Quarterly Journal of Experimental Psychology (1973) 25, 207–222

LEARNING 10,000 PICTURES

LIONEL STANDING†

Department of Psychology, Bishop's University, Lennoxville, P.Q., Canada

Four experiments are reported which examined memory capacity and retrieval speed for pictures and for words. Single-trial learning tasks were employed throughout, with memory performance assessed by forced-choice recognition, recall measures or choice reaction-time tasks. The main experimental findings were: (1) memory capacity, as a function of the amount of material presented, follows a general power law with a characteristic exponent for each task; (2) pictorial material obeys this power law and shows an overall superiority to verbal material. The capacity of recognition memory for pictures is almost limitless, when measured under appropriate conditions; (3) when the recognition task is made harder by using more alternatives, memory capacity stays constant and the superiority of pictures is maintained; (4) picture memory also exceeds verbal memory in terms of verbal recall; comparable recognition/recall ratios are obtained for pictures, words and nonsense syllables; (5) verbal memory shows a higher retrieval speed than picture memory, as inferred from reaction-time measures. Both types of material obey a power law, when reaction-time is measured for various sizes of learning set, and both show very rapid rates of memory search.

From a consideration of the experimental results and other data it is concluded that the superiority of the pictorial mode in recognition and free recall learning tasks is well established and cannot be attributed to methodological artifact.

Introduction

Human memory can store both abstract information (letters, words, numbers) and concrete stimuli (objects, scenes, sounds). Abstract memory can involve only that limited number of stimuli which through the subject's prior experience have already acquired a high degree of symbolic meaning. However, although the possible range of concrete stimuli appears to be much wider, nearly all studies of memory are traditionally confined to abstract material drawn from this restricted set.

The neglect of concrete learning tasks is unfortunate because experimental evidence suggests that picture memory, which represents one form of concrete learning, is a strikingly efficient process. Shepard (1967), extending a similar study by Nickerson (1965), has found that immediately following a single exposure of 612 picture stimuli, for about 6 s each, subjects could select the correct picture in two-alternative recognition tests with 98% success. (Similar tests using single words and short sentences as stimuli, produced 90% and 88% success, respectively.) Pictures also show excellent retention over time in memory, as Nickerson (1968) has demonstrated. Seeking the limits of picture memory,

† Now at Psychology Department, University of St Andrews, Fife.

Standing, Conezio and Haber (1970) gave subjects a single presentation of a sequence of 2560 photographs, for 5 or 10 s per picture. Their subjects then scored approximately 90% correct with pairs of photographs (one previously seen, one new), even when the mean retention interval was 1·5 days.

The first three experiments reported here examine systematically the storage capacity of picture memory, while the fourth examines the rate at which material may be retreived from it. A major consideration in each case is to evaluate picture memory against the standard of verbal memory.

In each case the subjects were experimentally naive students (non-psychologists), aged 18–25. The subjects were paid on an hourly basis.

Experiment I

It is desirable first to examine picture memory capacity in more detail than has been achieved by Nickerson, Shepard or Standing *et al.* in order to find the general relationship between the number of stimuli presented and the number retained in memory, which cannot be inferred from these studies. This experiment therefore follows the basic pattern of the above studies but systematically varies the number of stimuli shown to the subject over a wide range, the dependent variable in each case being the number of items that are retained in his memory. The experiment compares recognition memory for pictures (both vivid and normal) with that for words.

Method

Stimuli

A population of 11,000 photographic slides was first assembled. These were 35 mm transparencies; 92% were coloured. The method of assembling this population from various sources is outlined by Standing (1971*a*). No simple metric exists for specifying picture stimuli; however, these photographs may be characterized as resembling a highly variegated collection of competent snapshots, and will be referred to subsequently as Normal pictures. (The term is used simply as a convenient label.)

Another population, of 1200 striking pictures, was selected from an original pool of about fifteen thousand photographs by three judges. This process of selection is described by Standing (1971*b*). In brief, slides with definitely interesting subject matter (with or without technical excellence) were placed in this population, which will be referred to as Vivid pictures. Ordinary photographs of dogs appeared in the Normal population, but a picture of a dog holding a pipe in its mouth was assigned to the Vivid category; an aeroplane was generally Normal, but a crashed place Vivid.

A population of Word stimuli was produced by selecting English words randomly from a Merriam-Webster dictionary and printing them on 35 mm slides. This dictionary gives the 25,000 most common English words (e.g., salad, ton, station, landholder, cotton, zoology, camouflage, reduce, well-worn, somehow).

Procedure

The procedure involved a single-trial learning task, followed by a delayed recognition test. A different group of 5 subjects was used for each learning task, except when Vivid photos comprised the stimuli; the group size was always 10 in the latter case. All testing was performed under group conditions.

Before the learning task, the subjects were instructed to attend closely to the stimuli, and to try to learn them in preparation for a memory test. The importance of maintaining strict concentration even during long sequences of stimuli was strongly emphasised by the experimenter.

The subjects were then shown a set of stimuli randomly selected from one of the three populations (Vivid pictures, Normal pictures, or Words). This learning set was presented once only, at 5 s per item and with an interstimulus interval of 600 ms. A commercial Sawyer projector controlled by Hunter timers was used to present the slides under dark-room conditions; they subtended a maximum visual angle of approximately 14°.

Two days later the subjects performed a recognition test. On each trial two stimuli were presented side by side; one had been randomly selected from the learning set and then randomly allocated to the left-hand or right-hand position, while the other was new to the subject and had been selected randomly from the appropriate one of the three stimulus populations. Each subject wrote down an "L" or an "R" on each trial to indicate whether the left- or the right-hand stimulus looked most familiar to him. This task was forced-choice; unlimited time was allowed, but each trial in practice usually required only a second or two.

The size of the learning set for each type of material was 20, 40, 100, 200, 400, or 1000 stimuli; however when Normal pictures were used, additional groups were tested with 4000 and 10,000 items. The recognition test consisted of 80 trials, except when 4000 or 10,000 items were learned (160 trials), or with learning sets of twenty or forty (when 20 or 40 recognition trials were given). Rest pauses of 4 min were given during the learning task after every 200 items, and a 1 h break after 1000 items, where applicable.

While the recognition test was generally given exactly two days after the learning task, when 4000 or 10,000 stimuli were used only the *average* retention interval could be set at 2 days, since only 2000 slides per day were shown in these cases. This was achieved in the former case by testing recognition 1·5 days after the second learning session, and in the latter by giving the recognition test immediately after the fifth daily learning session.

Results and discussion

The mean number of errors occurring in the recognition task under each condition is given in Table I. Also given is an estimate of the number of items (M) that have been retained in memory in each case; making the usual guessing correction, this is calculated as $S(T\text{-}2E)/T$ where S is the size of learning set, E is the mean number of recognition errors, and T is the number of recognition test trials.

TABLE I

Mean errors in recognition test, Experiment I (standard deviations in parentheses).

		Material								
S	T	Vivid pictures			Normal pictures			Words		
20	20	0	(0)	20	0·2	(0·45)	19·6	2·0	(1·2)	16
40	40	0	(0)	40	1·8	(1·1)	36·4	5·6	(1·3)	28·8
100	80	1·6	(0·9)	96	4·0	(2·2)	90	12	(1·9)	70
200	80	2·0	(2·0)	190	6·8	(1·3)	166	16·8	(4·5)	116
400	80	1·9	(3·0)	381	11·4	(5·8)	286	16·6	(7·1)	234
1000	80	4·8	(3·3)	880	9·2	(3·0)	770	15·4	(5·5)	615
4000	160		—		30·2	(16·4)	2490		—	
10000	160		—		27·2	(6·1)	6600		—	

S is the number of stimuli presented in the learning set; T is the number of recognition test trials. The third value within each cell is the estimated number of items retained in memory (M). Each cell is based on 5 subjects (10 for Vivid pictures)

The mean number of items in memory (M) was then plotted against the number of stimuli presented (S) for each type of material, using log-log co-ordinates as shown in Figure 1.

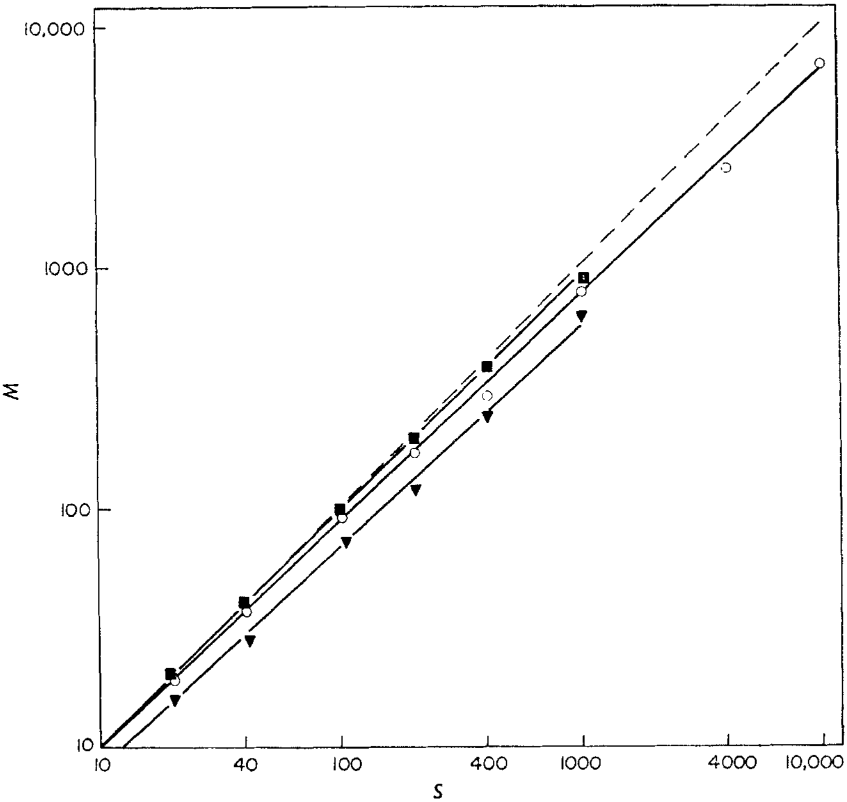

FIGURE 1. Number of items retained in memory (M), as a function of number presented for learning (S), Experiment I; the coordinates are log-log. The diagonal broken line represents perfect memory. ■ Vivid pictures (0·97 log S + 0·04); ○ Normal pictures (0·93 log S + 0·08) ▼ Words (0·92 log S − 0·01).

The data of Table I show that picture memory is consistently superior to verbal memory, particularly with the larger learning sets, and that Vivid pictures are better retained than Normal items. For all three types of material, Figure 1 shows that there is no upper bound to memory capacity; per cent retention gradually declines, but the absolute number of items retained always increases as the learning set is made progressively larger.

In practical terms, this satiation is appreciable when verbal items are used by the time the learning set reaches 1000, since performance has then fallen to 62%. But for Vivid pictures the gradient is 0·97 and memory capacity is almost limitless: extrapolation of the Vivid graph indicates that if one million items could be presented under these conditions then 731,400 would be retained. Naturally, better retention is obtained when an immediate rather than a delayed-recognition test is given: six additional subjects who were given 100 recognition trials each, immediately after viewing 1000 Vivid pictures under the same conditions as Experiment I, scored a mean of 99·6 correct (median 100), corresponding to retention of 992 items. Hypothetical extrapolation of this figure suggests that 986,300 items would be retained from a million stimuli.

Considering now the general question of a law for memory capacity, we note that all three materials yield straight-line functions when log M is plotted against log S, the correlation between these two variables in each case being above 0·99. This means that the data follow the power-law principle, which previously has often been noted in the study of sensory magnitudes (Stevens, 1961). Expressed simply, each time the size of the learning set is increased X times, then the number of items stored in memory increases X^n times, where n is the gradient of the appropriate line on a plot like that of Figure 1. The formal similarity to Stevens' law is close, despite the different tasks involved.

The generality of this law for memory capacity has been quantitatively established by the author by applying it to various data in the literature, involving diverse tasks and stimulus materials (e.g., Deese, 1960; Seibert, 1932; Strong, 1912; Woodworth, 1915). Two typical illustrations must suffice: Figure 2 shows the (unpublished) detail data for Experiment III of Murdock (1960) and observations of Binet and Henri (1894), both based on a verbal free recall task. Again, the correlation between log S and log M is above 0·99 in each case.

Returning to the present data, two aspects of the methodology were examined. First, the observations of Nickerson (1965), Shepard (1967) and Standing *et al.*

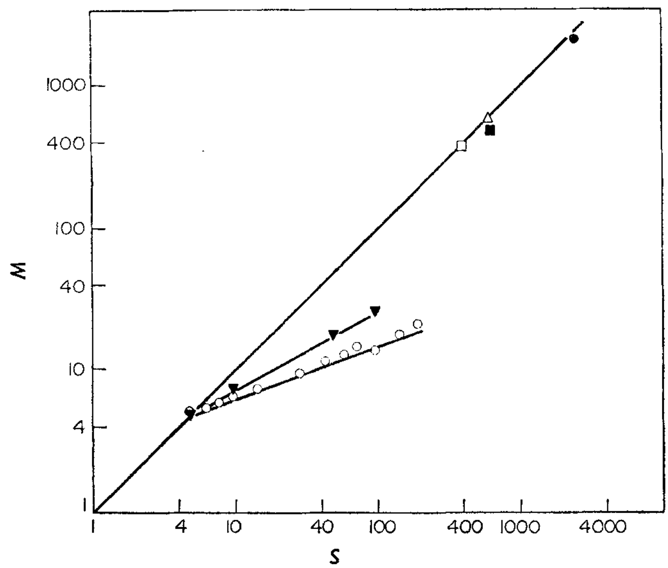

FIGURE 2. Items retained in memory (M) as a function of the number presented (S), in verbal free recall (Binet and Henri, 1894; Murdock, 1960). Also shown are estimated numbers of pictures retained in the recognition studies of Nickerson (1965), Standing *et al.* (1970), and Shepard (1967), and Shepard's value for words. The co-ordinates are log-log. □ Nickerson (pictures); △ Shepard (pictures); ■ Shepard (words); ● Standing *et al.* (pictures); ▼ Binet and Henri; ○ Murdock.

(1970) were plotted (in Fig. 2) to check that their findings resemble the present results. Comparison of Figures 1 and 2 supports this postulate. Second, the problem of retention intervals was considered by examining performance under the 10,000 Normal picture condition, for the items which had been held in memory for 0, 1, 2, 3, or 4 days at the time of the recognition test. Performance was found to be constant between these retention times, so that the decreased storage time in memory for the later items is offset by some other factor (e.g., cumulative fatigue). This conclusion is in accord with a similar observation of Standing *et al.* (1970).

Independently of the main study, some individual data were collected from a single clearly atypical subject (R). This subject was tested with both Words and Normal pictures, and with various sizes of learning set, but otherwise under the same conditions as the experimental subjects. Some characteristic data are shown in Table II; following the previous procedure, values for M were calculated from performance in the recognition task and are also given in Table II. This subject was quite different from all the others tested in that for learning sets over about 100 items, Words and Normal pictures consistently showed approximately equal (and low) capacities.

TABLE II

Errors made by subject R on recognition test; conditions as for Experiment I

		Material			
S	T	Normal pictures		Words	
10	20	2	8	8	2
20	40	8	12	15	5
40	80	17	23	23	17
100	80	13	67·5	33	17·5
200	80	27	65	27	65
400	80	27	130	28	120
1000	80	30	250	25	375

The second value in each cell is the estimated number of items retained in memory (M). Two separate learning sessions were employed for sets (S) of 10, 20 and 40 items to yield the ndicated total of test trials (T).

Experiment II

The purpose of this study was to test whether the apparent superiority of picture memory is maintained despite methodological changes in the experimental task. A possible objection to most previous picture memory studies is that they generally employ an atypical task to measure learning, i.e., a 2-alternative recognition task. The high performance observed with picture stimuli could simply reflect an advantage for a particular type of memory test with a specific type of material (an objection often raised in discussions).

A task was devised where following a single-trial learning session, given as before with Normal pictures or with Words, the subject performed a recognition test to pick out the previously-seen item from among 2, 4, 8, 16 or 32 alternatives. In two respects this represents a change from Experiment I: the alternatives were shown sequentially rather than simultaneously, and the larger number of alternatives made the task more difficult and proportionately closer to the recall type of task (following the reasoning of Davis, Sutherland and Judd, 1961). A third consideration is that repeated tests were performed on the same group of subjects, thus providing some index of the stability of the effect over time.

Method

Stimuli

Stimulus items for each learning session were selected randomly from the Normal picture or Word populations used before.

Procedure

Subjects were given a sequence of 16 learning sessions, at intervals of 2 days; each learning session was preceded by a test session for the items shown in the previous learning session. The number of items presented in each learning session was 100, at 2 s/item. Each test session consisted of two sequences each of 2, 4, 8, 16, and 32 items, given in random order. Within each sequence a single item had been seen by the subject in the previous learning session, the remainder being new. The subjects were shown these sequences at 5 s/item, and wrote down a "Y" or an "N" as each slide was shown to indicate whether they thought they had seen it before. If at the end of the sequence a subject had not made a "Y" response, or had made more than one, the sequence was repeated so that he could choose the most likely item. (This repetition was required on about 10% of the sequences.)

In this manner two subjects were given eight learning sessions for Words, followed by the same number for Normal pictures; the other two were tested in the reverse sequence.

Results and discussion

The mean number of error responses for Words and for Normal pictures was first determined per subject at each length of test sequence. A guessing correction was then applied: the number of errors occurring with a given number of alternatives in the recognition test sequence (*Na*) was multiplied by *Na*/ (*Na*-1) to yield the estimated number of guessed responses (*Ec*). The mean number of items retained in memory (*M*) was then estimated, as $S(T\text{-}Ec)/T$, for each value of *Na*. These values of *M* are given in Table III, with error scores.

TABLE III

Mean errors (E) in recognition test trials, Experiment II, when performed with various numbers of alternatives (Na).

Na	Normal pictures E	Normal pictures M	Normal pictures d′	Words E	Words M	Words d′
2	0·25	96·88	3·05	2·5	68·75	1·43
4	1·5	87·5	2·41	3·0	75·0	1·89
8	1·75	87·5	2·79	6·0	57·14	1·72
16	2·0	86·67	3·04	4·0	73·33	2·5
32	1·25	91·94	3·68	6·5	58·06	2·36
	Grand mean	90·1			66·5	
	S.D.	4·3			8·4	

Also shown are the estimated number of items retained in memory (*M*) from the learning set of 100 stimuli, and the detectability index (*d*′). The data are based on 4 subjects.

An alternative index of memory strength was obtained by converting the raw probability of correct response to the detectability index, d', for each value of *Na* (Swets, 1964); these values are also given in Table III.

The obtained values of *M* in Table III clearly show that the advantage of pictorial over verbal material is maintained when the memory test presents the subject with more alternatives from which to choose the target item. The largest difference between the two types of stimulus, in terms of *M*, is in fact noted with 32 alternatives. A similar conclusion follows from examination of the d' values which are higher for the pictorial stimuli for all values of *Na*. The stability of *M* with changes in *Na* for both types of stimulus is in accordance with the results of Davis *et al.* (1961).

Examination of the data showed that errors were constant over sessions; the schedule of repeated learning and test sessions produced no evident practice effects or interference phenomena. Subjects often described the recognition task as harder than a simultaneous comparison between paired items, but mean performance was almost identical to that obtained with 100 stimuli in Experiment I, where *M* for words and pictures was 70 and 90 items. The present values of 66·5 and 90·1 items were obtained despite the use of successive test stimuli, and despite reduction of the learning time to 2 s from 5 s per item.

Experiment III

A further study seemed desirable to establish the robustness of picture memory superiority under methodological changes and, employing a memory test still closer to more conventional learning paradigms, to test both recall and recognition performance for various forms of material. It has often been found that recognition performance exceeds recall, for a particular type of material and a given number of alternatives in the recognition test, by some characteristic factor (Luh, 1922). By measuring recall performance, as well as recognition, in a given task it is possible to test the hypothesis that pictorial material appears to be memorized easily, merely because it happens to be unusually well suited to recognition tests.

This study examined recall and recognition memory for Normal pictures, Words (both visual and auditory presentation) and Nonsense syllables (visual presentation). Some additional recognition data were also collected for Music segments. In each case the experimental paradigm involved a single-trial learning session, followed by a recall interval, followed by recognition tests.

Method

Stimuli

The stimuli consisted of 200 Normal pictures, Words (visual), Words (auditory), Nonsense syllables (visual) or Music segments. The visual stimuli were projected at 5 s intervals as before; the auditory Words were heard at 5 s intervals from a tape-recording. The Normal pictures and Word stimuli were drawn from the populations used for Experiment I; the Nonsense syllables were randomly sampled from Table VIII of Hilgard (1951), with a mean association value of 50%. The Music segments were tape-recorded but in this case the 5 s/item rate could not be used: due to the temporally-extended nature of music a 12-s duration per item was found necessary for subjects always to report a meaningful perception. It was also necessary to include a 3-s blank interval between items to prevent reports of discomfort. The Music segments were randomly sampled from classical, jazz and popular works, and always comprised the first 12 s of any piece.

Procedure

The subjects were tested in groups of 10, each used for one type of material. (No musically-trained subjects were included in the Music group.) Subjects were instructed to attend closely to the stimuli in preparation for an unspecified memory test, and were then presented with all 200 stimuli at 5 s/item (12 s in the case of Music). A 40-min recall period followed, during which the subjects wrote down as many items as they could remember; in the case of pictures, subjects wrote down a description of each item recalled. (No recall test was given in the case of the Music group.) Subjects were then given 40 forced choice recognition test trials, with two alternatives (one old, one new), presented successively in the modality used for the learning task.

No difficulties were encountered with the use of written recall responses for picture stimuli; the instructions for this group simply requested them to make each description sufficiently detailed to enable the picture described to be clearly identified among the total set of stimuli.

Results and discussion

The recognition data were scored as before, doubling the observed mean number of errors and using the resultant percentage of correct trials to estimate M, the number of items retained in memory.

The recall data were scored as the number of items recalled, minus errors. It is thus assumed that the probability of guessing a word or nonsense syllable correctly, is negligible. The mean number of erroneous recalls for visual words, auditory words and nonsense syllables respectively was 7·40, 4·88 and 11·45 items per subject.

To estimate the errors in the picture recall data, a fresh group of 5 subjects rated the written responses as being good descriptions, mediocre descriptions or erroneous descriptions of an item in the 200 learning stimuli. These subjects categorized an average of 4·5 and 0·6 descriptions (per original subject) as mediocre or erroneous. These were then taken to be guessed responses; however, allowance must also be made for guessed responses which happen to be correct by chance. Control tests using the experimental descriptions and a completely fresh set of 200 Normal picture stimuli showed that there was a mean probability of 0·13 that a description would, by chance, fit an item in this set well enough for a subject to rate it as a good description. Using this correction factor, the number of items guessed per experimental subject appears to be 5·9 rather than 5·1; subtracting this value from the total number of picture recalls per subject leads to the value given in Table IV.

The number of items stored in memory (M) in each task, and the recognition/recall ratio, is given in Table IV. Clearly, pictures are not only recognized better but are also recalled better than words, whether presented in the visual or the auditory modality, which in turn are recalled and recognized better than nonsense syllables. It is striking that although the recognition/recall ratio is roughly similar in all four cases (and resembles characteristic values in the literature, such as those of Luh), the lowest value was obtained with pictures; this

TABLE IV

Mean errors (E) in recognition task, Experiment III.

Material	Recognition			Recall			Recognition/Recall
	E		*M*	*R*		*FA*	
Normal pictures	2·5	(2·4)	175	51·6	(3·6)	5·9	3·4
Words (visual)	6·55	(2·7)	134·5	24·5	(10·0)	7·4	5·5
Words (auditory)	5·75	(3·4)	142·5	37·5	(17·7)	4·88	3·8
Nonsense syllables (visual)	15·15	(4·6)	48·5	11·25	(5·7)	11·45	4·3
Music segments	7·0	(4·4)	130	—			—

Standard deviations in parentheses, and estimated number of items in memory (*M*). Also given are *R*, the number of items correctly recalled (standard deviations in parentheses), and *FA*, the number of erroneous recalls made. The data are based on 10 subjects per cell, each learning 200 stimuli.

result clearly contradicts the hypothesis that with pictorial stimuli, recognition somehow becomes especially important.

Recognition scores with the other type of concrete stimuli employed (Music) were similar to those found with words. This is not due, apparently, to the longer time-per-item with Music than other stimuli since control tests showed that when nonsense syllables were presented on the schedule used for music, performance actually declined relative to that found with the 5-s rate; the advantage of more time per item was evidently outweighed by the increased retention time before the recall and recognition tests. Nor does this good performance with Music result from any known superiority of the auditory modality, as is shown by the present data for words and the results of previous investigators (Henmon, 1912). However, the comparison between Music and the other materials remains essentially qualitative rather than quantitative.

The subjects of this study were able, with surprising ease and consistency, to use stored information from pictures to provide short but generally accurate verbal descriptions of these stimuli. Some randomly selected examples are: "Two boys and man in field with oxen"; "Side-back view of an ox"; "Cliff with cliff dwellings"; "A mosque (white and dull red) with sky background"; "Old man with white beard, wearing advertising board"; "Train crossing a trestle". Although subjects were given no requirements concerning the descriptions' length, with some regularity they gave descriptions approximating six words in length. The grand mean and medium were 6·40 and 6 words respectively, while the within-subjects standard deviation was 2·37. The interquartile range of individual subject medians was 5–7 words.

Experiment IV

This study examines the rate at which picture information is retrieved from memory, compared with the rate for verbal information. To accomplish this, a learning session was followed by a reaction-time task in which the subject pressed one of two keys to indicate whether the stimulus was familiar (one of the learning stimuli) or novel. This task was performed with picture and word stimuli, for various sizes of stimulus set in the learning task.

Method

Apparatus and stimuli

Normal pictures and Words were used as stimuli. During the learning task they were shown as before, but projected to give a maximum visual angle of 8° for pictures and 3° for words. During the recognition test the stimuli were presented (with an 8° or 3° visual angle respectively) via a Marietta slide viewer/reaction time system (Model 14–1B). The subject's reaction time (to press one of two keys) was recorded on a centisec electric clock.

Procedure

Four subjects were tested individually in repeated sessions, In each learning session the subject was first shown a series of stimuli at 20 s/item; this series comprised 5, 10, 20, 40, 80, or 160 items. The type of material (words or pictures) and the sequence length was randomly varied between sessions. The recognition test was given 10 min after this learning task: test stimuli were items randomly selected from the learning stimuli and an equal number of previously unseen items, interspersed in random order. The subject was instructed that immediately upon seeing the test stimulus for each trial he was to press the left key, as quickly as possible, if the stimulus had been shown in the learning series or the right key if it had not. (These directions were reversed for two subjects.) Forty trials were then given of this reaction-time task. When the learning sequence was only 5 items, the task comprised 10 trials; however three additional sessions were given (with new stimuli) to yield 40 observations. Similarly, double sessions were given when the learning set was 10 items. Error trials were discounted and additional trials given later in the sequence.

Results and Discussion

The mean reaction time and standard deviation are shown in Table V for each type of stimulus and for each size of learning set, together with error rates.

TABLE V

Mean reaction times (ms) obtained in Experiment IV with various sizes of learning set (S).

S	Normal pictures		E	Words		E
5	710	(82)	3·8	567	(77)	3·8
10	845	(139)	6·9	675	(76)	13·1
20	830	(129)	9·4	738	(76)	16·3
40	901	(74)	13·1	758	(131)	15·0
80	950	(118)	10·0	780	(150)	18·8
160	988	(119)	13·1	825	(113)	18·1
			Grand means			
Positive trials	876	(140)		715	(133)	
Negative trials	864	(149)		738	(141)	

Standard deviations are given in parentheses. The mean percentage of errors (*E*) under each condition is also shown. The data are based on 160 trials per cell.

The mean reaction time for each condition is plotted against the learning set size in Figure 3, on log-log co-ordinates. From this figure it appears that the power-law principle again operates here; the correlation between these two variables is 0·95 for pictures and 0·93 for words.

Comparing the relative shapes of the two functions in Figure 3, we see that both follow the power law with similar exponents (0·086 and 0·096 respectively for pictures and words). The time to scan memory for an item in a given set of pictures is however greater than the time to scan for words (because of the higher overall level of the picture graph). In both cases, the time-per-item decreases drastically as the size of learning set is increased. Thus an extra picture added to a set initially consisting of a single picture would require an extra 39·9 ms of memory search time (according to the regression equation); added to a set of five items, it requires an extra 11·8 ms; and with a set of 160 items it involves only 0·54 ms of extra search time. The corresponding times for words are respectively 36·0, 10·8, and 0·51 ms.

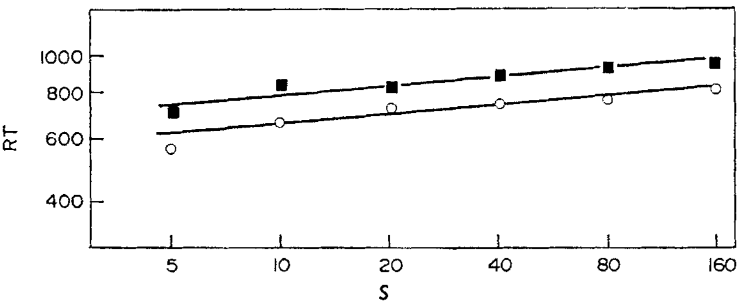

FIGURE 3. Log-log plot of mean reaction-time (ms) to categorize a picture or a word as previously seen/not previously seen, as a function of the number presented (S), Experiment IV. ■ Pictures; ○ Words.

The higher error rates found with words indicates that the faster verbal reaction times do not result from better learning. (These error responses were found to have approximately the same latency as correct responses.) However, the higher reaction times for pictures (Fig. 3) must be interpreted with some caution. Before a memory search can proceed, a percept of the test stimulus must presumably be established, and if perception for pictures is slower or more difficult than for words this would increase pictorial reaction times. Sternberg (1967) has shown this type of increase, by means of perceptually degraded stimuli, and the present data (like Sternberg's) do show an increase in reaction time for pictures over words which is approximately constant across different sizes of stimulus set (Table V).

We conclude that the overall difference in the memory task (Fig. 3) may involve a difference in perception time between pictures and words. This question requires careful experimental study, since picture perception obviously is not an all-or-none process (as with words) but may become progressively elaborated over time, as Gombrich (1969) has observed. Possibly, a rather more detailed perceptual process is needed to enable memory scanning for a picture than for a word.

Conclusions

One conclusion which has emerged during the present studies is that memory capacity in general follows a power law, which applies both to the number of items retained in memory and to the speed with which they may be retrieved (the independent variable being the number of stimulus items presented during learning). This law may be regarded as a reliable empirical generalization, at least for set sizes roughly within the range 10–10,000. The law often cannot be applied to very small sets because of a ceiling affect: the subject reaches perfect performance. This is well illustrated by Murdock's data in Figure 2.

The experimental data show that pictorial and verbal recognition memory possess many qualitative similarities; both follow a power law for capacity, both decline in terms of items correct (but not in terms of the detectability index) when the number of alternatives in the recognition test is increased, both show a comparable decline in performance when a recall task is substituted, and both follow a power law for retrieval time. However, with the exception of retrieval time, pictorial memory is quantitatively superior to verbal memory.

Naturally, this conclusion stands open to the objection that verbal memory may subsequently be found to show advantages over picture memory in other aspects than simple capacity. It is also possible that the superiority reflects in part a higher motivational level when pictorial material is used. However, the difference cannot be attributed to the use of a particularly "learnable" set of pictures, since the general finding of this and other studies is that essentially *any* set of pictures, arbitrarily chosen by the experimenter, is learned better than a set of words although the experimenter usually cannot say in advance whether he has chosen the best or the worst pictures. Apparently the only clear case of truly poor picture memory performance in the literature is that shown by Goldstein and Chance (1970), who carefully constructed extremely confusable items. While performance may be improved further by laboriously selecting striking items (as was done here with the Vivid set) this is not at all necessary to demonstrate pictorial superiority. Incidentally, even the Vivid set was far from perfect, with many items of indifferent quality.

The above conclusion, regarding recognition memory, agrees with other studies, particularly those of Nickerson, Shepard and Standing *et al.*, although Paivio and Csapo (1971) have limited its generality in showing that sequential memory for words is sometimes superior to that for pictures. The present finding that pictorial superiority is maintained when recall measures are used also agrees with some previous work (Paivio, Rogers and Smythe, 1968); the earlier finding of Jenkins, Stack and Deno (1969) that recall performance of seven-year-old children is no better with pictures than words clearly cannot be taken as definitive. Overall, pictorial superiority appears to be a robust phenomenon that is found under a wide variety of experimental tasks.

Only in the case of retrieval time is there a superiority for verbal stimuli, but this conclusion is not final due to the difficulty of allowing fully for perceptual factors. It is also noteworthy that both types of material are similar in following a power law, and in showing a very rapid rate of memory search. Extrapolation from the data (Fig. 3) suggests that one second of scanning time (beyond the time needed to respond to a single item) would allow the subject to search for 51,180 pictures (or 68,008 words) in memory. This figure differs markedly from Sternberg's (1967) estimate of 25–30 items per second, but is based on a much wider range of stimulus sets than was used by Sternberg. It also appears to be more compatible with memory search rates for the naming of objects (Oldfield, 1966). The present values do of course become progressively closer to Sternberg's if only very small stimulus sets are considered, and the regression equation suggests that to search for two pictures in memory rather than one requires an extra 40 ms, which coincides with Sternberg's estimate. Sternberg's linear principle may well be correct for the special case of small memory sets. The present data show agreement with Sternberg's principle of exhaustive memory search in that positive and negative trials gave the same reaction time (Table V).

It is probably erroneous to postulate that all memories are stored in the form of a verbal code (Paivio and Csapo, 1971), a conclusion which is reinforced by these studies for it is unlikely that requiring an extra stage of encoding of stimulus material during learning and during recognition could *facilitate* performance. (However, recoding into verbal form is readily performed when needed, as in Experiment III.) Instead, the converse appears more likely. Paivio (1969) has already shown that the best predictor of verbal learning performance is the imagery-producing property of the verbal stimuli employed; if we look on imagery as a type of internal picture, it seems that some type of pictorial coding is likely to be of more fundamental importance in learning than verbal coding.

There is an interesting contrast between the performance of R in Experiment I and that of the gifted mnemonist S studied in detail by Luria (1968). These two subjects are at opposite and extreme ends of the spectrum of learning ability: S displayed an enormous memory capacity (limitless, according to Luria), while R was noticeably poorer than several hundred other experimental subjects tested by the author. (Although in terms of intelligence he would clearly be rated as far superior to S.) Yet a curious similarity exists: unlike other subjects, both performed about equally well with any type of material, abstract or concrete. Luria describes S as possessing abnormally vivid and realistic imagery, which he was able to use with success in the learning situation; this of course corresponds to the ancient Method of Loci (Yates, 1966) and again suggests the primacy of pictorial coding. The subject R, conversely, possesses only very minimal imagery (visual or otherwise) and seems not to use visual thinking according to his own report. A gifted mathematician, he never employs pictorial methods of representation even in areas such as topology where they are indispensable to most individuals. It seems that S could visualize almost anything, and R almost nothing, while average individuals can visualize more concrete words—or actual pictures—more easily than abstract items. This overall pattern of results obtained by Paivio, Luria and the present investigator then becomes consistent if we accept the superiority of pictorial coding. It is also noteworthy that although short music sequences probably defy any verbal recoding (save possibly by trained musicians), they were found to show equivalent retention in memory to words and may perhaps be thought of as being roughly the auditory equivalent of pictures. In comparison, recognition memory for common (non-musical) sounds is lower than that for printed or spoken words—although verbal encoding of such sounds appears possible (Miller and Tanis, 1971).

The data show how different an estimate of memory capacity follows when free-recall or recognition tests are used rather than the classic "memory-span" test where performance seldom exceeds seven items. The span test measures primarily *order* information (since it makes little difference whether the subject knows in advance the set of stimulus items to be used), whereas the present study deals only with *item* information; many other differences also exist between the memory-span task and the present situation. However, it is curious that the length of descriptions spontaneously given in the verbal recall of pictures (Experiment III) consistently averaged six words, which equals the immediate memory span for (unrelated) words. This may simply reflect a particular type of linguistic habit, rather than implying that a common link exists between the two types of memory.

The main methodological problem encountered in the present studies is that the use of large learning sets precludes the use of a truly immediate test of memory, due to the considerable time needed to view the stimuli even once. This also leads to considerable fatigue, with sets over a few hundred items, and subjects clearly must make considerable efforts to maintain their vigilance. In the case of 10,000 items, the cumulative effects of five days' viewing, as checked by the author, are extremely gruelling and unpleasant. It is possible that some of the decline in performance found with larger sets is due to this factor rather than the informational load on memory as such. Despite such problems, picture memory is a suitable object for study, since presumably when the mechanism of memory is working well its full range of operating principles is most likely to be displayed and understood.

This research was supported by grants APA-0396 (National Research Council of Canada) and 9401-50 (Defence Research Board of Canada). The author is greatly indebted to J. Weller, I. Wallace and B. B. Murdock for various forms of assistance.

References

BINET, A. and HENRI, V. (1894). La memoire des mots. *Année Psychologique*, **1**, 1–23.

DAVIS, R., SUTHERLAND, N. S. and JUDD, B. R. (1961). Information content in recognition and recall, *Journal of Experimental Psychology*, **61**, 422–9.

DEESE, J. (1960). Frequency of usage and number of words in free recall: the role of association. *Psychological Reports*, **7**, 337–44.

GOLDSTEIN, A. G. and CHANCE, J. E. (1970). Visual recognition memory for complex configurations. *Perception and Psychophysics*, **9**, 237–41.

GOMBRICH, E. H. (1969). The evidence of images. In SINGLETON C. S. (Ed.), *Interpretation: Theory and Practice*, pp. 35–104. Baltimore: John Hopkins Press.

HENMON, V. A. (1912). Modes of presentation and retention. *Psychological Review*, **19**, 79–96.

HILGARD, E. R. (1951). Methods and procedures in the study of learning. In STEVENS, S. S. (Ed.), *Handbook of Experimental Psychology*, pp. 517–67. New York: Wiley.

JENKINS, J. R., STACK, W. and DENO, S. (1969). Children's recognition and recall of picture and word stimuli. *AV Communication Review*, **17**, 265–71.

LUH, C. W. (1922). The conditions of retention. *Psychological Monographs*, **31**, No. 142.

LURIA, A. (1968). *The Mind of a Mnemonist*. New York: Basic Books.

MILLER, J. D. and TANIS, D. C. (1971). Recognition memory for common sounds. *Psychonomic Science*, **23**, 307–8.

MURDOCK, B. B. (1960). The immediate retention of unrelated words. *Journal of Experimental Psychology*, **60**, 222–34.

NICKERSON, R. S. (1965). Short-term memory for complex meaningful visual configurations, a demonstration of capacity. *Canadian Journal of Psychology*, **19**, 155–60.

NICKERSON, R. S. (1968). A note on long-term recognition memory for pictorial material. *Psychonomic Science*, **11**, 58.

OLDFIELD, R. C. (1966). Things, words and the brain. *Quarterly Journal of Experimental Psychology*, **18**, 340–53.

PAIVIO, A. R. (1969). Mental imagery in associative learning and memory. *Psychological Review*, **76**, 241–63.

PAIVIO, A. R. and CSAPO, K. (1971). Short-term sequential memory for pictures and words. *Psychonomic Science*, **24**, 50–1.

PAIVIO, A. R., ROGERS, T. B. and SMYTHE, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science*, **11**, 137–8.

SEIBERT, L. C. (1932). A series of experiments on the learning of French vocabulary. *Johns Hopkins University Studies in Education*, No. 18.

SHEPARD, R. N. (1967). Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behavior*, **6**, 156–63.

STANDING, L. (1971*a*). Preparing stimuli for picture-memory studies. Research Report EP-1; Psychology Department, Bishop's University.

STANDING, L. (1971*b*). Categorizing stimuli for picture-memory studies. Research Report EP-2; Psychology Department, Bishop's University.

STANDING, L., CONEZIO, J. and HABER, R. N. (1970). Perception and memory for pictures: single-trial learning of 2500 visual stimuli. *Psychonomic Science*, **19**, 73–4.

STERNBERG, S. (1967). Two operations in character recognition: some evidence from reaction-time measurements. *Perception and Psychophysics*, **2**, 43–53.

STEVENS, S. S. (1961). The psychophysics of sensory function. In ROSENBLITH, W. A. (Ed.), *Sensory Communication*, pp. 1–33. Cambridge, Mass.: M.I.T. Press.

STRONG, E. K. (1912). The effect of length of series upon recognition memory. *Psychological Review*, **19**, 447–62.

SWETS, J. A. (Ed.) (1964). *Signal Detection and Recognition by Human Observers: Contemporary Readings*, pp. 682–3. New York: Wiley.

WOODWORTH, R. S. (1915). The influence on retention of conditions favouring quickness of learning. *Journal of Philosophy*, **12**, 246.

YATES, F. A. (1966). *The Art of Memory*. London: Routledge & Kegan Paul.

Received 25 *May* 1972