Journal of Comparative and Physiological Psychology, 1964, 58, 237-242.

Stellar, E., & Hill, J. H. The rat's rate of drinking as a function of water deprivation. Journal of Comparative and Physiological Psychology, 1952, 45, 96-102.

Teitlebaum, P. The use of operant methods in the assessment and control of motivational states. In W. K. Honig (Ed.), Operant behavior: Areas of research and application. New York: Appleton-Century-Crofts, 1966.

Semantically cued retrieval of words from long-term memory

D. J. MURRAY
Queen's University, Kingston, Ontario, Canada

Subjects were asked to give in 2 min as many words as possible which belonged to a given category such as "girls' names." It was found that the number of words reported was a power function of the number of words estimated to be in the category, and that high frequency words were given before low frequency words.

It has been found by Standing (1973) that the number of items which can be retrieved from memory is a power function of the number of items presented. If R is the number retrieved and P is the number presented,

$$R = kP^m \qquad (1)$$

where m varies with the experimental paradigm. Murray (1975) also examined retrieval by a "free output" method: subjects were asked to find as many words as they could, in 2 min, whose 1st, 2nd, 3rd, 4th, or 5th letter was A or B or . . . or Z. For all five levels of retrieval cue, the number of words emitted was a power function of the number of words having the required letter properties presumed to be known by the subject.

The purpose of the present paper is to examine the free output of words given semantic rather than graphemic retrieval cues. Although Loftus, Wiksten, and Abelson (1974) have shown that there may be differences in search techniques when one is attempting to retrieve a word cued graphemically as opposed to

semantically, we need to discover experimentally if free output cued semantically follows a power function in the same way as does free output cued graphemically. With graphemic cues, the index of vocabulary size (the number of words with the appropriate characteristics presumably known to the subject) was given by a count of the words in the Thorndike Lorge lists. With semantic cues, such a count, for example, of all the names of trees or animals, would be tedious and probably inaccurate. Instead, the category norms of Battig and Montague (1969) were used to estimate the probable numbers of words of a given category likely to be known to student subjects. These norms also give the relative frequencies with which each category member comes to mind given a category name, so than an examination of frequency effects over time can also be undertaken.

METHOD

Materials

Of the 56 categories in the Battig and Montague norms, 30 were selected on the basis of their giving a wide range of vocabulary sizes. In the norms, the words in each category are divided into two sections, those produced by 10 or more subjects in the original population, and those words produced by fewer than 10 subjects. In the present study, the former subset of words constituted the value of "vocabulary size." For example, there were 18 words given by more than 10 subjects in answer to the category "a precious stone" although a total of 68 words was given in all, including many words given by only one
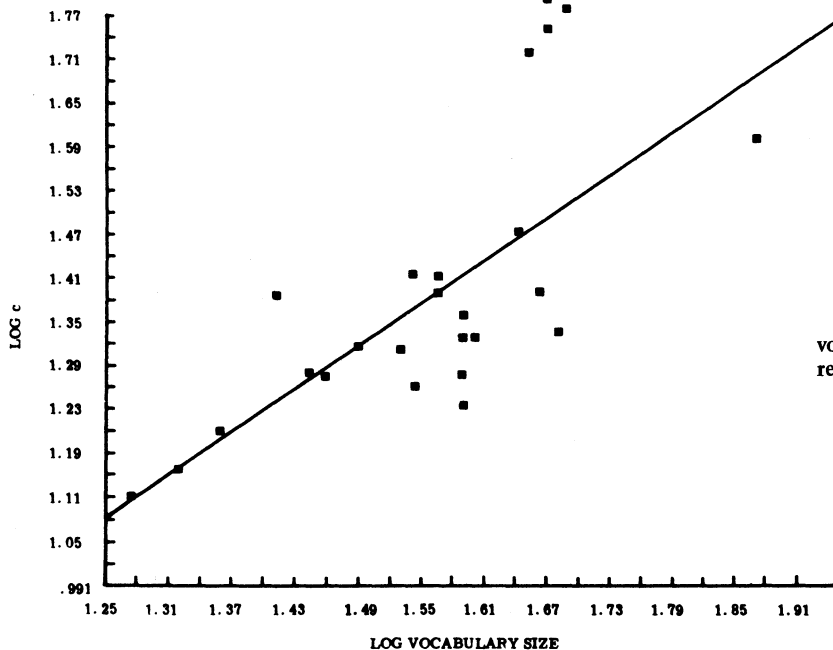
Figure 1. Log c plotted against log vocabulary size (see text). Each data point represents a single category.

subject. The vocabulary size for this category was therefore taken as 18; the vocabulary sizes ranged, across the 30 selected categories, from 18 (as this example) to 92 ("A female's first name"). The 30 categories chosen, ranked in order of increasing vocabulary size, were: a precious stone, a unit of distance, a memember of clergy, a unit of time, a military title, a color, an alcoholic beverage, a type of footgear, a crime, a flower, a type of human dwelling, a carpenter's tool, a disease, an article of furniture, a kitchen utensil, a tree, a type of dance, a type of vehicle, a type of music, a type of ship, a part of a building, an occupation or profession, a toy, a city, a part of the human body, a weapon, a country, a college or university, a male's first name, a female's first name. In addition three other categories were used for providing initial practice at the free output task: these were, a sport, a relative, and a kind of money.

Apparatus and Design

On arrival, each subjects was instructed that he would be asked to write down as many items from a given category as he could in 2 min. He wrote on a specially ruled piece of paper, which, when laid against a column of 12 small lights on a prepared wooden surface, allowed for 12 lines of writing, one line opposite each light. When subject felt ready to begin, he pressed a button. This caused the name of a category to appear in the viewing slit of a memory drum and at the same time caused the top light to come on. He wrote as many words as he could on the top line, then after 10 sec, the second light in this column came on with a click. He then switched the line he was writing on to the second line down. This procedure continued for 2 min. This device ensured that the subject's output protocol was divided into 12 10-sec units, a feature necessary for subsequent analysis of the data (see below).

Following the three practice categories, the subject worked through the 30 main categories. These appeared in one of two random orders in the memory drum. Typically, a test session took 90-120 min; the subject was told that he could work at his own rate, resting when he felt like it.

Subjects

Subjects were 32 volunteers from summer school in the Queen's University Psychology Department. They were paid $4 each for their services.

RESULTS AND DISCUSSION

As in Murray (1975), the measure of the number retrieved was taken to be asymptote c in the equation found by Bousfield and Sedgewick (1944) to be a good fit to free output data:

$$n = c(1 - e^{-\lambda t}) \qquad (2)$$

In this equation n is the total number of words found by time t, c is the asymptote of the negatively accelerated curve described by Equation 2, and λ is the number of words found in a very short interval of time dt. In the present study, the mean number of words obtained per subject for each 10-sec unit of time was plotted cumulatively and a computer then calculated, by a least-squares method, those values of c and λ which gave the best fit to the obtained data. Two categories gave cumulative output curves which were not in accord with the Bonsfield and Sedgewick equation and were therefore excluded from the analysis: these were "a type of human dwelling" and "a unit of time." For the remaining 28 categories, log c was plotted against the log of the vocabulary-size for each category. If Equation 1 holds, then a straight line should emerge on such a plot (c would be the equivalent of R, and the vocabulary size would be the equivalent of P). Figure 1 shows this plot. The value of k was .70, the value of m was .987, and the correlation between c and vocabulary size was .796. This last value is significant at the .01 level and suggests that the data are indeed described by a power function. The question may be asked whether similar results would hold if instead of taking as the measure of vocabulary size the number of words reported by at least 10 subjects in the Battig and Montague norms, we took the
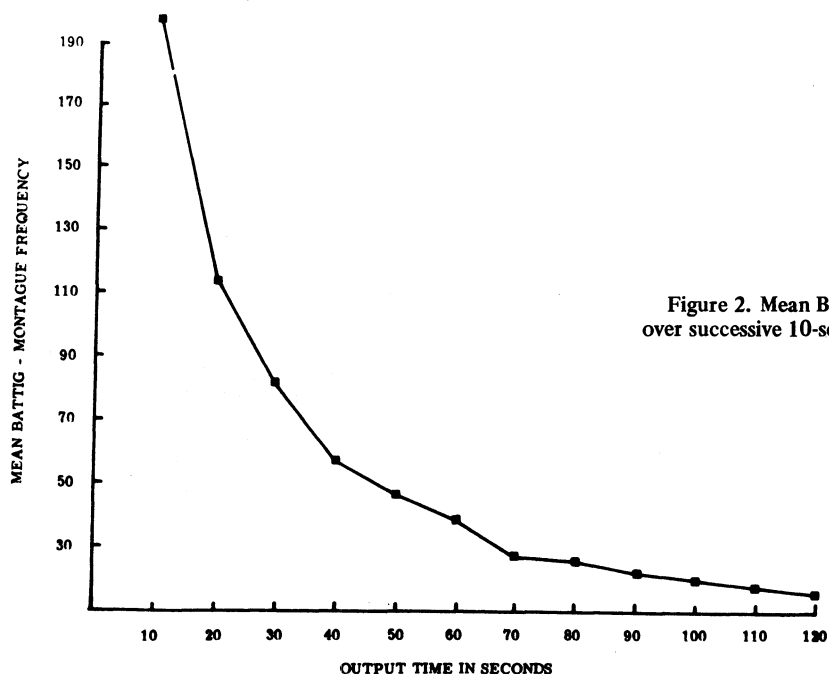
Figure 2. Mean Battig-Montague frequency of words produced over successive 10-sec intervals of output.

total number of words reported by all subjects. If so, since the latter value greatly exceeds the former value, we should expect the slope m of the resulting plot of c against the total vocabulary size to be less than the m obtained for the more conservative estimate of vocabulary size based only on popular responses. The value of k thus obtained was .83, the value of m was .765, and the correlation was .783. Again, a power function appears to be an adequate description of the data. However, the r values obtained here, for semantic retrieval, are in general lower than those values obtained either by Standing (1973) or by Murray (1975): there is clearly greater variability in the data when we are dealing with semantic cues. We believe that this is the result of greater variability between subjects in respect to the various numbers of items known in each category.

In Murray (1975) it was also found that, for easy retrieval cues, high frequency words were reported first, followed by rarer words. For difficult retrieval cues this did not hold, but it was argued that this finding was consistent with our knowledge of the distribution of word frequencies for words containing a given first letter as opposed to a given fifth letter. It seemed worth asking whether this evidence for a "spew hypothesis" (Underwood & Schulz, 1960) would also hold for the case of semantic retrieval. Accordingly each word in subjects' protocols was given a frequency rating, in this case, the number of subjects in Battig and Montague's population who also reported that word. The mean frequency of the words reported in our experiment was then plotted against retrieval time over the course of 2 min. Figure 2 shows the result, which clearly supports the "spew hypothesis." The smoothness of the curve also suggests that the Battig and Montague norms can be validly used as an estimate not only of vocabulary size

but of mean frequencies of usage and/or association, given a category name. Two categories were omitted from this analysis, namely "cities" and "colleges," because our Canadian population gave many answers not available in the Battig and Montague lists.

It may be concluded that graphemic and semantic retrieval are similar in two respects: both follow a power function with respect to vocabulary size and, provided the retrieval cue is a relatively easy one, both yield data in which high frequency words are more quickly found than are low frequency words. Although there may be qualitative differences in the way words are found, given graphemic as opposed to semantic cues, the two kinds of condition yield similar quantitative function. It may be inferred that the study of retrieval from long-term memory may be validly carried out using graphemic cues as well as semantic and that some of the problems in interpreting semantic data may be circumvented by the use of graphemic retrieval cues.

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 Monograph, 1969, 80, No. 3, Part 2.
Bousfield, W. A., & Sedgewick, G. H. W. An analysis of sequences of restricted associative responses. Journal of General Psychology, 1944, 30, 149-165.
Loftus, E. F. Wiksten, S., & Abelson, R. P. Using semantic memory to find versus create a word. Memory & Cognition, 1974, 2, 479-483.
Murray, D. J. Graphemically cued retrieval of words from long-term memory. Journal of Experimental Psychology: Human Learning and Memory, 1975 (in press).
Standing, L. Learning 10,000 pictures. Quarterly Journal of Experimental Psychology, 1973, 25 207-222.
Underwood, B. J., & Schulz, R. W. Meaningfulness and verbal learning. Chicago: Lippincott, 1960.