameter.

It can be seen that at the first recall position the parts of lists which were presented last are easily the best recalled, but that recall of this part suffers most from interference when other parts of the list are recalled. The part of a list which is presented first, although less likely to be correctly recalled if required immediately after presentation of a list, is little affected, when being retained, by the distraction of having to recall the remainder of the list. The fact that immediately after presentation the most recently presented items are the best recalled suggests that presentation of the letters occurring late in a sequence has an interfering effect on retention of the early items in a sequence, as Murdoch (1963) observed. On the other hand, it seems that after presentation of a list is complete early items whose retention has survived the effects of presentation of the remainder of the sequence are more consolidated in retention, as indicated by their greater resistance to recall activity, than items presented later in the same list. The degree of resistance or consolidation is indicated by the slopes on Figure 8 showing decrement in recall of items from a particular position in the presentation sequence, with varying amounts of interpolated recall. In Figure 8, where P1-3 represent presentation positions and R1-3 represent positions of three-letter parts in the recall sequence, the difference in recall before and after recall of other items is significantly ($\underline{p} < .01$) less for P1 consonants (P1R3-P3R3) than the corresponding decrement for recall of the last three items (P3R1-P3R3). Similarly P2R1-P2R3 is significantly less than P3R1-P3R3 ($\underline{p} < .001$) and P1R1-P1R3 is less than P2R1-P2R3 (n.s.).

An analysis of variance was performed to determine effects of subjects and recall orders on total recall of the nine-consonant sequences. There was no significant subject effect, but order of recall was found to be a significant source of variance ($F=7.40$, $p < .001$). For the greatest accuracy in reproduction of lists as a whole, the most efficient strategy in the present experimental situation was to recall the thirds of a list in a reverse order to that in which they had been presented. Table 12 shows the mean total number of consonants recalled at each of the six recall orders.

Condition	Recall order	Average number of consonants correct (out of 9)
1	123	5.00
2	132	4.80
3	213	4.68
4	231	5.29
5	312	5.56
6	321	6.12

t-tests for small correlated samples indicate significant differences between:

6 and 1 ($p < 0.01$).
6 and 2 ($p < 0.01$).
6 and 3 ($p < 0.01$).
5 and 3 ($p < 0.001$).

Table 12. Recall of nine-consonant sequences in Experiment 7 as a function of the order in which the three-item parts of a sequence were recalled.

9.3. Discussion

It is quite clear that varying order of recall in the present experiment has profound effects. The results are generally in line with those described by Tulving and Arbuckle (1963) and show that the difference which they found between remembered parts of lists in degree of resistance to interference can be observed within sequences from which no items are stored for more than nine seconds. The crossover in Figure 7, from which it can be seen that P3 items are better recalled than P1 items at R1, whereas the reverse is the case at R3, indicates that Murdoch's (1963) explanation is not adequate for the present results. There is a definite difference between parts of a list in the amount of forgetting which occurs after presentation is complete and the simplest explanation of this fact is that early items are more consolidated in memory, and hence more resistant to distraction by the activity of interpolated recall. This may not be an entirely correct description of events, since the interpolated activity between presentation and recall of P1 items is not the same as that for P3 consonants. For instance, consider items recalled last (R3). If these items were presented first (P1) then recall of the consonants in parts P2 and P3 constitutes the interpolated activity. But if the items which are recalled last were also presented last, then the interpolated items will be P1 and P2. It might be the case that attempting recall of P1 in recall position R1 requires more or less mental capacity than recalling P2 or P3 in position R1, and hence has a more distracting effect on recall of other items. It is just possible that the time taken to recall items is affected by their position in the presentation of a list, and such a difference could conceivably influence the accuracy of reproduction of subsequently recalled parts. However, although there is some evidence (Experiment 4) that the distracting effect of interpolated recall is a function of certain aspects of the interpolated material, it seems unlikely that effects of the above type could alone account satisfactorily for the fact that the R1-R3 decrement for P3 items is almost three times as large as that for P1 items, particularly since half of the interpolated recall material (the P2 consonants) is equivalent between the two conditions. A test would be to compare the effects of interpolating recall of P1 and P3 on reproduction of P2 to determine whether recall at P2R2 is affected by the presentation position of items previously recalled (at R1). Examination of the experimental data showed that there was no difference in recall of P2 items between the two relevant conditions, and it can therefore be concluded that the observed difference in R1-R3 decrement between P1 and P3 items reflects a genuine difference in the extent to which those items were consolidated on completion of presentation of the nine-consonant sequence. Thus the observation that immediately after presentation of a sequence of items in a short-term memory trial, differences may exist within a list not only in the probability with which particular items will be correctly recalled if required, but in the degree to which stored items are resistant to subsequent distracting events, is valid for the memory over the very brief periods studied in the present experiment. Jost's Law, which states that old memories are more stable than new ones, holds good within the confines of memory over a few seconds.

The results of Experiment 7 raise certain questions. For instance:

(1) What are the implications of these results with respect to the nature of short-term retention?

(2) How can the intra-list differences in consolidation be explained?

(3) What are the implications of the finding that total recall of a sequence is affected by recall order?

Of these, (1) will be discussed next, and questions (2) and (3) will be raised in subsequent chapters.

Since, as has been shown, the items presented over a short period are not all equally well retained, it follows that either (a) more than one store is being used (using the word "store" to indicate a stage at which all items have been subjected to equivalent coding processes, and are equally liable to be affected by subsequent experimental variables, such as rehearsal or interfering activity), or (b) all items are in the same store, but some are better "hooked up" or integrated in that store than others. It is possible, as Sanders (1961) indicates, to conceive of a system in which materials in memory can differ both in the stores within which they are retained, and in the efficiency with which different items are retained within a particular store. As was suggested earlier, a model in which there is homogeneity of retention within the different stages or stores, if it were consistent with the available experimental evidence, would seem to have more precision and hence greater explanatory value than one in which retention of items could vary not only between but also within stages representing different levels of pre-storage coding. For instance, if an item X was better retained than item Y, a model using only assumption (a) would describe the underlying events in terms of a more efficient or advanced coding process corresponding with retention of X than was the case with Y. In a model incorporating either assumption (b), or both (a) and (b), as suggested by Sanders (1961), some precision is lost, since there would be a larger number of possible explanations in terms of the model of an observed difference in retention.

Some possibly misleading simplifications are implicit in the above remarks. First, retention of a sequence of items is not just a matter of storing a number of individual parts; the sequence is inherent in the material as retained, a fact which was emphasized in the previous section. Secondly the word "store" suggests a sort of passive repository. In fact, mechanisms which function actively to select and code presented items in various ways are essential for the processing and retention of verbal sequences. A coded representation of the sensory data has to be manufactured before it can be retained. Transforming, or coding, the raw data into forms in which it can be stored probably requires more complicated operations than the actual storage process, once the material is coded. It is natural to think of a single store as something discrete; two stores are most simply imagined as physically separate. However, considerations of accessibility alone make it unlikely that this corresponds with reality.

It seems probable that in the memory process different materials are coded on the basis of characteristics varying in degree of abstraction. Sperling (1963) suggeststhat a visual image is stored over a few milliseconds; and results obtained by Conrad (1964) and by Wickelgren (1965) indicate that retention over periods of a few seconds may make use of auditory characteristics of verbal material. For retention over long periods (Bartlett, 1932) a coded record of the "meaning" may be required. Some sort of scanning system would be necessary at each stage to determine which of the material could be accommodated in terms of the past experience, and this could be done more easily and at higher levels of abstraction with data which is immediately recognized as being "meaningful" than with other material. It might be possible to build an analogue of some of the simpler functions of short-term memory, incorporating various rules for coding and scanning at successive stages. A great difficulty, however, would be to incorporate in the model the accessibility which is characteristic of human memory. Memory tasks, as the present results suggest, may require material to be retrieved from more than one store at a time. The problem is to get a batch of material out of storage when it is not all at the same stage of coding. It might be possible to build into the analogue a system by which all items are tagged on the basis of the time of entering storage, and this tagging could be independent of the other ways in which materials are coded and categorized. However, accessibility does remain a problem for any hypothesized multistage memory system.

It is probably true that the mechanisms used by the human brain to store information are much more complex than is generally imagined. The demonstration that there can be large differences in memory within apparently homogeneous sequences which are presented and then retained over a few seconds, is one indication of this complexity.

SUMMARY OF CHAPTER 9

Experiment 7 is described, in which subjects attempted to recall lists of nine consonants immediately after presentation. Using various recall orders, it was found that recall of part of a list interfered with retention of the other parts, but that memory for items presented early in a sequence was considerably less affected by such interference than was memory for items presented later. Some implications of this result are discussed.

CHAPTER 10

EXPLANATION OF RECALL ORDER EFFECTS

10.1. Introduction

The experiments to be described in the present chapter were designed to explain the intra-list differences in consolidation of retained materials observed in the results of Experiment 7. They attempt to answer the question, why is the retention of items presented at the beginning of a sequence more resistant to distracting activities than that of items presented later in the same sequence?

A possible explanation is that items at the beginning of a list can be rehearsed during presentation of the remainder of the same list. When consonants are presented at the rate of one per second, subjects report that some rehearsal during presentation is possible. This favours early items. Subjects may prefer to rehearse the whole sequence of items so far presented, and since the interval between items remains constant it becomes more difficult to do so as a sequence progresses. Also, if rehearsal at every opportunity starts at the beginning of the sequence, by the time an entire sequence has been presented early items will have been rehearsed more often than late items.

If it is true that the greater consolidation of early items is due to rehearsal, then reducing the opportunity for rehearsal to occur should reduce the difference between early and late items in the effects of distracting activities upon retention, as measured by recall. An experiment was designed to test whether this is the case.

10.2. Experiment 8.

The experimental design was similar to that used in Experiment 7, except that two rates of presentation was used, and there were only two orders of recall.

Method

Subjects and Materials Sixteen unpaid undergraduate subjects were tested separately. The lists to be recalled consisted of the nine consonants C, D, F, H, L, N, R, S, T ordered randomly, but excluding alphabetical orderings of consonants adjacent in the alphabet e.g., CD, RST, and digrams with a frequency over 500 in the Underwood and Schulz (1960) digram count. The lists were presented by tape-recorder, at one of two speeds, one letter per second, or two letters per second.

Design and Procedure Two orders of recall were used. These were equivalent to the P1, P2, P3 and P3, P2, P1 orders in Experiment 7, and will be called the "Forward" and "Reverse" orders respectively. The combinations of two recall orders and two rates of presentation made four experimental conditions. Subjects were told in advance the rate at which a particular list would be presented, but not the recall order. Immediately after the presentation of a list, one of two arrows was illuminated. Of these, one pointed right, the other left. Subjects were told that when the right-pointing arrow was illuminated they were to attempt recall of the preceding list in the forward order. The other arrow was the cue for recall in the reverse order. As in Experiment 7, recall was dictated, and subjects were encouraged to guess if uncertain of a letter, and to indicate the position of an item which they were not attempting by the word "Blank". There were 40 experimental lists, ten in each condition. A three minute rest was allowed after 20 of the lists had been attempted, and within each successive group of 20 lists there were five in each condition, their positions varying among subjects, except that batches of five successive lists were presented at the same rate. Within each batch, each recall order was used at least twice; otherwise the different recall orders were used randomly. Prior to the experiment each subject attempted eight practice sequences. The method of scoring was identical to that used in Experiment 7, except that all letters recalled in their original positions were scored correct, regardless of whether previous items had been attempted.

Results

Table 13 shows the average total number of letters correctly recalled per list, as a function of rate of presentation and order of recall. t-tests showed that among lists presented at the fast rate there was no significant effect of varying recall order. However, when the slow rate of presentation was used, lists were better recalled in the reverse order than in the forward order ($t=2.65$, $p < .02$). These results are

		Recall Order	
		Forward	Reverse
Presentation Rate	Fast	4.19	4.36
	Slow	4.83	5.72

Table 13 Average number of consonants correctly recalled per nine-item list in Experiment 8, as a function of rate of presentation and recall order.

consistent with the suggestion that the facilitating effect of reverse recall, observed in Experiment 7, was due to the rehearsal of early items during the period in which a list was being presented. Retention of the early items was little affected by their recall position, so that a strategy of recalling first the less consolidated late items led to the highest overall accuracy in reproduction. A between presentation-rates comparison was made of correct reproduction in items recalled early and items recalled late. If R1, R2, R3 are used to indicate the recall positions they represented in Experiment 7, then with recall in the reverse order the greater the accuracy with which early items were reproduced, relative to recall of late items, the larger the value of $\frac{P1R3 + P2R2}{P3R1}$ The magnitude of this function for lists presented at the slow rate is 0.96, which is significantly higher ($t=3.55$, $\underline{p}<.01$) than the corresponding value (0.70) when presentation rate is fast. Thus it is the early items which benefit most from increased presentation-time, and this finding is consistent with the suggestion that consolidation of early items occurs because there is an opportunity to rehearse them.

Discussion

Some results obtained by Posner (1964) provide support for the present findings, Posner's subjects listened to lists of eight digits, presented at either of two rates, 96 items per minute and 30 items per minute. Recall order, which was specified after presentation of a list, was either identical to order of presentation, or the last four digits were required before the first four. Reproduction in the latter condition was more accurate than with forward recall if the list had been presented at the slow rate, but not when the fast rate of presentation had been used. Examination of his results led Posner to suggest that, "with presentation at the slow rate, rehearsal during the first four items preserves them during the subsequent presentation and recall of the last four". Additional evidence of the effects of rehearsal is available from the results of a number of studies (e.g., Brown, 1958; Conrad, 1960; Sanders, 1961; Pollack, 1963) in which presentation of a series of items is followed by a variable period during which rehearsal can occur. An interfering task follows the rehearsal period, and finally recall is attempted. In experiments of this kind accuracy of recall is generally found to vary positively with duration of the rehearsal period.

Thus there is substantial evidence consistent with the suggestion that greater opportunity for rehearsal underlies the higher resistance to recall interference of early than of late items in the lists presented in Experiment 7. However, an alternative explanation may be correct. It is possible that the degree of resistance to distraction in retained material is positively related to the <u>time</u> over which that material has been stored, regardless of what has gone on during that time. It is conceivable, for instance, that involuntary circulation of stored material occurs, in the process of which the material becomes more efficiently retained. Some results (e.g., those of Brown, 1958) have been interpreted as showing that in the absence of rehearsal decay occurs, which would suggest that active rehearsal is essential if consolidation is to occur on material being retained. However, to prevent rehearsal in short-term memory experiments it is customary to interpolate additional tasks, and it is possible that the activity of performing such tasks, rather than the resulting prevention of rehearsal, may have led to the observed loss in accuracy of recall.

10.3. Experiment 9.

It is conceivable that time alone may lead to consolidation in random verbal sequences which a person is trying to retain. Alternatively both time and rehearsal may be required. An experiment was designed to investigate the effects of separately manipulating time and rehearsal variables. The design enabled the experimenter to determine the effects on resistance to interference of material being retained of varying amounts of rehearsal, given a constant time in which consolidation could occur, and to observe the effects on post-interference recall of varying length of time in store, when the opportunity for rehearsal is constant. Lists were presented which were well below the memory-span length. It was assumed that the time taken to rehearse a verbal item is independent of the number of items in a list, so that in a given time the ratio of the number of rehearsals of a list containing x items to the number of rehearsals of a list of y items is y:x. Similarly, the time taken to rehearse a list of x items a given number of times is x/y the time required for the same number of rehearsals of a y item list. Thus it is possible to compare the extent to which retained items have become consolidated in different lengths or sequences, (a) when the materials have been stored over the same period of time but differ in the number of rehearsals possible, and (b) when the items have had the opportunity for an equal number of rehearsals, but the lists have been retained over different lengths of time.

Method

Subjects and Materials Twenty undergraduates were paid to act as subjects. The material to be remembered consisted of four- and six-digit lists which were presented by tape-recorder. The digit sequences were random, except that no item occurred more than once in a list. Each subject was tested separately on 32 lists. For the interference task, subjects had to subtract aloud by sevens from a three-digit number, as fast as possible, for ten seconds. The initial number was illuminated on Nixie tubes, and a different number was provided after each list.

Design and Procedure There were 16 lists at each of the two lengths, presented in batches containing four sequences of the same length. The orders in which the batches were presented were balanced across subjects. Four rehearsal periods were used, 4, 8, 16 and 24 seconds. Within each successive 16 lists each rehearsal period occurred four times, combined twice with each list length. As far as was consistent with these requirements the ordering of the various rehearsal periods was random, and the combinations of particular lists with lengths of rehearsal period differed among subjects. Subjects were told in advance the length of a list, but they were given no advance information about the duration of the succeeding rehearsal period. There was a two minute rest pause in the middle of the experiment, which was preceded by eight practice lists.

Subjects listened to a list, after which came the variable rehearsal period. The instructions were to use this period for silent rehearsal. Knowledge of the duration of a rehearsal period was not provided until it ended, and this was indicated by a light being extinguished. At the same time the Nixie tubes were illuminated, showing a three-digit number. Subjects were told to read the number out aloud immediately it appeared, and then to subtract aloud as fast as possible for ten seconds. To mark the end of the period taken up by the interference task, the light which had been extinguished was re-illuminated, and the Nixie tubes went out. Subjects then attempted spoken recall of the original sequence, and digits were scored correct if recalled in their original position.

Results

Figure 9 shows the average number of digits recalled per list as

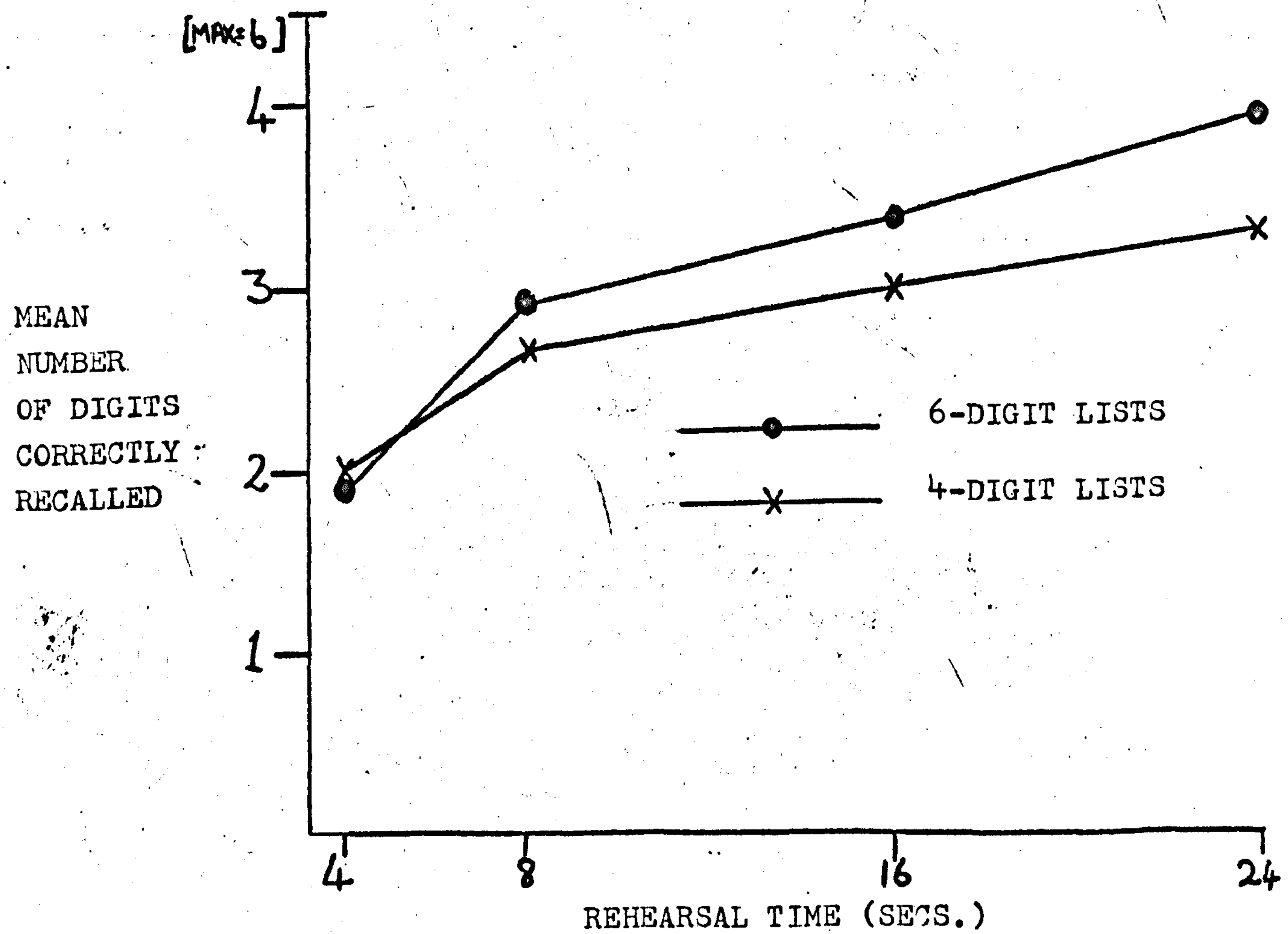

Figure 9 Average number of digits correctly recalled in Experiment 9 as a function of duration of the interval for rehearsal between presentation of a sequence and the ten-second interfering task.

a function of time available for rehearsal between the end of presentation of a list and the start of the interfering task. Length of list is the parameter. The results of an analysis of variance on the data are shown in Table 14. The experimental variables clearly are significant

Source	Sum of Squares	d.f.	Variance	F	p<
Length of List (L)	209.86	1	209.86	49.97	.001
Rehearsal Time (T)	598.21	3	199.40	47.48	.001
Subjects (S)	466.96	19	24.58	5.85	.001
LT	96.34	3	32.11	7.65	.001
LS	262.14	19	13.80	3.29	.01
TS	289.04	57	5.07	1.21	.1
Remainder	239.26	57	4.20		
Total	2161.81				

Table 14 Results of an Analysis of Variance on the data from Experiment 9, as shown in Figure 9.

determinants of the results.

It will be recalled that the reason for performing Experiment 9 was to compare the effects of (a) total time stored, and (b) time per item available for rehearsal, on the resistance of items being retained to a subsequent distracting task. An obvious problem is that resistance, or consolidation, cannot be measured directly, it can only be inferred from post-interference recall scores. Figures 10a and 10b show the relationship between increase in probability of correct recall and (a) given durations of rehearsal time, and (b) given amounts of rehearsal time per item, where both take the four seconds rehearsal time condition as a baseline.

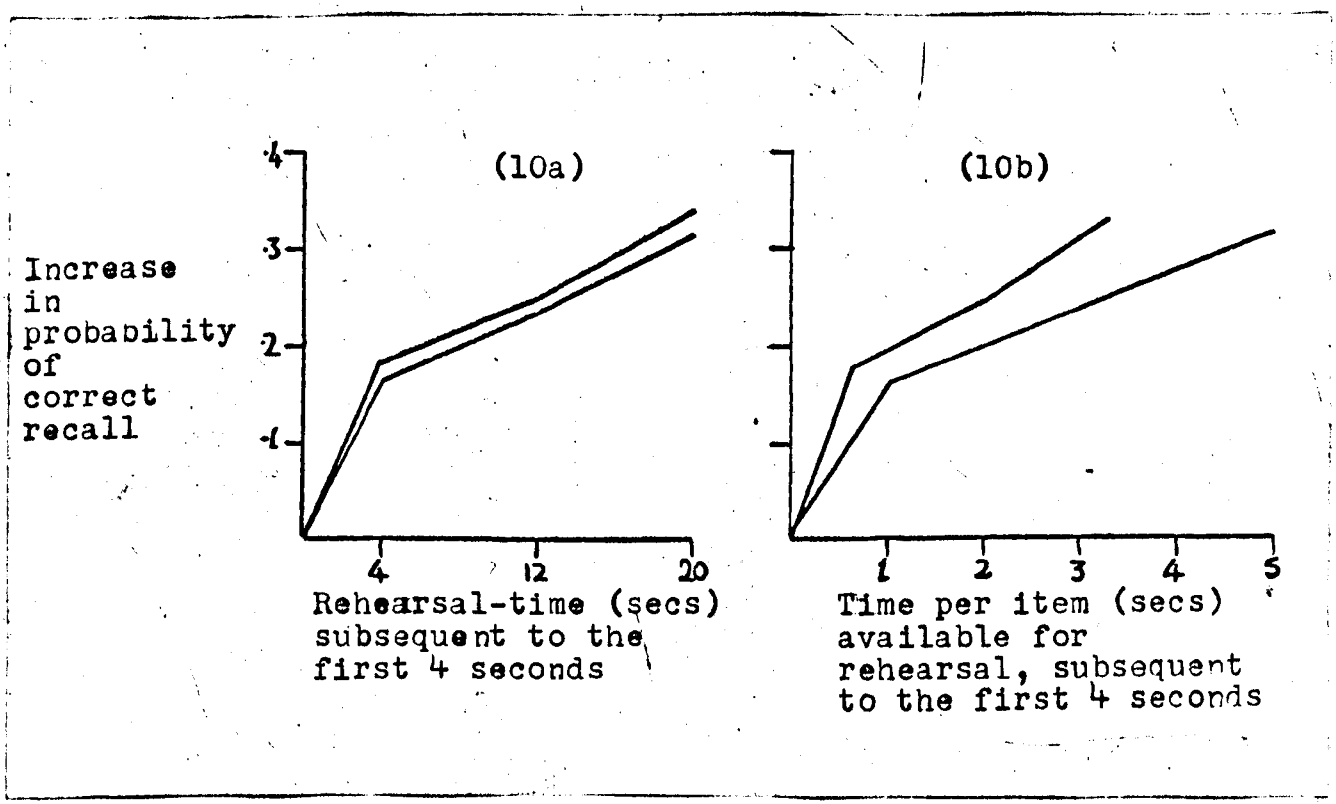

Figure 10. Mean increases in probability of correct recall in Experiment 9 when rehearsal time exceeds 4 seconds, (a) as a function of total rehearsal time (subsequent to 4 secs.) and (b) as a function of rehearsal time (subsequent to 4 secs.) per item.

It is clear that when the time available for rehearsal exceeds four seconds the effects of providing a given additional period in which rehearsal can occur are practically identical for the two lengths of list. However, the effects of given additional periods of time per item differ between lists of the two lengths. This suggests that in the present experiment, total time available for rehearsal, subsequent to the first four seconds, is a better predictor of resistance to subsequent interference than is time for rehearsal per item. It does not, however, follow that consolidation occurs through time alone, irrespective of whether or not rehearsal occurs. Hindsight suggests that the provision of values for the probability of recall being correct when rehearsal time is zero would have enhanced the utility of the results, since a fuller description of the effects of rehearsal over very short periods would then have been provided.

Discussion

The present results indicate that conscious rehearsal is not the sole determinant of resistivity to interference in verbal materials being retained over short periods. If it was, the two lines on Figure 10b would have been identically placed. Various factors could have influenced the results; for instance, there is much greater room for improvement in accuracy of recall of the six-digit lists than of the four-digit sequences. The decision to use probability of correct recall as the index on which to base inferences about pre-interference consolidation was made after some consideration. There are some reasons favouring measurement of the average number of items retained, and it is not possible to say definitely whether a measure based on the proportion of items recalled or one based on the amount of items recalled is the better index for comparison of the efficiency with which items have been retained in sequences of different lengths without making some assumptions about the nature of the mechanisms underlying retention. In order to obtain more evidence about the comparative effects of time and rehearsal on consolidation of retained verbal materials, a second approach was used.

10.4. Experiment 10.

The rationale behind the present experiment is that the effects of time on the resistance of items being retained for short periods to distracting activities, can be determined if all items are given the same amount of rehearsal. The latter is then a controlled variable. As has been mentioned, experiments in which rehearsal has been prevented for various periods of time are subject to the objection that the interference task introduced between presentation and recall may have effects on retention processes other than that of preventing rehearsal. This difficulty can be overcome by filling the variable presentation-recall interval with rehearsal of items, but ensuring that no item is rehearsed more or less than others.

To prevent subjects from concentrating their rehearsal on particular items a task was designed in which the subjects were kept fully engaged during presentation. Nine-consonant lists were presented on cards, which each contained three items. A subject had to speak the contents of the cards aloud three times, as fast as possible, thus pacing himself at his optimum rate. After presentation of all three cards constituting a list, recall could be required of the contents of ~~and~~ one card, and since they had all been equally rehearsed, any effects of presentation position on the resistance of the stored material a distracting task that was interpolated between presentation and recall could reasonably be attributed to time in store, if the decrement in accuracy of retention observed when a distracting task is interpolated between presentation and recall is used as a reciprocal measure of consolidation.

Subjects and Materials There were 16 paid undergraduate and postgraduate subjects, who were tested individually. Forty-eight lists were constructed, each of which contained the nine consonants C, D, F, H, L, N, R, S, T, ordered randomly, as in Experiment 8. The lists were typed on white cards, three letters per card. The distracting task consisted of a sequence of eight digits, which a subject had to read aloud, as quickly as possible.

Design and Procedure. There were six conditions, since recall could be required of any one of the three cards, and the distracting task was interpolated between presentation and recall of half the lists. The 48 lists contained eight combinations of each of the three partial-recall conditions with both the interference and no-interference conditions. The order in which the various conditions were used was random, and varied among subjects, except that no particular value of either of the two independent variables occurred in three or more successive lists, and each condition occurred twice in every successive group of twelve lists.

For presentation of a list, a subject turned over a blank card, exposing the first three letters. He immediately read these letters aloud, and then repeated the three-letter sequence twice more, as quickly as possible. Then he immediately turned over the card, exposing the next three letters in the list, which he read and twice repeated, as with the sequence on the first card. The second card was next turned over, to expose letters 7, 8, 9, and the reading process was repeated as before, thus completing the presentation of a list. The subject then immediately turned over the card containing the final three consonants. In half the lists the next card was blank, but the card which directly succeeddthe other lists contained a sequence of eight digits, a different sequence being used with each list. A blank card indicated that no intervening activity was required before recall of the list. If the card contained digits, the subject read them aloud once, as quickly as possible. This constituted the interfering task. The next, and last card in each trial contained three dashes, typed side by side. Above one of them was typed another, parallel, line in red. This was the cue for recall. If the red line was above the left hand dash, recall was required of the letters typed on the first card, i.e. consonants 1, 2, 3; if the red line was in the middle, letters 4, 5, 6 were to be recalled, and if the red line was above the dash on the right of the card, the subject had to attempt recall of the letters typed on the third of the cards on which the list was presented. Subjects dictated their attempts at recall, which were written down. The word "Blank" was used to attempt to indicate a letter not being attempted, and letters recalled in their original position were scored correct. The experiment was preceded by six practice trials. Subjects worked through card packs, each of which contained the cards constituting 12 lists, and apart from the natural breaks which occurred between packs there was a two minute rest period in the middle of the experiment.

Results

Table 15 shows the average number of letters correctly recalled

	Positions in Presentation		
	1,2,3	4,5,6	7,8,9
No interference	0.95	0.97	1.91
Interference	1.39	1.93	2.72

<u>Table 15</u> Average number of consonants correctly recalled per card (out of three) in Experiment 10, as a function of (a) inter-list position and (b) whether or not there was an interfering task between presentation and recall.

per three-item sequence as a function of position of the sequence within a list, and of the presence or absence of a distracting task between presentation and recall.

Discussion

It can be seen that the part of a list presented last is best recalled, whether or not a distracting task intervened between presentation and recall. Less of the middle items are correctly recalled, and the initial three letters in a list are recalled least accurately of all. In conditions where there was no distracting task the distribution of results is similar to that occurring in the R1 conditions of Experiment 7. However, when a distracting task is introduced, the pattern of the present results is quite unlike the corresponding findings for R3 in Experiment 7. In the latter, recall of letters 1, 2, 3, (P1) was very little affected by the distraction caused by recalling other letters; in the present experiment the distracting task cuts down the average number of letters correctly recalled by about a third. It is interesting to compare the effects of the interfering task on recall of each of the different parts. The figures below show the percentage drop in probability of correct recall of items when subjects are required to carry out the interpolated distracting activity.

Presentation Position

1,2,3	4,5,6,	7,8,9
31.55%	49.23%	29.78%

It is interesting to note that there is not a significant difference in the effect of a distracting task on probability of correct recall between the first three and the last three consonants, but having to recite eight digits considerably influences retention of the middle three items. Comparing the effects of a distracting activity on percentage recall of the first and third parts suggests that in the present experimental situation, the probability of recalling items which have been equally rehearsed during presentation is unrelated to the time over which items have been stored. However, when the amount by which the distracting task affects recall is ~~calculated~~ measured by the number of items forgotten, rather than the drop in probability of recall, the inter-position distribution of results is rather different. The following figures represent the average decrement in the number of items correctly recalled per three-consonant sequence, when an interfering activity occurs between presentation and recall.

Presentation Position		
1,2,3	4,5,6	7,8,9
0.41	0.96	0.81

Clearly the distribution of results is strongly influenced by the particular measure of recall used. If the number of items forgotten is used as the measure of the effect of distracting activity, then as the above figures indicate, the early (1,2,3) items in the present experiment possessed greater resistance in retention than the later items. But, as has already been said, it cannot be decided with certainty which of (a) the amount by which recall drops, or (b) the decrement in percentage, and hence in probability of recall, is the more meaningful measure of vulnerability in stored items to interfering activities. A probability measure seems most apt in the present context. When a measure of the decrement in probability of correct recall is adopted, the present experimental results suggest that stored items which have been equally rehearsed are equally resistant to interference, irrespective of the time for which they have been retained, at least within the time values considered in this study.

Thus it appears that the intra-list differences in consolidation of consonants, which were observed in Experiment 7, may be due to inequal opportunities for rehearsal during presentation of a list. In Experiment 7, the measure adopted was average number of correct items, rather than the average probability of an item being correctly recalled. However, the cross-over phenomenon which can be seen in Figure 8 demonstrates that an intra-list difference in resistance to interference is evident, irrespective of whether the retention measure used scores probability of correct recall, or number of items correctly recalled. Thus the differences between the experiments in the way recall is scored do not negate the above conclusions. But this conclusion is not supported by the results of Experiment 9, and there is no known reason for the apparently conflicting results. There are a number of differences both between the materials and between the procedures of Experiments 9 and 10. First, subjects had to remember digits in one of the experiments and letters in the other. Also, the interference tasks differed between experiments, both in nature and in duration. Thirdly, the periods of time available for rehearsal differed between the two experiments. However, it is not at present possible to say which, if any, of these and other differences between the experimental procedures adopted in Experiments 9 and 10 contributed to the conflicting findings about the separate effects of rehearsal and time in store on the resistivity of interference possessed by verbal items in short-term memory.

SUMMARY OF CHAPTER 10

Three experiments are described, which were designed to explain intra-list differences in consolidation of verbal material. The results of Experiment 8 indicate that rehearsal is an important determinant of consolidation, and Experiment 10 produced results consistent with this view. However, Experiment 9, in which the number of rehearsals is held constant, indicates that consolidation is also a function of the time over which items have been stored.

CHAPTER 11

RECALL OF SEQUENCES AS A FUNCTION OF THE ORDER IN WHICH ITEMS ARE REPRODUCED.

11.1. Interpreting some results

A result of Experiment 7 was that the accuracy of recall of nine-consonant lists as wholes varied with the order in which reproduction of the items was required. The lists were best recalled (see Table 12) when the final three letters were required first, then the middle items, and the first three letters were recalled last of all. It is possible that there are practical applications of this result. The situation which first comes to mind is telephone dialling; if more items can be reproduced in this situation when the late items are recalled first, it might be advantageous to design a dialling system in which the order of dialling takes this into account. It is possible that with other situations in which lists have to be retained over short periods, accuracy could be improved by reproducing the items in an order different to that in which they were presented. Also, if items can be most accurately reproduced in other than the presented order, it is not impossible that the human organism may use strategies which take advantage of this, albeit unaware to the individual, in processing information. Thus there is considerable interest in the finding that verbal sequences may be most efficiently reproduced in a different order ~~to~~ from that of presentation.

However, the experimental situation of Experiment 7 was somewhat complicated, and hence atypical of experiments on memory, and dissimilar to activities requiring the processing of information in everyday life. It would be interesting to know whether a similar result would occur in a situation where the experimental procedure is less complex. In Experiment 7, subjects were given no information about order of recall until after presentation of a sequence, and then only part by part. In other experiments, for instance Experiment 8 (10.2. above) and Posner (1964), the entire recall order was specified immediately after presentation, and in both these studies it was observed that, at certain presentation rates, lists were most accurately retained when items presented late in the sequence were recalled first. Results obtained by Kay and Poulton (1951) and by Brown (1954) also indicate that recall order may be an influential variable when it is completely specified immediately after the material has been presented.

The observation that recall order affects total reproduction would have more generality if it were found also to occur in situations where order of recall is known prior to the to-be-retained material being presented. The available experimental evidence on this point is limited, but results obtained by Brown (1954) show that when sets of arrows and digits are presented simultaneously, the order in which they are to be recalled being specified in advance, order of recall is still an important variable. It would be interesting to know whether this occurs when verbal items are presented in a straightforward sequence.

11.2. Experiment 11

The object of the present experiment was to observe the effects of varying the recall order of a sequence of letters when the experimental situation is simple and straightforward, and the complete instructions for order of recall are given at one time, and prior to the sequence being presented.

Method

Subjects, Materials and Experimental Design Sixteen paid undergraduates were each tested individually on 20 nine-consonant lists. The letters C, D, F, H, L, N, R, S, T were used, presented randomly, but excepting the combinations excluded in Experiment 8. The lists were presented at a rate of one item per second, by tape-recorder. There were recall orders. With one, "Forward Recall", the letters had to be reproduced in the order 1, 2, 3, 4, 5, 6, 7, 8, 9, the digits representing the positions of items in the presentation sequence, and in the other condition ("Reverse Recall") the consonants were recalled in the order 7, 8, 9, 4, 5, 6, 1, 2, 3. The conditions occurred at different random positions within the experimental trials for each subject, except that each recall order was used five times within each of the first and second ten lists.

Procedure Immediately before presentation of a list the illumination of one of two lights, which were placed in front of a subject, indicated the order in which recall of the succeeding list was to be attempted, the light remaining on throughout presentation of the list. Subjects dictated their recall to the experimenter. They were encouraged to guess when uncertain, and to eliminate ambiguity about positions of items by indicating any consonant not being attempted by the word "Blank". There were five-second intervals between recall of a list and presentation of the next. Prior to the experiment subjects attempted six practice lists. To be scored correct a consonant had to be recalled in the appropriate position.

Results

Table 16 shows the average number of consonants recalled in their correct positions. The prediction from previous results that recall of

	Order of Recall	
Item Numbers	Forward	Reverse
1 2 3	2.36	1.89
4 5 6	1.98	2.06
7 8 9	1.82	2.75
Whole lists (1-9)	6.16	6.70

Table 16 Mean number of consonants correctly recalled per part (out of three) and per list (out of nine) as a function of recall order, in Experiment 11.

whole lists would be significantly better in the reverse than in the forward recall order condition was confirmed ($t=2.28$, $p<.025$, one-tailed). Reproduction of the first three items in lists was more accurate with forward than with reverse recall ($t=4.69$, $p<.01$). The last three consonants were better retained in the reverse recall condition ($t=10.46$, $p<.001$). Retention of consonants 4, 5 and 6 was not significantly influenced by the order in which lists were recalled.

Discussion

Since the previous findings about the effects of recall order are confirmed in a less complicated experimental situation it may be worthwhile to look for practical applications of such variations in order of recall to activities requiring retention over brief periods, and to be open to the possibility of their present use in information processing by the human nervous system. However, it should be pointed out that any advantage in the amount recalled when orders of presentation and recall are non-identical is naturally worthless when the precise original sequence must remain intact. Caution is required in making generalizations on the basis of these results, especially since data obtained by Posner (1964) and from Experiment 8, above, suggest that the present result would not occur when sequences are presented at rates not identical to those used in this Experiment, even when identical items and mode of presentation are used.

It seems possible that a strategy of recalling first the items presented late in a sequence might improve performance at learning tasks. In many verbal learning experiments, a list of items is presented to a subject, who then has to record what he can remember of them. The materials are then presented again, and the processes are repeated either a given number of times, or until all the items can be correctly reproduced after presentation. The list is then said to be learned, and the number of trials necessary before an entirely correct reproduction is made can be used as an index of the difficulty to the task. An experiment was performed to determine if a strategy of recalling first the final items in a list would improve performance in a task of this sort. A list containing 14 common nouns was presented by ear at a rate of two seconds per item. Subjects were divided into two groups. After the list had been presented, one group had to attempt to write down the words in the order in which they had been presented. The other group had to write first the last four words in the list, and then the remainder, in order. The lists were presented five times in all, and subjects' average performance improved from about seven words correctly recalled after the first presentation, to about thirteen words after the fifth presentation of the list. The results of the two experimental groups were compared, to determine whether recall order had any effect on performance af the learning task. There was no significant difference between groups, which suggests that the effects of recall order on immediate reproduction cannot be successfully utilized in a learning task. In a modified version of the experiment just summarized, rate of presentation was increased to one word per second, and both the list lengths and number of trials were reduced, to make conditions more similar to those of Experiment 11. However, even in this modified learning experiment, there were no differences in results between groups. This result supports the previous conclusion that performance in some learning tasks is not affected by recall order strategy.

SUMMARY OF CHAPTER 11

Recall order was found to affect accuracy of reproduction in a short-term memory task (Experiment 11) where the procedure and conditions were as simple as possible. However, order of recall was not a significant determinant of performance at a five-trial learning task.

CHAPTER 12

CONCLUSIONS TO SECTION 2

12.1. Concluding Remarks

Summaries of the individual chapters in Section 2 have already been provided. The succeeding pages will contain a brief survey of the main conclusions reached in Section 2, followed by some remarks about possible directions for future research, as suggested by the present results.

It was observed that memory scores can be influenced by the activity of recalling part of a stored verbal sequence. Recall may interfere with the retention of the remainder of the material, as is apparent from the finding that if a list of items is presented, and recall then required of any part of it, recall of that part is more accurate than it would be if the whole sequence had to be recalled. Related to this is the observation that recall of a particular part is affected by the order in which the sequence is recalled in relation to order of presentation. Other things being equal, the first items to be recalled are reproduced most accurately. Experiment 7 was designed to investigate further the effects of recall on retention. It was found that whereas the activity of interpolated recall seriously interfered with the retention of items presented at the end of a sequence, retention of items which had been placed early in a sequence was little affected by the position in which they were recalled.

The interpretation given to this finding was that some consolidation occurs during the presentation of the material. By the time representation is complete, stored early items have acquired a greater degree of resistance to subsequent distracting activities than have late items. Thus there are differences in the efficiency with which verbal materials are stored, even within sequences presented over periods of several seconds. This observation has certain implications relevant to consideration of the mechanisms underlying short-term retention for instance, it is clear that the present experimental results could not have occurred if the operations underlying short-term memory were carried out by a system in which all items were stored in the same manner.

Subsequent experiments were designed to explain the observed intra-list differences in relative consolidation between retained verbal items. Consolidation could be caused by rehearsal, or time itself might have effects on retention over and above that of allowing rehearsal to occur. The results of Experiment 10 indicated that the amount of rehearsal was an important variable, when storage time was controlled, and the results of Experiment 8 were consistent with this view. However Experiment 9 produced findings suggesting that time _per se_ is important, and it was therefore concluded that active rehearsal is a determinant of consolidation, but that time itself has effects additional to that of enabling rehearsal to occur.

The observation that recall order does affect the accuracy of reproduction of whole lists is interesting in itself, since it follows that the most efficient strategy for human transmission of retained verbal material may require reproduction of the material in an order other than that in which it was originally prevented. Since Experiment 7, in which this observation was made, was somewhat complex in design and procedure, it seemed possible that the result was an artefact of the particular experimental conditions. However a check, using a very simple experimental design (Experiment 11) showed that this was not the case.

12.2. Suggestions for further research.

The series of experiments described in Section 2 has naturally given rise to a number of ideas about possible lines for further research.

1. The basic experimental design of Experiment 7 might be used in conjunction with a rage of presentation variables. Among those whose effects it would be interesting to observe are:

a) Presentation time per item.

b) Number of items in a sequence.

c) Basic units in the sequences presented (e.g., digits, colours, shapes, words).

d) Mode of presentation (e.g., visual compared with auditory).

e) Simultaneous compared with successive visual presentation.

Investigation into the effects of variables such as these would provide information about the influence of recall on retention in a number of conditions, and the results might suggest ways in which the mechanisms underlying retention are actually affected by the activity of recalling.

2. A fair amount of work has been carried out (see, e.g., Mackworth, 1962) on the effects of varying mode of recall in short-term memory. It would be valuable to incorporate variations in recall mode (e.g., written, dictated, keyboard) in an experiment based on the design of Experiment 7, with a view to determining the extent to which physical aspects of recall underly the distracting effects of recall on retention. Distracting activities might interfere with retention by disrupting the available capacity. On the other hand, the effect may be less specific, and forgetting may be caused by a shift in attention which results in processing capacity which is normally used for retention becoming no longer available. Experiments in which recall mode is varied might make it possible to decide between these alternatives.

3. There is a need for fuller quantitative information concerning the relationship between the proportion of the original sequence which has to be recalled, and accuracy of reproduction. The finding by Anderson (1960) that after presentation of 12-digit lists recall is proportionately more accurate when eight items are required than when all twelve have to be reproduced, and when four than when eight items are required, might suggest that the smaller the part of the stored material to be recalled, the more accurate is reproduction. However, it is known (see, e.g., Martin and Fernberger, 1929) that recall is improved by subjects grouping the items during presentation, and it might well be that the number of items at which partial recall brings about the greatest improvement in accuracy of recall

is equivalent to the optimum size for grouping. When total recall is required, it is probably important that any re-ordering does not break up the groupings made by subjects at the time of presentation.

4. In a study which formed part of the previous section (Experiment 6) response latency was used, rather than recall, as a measure of retention. There are reasons for using a latency measure in conjunction with experiments based on the present series. For instance, it would not be necessary to use long interference tasks, which lengthen the time over which items are stored and tend to increase the variance in experimental results. A worthwhile project would be to calculate the correlations between a) response latency, b) number of rehearsals, c) time over which material is stored and d) consolidation in storage (inferred from post-interference recall), with a view to understanding something about retrieval in relation to retention.

5. Further detailed investigation of the practical effects of various recall strageties is required. This could take several forms. For instance, the following questions need answering:

a) What are the effects of varying recall order upon accuracy or latency of reproduction in an everyday situation, such as telephone dialling?

b) Can order of recall affect performance at multi-trial learning tasks? The experiments mentioned in Chapter 12 produced a negative answer to this question, but it is nevertheless possible that there are some learning situations in which varying the order of reproduction would have a significant effect.

REFERENCES

ABORN, M., RUBENSTEIN, H., & STERLING, T.D. (1959) Sources of contextual constraint upon words in sentences. J.exp.Psychol., 57, 171-180.

ANDERSON, N.S. (1960) Poststimulus cuing in immediate memory. J.exp.Psychol., 44, 260-266.

AVERBACH, E., & CORIELL, A.S. (1961) Short-term memory in vision. Bell Syst.Tech.J., 40, 309-328.

BADDELEY, A.D. (1963) Some psychological aspects of the coding of information. Unpublished Ph.D. dissertation, Cambridge University.

BADDELEY, A.D. (1964a) Immediate memory and the "perception" of letter sequences. Quart.J.exp.Psychol., 16, 364-367.

BADDELEY, A.D. (1964b) Language-habits, S-R compatibility and verbal learning. Amer.J.Psychol., 3, 463-468.

BADDELEY, A.D., CONRAD, R., & HULL, A.J. (1965) Predictability and immediate memory for consonant sequences. Quart.J.exp.Psychol., 17, 175-177.

BADDELEY, A.D., CONRAD, R., & THOMSON, W.E. (1960) Letter structure of the English language. Nature, 186, 414-416.

BARTLETT, F.C. (1932) Remembering: a study in experimental and social psychology. Cambridge University Press.

BROADBENT, D.E. (1957) Immediate memory and simultaneous stimuli. Quart.J.exp.Psychol., 9, 1-11.

BROWN, J. (1954) The nature of the set to learn and of intramaterial interference in immediate memory. Quart.J.exp.Psychol., 6, 141-148.

BROWN, J. (1958) Some tests of the decay theory of immediate memory. Quart.J.exp.Psychol., 10, 12-24.

CHAPANIS, A. (1954) The reconstruction of abbreviated printed messages. J.exp.Psychol., 48, 496-510.

COLEMAN, E.B. (1962) Sequential interferences demonstrated by serial reconstructions. J.exp.Psychol., 64, 46-51.

COLEMAN, E.B. (1963) Approximations to English: some comments on the method. Amer.J.Psychol., 76, 239-247.

CONRAD, R. (1960) Very brief delay of immediate recall. Quart.J.exp.Psychol., 12, 45-47.

CONRAD, R. (1964) Acoustic confusions in immediate memory. Brit.J.Psychol., 55, 19-30.

CONRAD, R., FREEMAN, P.R., & HULL, A.J. (1965) Acoustic factors versus language factors in short-term memory. Psychon.Sci., 3, 57.

DAVIS, R., SUTHERLAND, N.S., & JUDD, B. (1961) Information content in recognition and recall. J.exp.Psychol., 61, 422-428.

DEESE, J. (1961) From the isolated verbal unit to connected discourse. pp. 11-31 in Cofer, C.N. (Ed.) Verbal learning and verbal behavior. New York, McGraw-Hill.

DEESE, J., & KAUFMAN, R.A. (1957) Serial effects in recall of unorganized and sequentially organized verbal material. J.exp.Psychol., 54, 180-187.

DIEBOLD, A.R. (1965) A survey of psycholinguistics research, 1954-64. Published with Osgood, C.E. & Sebeok, T.A. Psycholinguistics: a survey of theory and research problems. Bloomington. Indiana University Press.

EPSTEIN, W. (1961) The influence of syntactical structure on learning. Amer.J.Psychol., 74,80-85.

GARNER, W.R., & CARSON, D.H. (1960) A multivariate solution of the redundancy of printed English. Psychol.Rep., 6, 123-141.

GOLDMAN-EISLER, F. (1958) Speech production and the predictability of words in context. Quart.J.exp.Psychol., 10, 96-106.

GOMULICKI, B.R. (1956) Recall as an abstractive process. Acta Psychol., 12, 77-94.

HABER, R.N. (1964) A replication of selective attention and coding in visual perception. J.exp.Psychol., 67, 402-404.

HARRISON, G.J. (1964) Effect of conceptual cuing on the recall of order. Nature, 202, 1365.

HOGAN, J.A. (1961) Copying redundant messages. J.exp.Psychol., 62, 153-157.

HUNTER, I.M.L. (1961) Further studies on anagram solving. Brit.J.Psychol., 52, 161-165.

JAHNKE, J.C. (1963) Serial position effects in immediate serial recall. J.verb.Learn.verb.Behav., 2, 284-287.

JAMES, W. (1890) The principles of psychology. New York, Holt. Vol.1.

JENSEN, A.R., & RODEN, A. (1963) Memory span and the skewness of the serial-position curve. Brit.J.Psychol., 54, 337-349.

JOOS, M. (1950) Description of language design. J.acous. Soc.Amer., 22, 701-708.

KAY, H. (1955) Learning and retaining verbal material. Brit.J.Psychol., 46, 81-100.

KAY, H., & POULTON, E.C. (1951) Anticipation in memorizing. Brit.J.Psychol., 42, 34-41.

KEPPEL, G. (1965) Problems of method in the study of short-term memory. Psychol.Bull., 63, 1-3.

KEPPEL, G., & UNDERWOOD, B.J. (1962) Proactive inhibition in short-term retention of single items. J.verb.Learn. verb.Behav., 3, 362-368.

KLEMMER, E.T. (1961) The perception of all patterns produced by a seven-line matrix. J.exp.Psychol., 61, 274-282.

LACHMAN, R., & FELD, W.H. (1965) Recognition and recall of verbal material as a function of degree of training. Psychon.Sci., 2, 225-226.

LACHMAN, R., & TUTTLE, A.V. (1965) Approximations to English (AE) and short-term memory: construction or storage? J.exp.Psychol., 70, 386-393.

LASHLEY, K.S. (1951) The problem of serial order in behavior. Pp.112-136 in Jeffress, L.A. (Ed.) Cerebral mechanisms in behavior. New York, Wiley.

LAWRENCE, D.H., & LABERGE, D.L. (1956) The relationship between recognition accuracy and order of reporting stimulus dimensions. J.exp.Psychol., 51, 12-18.

LAWSON, E.A. (1961) A note on the influence of different orders of approximation to the English language upon eye-voice span. Quart.J. exp.Psychol., 13, 53-55.

LLOYD, K.E., & JOHNSTON, W.A. (1963) Short-term memory as a function of contextual constraint. J.exp.Psychol., 65, 460-467.

McKNIGHT, J.L. (1959) The quantum theoretical concept of measurement. In Churchman, C.W., & Ratoosh, P. (Eds.) Measurement, Definitions and Theories. New York, Wiley.

McNULTY, J.A. (1965) An analysis of recall and recognition processes in verbal learning. J.verb. Learn.verb.Behav., 4, 430-436.

MACKWORTH, J.F. (1962) The effect of the response upon the immediate memory span. Canad.J.Psychol., 16, 120-127.

MAYZNER, M.S., & ADLER, S. (1965) A further study of information "chunking" and short-term retention. J.Psychol., 59, 125-130.

MAYZNER, M.S., & GABRIEL, R.F. (1964) Short-term retention of paired-associate material. J.Psychol., 58, 3-7.

MAYZNER, M.S., & SCHOENBERG, K.M. (1964) Single-letter and digram frequency effects in immediate serial recall. J.verb.Learn.verb.Behav., 3, 397-400.

MAYZNER, M.S., & TRESSELT, M.E. (1959) Anagram solution times: a function of transition probabilities. J.Psychol., 47, 117-125.

MAYZNER, M.S. & TRESSELT, M.E. (1962a) Tables of digram and single-letter frequency counts by word length and letter-position for a sample of 20,000 words. Mimeographed. New York University.

MAYZNER, M.S., & TRESSELT, M.E. (1962b) Anagram solution times: a function of word transition probabilities. J.exp. Psychol., 63, 510-513.

MAYZNER, M.S., & TRESSELT, M.E. (1962c) The ranking of letter pairs and single letters to match digram and single-letter frequency counts. J.verb.Learn.verb.Behav., 1, 203-207.

MAYZNER, M.S., & TRESSELT, M.E. (1963) Anagram solution times: a function of word length and letter position variables. J.Psychol., 55, 469-475.

MAYZNER, M.S., TRESSELT, M.E., & HELBOCK, H. (1964) An exploratory study of mediational responses in anagram problem solving. J.Psychol., 57, 263-274.

MARKS, M.R., & JACK, O. (1952) Verbal context and memory span for meaningful material. Amer.J.Psychol., 65, 298-300.

MARTIN, P.R., & FERNBERGER, S.W. (1929) Improvement in memory span. Amer.J. Psychol., 41, 91-94.

MILLER, G.A. (1958) Free recall of redundant strings of letters. J.exp.Psychol., 56, 485-491.

MILLER, G.A., BRUNER, J.S., & POSTMAN, L. (1954) Familiarity of letter sequences and tachistoscopic identification. J.Gen. Psychol., 50, 129-139.

MILLER, G.A., & FRIEDMAN, E.A. (1957) The reconstruction of mutilated English texts. Information and Control, 1, 38-55.

MILLER, G.A., HEISE, G.A., & LICHTEN, W. (1951) The intelligibility of speech as a function of the context of the test materials. J.exp.Psychol., 41, 329-335.

MILLER, G.A., & SELFRIDGE, J.A. (1950) Verbal context and the recall of meaningful material. Amer.J.Psychol., 63, 176-185.

MILLER, I. (1957) Perception of nonsense passages in relation to amount of information and speech-to-noise ratio. J.exp.Psychol., 53, 388-393.

MORAY, N., & BARNETT, T. (1965) Stimulus presentation and methods of scoring in short-term memory experiments. Acta Psychol., 24, 253-263.

MORAY, N., BATES, A., & BARNETT, T. (1965) Experiments on the four-eared man. J.acous.Soc.Amer., 38, 196-201.

MORAY, N., & TAYLOR, A.M. (1958) Statistical approximations to English and French. Lang.Speech., 1, 102-109.

MORRISON, H.M., & BLACK, J.W. (1957) Prediction of missing words in sentences. J.Speech Hearing Disorders, 22, 236-240.

MURDOCH, B.B. (1963) Short-term retention of single paired-associates. J.exp.Psychol., 65, 433-443.

NEWMAN, E.B., & GERSTMAN, L.J. (1952) A new method for analysing printed English. J.exp.Psychol., 44, 114-125.

O'NEILL, J.J. (1957) Recognition of intelligibility test materials in context and isolation. J.Speech Hearing Disorders, 22, 87-90.

PIERCE, J.R., & KARLIN, J.E. (1957) Reading rates and the information rate of a human channel. Bell Syst.tech.J., 36, 497-516.

POSNER, M.I. (1963) Immediate memory in sequential tasks. Psychol.Bull., 60, 333-349.

POSNER, M.I. (1964) Rate of presentation and order of recall in immediate memory. Brit.J.Psychol., 55, 303-306.

POSNER, M.I., & ROSSMAN, E. (1965) Effect of size and location of informational transforms upon short-term retention. J.exp.Psychol., 70,496-505.

POSTMAN, L. (1964) Short-term memory and incidental learning. pp.146-201 in Melton,A.E.(Ed.). Categories of Human Learning, New York, Academic Press.

POSTMAN, L., & ADAMS, P.A. (1960) Studies of incidental learning: VIII. The effects of contextual determination. J.exp.Psychol., 59, 153-164.

POSTMAN, L., & RAU,L. (1957) Retention as a function of the method of measurement. Univ.California Publ. Psychol., 8, 217-270.

PRATT, F. (1939) Secret and urgent, the story of codes and ciphers. London, Robert Hale.

RABBITT, P.M. (1962) Short-term retention of more than one aspect of a series of stimuli. Nature, 195, 102.

RICHARDSON, P., & VOSS, J.A. (1960) Replication reports: verbal context and the recall of meaningful material. J.exp.Psychol., 60, 417-418.

RUBENSTEIN, H., & POLLACK, I. (1963) Word predictability and intelligibility. J.verb.Learn.verb.Behav., 2, 147-158.

SANDERS, A.F. (1961) Rehearsal and recall in immediate memory. Ergonomics, 4, 25-34.

SHANNON, C.E. (1951) Prediction and entropy of printed English. Bell Syst.tech.J., 30, 50-64.

SHARP, H.C. (1958) Effect of contextual constraint upon recall of verbal passages. Amer.J.Psychol., 71, 588-572.

SIMPSON, W.E. (1965) Effects of approximation to sentence word-order and grammatical class upon the serial learning of word lists. J.verb. Learn.verb.Behav., 4, 510-514.

SPERLING, G. (1963) A model for visual memory tasks. Human Factors, 5, 19-36.

SUMBY, W.H., & POLLACK, I. (1954) Visual contribution to speech intelligibility. J.acous.Soc.Amer., 26, 212-215.

TANNENBAUM, P.H., WILLIAMS, F., & HILLIER, C.S. (1965) Word predictability in the environments of hesitations. J.verb.Learn.verb.Behav., 4, 134-140.

TAYLOR, A.M., & MORAY, N. (1960) Statistical approximations to English and French. Lang.Speech, 3, 7-10.

THORNDIKE, E.L. & LORGE, I. (1944) The teacher's word book of 30,000 words. New York: Teachers College, Columbia University.

TREISMAN, A.M. (1964) Monitoring and storage of irrelevant messages in selective attention. J.verb. Learn.verb.Behav., 3, 449-459.

TREISMAN, A.M. (1965) The effects of redundancy and familiarity on translating and repeating back a foreign and a native language. Brit.J. Psychol., 56, 369-379.

TULVING, E., & ARBUCKLE, T.Y. (1963) Sources of intratrial interference in immediate recall of paired associates. J.verb.Learn.verb.Behav., 1963, 1, 321-334.

TULVING, E., & ARBUCKLE, T.Y. (1966) Input and output interference in short-term associative memory. J.exp.Psychol., 72, 145-150.

TULVING, E., & PATKAU, J.E. (1962) Concurrent effects of contextual constraint and word frequency on immediate recall and learning of verbal material. Canad.J.Psychol., 16, 83-95.

UNDERWOOD, B.J., & SCHULZ, R.W. (1960) Meaningfulness and Verbal Learning. Philadelphia, Lippincott.

WICKELGREN, W.A. (1965) Acoustic similarity and intrusion errors in short-term memory. J.exp.Psychol., 70, 102-108.