The length-difficulty relation in serial memorization of words, pictures, and snowflakes

PETER L. DERKS
College of William and Mary, Williamsburg, Virginia 23185

Arrays of 3, 6, 9, or 12 items were studied until they could be ordered from memory without error. A power function described the relation between array size and study time for words, pictures of objects, and pictures of snowflakes. Under two payoff conditions, the exponents for words and snowflakes were around 2.3 and for pictures, 2.0.

One of the oldest exercises in the study of memory is the determination of the functional relation between the amount of material to be learned and the time required to learn it (Ebbinghaus, 1885/1964). Although the measured relations have been fairly consistent, different procedures have produced different estimates (Woodworth & Schlosberg, 1954). In a relatively recent series of studies, an attempt was made to determine a length-difficulty relation that depended primarily on the subject's behavior and not on the restrictions of the experimental method (Derks, 1974; Martin, Fleming, Hennrikus, & Erickson, 1977). Subjects were presented with all the material to be learned and were permitted to study it until they felt they knew it. In this way, no studying could be concealed on presentations of known items, during the time between trials, or during recall, possibilities uncontrolled in most earlier experiments.

Under these conditions, the time the subjects used to learn a list of words was best described by a power function, $D = iL^{2.6}$, where D was the difficulty estimated by study time, i was a constant of intercept that generally increased with single-item difficulty, and L was the list length or number of items. In other words, over the range tested, a doubling in list length resulted in approximately a sixfold increase in study time. Although changes in subjective organization that correspond to list length changes might have accounted for this ratio relationship, Martin et al. (1977) manipulated subjective organization and found that the power function, or something very much like it, continued to appear. They concluded that "the length-difficulty relation is not decomposable" and, on the basis of some of the same data, that this function "has no universal shape." Since it is a rare phenomenon that is invariant through

all transformations and since the power function seems to describe more than its share of the length-difficulty relations, the following study was performed to test the limits of this relation and to examine possible sources of its modification.

Two kinds of materials were chosen as likely candidates to disrupt the function that best describes the length-difficulty relation for words. They were pictures, which are easier to remember than words, and pictures of snowflakes, which are much harder (Bevan & Steger, 1971; Goldstein & Chance, 1971). Furthermore, the difference between memory for pictures of things and memory for shapes like snowflakes is related to the ease with which they can be given labels or names (Kelly & Martin, 1974; Paivio, 1971). Since pictures seem to be remembered by two "codes," the object represented and the label of the object, this dual coding plus the discriminability between the pictures themselves makes longer lists of pictures easier to remember than longer lists of words. As a result, the length-difficulty function for pictures might be concave downward on a log-log plot. These multiple cues might increase opportunities for clustering and yield groups of items that would be recalled as individual items. As a result, increased length might result in decreased difficulty per item. Snowflakes, on the other hand, should be unlabelable and relatively hard to discriminate. Consequently, the difficulty of determining and retaining an array of snowflakes would be increased exponentially with each additional snowflake. If this were the case, it would be reflected in a length-difficulty relation that would be concave upward on a log-log plot.

METHOD

A set of 196 names of objects was selected to minimize membership in common conceptual categories (Battig & Montague, 1969). Another set of stimuli was prepared with detailed line drawings of the same objects. The snowflake stimuli were 120 items selected from Bentley and Humphrey (1931/1962) to be discriminable in specific content but similar in general outline. The names, pictures, and snowflakes were mounted on 6.5 x 8.5 cm cards for easy handling.

For testing, the arrays of stimuli were arranged in a 30 x 45 cm "book" that could be opened and closed by the subject. The duration of studying an array was timed by a Meylan .1-sec digital timer activated by the opening of the book. Inside the book, a Plexiglas shield held the cards in place when the cover was opened or closed.

The stimulus arrays were of 3, 6, 9, or 12 cards. Although this is a relatively narrow range, it includes two doublings of array size and crosses the classic boundary of memory span (Miller, 1956). All subjects received at least 12 different arrays. The different lengths were randomly presented except that all lengths were used once per block of four presentations.

On each trial the experimenter would arrange an array in columns of three, and in one, two, three, or four rows, depending on array size, and close the book. The subject would open the book, study the array until it was presumed known, and close the book. The experimenter then collected the cards in a pile and handed them to the subject, who put them back in their original order. Thus the subjects did not have to learn the individual items, only their order.

When an error was made in the item order, that array size, but not the specific array, was repeated in an effort to obtain at least three estimates of study time for accurate memory at each array size. Since the subjects were informed that they could leave at any time and were already guaranteed payment for their participation, three accurate estimates were not always obtained. In fact, the lack of complete results for several subjects led to a replication in which the subjects were given payment at the beginning of the session and then lost a quarter for each error.

In the original (flat-rate) study, payment was $2 for pictures and words and $4 for snowflakes, pilot research having shown that snowflakes were about twice as hard to learn as words or pictures. (The subjects usually took two sessions of about 40 min each to participate in the snowflake task and one session for pictures and words.) For the replication (contingency) study, the initial sum was $3 for pictures and words and $6 for snowflakes. No subject knew what other subjects received. Six males and six females were tested on each of the three stimulus types in the two experiments. One male did not return for his second snowflake session in the flat-rate condition, leaving a total of 71 subjects.

RESULTS AND DISCUSSION

In previous research using this method, geometric means of latencies for correct responses have been used as representative of the subject's study times. Geometric means were used to reduce the effects of long latencies that could come about for a variety of reasons that have nothing to do with study time. Nevertheless, geometric means still include latencies that may have resulted from fatigue, inattention, or idiosyncratically difficult lists. Since the probability of getting a list correct by chance and producing an artificially short latency is small, minimum latencies were used as the estimate of study time. Figure 1, then, is based on the average of each student's shortest latency without an error (D) at each array size (L). Geometric means were also calculated and analyzed. Their values were similar but, of course, slightly larger. The pattern of statistically significant effects was the same, with one exception to be noted later.

The length-difficulty functions are plotted on log-log coordinates to emphasize any departures from a power function, which is linear on such a graph. Words
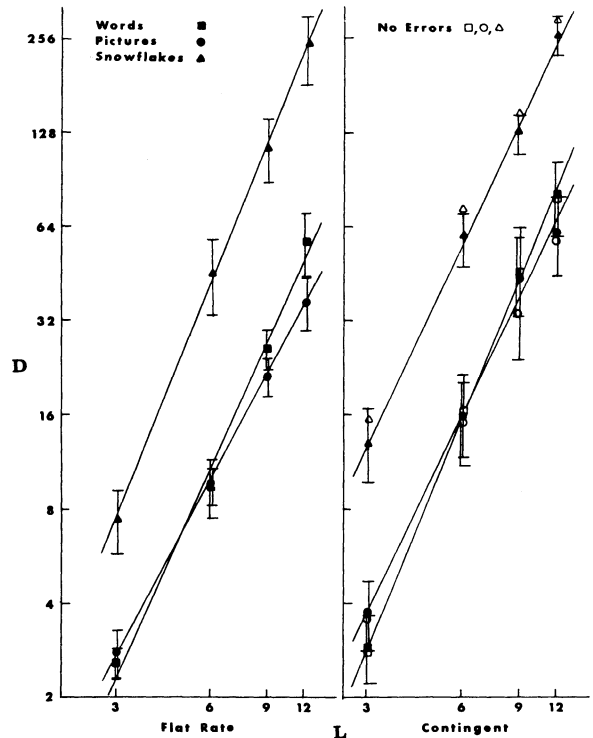

Figure 1. Average minimum study times with all items correctly placed in recall (D) by number of items in the studied array (L).

memorized in the contingency condition appear slightly concave upward, and pictures bend up a little for 12 items in the contingency condition and down a little in the flat-rate condition. The standard deviations of these points all encompass the best-fit power function, however, so the predicted departures from this function have not been obtained.

The power functions that best describe the six conditions are, for words, flat-rate $D = .210L^{2.21}$ and contingent $D = .204L^{2.42}$; pictures, flat-rate $D = .357L^{1.85}$ and contingent $D = .391L^{2.07}$; snowflakes, flat-rate $D = .496L^{2.48}$ and contingent $D = 1.25L^{2.13}$. The exponent for the word acquisition condition is less than the 2.6 of earlier studies. This may be because the subject did not have to remember the identity of the items, only their order. In the earlier studies using this method, a recognition task without ordering requirements reduced the exponent to 1, that is, a linear function. A poststudying cue task that required the subjects to recall the location of just a single item reduced the exponent to 2.4. Here, the exponent for pictures is 2, that is, study time goes up as the square of the number of pictures, and words and snowflakes are about 2.2 − 2.4, that is, study time increased about fivefold for a doubling of list length.

An analysis of variance was performed on these data for the combined payoff conditions, even though they were not directly comparable, in order to estimate their relative consistency. Array size significantly increased difficulty $[F(3,195) = 204.79, p < .001]$ and

stimulus type influenced difficulty [$F(2,65) = 81.82$, $p < .001$]. Of greatest interest was a significant interaction between array size and complexity [$F(6,195) = 55.36$, $p < .001$]. Therefore, the slopes of the length-difficulty relations are significantly different. Furthermore, the only significant difference between minimum study time for flat-rate payoff and the contingency payoff replication was a main effect between conditions [$F(1,65) = 4.72$, $p = .033$]. Neither the overall slopes of the function [$F(2,65) = .003$] nor the interaction of slope with complexity of material [$F(6,195) = .121$] changed significantly between payoff conditions.

Since power values of the slopes for snowflakes bracket those for words and the picture slopes were the least of all, the length-difficulty relation is more shallow for pictures than that for words and snowflakes. Nevertheless, the function was not concave downward on the log-log plot. Pictures also had a larger constant of intercept than words. This increased latency suggests that the individual picture takes longer for the possible "dual" coding in this task than a word for verbal coding alone.

Although snowflakes were more difficult to learn than words, their relative difficulty remained constant with array size. Estimated from Figure 1, a snowflake is worth a little less than two words for the array sizes tested here. The constant of intercept also suggests that a snowflake takes more processing than a word, but the ratio is somewhat larger than two to one. Since they did not have ready-made names, subjects may have had to discriminate and label a couple of critical features for each flake to remember their order.

As indicated in Table 1, the contingency condition significantly reduced error rate [$F(1,65) = 102.25$, $p < .001$]. Error rates were also significantly affected by array size [$F(3,195) = 16.66$, $p < .001$] and nonsignificantly by complexity [$F(2,65) = 2.78$, $p = .069$]. The contingency condition reduced these effects somewhat, but the interaction between payoff and array size was not quite significant [$F(3,195) = 2.55$, $p = .057$], nor was the interaction between payoff and complexity [$F(2,65) = 2.66$, $p = .077$].

Nevertheless, the contingency condition produced subjects who made no errors: six subjects for words, seven for pictures, and five for snowflakes. In the

flat-rate condition, only 2 of 36 subjects learned their 12 arrays without an error. The power functions that describe the average minimum study times for the subjects who made no errors in the contingency payoff condition are, for words, $D = .200L^{2.42}$, for pictures $D = .390L^{2.01}$, and for snowflakes $D = 1.532L^{2.11}$. These averages for the errorless subjects are included in Figure 1. Since these values are very similar to those for the group as a whole, it can be inferred that the use of minimum study time gives a stable estimate of task difficulty regardless of strategic speed-accuracy tradoffs individual subjects may use.

It could also be argued that the length-difficulty functions for the contingency condition are more nearly the "truth," since error rates were reduced to more equivalent values by contingency payoff and had no effect on the length-difficulty functions. Certainly, with no errors there is no possibility of easier lists being introduced on trials substituted for errors. In this case, the slope for snowflakes is more similar to the slope for pictures, and the constant of intercept indicates much more processing per item. It could be argued that under the pressure of contingency payoff, the subjects treated the snowflakes more like pictures and coded an image and a label that produced better memory for longer lists. In this case a snowflake might be considered worth about two pictures instead of two words. Such an argument is suspect, however, since the interaction between these two experimental conditions was not significant.

This interaction was slightly significant, however, when geometric means were compared [$F(6,195) = 2.31$, $p = .035$], the one difference between the patterns of minimum and geometric means. Perhaps the effect of increased motivation was primarily on the longer latency responses, which would have been included in the geometric mean measure. Nevertheless, the idea that the subject turns the snowflake into a meaningful picture when accuracy of recall is important must remain a hypothesis for future testing since the present study manipulated motivation only incidently. Furthermore, the decision to use minimum latencies or geometric means seems to depend on whether stable processes or external manipulations are being studied.

In any case the main conclusion to be drawn from this study is that the power function that describes the relation between length and difficulty for words holds equally well for pictures of objects and of complex shapes. Although the slopes and intercepts of those functions yield information that seems consistent with other conceptions of memory for various kinds of stimuli, the underlying cause of this particular phenomenon remains unclear.

Table 1
Error Rate (Errors/Total Responses)

	3	6	9	12	By Condition
Flat Rate					
Words	.00	.16	.26	.38	.262
Pictures	.04	.04	.12	.18	.116
Snowflakes	.10	.17	.37	.37	.290
Contingency					
Words	.00	.08	.06	.13	.086
Pictures	.00	.04	.13	.13	.081
Snowflakes	.00	.13	.04	.15	.102

REFERENCES

BATTIG, W. F., & MONTAGUE, W. E. Category norms for verbal items in 56 categories: A replication and extension of the

Connecticut category norms. *Journal of Experimental Psychology Monographs*, 1969, **80**(3, Part 2).

BENTLEY, W. A., & HUMPHREYS, W. J. *Snow crystals*. New York: Dover, 1962. (Originally published, 1931.)

BEVAN, W. A., & STEGER, J. A. Free recall and abstractness of stimuli. *Science*, 1971, **172**, 597-599.

DERKS, P. L. The length-difficulty relation in immediate serial recall. *Journal of Verbal Learning and Verbal Behavior*, 1974, **13**, 335-354.

EBBINGHAUS, J. [*Memory: A contribution to experimental psychology*] (H. A. Ruger & C. E. Bussenius, trans.). New York: Dover, 1964. (Originally published, 1885.)

GOLDSTEIN, A. G., & CHANCE, J. E. Visual recognition memory for complex configurations. *Perception & Psychophysics*, 1971, **9**, 237-241.

KELLY, R. T., & MARTIN, D. W. Memory for random shapes: A dual-task analysis. *Journal of Experimental Psychology*, 1974, **103**, 224-229.

MARTIN, E., FLEMING, F. G., HENNRIKUS, D. J., & ERICKSON, E. A. Studies of the length-difficulty relation in serial memorization. *Journal of Verbal Learning and Verbal Behavior*, 1977, **16**, 535-548.

MILLER, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, **63**, 81-97.

PAIVIO, A. *Imagery and verbal processes*. New York: Holt, Rinehart, & Winston, 1971.

WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology*. New York: Holt, 1954.