Gazzaniga-
Textbook Chapter

Geoffrey R Loftus

The Experimental Analysis of Memory: **GEOFFREY LOFTUS**

Previous chapters of this book have viewed man and man's behavior primarily from a biological point of view. However, another way to conceptualize a number of psychological phenomena - in particular, phenomena having to do with memory - is in terms of <u>information</u>. Psychologists who view man as an information processor generally think in terms of computers; the old cliche that a computer is a "giant brain" is turned around and the brain is thought of as a "compacted computer" (although a considerably more complex one). As we know, computers accept information from the outside world (for example, in the form of IBM cards) and then process this information to make decisions on the basis of it and generate whatever it is that the computer user wishes to know. Similarly, human beings accept information from the outside world (through their eyes and ears for example) act upon this information to produce appropriate responses and thereby interact with the environment.

The same information may be represented in a variety of different ways, Suppose, for example, one wants to convey the configuration of checkers on a checkerboard. One way to do this would be to write down on a list all 64 squares, noting what piece was on each square. Another way would be to draw a diagram of the board with X 's in the squares that contained black checkers and O's in the squares which contained white checkers. Notice, of course, either of these representations of the board could be used to produce the other. Such a transformation of information from one form to another is called <u>recoding</u> which is a very important concept in many theories of how people process information. Recoding may preserve all the original information, as in this example, or some information may get lost in the process.

If in the checkerboard problem the recoding had taken the form "there are eight white checkers and ten black checkers on the board", a good deal of the original information - namely the position of the checkers - would be lost.

Many psychologists take the view that in a human being, information about the environment is constantly being taken in through the senses and is then succesively recoded and stored for varying periods of time in a series of stages. Stop for a moment and (without looking back) try to remember what was said in the first few sentences of this chapter. Perhaps you can recollect that it had something to do with physiological as opposed to informational views of man. But notice a number of things: first, rather than <u>seeing</u> the words in your "mind's eye" you probably, if anything, heard them in your "minds ear". This means that information which was originally <u>visual</u> (words printed on a page) somewhere along the line got recoded into <u>acoustic</u> information. Notice also that you can't remember the exact words of the sentences, but rather you can only remember the gist. This means that at some stage, a good deal of the information got lost. The general questions dealt with by information processing theories are the following: (a) What information from the environment does a person take in? (b) What are the stages within the person through which the taken-in information passes? (c) How is the information represented in these various stages and (d) how much and what kind of information is lost when it is transferred from stage to stage?

We have been using the expression "stage" very loosely. In physiological terms, stages may be defined with reference to specific parts of the nervous system - e.g., one stage might correspond to the retina, one to the lateral geniculate body and so on. In information-processing terms, however, a stage is thought of somewhat differently: it is conceptual rather than physical. Stages through which information passes do not necessarily have physiological referents. As will be seen in the next section, they are most conveniently thought of as stages of <u>memory</u>.

Memory

We know that people have memory because we can convey some information to them and then later on, they are able to report back at least some of the information. During the intervening time, the information must have been stored somehow inside the person, and this is what is meant by memory.

Short and long-term memory. Intuitivly, it seems that there is more than one kind of memory. If someone reads you a telephone number and then asks you to repeat it back, you can generally do it. You must have had the number stored in memory for a few seconds, but were you using the same kind of memory in which your own telephone number is stored? Probably not. William James, in the 19th century proposed two distinct memories. He defined "primary memory" as containing that information which has never left consciousness, and "secondary memory" as containing information which is stored more or less permanently (e.g., our names, our ability to speak a language, the multiplication table, etc). Thus we can intuit at least these two types of memory, which are in current jargon are called short-term memory (STM) and long-term memory (LTM). The former is a transient sort of memory, the latter much more permanent. Later emprical evidence will be presented in support of this dichotomization of memory. First, however, there is one other kind of memory to deal with.

The iconic store. Suppose a subject in an experiment is shown an array of ten letters for a very brief period of time - say 50 milliseconds - and is then required to write down all the letters he can remember. Typically, people can only report four or five of the letters; interestingly, however, subjects in such an experiment generally claim that they can actually remember much more than that - they have a visual image of the array of letters, but that this image fades away as they are making their report, so that as soon as they have reported the first few letters, the image of the rest has disappeared. In fact, naive subjects believe that the physical stimulus is actually fading away when in reality it is no longer present. ~~These~~ Such introspective reports suggest that immediately after a visual stimulus is presented, there is an extremely brief memory of it, and that this memory, while it is there, contains all the information in the original stimulus. This possibility was tested in a series of experiments by George Sperling (Sperling, 1960). The question to which Sperling addressed himself was: how much information is available to a subject immediatly after the stimulus is turned off, and how does this information decay with time? To get at this, Sperling devised a very clever technique called partial report. The stimuli used were arrays of letters presented for 50 msec (see Figure 1). The twist was that subjects were only required to report the contents of one row of letters rather than having to report the entire array; however, they were not told which row to report until after the stimulus had disappeared. This was achieved by playing the subject a high, medium or low frequency tone immediately after the stimulus was turned off. A high frequency tone meant that the subject was to report the contents of the top row; similarly, a medium or low frequcny tone signaled that a report of the middle or bottom row was to be made. Almost invariably, subjects were able to report back almost all the letters of the appropriate row. But since a subject didn't know at the time the stimulus was on which row he was to report, this must mean that he must have had almost all the letters of the array at his disposal at the time the stimulus disappeared. Apparently, immediatly after a visual stimulus is terminated, a memory remains containing all the information of the stimulus. We shall call this memory the iconic store, *using the terminology of Neisser (1967).

Insert Figure 1-1 about here

*Sperling called it visual information store. Other people call it the sensory register.

We know that this memory is very brief, since, as we said above, if subjects are asked to report the <u>whole</u> array, the best they can do is about 4 or 5 letters; thus in the short time it takes to report these 4 or 5 letters, the iconic store has faded away. Exactly how long <u>does</u> the iconic store last? Using Sperling's technique, this was an easy question to answer. All it requires is that the interval between the offset of the stimulus and the onset of the signlaing tone be varied. Presumably, the longer this interval is, the more the iconic store will have faded away, and the poorer will be the report. The results of such an experiment are shown in Figure 17-2. Notice that the ordinate represents the estimated total letters available to the subject; this is obtained by simply multiplying the number of remembered letters per row by the number of rows. As can be seen, when the interval is zero, almost all letters are available; this is the result described above. However, after about a second, there are only about 5 letters available which is the number which can be reported when the subject is asked to give back the entire array. Evidently, in this experiment iconic store fades away completely by a second. (Other experiments have placed this valve atabout quarter of a second.). In a later section, we will consider what happens to the information that had been in the store.

Insert Figure 17-2 about here

To reiterate, we have enumerated three types of memory: the iconic store, STM, and LTM. Other types of memory (eg., "intermediate-term memory") have been postulated by various investigators, but these three constitute the general bases of most integrated theories of memory (for example, Waugh & Norman, 1965; Atkinson & Shiffrin, 1968). The remainder of this chapter will be aimed at describing the characteristics of these types of memory in somewhat more detail.

For both STM and LTM a number of questions must be dealt with: How is information stored? In what form is the information stored? How is the information retrieved? How is the information lost?

Figure 17-2: The estimated amount of information (letters) available to the subject at varying intervals after the stimulus has been turned off.

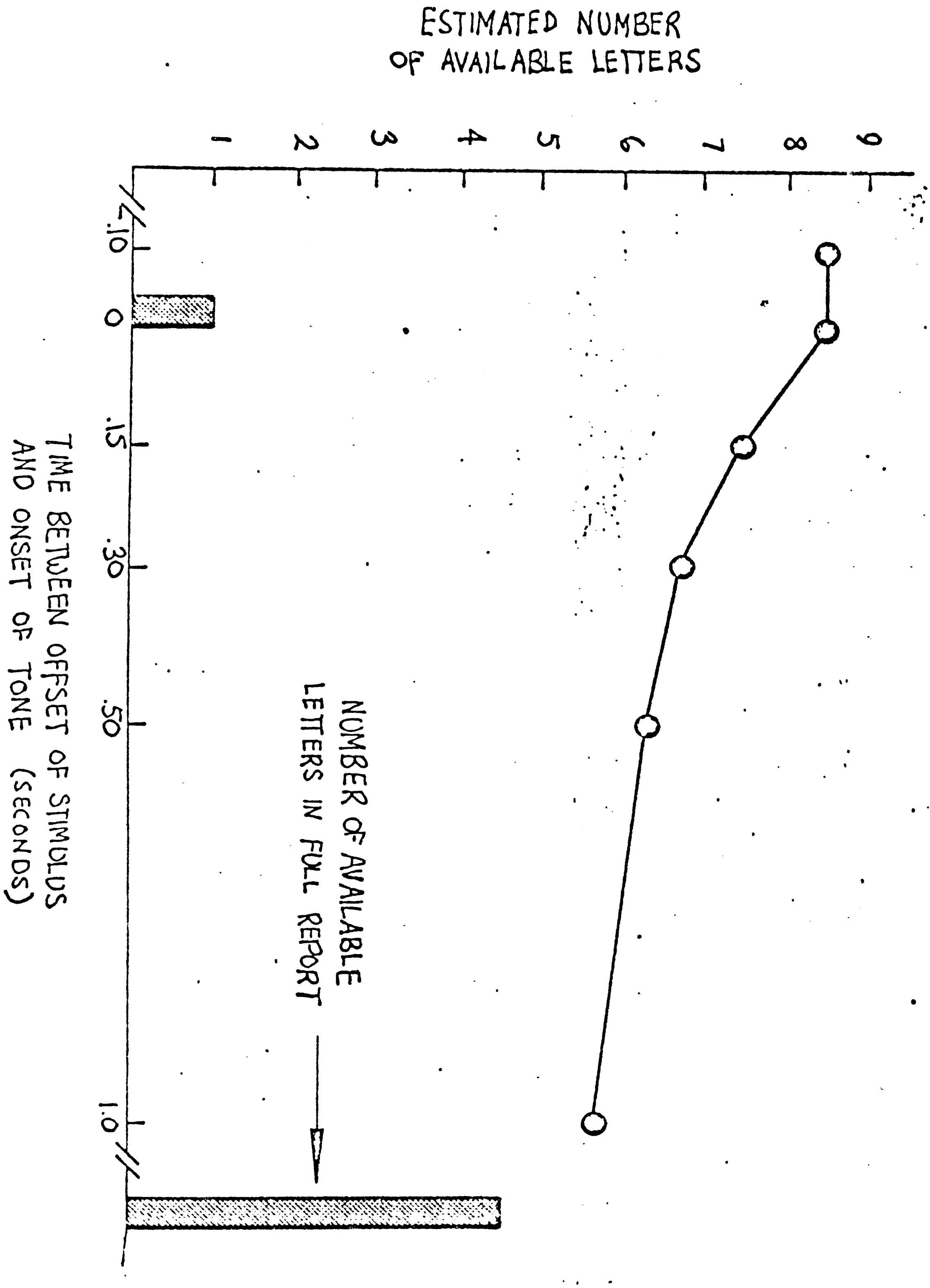
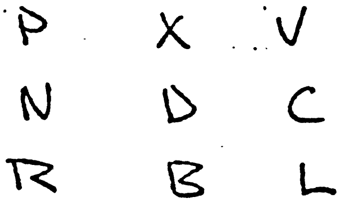

ESTIMATED NUMBER
OF AVAILABLE LETTERS
1
2
3
4
5
6
7
8
9
-.10
0
.15
.30
.50
1.0
NUMBER OF AVAILABLE
LETTERS IN FULL REPORT
TIME BETWEEN OFFSET OF STIMULUS
AND ONSET OF TONE (SECONDS)

Figure 17-1: An example of the stimuli used in the Sperling (1960) experiments. Arrays such as this were presented for 50 msec. After it was turned on, a tone



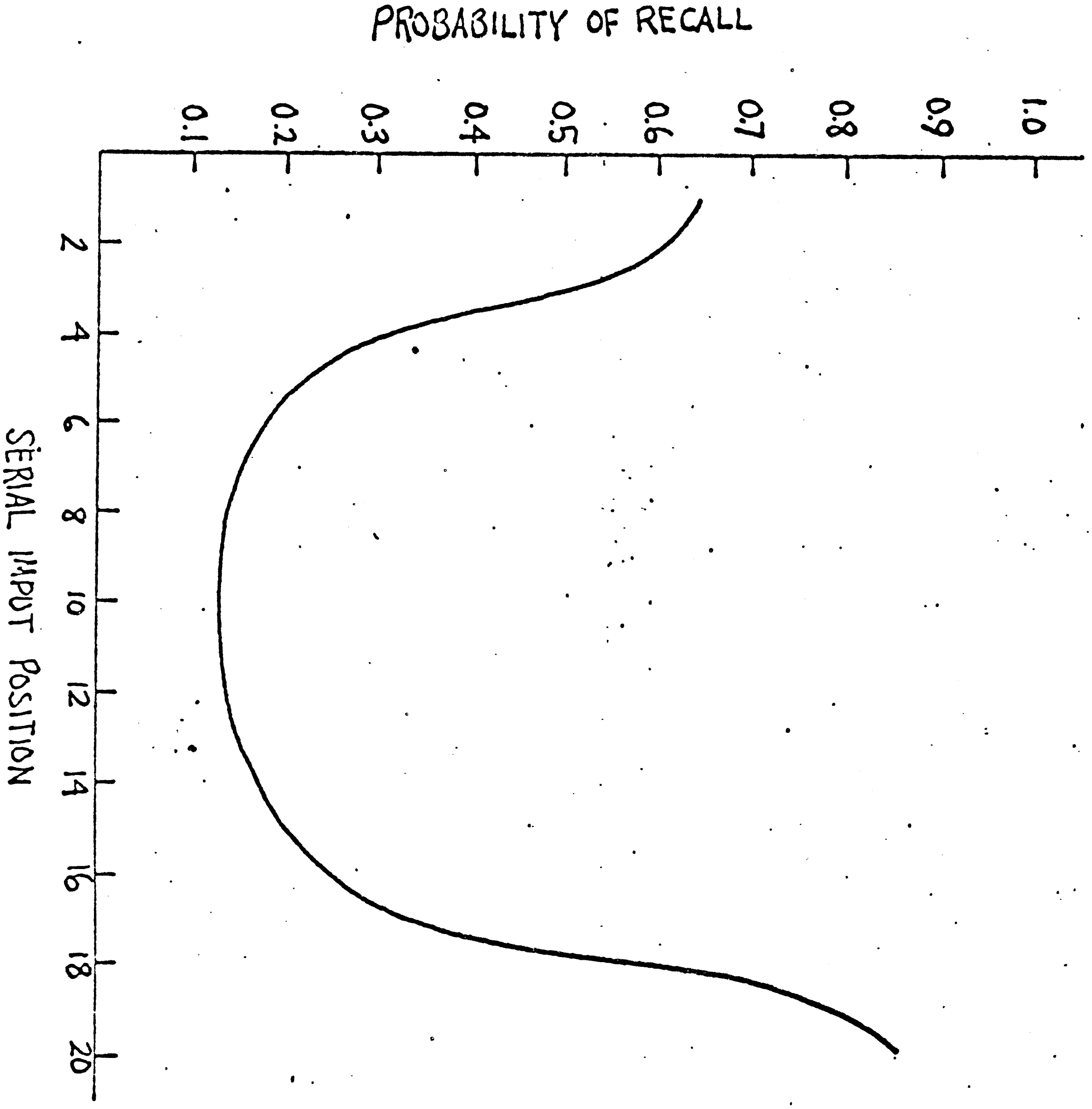
Short-term store. Before presenting a detailed description of STM, we are compelled to face the question: does it really exist? Or, to be more precise, why include it as part of a model of memory? Ten years ago, the majority of memory theorists considered the postulation of STM to be an unnecessary frill in any theory of memory, and claimed that existing data could be handled just as well, and more parsimoniously, by a theory which postulated a unitary memory. Today, however, this viewpoint is no longer viable. There is a large number of results which can only be explained by assuming two separate and distinct types of memory. For the present, we shall describe two results which provide support for such a dual memory theory.

Milner's patients. The first type of evidence is of a clinical and antecdotal rather than an experimental nature. Brenda Milner, a McGill University neuropsychologist, has worked with several patients who have a rather strange disorder. Originally these patients were epileptic; as a means of stopping their seizures, certain types of lesions (bilateral hippocampal) were made in their brains. At first, it seemed that these operations had had no adverse memory effects - the patients were able to remember things that had happened to them in the past, and performed normally on several types of memory tests such as digit span (i.e., they had no trouble remembering telephone numbers). However, a serious deficit soon became apparent - these patients could not learn anything new! If a new doctor came in to talk with such a patient, the two could converse normally. However, if the doctor left the room for five minutes, the patient would have no memory of every having seen him before, when he returned. This syndrome can easily be explained in terms of a dual process theory of memory - it need only be assumed that the lesions had been made in that region of the brain involved in transferring information from STM to LTM. As we pointed out, STM and LTM themselves are quite normal for these patients. The patients can retrieve things from LTM such as their names, the multiplication table, childhood events - in short, everything they had learned and stored in LTM prior to the operation. Similarly, situations which only involve the use of STM cause no difficulty - the patients can remember small amounts of information for a short time, or they can rehearse information and keep it available indefinately. The difficulty is that no new information can be transferred to LTM. Once the information leaves STM, it is lost forever, for it has nowhere else to go.

Free recall. An experimental paradigm which is very useful for studying memory phenomena is that of free recall. A free recall task goes as follows: a list of words (say twenty words) is presented to a subject, one word at a time. The subject is then required to repeat back as many of the words as possible. One result which invariably shows up is known as the *serial position effect*. Suppose we plot the probability of a word's being recalled as a function of its serial input position (By "serial input position" we mean the first word that had been presented to the subject during study, the second, third and on up to the last word). The resulting curve looks like that shown in Figure 3. The first and last few words presented are remembered best, while the words in the middle of the list are remembered the poorest.The reasons why the first few words are remembered so well are fairly complex and will be discussed in a later section. For the moment, consider why the subject is so good at remembering the last few words. Proponents of a single-memory theory would make the following argument: Items which were presented most recently are remembered the best. Therefore, the last word, which was presented the most recently will have the highest probability of being recalled, the second-to-last word presented will have the second highest probability, and so on. However, those arguing in favor of two separate memories would say the following: Words which are in STM are recalled perfectly, whereas words which are not in STM,

Insert Figure 7-3 about here

Figure 17-3: For a free-recall paradigm, the probability of a word's being recalled as a function of its serial input position


PROBABILITY OF RECALL
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
2
4
6
8
10
12
14
16
18
20
SERIAL INPUT POSITION

but only in LTM (if they were stored at all) have a much lower probability of being recalled. The closer a word is to the end of the list - i.e., the more recently it was presented - the higher the probability that it is still in STM, and thus, the higher its probability of being recalled.

A simple experiment can be run to distinguish these two hypotheses. Suppose that after the subject has been presented the list (but before he recalls it) he is required to perform some complex mental activity (such as a difficult arithmetic task). A one-memory theory would say that after this activity, the last words in the list should still be remembered the best (since, after all, they still were the most recently presented words). Figure 4a shows what the serial position curve of the experiment should be according to this hypothesis. However, a dual trace theory would predict that the mental activity would clear STM of the words in the list; thus when the subject recalls the list, *everything* would be recalled from LTM and the probability of recalling the last few words should be no greater than the probability of recalling words from the middle of the list. Figure 4b shows what the serial-position curve would look like according to a dual-memory hypothesis.

Insert Figure 3-4 about here

This experiment has been performed (coincidentally, it was performed independently in two different laboratories at the same time and is reported by Glanzer & Cunitz, 1966; and Postman & Phillips, 1965). The results are clear-cut and look like those depicted in Figure 4b. Again strong support was given for hypothesizing a STM which is separate from and independent of LTM.

We therefore have fairly strong experimental evidence that there exists at least two separate memory stores. A very large body of research has been aimed at making a detailed examination of the characteristics of these stores. Let us now explore some of these characteristics.

Storage of information in STM - transfer from iconic store. It will be recalled that George Sperling performed experiments to demonstrate the existence of the iconic

Figure 17-4: Hypothesized outcomes of the intervening-activity experiment (see text) according to two different theories of memory.

store. Sperling did not stop there, however, and he next attacked the question of how information is transferred from iconic store to STM. The first question was: at what rate is information transferred? This, of course is an extremely important question reflecting on man's basic ability to handle information. Apparently, all information from the outside world impinges initially upon the sense organs. The capacity of a human being to informationally interact with the environment is thus initially limited by the rate at which he can extract information from iconic store and enter it into STM. To examine the transfer rate, an obvious experiment suggests itself: simply present information (say arrays of letters as in Sperling's experiments) for varying periods of time and see how much of the information can be reported. Since the information has to be transferred to STM before it can be reported, we can infer from such an experiment how much information per unit time gets transferred. Suppose, for example, that when we present an array of letters for 20 msec two letters can be reported and that when we present the array for 30 msec, three letters can be reported. We would then assume that one letter gets transferred to STM every 10 msec. However, this obvious experiment has an obvious difficulty. Remember that a major characteristic of iconic store is that it lasts lasts about 250 msec.; thus, there seems to be no way to present information for less than this time. Fortunately, there is a solution to this dilemma. It turns out that if any bright stimulus is flashed to a subject, it destroys the current contents of iconic store. Thus, if the array of letters is presented for say 30 msec and then is immediately followed with some other bright visual stimulus (called a "masking stimulus") we can be sure that the letters were available to the subject for only the 30 msec.

The following experiment was thus performed: letter arrays were presented for periods of time varying from 5 to 60 msec. As soon as the array was turned off, it was followed with a masking stimulus (for various reasons, Sperling used a field of bits and peices of broken letters as a masking stimulus). Now we know that any letters reported by the subject must have been transferred to STM, and furthermore, we know exactly how long he had to transfer them. The results of the experiment are shown in Figure 17 5. One letter can be transferred to STM for every ten msec the stimulus was on, up to about four letters. Apparently, then, STM is filled up from iconic store at the rate of about one letter every ten msec. Why does the number of reported letters level off at about four? Remember that STM has a very limited capacity, and it seems that in this experiment, its capacity is about four letters. Capacity of STM will be discussed in more detail in a later section.

Insert Figure 13-5 about here

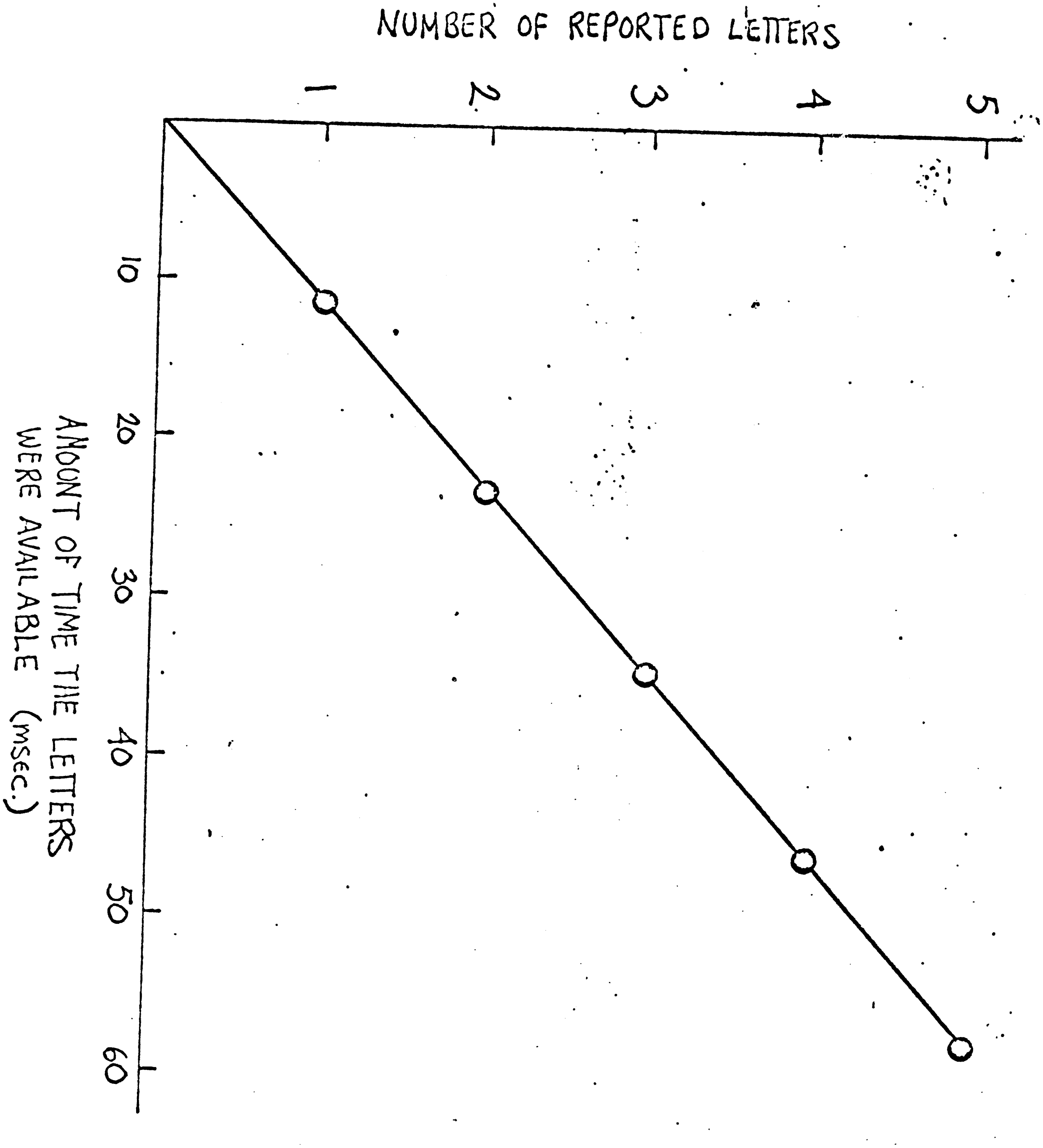
So far, our discussions of iconic store have concerned memory that is visual. The question may come up: what about iconic stores or their analogues in other modalities? There may be such stores but relatively little is known about them. Some work has been done (Moray, 196 ; Crowder, 1972) showing the existence of such a store in the auditory store (termed "echoic store" by Neisser, 1967). We have, however, chosen to focus on vision because a good deal more is known about it than about other modalities.

Form of information storage in STM.

In what form is information stored in STM? Remember that the information in iconic store is visual, faithfully reflecting the original stimulus. However, the situation is quite different in STM. There is a good deal of evidence that while information is being transferred to STM it is also being recoded into acoustic form, and that the representation of information in STM is acoustic, or verbal, in nature.

Evidence for this contention comes from studies of STM in which strings of letters are visually presented to a subject, who is required to repeat them back. These letter strings are either acoustically distinct (e.g., ENXWQHM) or acoustically

Figure: Number of letters transferred from the iconic store to star as a function of the amount of time the letters were available. About one letter is transferred every 10 msec. (After Sperling 1960)


NUMBER OF REPORTED LETTERS
1
2
3
4
5
10
20
30
40
50
60
AMOUNT OF TIME THE LETTERS WERE AVAILABLE (MSEC.)

confusable (e.g. ETVPDZCG). Subjects remember the acoustically distinct lists considerably better than the acoustically confusable ones (Conrad & Hull, 1964). Other evidence for acoustic encoding in STM comes from an examination of errors in short-term recall. Both Conrad (1964) and Sperling (1967) have pointed out that when a subject makes an error in a short-term recall task, his response is likely to sound like (as opposed, for example, to look like) the correct response. Thus, if the subject is trying the remember the letter "B" he is much more likely to erroneously respond "T" which sounds but does not look like B than "R" which looks like but does not sound like B. This is the case even when the original presentation of the letters was visual.

Capacity of Short-term memory. Throughout this chapter, it has implicitly been assumed that STM is of limited capacity. We now address ourselves to the question: "How limited is limited?" or, "Just what is the capacity of STM, anyway?" Before we can measure capacity, however, it is necessary to consider what actually goes into STM. That is to say, we cannot determine how many things we can put into STM without first knowing what those things are. In measuring the capacity of a bathtub or a thimble or a cement truck, we can express the result of our measurement in quarts or cubic centimeters or whatever. Unfortunately, as we shall see, finding such a unit of measurement for memory capacity is not such an easy task.

At first glance, the solution seems straightforward. We have been claiming that information is the stuff that fills up memory, and information has a unit of measurment, namely a bit (see chapter XX). A good hypothesis, then, would be that the capacity of STM is some contant number of bits.

. An experiment which provides conclusive evidence that STM capacity cannot be measured in bits was carried out by Hayes (1952) and replicated by Pollack (1953). The experiment is simple. Consider a task in which a list of things is read to a subject whose job it is to simply remember and report as many of them as he can. (This is called a "memory span" experiment, and if you have ever taken an IQ test, your memory span has probably been measured). Of interest is the number of such things that the subject can remember. Now consider what the "things" may be. They can be binary numbers (i.e., ones and zeros) or they may be letters of the alphabet. Notice that these classes of things differ in the amount of information that each member of the class contains. Each binary number contains one bit of information, whereas a letter contains about 4.7 bits of information. It follows that if STM capacity is constant in terms of information, then a subject should be able to remember 4.7 times as many binary numbers as letters. (That is, suppose that the capacity of STM were 9.4 bits. Then it should be able to hold two letters, at 4.7 bits per letter, or 9.4 binary numbers at one bit per binary number). However, this turns out not to be the case, as Figure 5.5 shows. This figure depicts the results of the Hayes (1952) experiment. On the abscissa is the information per item, and on the ordinate is the number of items retained. As can be seen, approximately the same number of items is retained no matter how much information is contained in each item. (Although the curve drops slightly, it is not even in the same ballpark as a result which would support a constant-information hypothesis.) The main point of the experiment is that a person can retain on the order of about seven (plus or minus two) things, whether the things are binary numbers, decimal numbers, letters or words. Thus, a person retaining seven binary numbers is retaining seven bits of information, whereas a person retaining seven letters is retaining about 38 bits of information. Clearly, then STM capacity is not a constant number of bits.

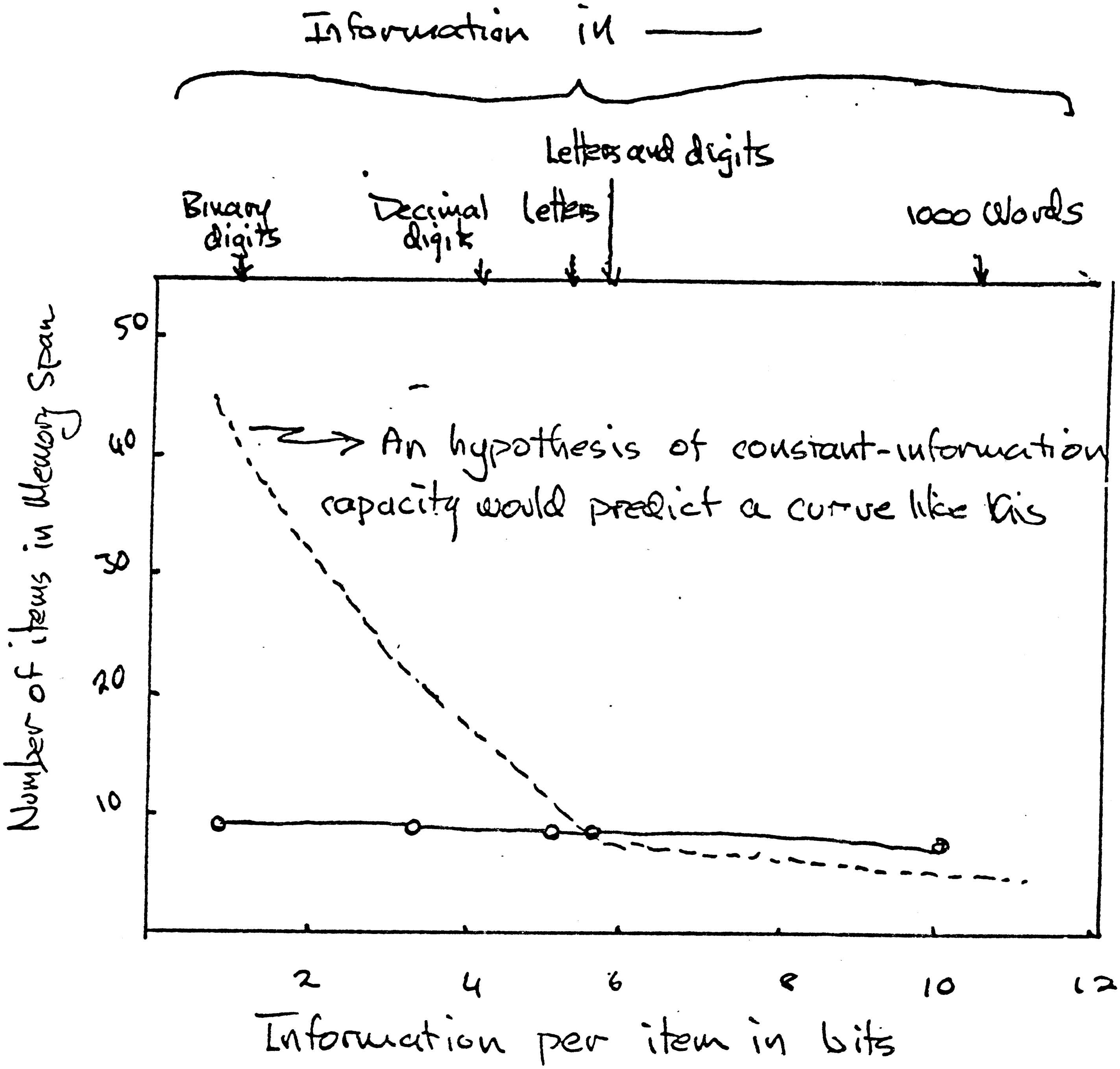
Insert Figure 7-6 about here

It may therefore be argued that STM capacity is constant, but it is constant in a paradoxical sort of way - it can retain about seven of anything! Miller coined the word "chunk" in referring to that which STM can hold about seven of. A chunk can contain a highly variable amount of information ranging from one bit if the

Figure 17-6: Number of items in the memory span as a dunction of amount of information per item. The observed curve is fairly flat at about seven items. The dashed curve depicts the kind of prediction a constant-information sts capacity hypothesis would make. (From Hayes, 1952)

17-6

Figure ~~5~~



Information in ——
Letters and digits
Binary digits
Decimal digit
letters
1000 Words
50
40
30
20
10
Number of items in Memory Span
An hypothesis of constant-information capacity would predict a curve like this
2
4
6
8
10
12
Information per item in bits

chunk is a binary number to many bits if the chunk is a letter or a word.

This finding is paradoxical because it must, by appropriate recoding schemes, be possible to increase the number of low-information chunks one can retain. A possible recoding scheme would be the following: Suppose you are given a series of binary numbers to remember. Suppose further that you had decided beforehand that any particular string of four binary numbers would correspond to a particular letter, as shown in Table 1. Now as you are presented with the binary numbers, you simply take each sequential group of four and recode them into the appropriate letters. Now how many binary numbers can you retain? You can retain seven letters, and each letter corresponds to four binary numbers, so you can retain a total of 28 binary digits, which is quite an improvement over seven! That this can indeed be done was shown in a demonstration by Smith (described in Miller, 1956) who, using recoding schemes such as the one described above, was able to remember up to 40 binary numbers! These types of schemes are also used by professional "memory experts" and are taught in the "memory improvement schemes" that you so often see advertised in your Sunday newspaper.

Lest you think that such recoding schemes have little relationship to what goes on in everyday life, notice that you are using exactly such schemes when you remember words. Let's assume that you are being presented strings of words to remember. As we know, the number of words you can remember is about seven. Suppose now that each word is five letters long. If asked to spell the words as you were recalling them, you could probably do so. But then you would be remembering 35 letters! How can this be reconciled with the notion that only seven letters can be remembered? The answer is that only seven letters can be recalled when the letters are <u>unrelated</u>. When the letters are organized into words, then we are recoding the letters into chunks which are informationally richer. We can do this, however, only because we have already learned the recoding scheme - that, for example, the letters C, L, O, W and N can be recoded into the word "clown" and then remembered as one word rather

Table 17-1: A possible recording scheme for remembering strings of binary numbers. The recording scheme is memorized and then each string of four binary numbers is recoded into a letter, as shown in the table.

String of Binary Numbers	Letter
0000	A
0001	B
0010	C
0011	D
0100	E
0101	F
0110	G
0111	H
1000	I
1001	J
1010	K
1011	L
1100	M
1101	N
1110	O
1111	P

than as five separate letters. This is no different from the recoding scheme described above where the numbers 1, 0, 0 and 1 are recoded into the letter "J" and remembered as one letter rather than as four separate binary numbers.

Thus, we can see that the "capacity" of STM is an elusive quantity. We say that it can hold seven "chunks;" however, this is a loose and qualitative specification, since a "chunk" is so ill-defined. The issue of STM capacity is confused due to the fact that what we are remembering out of STM is combined with recoding schemes which we have learned in the past and have stored in LTM. Hopefully these processes can, at some point, be un-confounded and STM will be specifiable in informational terms.

A computer analogy may be helpful in understanding the distinction between "pure" STM capacity and capacity which is enhanced by recoding schemes built into LTM. Let us consider a computer which has a large amount of memory but which has seven special words of its memory reserved and designated as "STM". If we are going to use only those seven words then the amount of information which we can put into them is limited and is exactly specified (in fact, in bits, it is seven times the number of bits per word). But now let us assume that we can use the rest of memory as well. We can then fill up our seven slots of STM with addresses in LTM, and at those addresses might be found arrays of words which, depending on the lengths of the arrays, contain an indefinite amount of information.

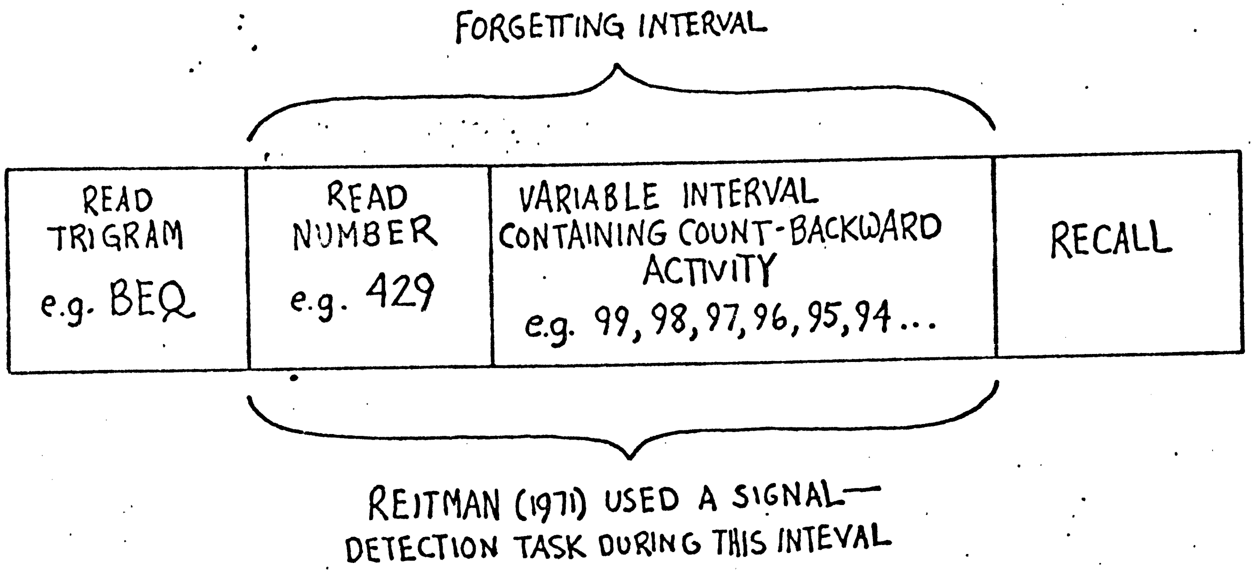
Thus, by using the seven slots of STM as pointers to other information (as opposed to containing the information proper) we can greatly increase the amount of information that those slots could contain. If we are trying to remember seven binary digits, we could have in each of the seven slots of STM a pointer to a location in LTM that would contain either the information "one" or "zero". If we are trying to remember seven words, each of the STS slots could point to a location in LTM which would specify the word. This scheme can be pushed to any length. Suppose we are asked to remember seven Shakspeare solloquies. We could put in each STM slot an address which could point in LTM to a place where a particular Shakespeare solliquy is stored. But notice that, as we have been stressing, this scheme depends on having the other information already carefully organized in LTM. For most people, words are already stored as units. If we really wanted to see a case where people could remember seven Shakespeare solliquies "in STM" we would have to choose people who had already memorized all of Shakespeare, and thus had each solliquy stored as a unit.

Forgetting from STM. Another basic characteristic of STM is that unrehearsed information decays away from it very rapidly. (By "rehearsed information" is simply meant information which is repeated over and over as the means of maintaining it in STM. . We shall discuss the concept of rehearsal in more detail in a later section). An experiment which showed exactly how fast this decay actually is was performed by Peterson & Peterson (1959). Their procedure was quite simple: first, a three-consonant trigram (such as BZQ) was presented to a subject for two seconds. Then at intervals ranging from zero to 32 seconds later, the subject was asked to recall the trigram. A problem in this experiment was to prevent the subject from rehearsing the trigram during the forgetting interval. To accomplish this, a three-digit number was read to the subject immediately after he had studied the trigram, and he was asked to begin counting backward by threes from that number. The entire design of the experiment is schematized in Figure 13-7.

Insert Figure 13-7 about here

The results of this experiment are rather interesting, and are shown in Figure 13-8. This figure depicts the probability of correctly recalling a trigram as a function of the interval during which the memory of the trigram had been allowed to decay. This probability is over 0.90 when zero time has elapsed, and drops to

Figure 17-7: The basic design of a Peterson & Peterson type experiment.


FORGETTING INTERVAL
READ TRIGRAM e.g. BEQ
READ NUMBER e.g. 429
VARIABLE INTERVAL CONTAINING COUNT-BACKWARD ACTIVITY e.g. 99, 98, 97, 96, 95, 94...
RECALL
REITMAN (1971) USED A SIGNAL— DETECTION TASK DURING THIS INTEVAL

about 15% after only about 15 seconds. From then on, recall probability is about constant, at least out to 32 seconds. Why does recall probability not decay away to zero? The answer to this question is probably that there is a little information about the trigram that has been transfered into LTM which remains indefinitely, and this information can be used for responding if the short-term memory has decayed away. Thus the point at which this forgetting function reaches asymptote - about 15 seconds - is probably about the point at which STM traces of the information have completely disappeared. Further evidence for this contention will be discussed in a later section.

Insert Figure 17-8 about here

This experiment thus provides a rather clearcut view of the time course of information decay from STM. The experiment has been replicated many times, and the function is always very close to that shown in Figure 17-8. A replication by Murdock (1961) used triads of words rather than trigrams of letters, and he still found the same phenomenon. This result provided evidence not only that the Peterson & Peterson finding is very stable, but it also fits in very nicely with the chunking notion of STM capacity discussed in the last section.

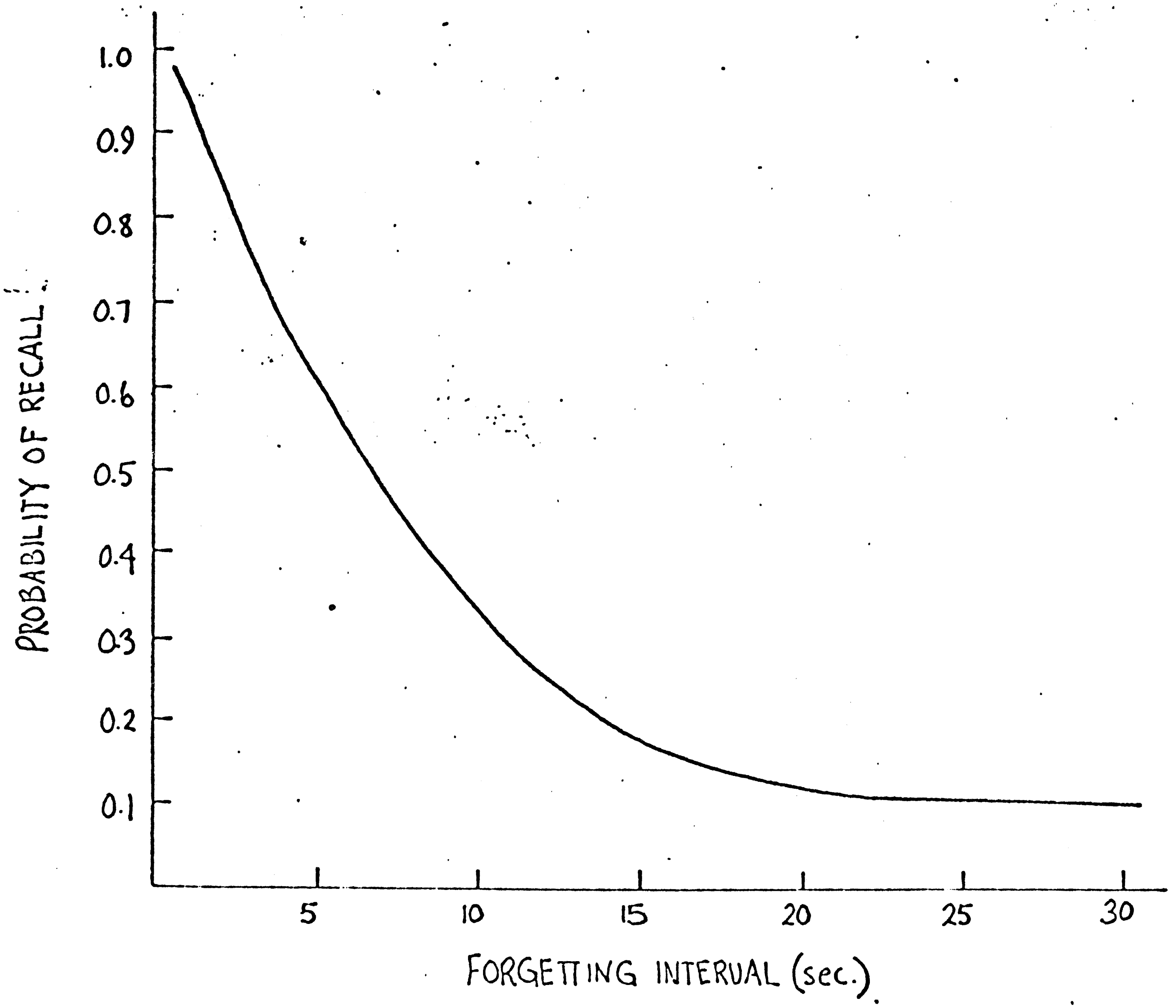
What caused decay of information from STM? Two basic hypotheses, the decay hypothesis and the interference hypothesis have been postulated. The decay hypothesis states that any information in STM which is not rehearsed will decay away of its own accord, much the same way as a swing will stop swinging if it is not pushed or otherwise propelled. The alternative hypothesis maintains that forgetting is a byproduct of the limited capacity of STM. According to this hypothesis, information is only lost from STM because it gets shoved out, so to speak, by other information. The interference notion would explain the Peterson & Peterson results by assuming that the count-backward task caused new information to be entered into STM which eventually displaced the information necessary to remember the trigram.

A clever experiment to distinguish between these two hypotheses was performed by Reitman (1971). Reitman used a Peterson & Peterson task; she reasoned, however, that instead of counting backward during the forgetting interval, some kind of intervening task was needed which prevented rehearsal, but which did not cause new information to enter STM. One task which fits this bill is called a "signal detection task." Throughout the forgetting interval, from zero to five tones were played to the subject through earphones and the subject's job was to detect these tones (signals) and push a button each time he heard one. The tones were so soft that the subject could only detect them about half the time. The task was thus very difficult, requiring sufficient concentration that any other activity such as rehearsing the trigram was effectively prevented. Notice, however, that this task does not require entering any new information into STM, as does a count-backward task.

The design of the experiment is, then, like that of Peterson & Peterson (see Figure 17-7) except that a signal-detection task replaces counting backward as the activity performed during the forgetting interval. What predictions do the two forgetting hypotheses make about the outcome of this experiment? The decay hypothesis would say that since rehearsal has been prevented, the information should decay away as usual, and the forgetting curve should be no different from that obtained by Peterson & Peterson. However, the prediction of the interference hypothesis would be quite different. It would state that since no new information has been entered into STM during the forgetting interval, no forgetting should take place.

The results of the experiment were dramatic--after 15 seconds, median recall was 98%! The information had not been at all degraded

Figure 17-8: Results of Peterson & Peterson (1959). The trigram is forgotten very quickly as the forgetting interval increases up to 15 seconds. The asymptote of the curve represents long-term memory strength.


PROBABILITY OF RECALL
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
5
10
15
20
25
30
FORGETTING INTERVAL (sec.)

after an interval during which a count-backward task reduced recall to 15%. The subjects reported that during the interval, they were concentrating on the detection task, they were not rehearsing, and were not thinking of the trigram they were supposed to remember, but that when they were asked to recall it it was "just right there, hanging around." Apparently, there is no such thing as "atonomous decay" from STM and it is indeed the case that forgetting takes place simply due to the fact that information is only lost from STS when other, new information is entered in.

Retrieval from STM - So far, we have discussed how information is entered into STM, what form the information takes, and how the information is forgotten. Now suppose we need to use some of the information in STM - that is, we need to retrieve it. This section deals with characteristics of the retrieval process.

Earlier, it was asserted that information from STM is retrieved rapidly and completely. Several questions now arise about the retriveal process. First, how fast is retriveal? Secodnly, what is the nature of the process? In particular, suppose we have information in STM and we wish to retrieve some part of it. Can the desired information be simply "plucked out" or is it necessary to perform some kind of a search to find what we are looking for? If the latter turns out to be the case, what kind of search is made? A series of experiments by Saul Sternberg provides some provocative answers to these questions.

Sternberg uses the following general task for the experiments presented here. A set of from one to six digits (called the memory set) is first presented to a subject for a few seconds, to be entered into STM. Shortly thereafter, a single digit (the test digit) is presented. The task of the subject is to decide whether or not the test digit was a member of the memory set, and to respond by pulling one lever if the answer is "yes" and another lever if the answer is "no". Subjects are urged to respond as rapidly as possible without making errors (in practice, the error rate is very low: the subject pulls the wrong lever only about two percent of the time). The main dependent variable is the reaction time (RT) measured from the onset of the test digit to the subject's response.

17-9

Insert Figure 9 about here

Typical results of this kind of experiment are shown in Figure 17-9. The reaction time to make either a "yes" or a "no" response is related in a very systematic way to the size of the memory set: the function is linear. For each additional member of the memory set, RT is about 38 msec longer. Thus when the memory set size is one, RT is about 435 msec, when the set size is two, RT is 435+38 = 473 msec, when set size is three, RT is 473+38 = 511 msec, and so on. This result is very important, for it suggests that a specific process is taking place during retrieval from STM - namely, that the test digit is being sequentially compared to each member of the memory set and that each comparison takes the same amount of time.

Let us examine in more detail what may be going on when a person performs the type of retrieval task we have just described. Sternberg postulated that the total reaction time is the sum of times taken by a series of information-processing stages, and he attempts to define these stages and determine how long each one takes. Sternberg's stages are the following:

Stage 1. The dest digit is read and encoded into some suitable form. Assume this takes $\underline{a}$ seconds.

Stage 2. The encoded form of the digit is compared sequentially with each of the members of the memory set. Assume that this comparison time (per member) is $\underline{c}$. Thus, if there are n members of the memory set, the total time for comparison will be $\underline{cn}$.

Stage 3. A decision is made to say "yes" or "no" depending upon whether any of the comparisons yields a match. Assume that the decision takes $\underline{d}$ msec.

Stage 4. A response is made. Assume that this takes $\underline{r}$ msec.

Thus, the total reaction time can be expressed as the sum of all the individual times. When there are n members of the memory set,

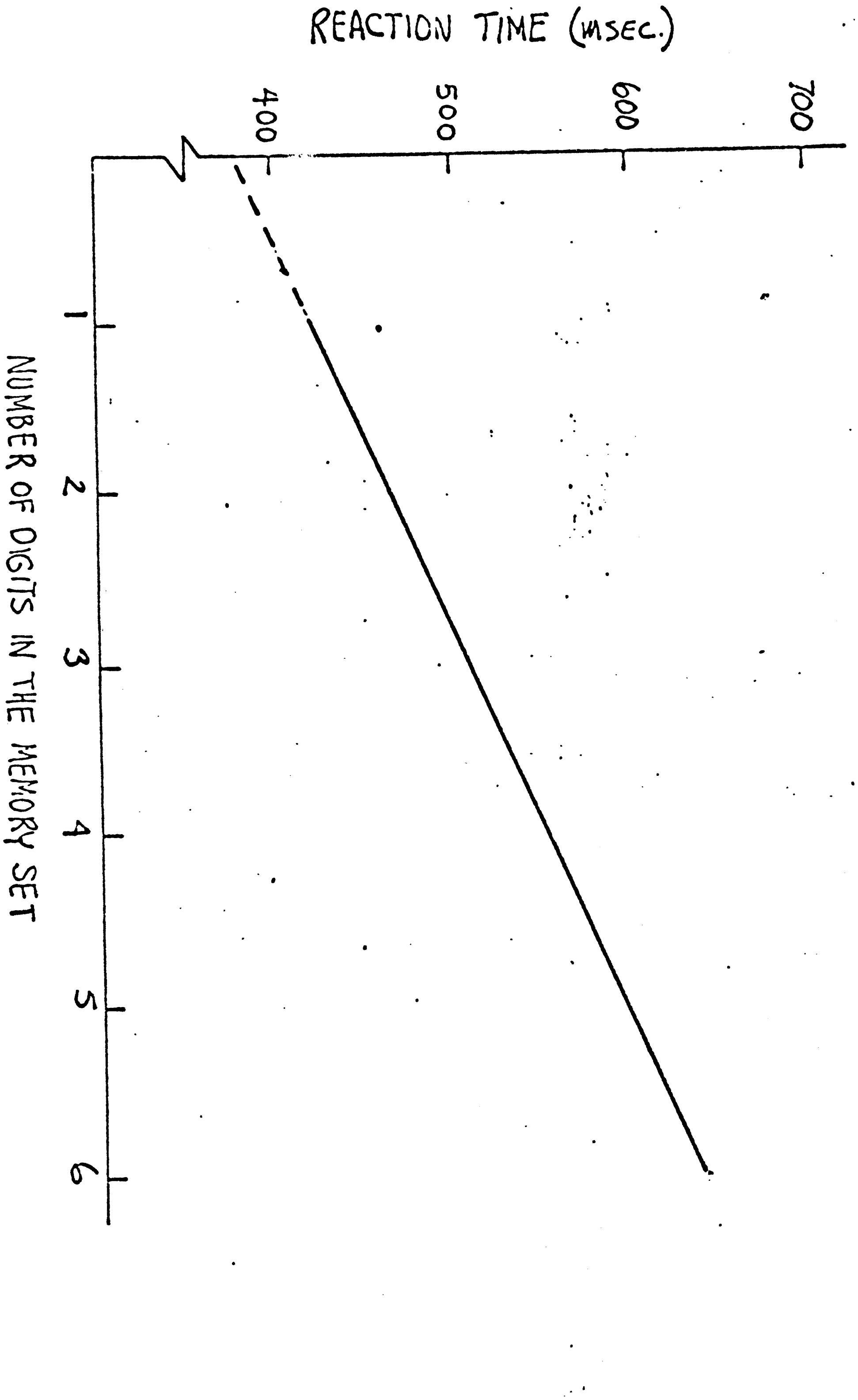
$$RT = \underline{cn} + \underline{a} + \underline{d} + \underline{r} \qquad (1)$$

Remember from algebra that when we have a function of the form, $y = ax + b$, we obtain a straight line if we plot y as a function of x where a is the slope of the line, and b is the y-intercept. This is the form of equation (1) which can be seen more easily if we lump together stages 1, 3 and 4, call this time b (where, of course, $b = a + d + r$) and rewrite equation (1)

$$RT = \underline{cn} + b$$

Consider again, the straight-line function shown in Figure 17-9. Remember, this is a graph of RT as a function of n, the memory set size. The slope of this function is 38 msec., and the intercept is 397 msec. Thus, equation (2) can be rewritten with specific quantities substituted for

Figure 17-9: Reaction time to say "yes" or "no" to a test as a function of the size of the number of digits in the memory set


REACTION TIME (msec.)
400
500
600
700
NUMBER OF DIGITS IN THE MEMORY SET
1
2
3
4
5
6

c and b, or

$$RT = 38n + 397. \tag{3}$$

What does equation (3) mean? It means that, according to Sternberg's results it takes 397 msec for stages 1, 3 and 4 (encoding the test digit, making a decision, and outputing a response). It takes 38 msec to make one comparison of the test digit with a member of the memory set.

One other point should be made about Sternberg's findings. The question arises: Why are the functions relating RT to memory set size the same for "yes" and "no" responses? The reason why this question arises is that one might wish to make the following argument: When the answer is "no" the subject must search through <u>all</u> the n digits in the memory set. He therefore must make n separate comparisons, and the slope of the functions should be the average time per comparison, or c as stated above. Consider, however, a "yes" response. If the subject searches through the memory set until he finds the target digit and then immediately responds, what does this suggest about the number of comparisons which must be made? It means that on the average, <u>fewer</u> than n comparisons will have to be made, since in general, the target digit will be in the middle of the list somewhere. In fact, since on the average, the target will be in the middle of the list, the average number of comparisons made will be $n/2$. The slope of the function relating RT to n will then be $c/2$ rather than c - i.e., the slope of the "yes" function should be only half that of the "no" function.

What are we to make of the fact that the yes and no slopes are the same? We must conclude that even when the subject <u>finds</u> the target somewhere in the middle of the list, he continues scanning until the end of the list before outputing a response. Why do subjects use this seemingly inefficient strategy? It may be that making a decision about whether a particular member of the memory set is or is not the target is a separate process from actually making a comparison of the two. If decision time were very long relative to comparison time, then it may be more efficient to make n comparisons followed by <u>one</u> decision, rather than to make <u>both</u> a comparison and a decision for each member of the memory set. This explanation is, of course, rather post-hoc and tentative. To be convincing, it must be made more explicit, and tested directly.

Sternberg's basic experiment has been replicated many times under varying conditions, and using different types of test material. In all cases, the same qualitative result has been found: a linear function relates RT to memory set size. For most types of material, the time to compare a test stimulus to a member of the memory set is about constant, ranging between 30 and 50 msec. Again, we see a certain degree of consistency regarding short-term memory. Capacity is fairly constant around seven chunks (see last section) and in addition, the time to compare two chunks to see if they are the same is constant at around 30-50 msec.

<u>Short-term memory - a brief recapitulation.</u> Before moving on to a discussion of long-term memory, let us summarize what is known about STS.

1. First, short-term memory is still basically conceptualized in terms used by William James - that is, it is a limited-capacity store, used in a transient way to hold information.
2. Information enters STM from a high-capacity sensory store, and information can be read in very quickly - in some cases, at the rate of one letter every 10 msec, up to about 4 or 5 letters.
3. The maximum amount of information which can be held in STS is

ground seven chunks, where a "chunk" is anything which is organized as a whole in long-term memory (e.g., a digit, a letter, a word, etc.) Thus, in general, it is a poor idea to think of STM as a box in which "real" information is stored; rather it is probably better to conceptualize it as a box containing a list of directions to the places in which the real information is stored.

4. It is usually the case that unrehearsed information is forgotten from STM. However, it appears to be the case that the reason for this forgetting is that non-rehearsal of information is, in general accompanied by the entrance into STM of new information. It is the displacement of the old information by the new which causes forgetting.

5. When information is retrieved from STM, it is retrieved in a very systematic fashion: all the information in STM is searched through in a serial order and the desired information is thereby found. Thus, the more information resident in STM, the longer the process of retrieval.

Long-term Memory

Long-term memory, our repository of permanent or semipermanent information is, of course, a <u>sine qua non</u> of our existence. Without LTM, we would be unable to learn anything, and unable to interact in any kind of dynamic way with our environment. In this section; we shall systematically examine the properties of LTM inquiring as to how information gets there, how it is stored and forgotten and how it is retrieved.

<u>Entry of information into LTM - rehearsal</u>. Atkinson & Shiffrin (1968) have proposed a theory of memory which places a heavy emphasis on the process of <u>rehearsal</u>. Rehearsal of information simply means repeating the information over and over again in STM, and serves two purposes: 1) maintains information in STM for an indefinite period of time and 2) rehearsal serves as a process by which information is transferred from short to long-term memory, i.e., the amount of information in LTM is some monotonic function of the number of times it was rehearsed while resident in STM. At the present time, there is a good deal of evidence to support this notion of rehearsal as an information-transfer mechanisms.

The first study to demonstrate directly the relationship between rehearsal and LTM strength was performed by Hellyer (1962). Hellyer used a Peterson & Peterson task (see figure 17- (7) with the following innovation: when the trigram was first presented for study, the number of times the subject was allowed to rehearse it was controlled. Either one, two, four or eight repetitions of the word were allowed. Otherwise, the task was exactly the same as that of Peterson & Peterson. The results of Hellyer's experiment are shown in Figure 10. Forgetting still takes place; however, the asymptote is higher the greater the number of rehearsals allotted the trigram. What does this mean? Remember that the asymptote of the forgetting curve is thought of as representing the amount of long-term memory strength which has accrued about the trigram. Thus, Hellyer's results suggest that more rehearsal leads to more ~~strength, or~~ information. in LTM.

Insert Figure 17-10 about here

A similar result was found by Hebb (1961). Hebb's paradigm was as follows: Subjects were read a series of nine-digit numbers, and after hearing a number were required to repeat back as many of the digits in the number as they would remember. Unbeknownst to the subjects, however, every third number was the same! Over trials, the number of digits remembered from the non-repeated changing lists remained constant, whereas the amount remembered from the repeated lists gradually increased, as shown in Figure 11. The implication of this result is that, each time that a number is presented, some information about that number is transfered to LTM. If the number is presented and tested again, this LTM information is useful; thus the increase in performance for the repeated numbers. Naturally, if a number is never tested again, then the information about it available in LTM is of no value.

Insert Figure 17-11 about here

Recently, Rundus (1971) has performed a series of experiments which provide major evidence concerning the relationshop between rehearsal and long-term memory strength. Rundus used a free-recall paradigm: lists of 20 words were shown to a subject at about one word every five seconds, and the subject recalled back as many of the words in the list as he could, in any order. Rundus's innovation was that during the time the words were being presented, subjects were required to rehearse out loud any words they wanted from the list they were currently studying, and this overt rehearsal was tape-recorded. Figure 12 shows a typical protocol of a subject's rehearsal. Thus, for a given word, Rundus obtained (a) the number of times it was rehearsed, and (b) its probability of being recalled. It was therefore possible to plot the letter as a functio of the former; this function is

Insert Figure 17-12 about here

labeled "recall"

Figure 17-11. Results of Hebb's (1961) experiment. A series of nine-digit numbers was read to the subject whose task was to repeat them back. Every third number was the same, thus dividing the numbers into unchanging and changing numbers. Over trials, the amount of information recalled from the changing numbers remains constant, whereas the amount recalled from the unchanging digits increases.

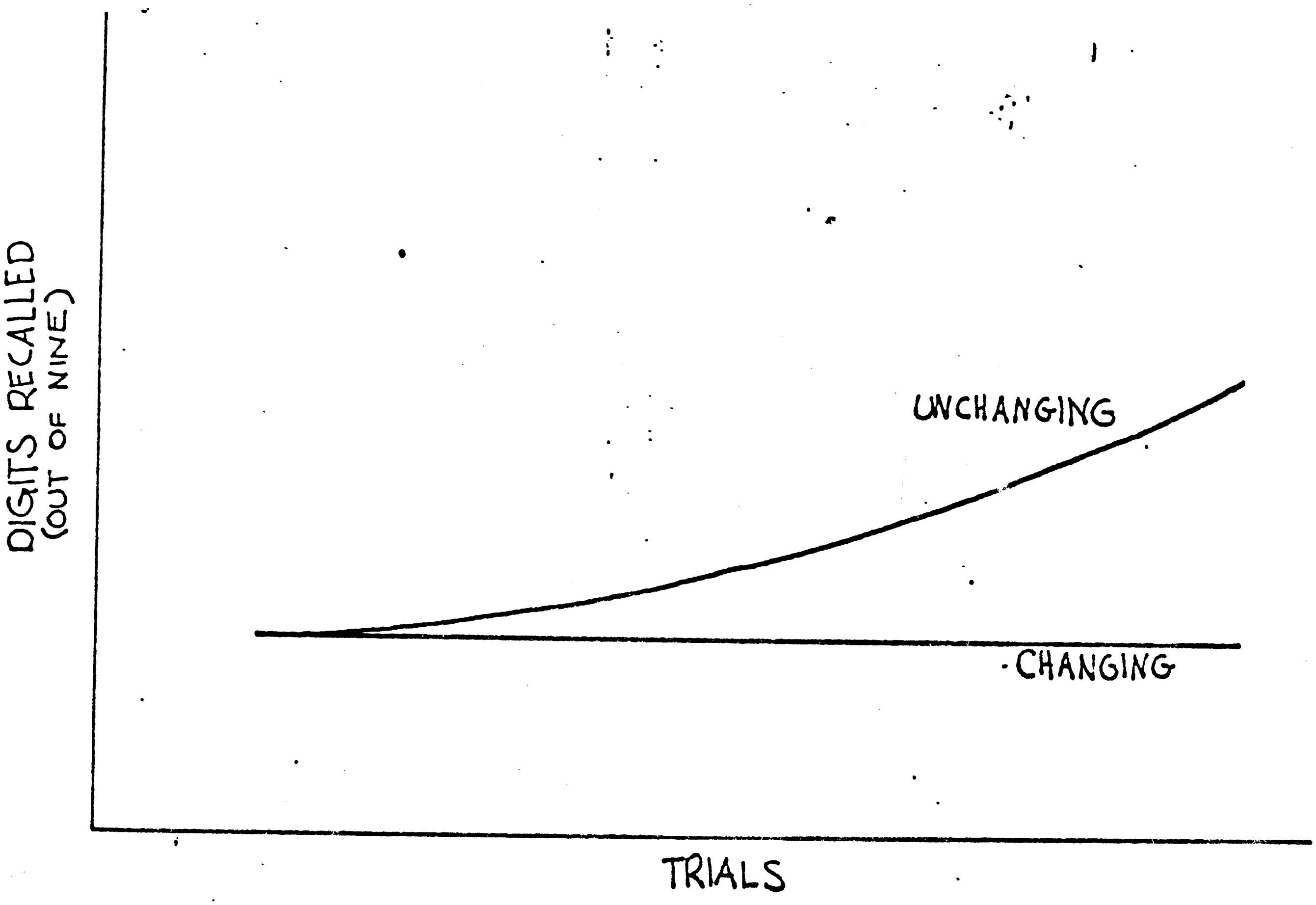

DIGITS RECALLED
(OUT OF NINE)
UNCHANGING
CHANGING
TRIALS

Figure 17-10: Retention as a function of forgetting interval for items rehearsed or 8 times.

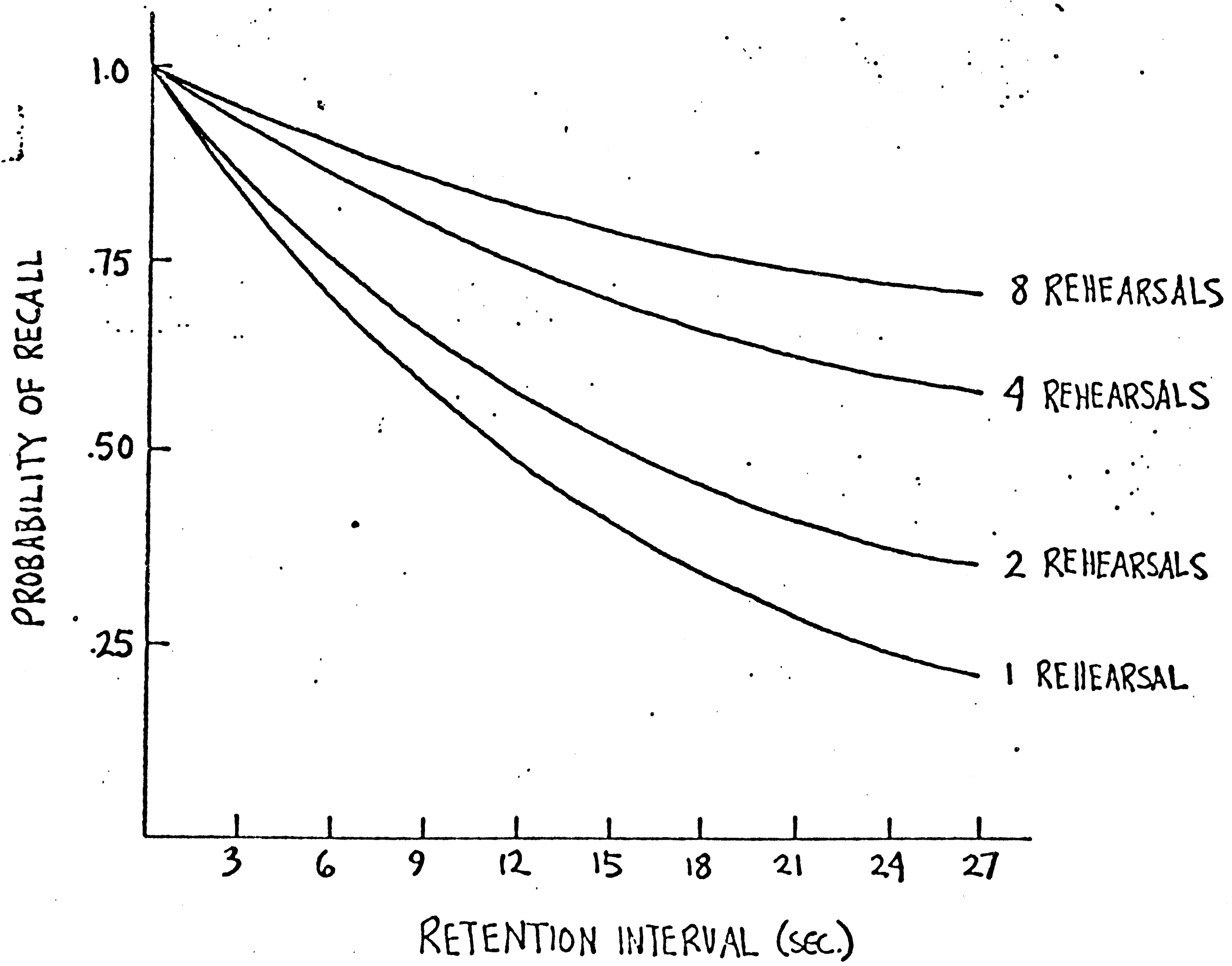

PROBABILITY OF RECALL
1.0
.75
.50
.25
3
6
9
12
15
18
21
24
27
RETENTION INTERVAL (sec.)
8 REHEARSALS
4 REHEARSALS
2 REHEARSALS
1 REHEARSAL

Figure 17-12: Words presented in a list from Rundus's (1971) experiement and a typical rehearsal protocol for the first seven words.

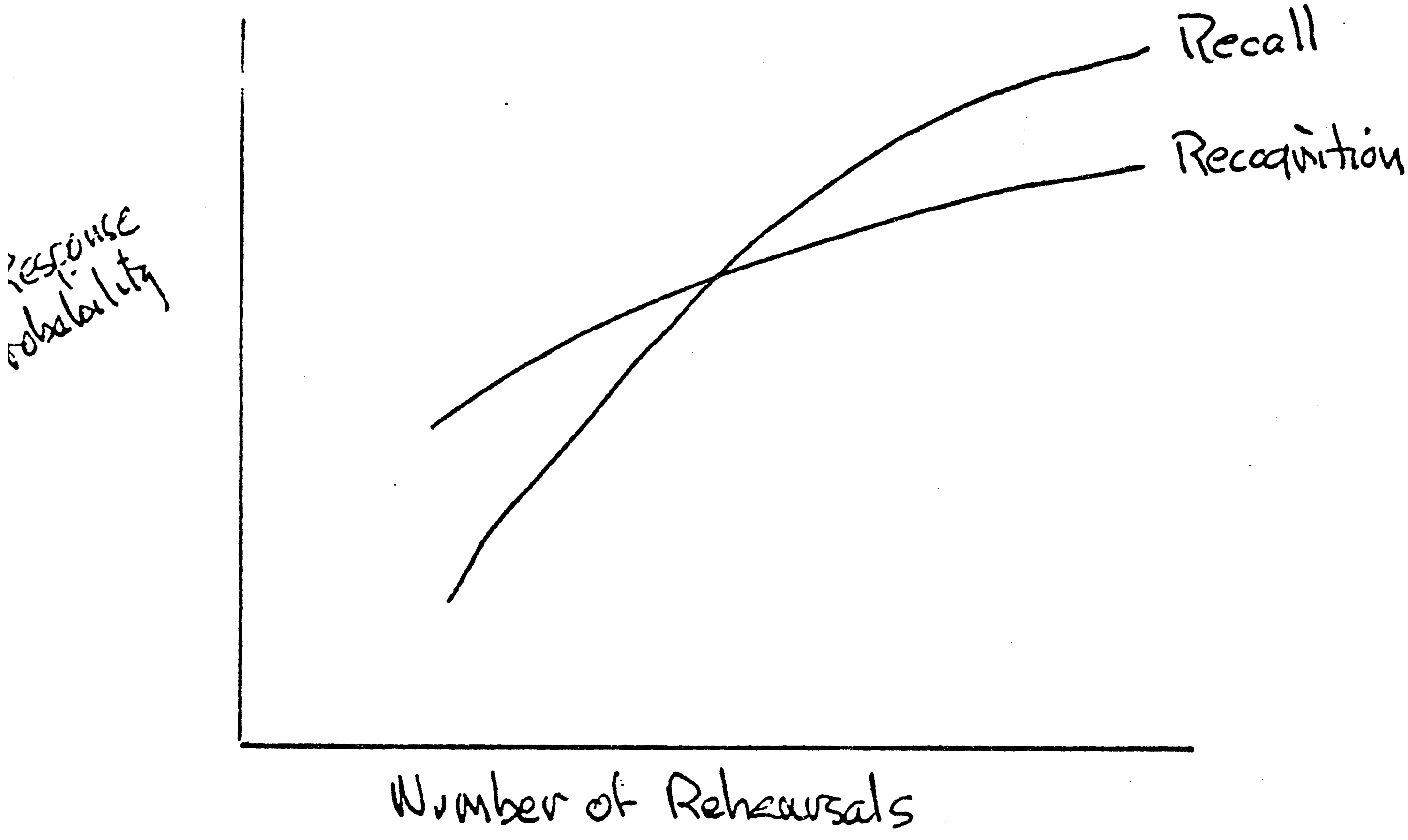

PSYCHOLOGY FUNNIES
OK Jon, I'm going to read you a list of words very slowly and I want you to remember them and say them back to me when I'm done.
Soap...
No sweat Doc, shoot!
Soap, soap, soap, soap, soap, soap...
...bird...
...bird, bird, soap, soap, bird, soap...
...chair...
...chair, soap, chair, bird, bird, chair, soap, soap, bird...
...king...
...king, king, soap, chair, king, bird...
TESTING IN PROGRESS
DO NOT DISTURB
SUBJECT POOL SIGN-UP
...book...
...book, book, king, chair, bird, book...
...shoe...
...shoe, king, book, shoe, chair, shoe...
...pen...
...pen, shoe, chair, pen, book, shoe...
Why did I ever enroll in Psychology?

in Figure 17-13. This result demonstrates that the number of rehearsals given a word is a very strong predictor of the long-term memory strength for the word, giving substantive support to the assumption of Atkinson & Shiffrin. In an extension of this result, subjects from Rundus's experiment were called back three weeks later and given a recognition test for all the words they had studied. The curve labeled "recognition" in Figure 17-13 shows that recognition probability for a word also is greater the more times a word iw rehearsed.

In addition to underlining the importance of rehearsal in transferring information from short- to long-term store, Rundus's work helped to clarify several hietherto murky phenomena. Consider, for example, the serial position curve shown in Figure 17-3. We have already seen that the recency (right-hand) portion of the curve is due to the fact that the most recent items have the highest probability of being in STM. However, what causes the primacy effect? Why are the first few items presented remembered so well? Atkinson & Shiffrin, (1968) in discussing the phenomena, suggest that it may be a rehearsal effect. Consider Figure 17-12, suppose that you can rehearse four or five words at once. When the first word of a list is presented to you, it is all you have to rehearse, and it will get your undivided attention--it will be rehearsed many times. However, when the <u>second</u> word of the list is presented, it will be rehearsed along with the first word. Thus, it will not get your undivided attention, but rather it will get your divided attention and it will not be rehearsed as much as the first word. Similarly, the third word in the list will have to compete with the first and second words for attention, and

Figure 13: Probability of recall and of delayed recognition as a function of the number of times a word was rehearsed. (From Rundus, Loftus & Atkinson, 1970)

Figure ~~12~~ 13. Probability of recall and of delayed recognition as a function of the number of times a word was rehearsed. (From Rundus, Loftus & Atkinson, 1970)

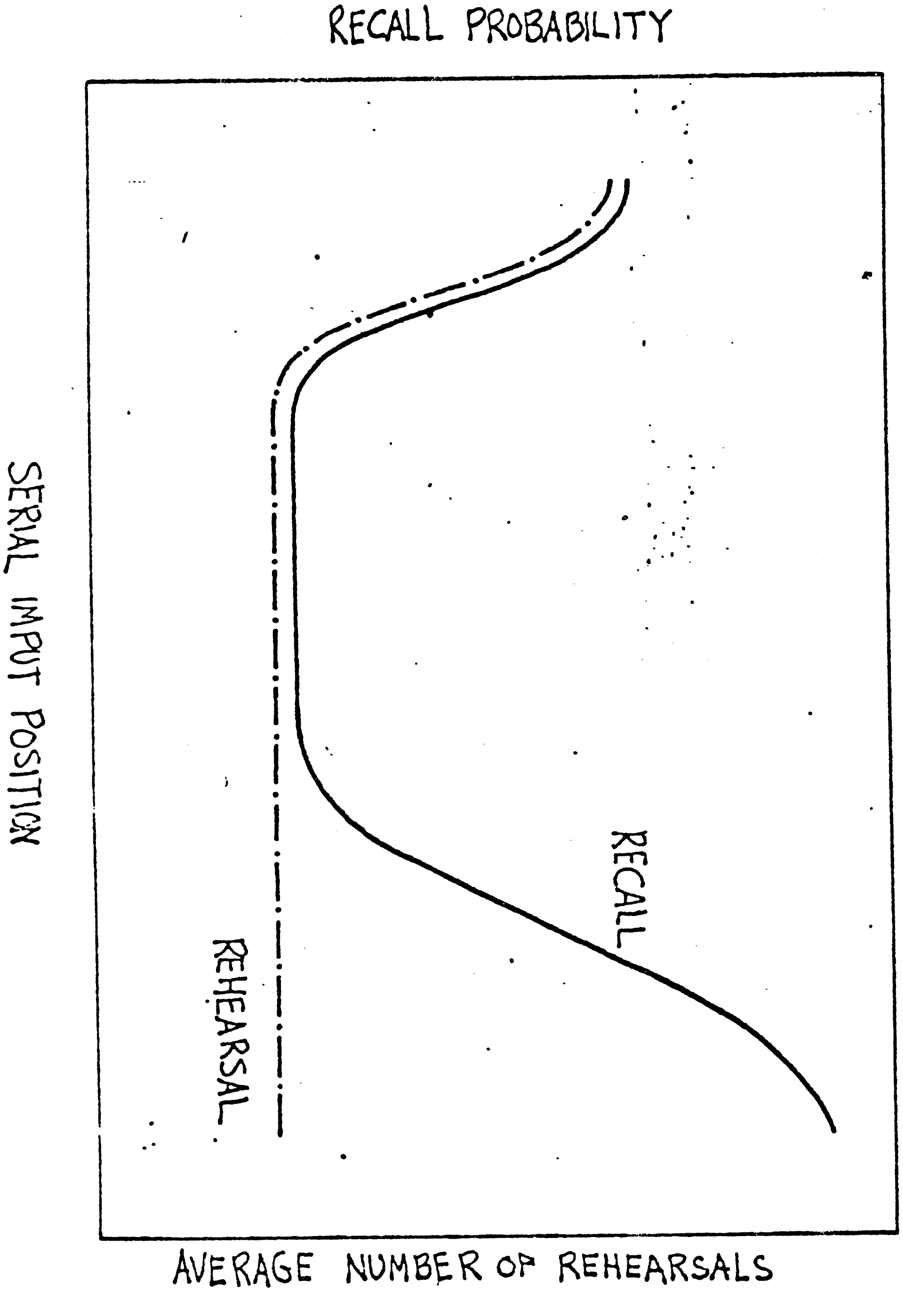


it will not be rehearsed as much as either of them. This process of diminishing attention (rehearsal) for each succeeding word will continue until the fourth or fifth word when the rehearsal buffer will be full, and some kind of a "steady state" will have been achieved. With Rundus's procedure, this hypothesis can be tested directly. Figure 17-14 shows recall probability, and along with it, number of rehearsals for a word as a function of the word's serial input position. As can be seen from Figures 11 and 13, it is indeed the case that the first few words receive more rehearsals, and the rehearsal curve nicely reflects the recall probability curve, up to the recency portion. Apparently then, the primacy effect can be at least partially explained by the fact that the first few words are simply rehearsed more. Can the primacy effect be <u>fully</u> explained by rehearsal? That is, is there anything <u>else</u> about the first few words that makes them easier to remember? It turns out that words from the middle of the list which happen to receive the same number of rehearsals as words at the beginning of the list were remembered just as well. The implication of this result is that the primacy effect is due <u>only</u> to the fact that the initial words are rehearsed more. Other free recall phenomena too complex to delve into here were similarly accounted for by rehearsal as found by Rundus (1971).

<u>Entry of information into STM - organization</u>. The notion which we have presented above - that transfer of information from short to long term memory depends only on the number of rehearsals - is, however, undoubtedly oversimplified, <u>vis a vis</u> everyday life. People do not in general go around rehearsing information out loud, and other factors which, on the surface, have little to do with rehearsal, can be shown to be very important in generating long-term information. For example several investigators have stressed <u>organization</u> of verbal material.

<u>Insert Figure 17-14 about here</u>

Figure 14: Recall probability (solid curve, left-hand ordinate) and average number of rehearsals (dashed curve, right-hand ordinate) as a function of serial input position. Up until the recency portion, the two curves are exceedingly similar, indicating that the primary effect is due, at least in part, to rehearsal factors.


RECALL PROBABILITY
SERIAL INPUT POSITION
RECALL
REHEARSAL
AVERAGE NUMBER OF REHEARSALS

Most of the work done on organization has been done in the context of free-recall experiments, and has stressed a chunking notion. Suppose that a list of words is presented several times in a study-free recall sequence. That is, a list is studied and then recalled, studied and recalled again, this procedure continuing for several trials. Over successive trials, more and more words are recalled, as might be expected. Work by Mandler (1967) and Tulving (1968) has investigated the question of _why_ there is this increase in recalled words. The gist of their findings is that subjects using some (unspecified) scheme, manage to set up _groupings_ of words in the list.

The organizational principle used to remember more words consists of adding more and more words to the basic groupings. Retrieval is accomplished by recalling what a particular grouping is, and then reporting the words in the grouping. Several lines of evidence are offered in support of this hypothesis. First, Tulving has examined the order in which subjects report words in a multitrial free recall experiment; such an examination reveals the presence of groupings which are either remembered _in toto_ or not at all. As the trials progress it turns out that, in fact, each grouping incorporates successively larger numbers of words.

To bring the phenomenon under greater experimental control, a free recall list may be so constructed as to consist of instances of taxonomic categories. Thus, for example, a thirty-word list might consist of six birds, six metals, six fruits, six cities and six sexes. Tulving & Pearlstone (1966) presented such lists to subjects, varying two things: the length of the list was either 12, 24, or 48 words, and subjects were either tested normally or in the presence of the category names as recall cues. Two major results were found: To quote Tulving (1968, p.28) "First, the presentation of the category names as retrieval cues affected only the number of accessible categories, but not the number of words within accessible categories. Thus, given that the subject could recall at least one word from a category, the number of additional words recalled for that category was the same regardless of whether he remembered the category on his own or was reminded of it. Second, the number of accessible categories was considerably greater for the 48-word list than for the 24-word lists, but again, the number of words within accessible categories was indentical for both list lengths." Thus the groupings of words seem to be the critical thing - grouping again were remembered in an all-or-none fashion.

Mandler (1966) has performed the following interesting experiment: First, a list of unrelated words was presented to a subject who simply had to sort the words into as many categories as he wished. After the first sorting, the words were reshuffled and another sort was required. This continued until the subject was sorting consistently - that is until the produced two identical sorts in a row. Then a free recall test was given on the words. Mandler's finding was that the number of words recalled was a linear function of the number of categories into which the words had been sorted. The more categories - the better the recall. Although these results are correlational, they provide support for the notion that organization of memory into categories is instrumental in recall.

Evidence for memory organization of a somewhat different sort has been adduced by Bower (1970). Bower presented 12-digit number strings for recall in a Hebb paradigm (see above). However, instead of simply reading the digits one-by-one, a string would be grouped; thus for example, the string 271828182890 might be read as "two seventy one, eithy-two, eight-thousand one hundred eithy two, eight, ninety). As would be expected from Hebb's results, when strings were repeated, their probability of being recalled increased over trials. However, for some of the strings, the grouping changed over trials. Thus, the above string might be read as "Two, seven hundred eithteen, twenty-eight, two eighty-one, eith-thousand two hundred eighty-nine, zero" the second time it was presented. For these "same-sequence-but-changed-organization" strings recall did not improve over trials.

Again it appears to be the case that organization is a primary determinant of memory performance.

Which theory is correct? We have discussed two ways of conceptualizing transfer of information from short-term to long-term memory; the first emphasizes *rehearsal*, while the second stresses *organization*. Is it possible to ascertain which, if either of these conceptualizations is the correct one? Or could it be both amount to the same thing? Although the concept of organization seems to explain a good deal of data, rehearsal, which explains much fo data in memory experiments is a considerably simpler concept. Since it is a goal of science to find simple mechanisms which underlie more complicated phenomena, we may hopefully sometime be able to explain organization in ~~principles in~~ terms of something like rehearsal - or better yet, to try to explain both in terms of something simpler still.

The information stored in long-term memory. Remember that the nature of information stored in *short*-term memory is thought to be mostly acoustic - when errors are made, the errors are likely to sound like the correct response - and that this information had been recoded from information in the iconic store which was basically *visual* in nature. What about long-term memory? When information is transferred from STM to LTM what kind of recoding does it go through? What kind of information eventually is stored in LTM? A popular veiw has been that

LTM information is basically semantic--that when a person is presented with words to remember, he takes the acoustic representation in STM, and extracts information having to do with the meaning of the word and this is what gets transferred to LTM. As we shall presently see, this view is a bit too simple to be tenable. First let us briefly examine the evidence for semantic encoding in LTM.

Like the studies showing acoustic encoding in STM, it is mostly an examination of the errors people make that has led to the postulation of semantic encoding in LTM. A series of studies has indicated that when errors are made in LTM, the error is more likely to be semantically related to the correct information than it is, for example, to be acoustically related. Suppose, for example, that you are searching for the word "car" which is the correct response to some query of an experimenter. If you make an error, you are more likely to say something like "auto" or "Porsche" or "vehicle" or even "motorcycle"--words which have some connection in meaning--than you are, for example, to say "bar" or "star" which sound like the correct answer but which are semantically unrelated to it. (Baddely, 1966; Anisfeld & Kanpp, 1968).

Lately, however, a variety of work has demonstrated that somewhat more than semantic information is stored in LTM. Posner (1971) has raised the logical question: "If, only semantic information were stored in LTM, how would we ever learn to recognize accents?" And, in the past year, an ingenious series of studies by Thomas Nelson, (Nelson, 1971; Nelson & Rothbart, 1972; Fehling & Nelson, 1972) has provided an apt demonstration that many kinds of information are stored in LTM. Nelson was basically concerned with forgetting from LTM, and he raised the question: Suppose some information is stored in LTM and then is forgotten. Is it the case that all the original information is lost, or does some remain? Nelson's results indicate that some information is not lost in an all-or-none fashion. In itself, this finding (known as the "savings effect") is not particularly earthshaking. The interesting parto of Nelson's work is that he then goes on to investigate what kind of information remians: he finds that some of it is semantic, but some of it is certainly acoustic. Note that if there is acoustic savings, then acoustic information must have been originally stored. We shall describe one of Nelson's experiments (Nelson & Rothbart, 1972) which illustrates the existence of acoustic savings.

Nelson's paradigm is the following: Information is originally learned in a paired-associate task. (In a paired-associate task, a list of stimuli is paired w with a list of responses. Nelson used two-digit numbers as stimuli and nouns as responses. So, for example, the stimulus"37" might be paired with the response "DOE", 19 might be paired with EGG, and so on. The subject first studies the pairing and then is required to produce the response when given the stimulus.) Nelson & Rothbart had their subjects learn the list until all the pairings were stored in LTM. The subjects then left, only to be asked to return four weeks later for the second half of the experiment. At this time, the subject first went through the list of stimuli once and tried to remember the responses. Some of the responses had, of course, been forgotten, and the concern was now with these forgotten items. The stimuli from these forgotten pairings were re-paired with new responses which bore one of two relationships with the old (forgotten) responses. In the first condition (experimental condition) the new responses were acoustically similar to the old responses (they were homophones.) So, if the forgotten pair had been 37-DOE, the new pairing might be 37-DOUGH. In the second (control) condition, the new responses had no acoustic similarity to the old response. Thus in this condition, 37 which had been paired with "DOE" might now be repaired with "KETCHUP." The subject was then required to relearn these new pairings. Nelson & Rothbart reasoned that if there is acoustic information about the old response remaining in LTM, then the acoustically similar re-pairings should be easier to relearn than the acoustically unrelated re-pairings. This turned out to be the case. When the new response was acoustically similar to the old, subjects were about 50% correct on the relearning trials, as opposed to about 10% correct for the acoustically unrelated re-pairings. This is impressive evidence for the notion that information stored in LTM is at least partially acoustic.

Using this paradigm, Nelson made a detailed investigation of the exact nature of semantic information stored in LTM. For example, it has been discovered (Fehling & Nelson, 1972) that superordinate information is stored about words--that is, if a subject has to remember a paired-associate such as "22-BUICK" part of the information that is stored is that a Buick is a car. Very active research is still being carried out to discover what other kinds of semantic information (e.g., synonymity, antynomy, etc.) is stored.

This type of finding, along with a variety of other studies, indicates that the information stored in LTM cannot be neatly classified into one type--rather, a large constellation of information, partly semantic, partly acoustic, and possibly partly visual is stored. Even when this information is "forgotten" a good deal of it apparently remains--and using techniques such as the Nelson paradigm, the nature of this information can be examined in some detail.

Retrieval from long-term memory. In the past few years, there has been a good deal of interest in how information--in particular, semantic information--is retrieved from LTM. This interest has been sparked in part by Sternberg's findings on retrieval from STM. Remember that Sternberg provided fairly conclusive evidence for a serial scanning process--the items in STM are examined one-by-one until the desired information is found. An obvious question is whether this is true for LTM--that is, if you are asked to retrieve information from your long-term memory, to what extent do you go through, consulting successive items in your quest for the desired information? It seems pretty safe to assume that people do not randomly search through all of their memory when retrieving information. If I ask you for your mother's name, you do not, in the process of finding it come up with the information that "Paris is the capital of France" (unless you have a very strange mother). It might, however, be plausible that you have stored in LTM, a category called "family names" and that you scan through this information to produce the desired fact.

There is a major problem in studying how semantic information is retrieved from LTM. The problem is that the experimentor has no control over how the information gets <u>into</u> LTM, and in order to postulate a retrieval scheme it is necessary to also postulate a structure for LTM. To see why this is so, imagine two situations: the first, trying to retrieve a subway token from a woman's pocketbook, and the second, trying to retrieve a book from a library. In the former case, the contents of the pocketbook are apt to be arranged somewhat randomly; a reasonable search strategy might thus be to go systematically from one side of the pocketbook to the other. In the latter case, however, books are rather well organized in a library; the general subject matter of the book you were looking for would tell you the approximate location in which it would be likely to be found. So you see that in whatever structure one postulates about anything--in this case, long-term memory--determines to a large extent what a reasonable retrieval scheme would be.

The research strategy followed by several investigators has thus been to postulate <u>both</u> an information retrieval scheme and a structure for LTM. Then to the extent that experimentation supports the theory, it supports both the structure and the retrieval scheme. However, if experiments do <u>not</u> support to theory, it is impossible to judge which component is at fault.

A popular conceptualization of long-term memory structure is that it is <u>hierarchical</u>. Thus, part of the structure might resemble what is depicted in Figure 17-15. Here, the information about "living thing" is subdivided into "plants" and "animals". Animals, in turn is subdivided into Birds and non-birds, and so on. A scheme such as this has been proposed by Collins & Quillian (1969). An important part of the memory structure posutlated by Cillins & Quillian is the assumption that information germain to a particular class of things is stored only at the level of the hierarchy corresponding to that class. Let us examine what is meant by this assertion. Consider a part of the hierarchy involving "animals", "birds" and "canaries". Coliins & Quillian assume that with "animals" is stored information common to all animals such as "drinks", "eats" and so on. Similarly, with "birds" is stored information relevent to all birds such as "has feathers" and "flys". Finally, at the level of "canary" there is information relevent to all canaries, such as "is yellow" and "sings". Suppose we perform an experiment in which a statement is made about canaries and the subject has to respond "yes" or "no" as quickly as possible corresponding to whether the statement is true or false. Now consider these three types of statemtns:

1) A canary is yellow
2) A canary has wings
3) A canary has skin

What kinds of information processing are involved in deciding that each of these statements is true? Here is where the retrieval scheme comes in. Collins & Quillian assume the following: When information is required about a noun (such as canary) the subject commences a search for memory. The seach consists of entering the place in the hierarchy where information about the noun is stored, and from there, whatever searching necessary to retrieve the desired information is carried out. Thus, to answer the first question, the subject would enter the place in memory corresponding to "canary" (see Figure 15) where he immediately would find the information that "canaries are yellow." He therefore should be able to produce the response "yes" rather quickly. Now consider question (2). In this case, the subject will go to the "canary" location, but will find there no information relevant to whether canaries have wings. Instead, he will have to move up one location and interrogate "birds". There, he will find that birds have wings and since a canary is a bird, he will then be able to produce the correct response. Similarly, for question (3) our subject will have to start at canary and then go up two levels in the hierarchy to determine that a canary, which is a bird, which is an animal has skin. To the extent that moving around in the hierarchy has to take place during retrieval, the process will take more time.

Insert Figure 13-15 about here

Insert Figure 13-16 about here

When this experiment was performed, the results shown in Figure 16 were obtained. As you can see, these results support Collins & Quillian's memory-structure retrieval scheme. Reaction time to respond to "a canary has skin" is about 90 msec longer than RT to respond to "a canary has wings" which in turn is about 90 msec longer than the response to "a canary is yellow". According to Collins & Quillian's model and data, then, it takes a little under a tenth of a second to move from one level of a semantic hierarchy to another.

Figure 15: Part of a hierarchically organized memory structure. Collins & Quillian assume that information about particular classes is stored <u>only</u> at the level of that class. (after Collins & Quillian, 1969)

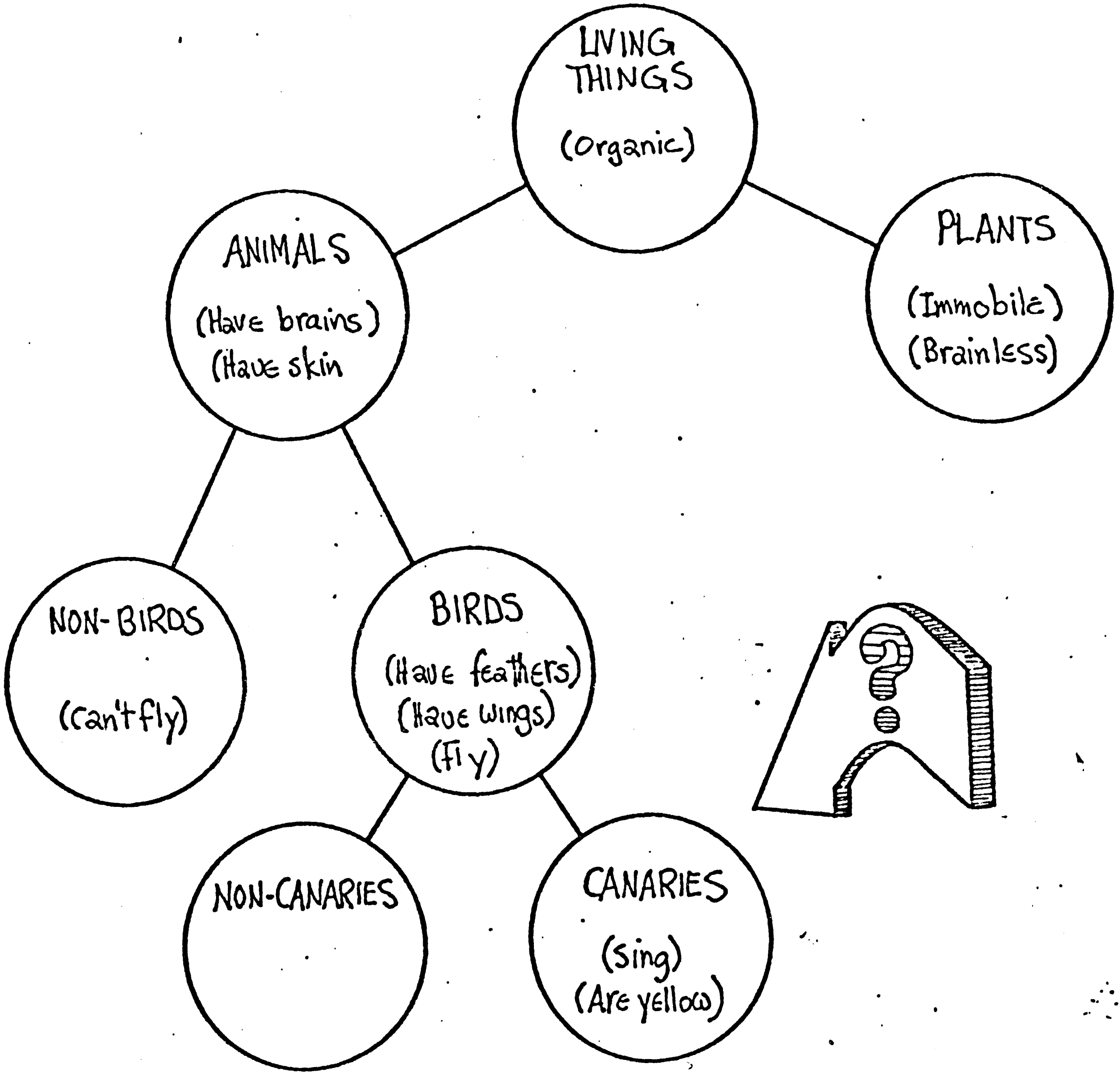
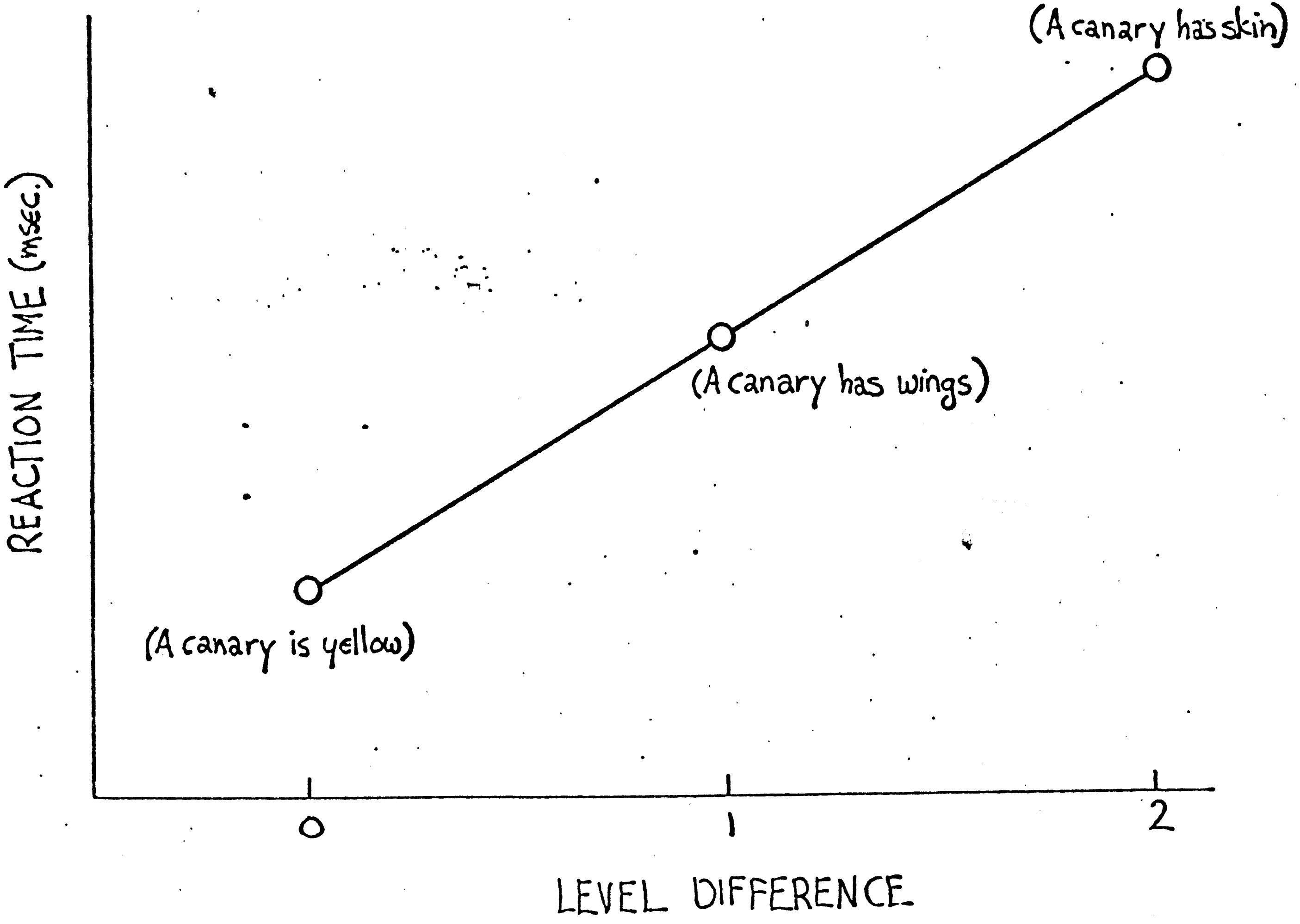

LIVING THINGS
(Organic)
PLANTS
(Immobile)
(Brainless)
ANIMALS
(Have brains)
(Have skin
NON-BIRDS
(can't fly)
BIRDS
(Have feathers)
(Have wings)
(Fly)
NON-CANARIES
CANARIES
(Sing)
(Are yellow)

Figure 16: Reaction time taken to answer various questions about nouns. The abscissa (level difference) indicstes how far from the noun (in hierarchical levels) the necessary information is stored.


(A canary has skin)
(A canary has wings)
(A canary is yellow)
REACTION TIME (msec.)
0
1
2
LEVEL DIFFERENCE

Long-term memory - a recapitulation. Let us briefly consider the things we now know about LTM.

1) LTM is a reasonable permanent repository of the information that we need to survive in our environment. Capacity of LTM is very large - no limit has yet been found on the amount of information that can be stored in it.

2) information is put into LTM from STM. Although there are many ways of describing STM - LTM information transfer - organization, imagery, etc., it appears that a simple theory using only the concept of *rehearsal* provides the best, most parsimonious account of such transfer.

3) Information in LTM is coded in many ways - semantically, acoustically and probably visually as well.

4) A good description of LTM structure is that it is hierarchical. Retrival of information consists of entering the hierarchy in some appropriate place and casting about fromk there for the desired information.

Suggested readings

Neisser, U. Cognitive Psychology Λ
Appleton - Century - Crofts, (1967) New York

Norman, D. Memory And Attention
Wiley, New York (1968)