Journal of Experimental Psychology:
Human Learning and Memory
1980, Vol. 6, No. 4, 370–378

Automatic Encoding of Category Size Information

Joseph W. Alba, Walter Chromiak, Lynn Hasher, and Mary S. Attig
Temple University

The memory codes of environmental events are known to contain information about the frequency of occurrence of those events. Two studies demonstrated that adults could estimate the frequency of implicitly referenced events (category names) when they were presented with individual events (category instances). In Experiment 1 it was found that under incidental instructions, adults are sensitive to the frequency with which category exemplars occur. It was also suggested that such information is encoded at the time the exemplars are presented. The results of Experiment 2, by varying instructions and mode of exemplar presentation, suggested that category-frequency information, like individual event frequency, may be encoded automatically.

Current evidence suggests that people are extremely sensitive to the rates with which environmental events occur. For example, they are sensitive to the relative frequency of diseases and accidents, although their knowledge is far from perfect and is skewed by unrepresentative media coverage (Lichtenstein, Slovic, Fischhoff, Layman, & Combs, 1978). A good deal of environmental information is linguistic in nature, and there are a number of findings that suggest that people process frequency information for naturally occurring linguistic events such as utterances (Gude & Zechmeister, 1975), as well as for the components of utterances—words (Carroll, 1971; Shapiro, 1969), pairs of letters (Underwood, 1971), and single letters (Attneave, 1953).

The question explored in the two studies reported here is whether people also abstract occurrence-rate information for the higher order, superordinate level information that is referenced, often only implicitly, by an event. For example, one could have had conversations on the following three seemingly different topics: (a) the state of the economy, (b) the role of nuclear energy in the next decades, and (c) federal funding for abortion. In each case, an implicit reference or "generalization" (Lindsay & Norman, 1977, p. 403) might be made to the incumbent president's chances for reelection. Are frequency counts kept for such higher order information?

In addition, the two studies reported here were also concerned with the suggestion that information regarding the frequency of occurrence of elements is encoded automatically or obligatorily (e.g., Hasher & Zacks, 1979; Hintzman & Stern, 1978). Previous research has shown that sensitivity to frequency information remains stable across a variety of experimental manipulations. The inference has been that frequency information is one of a number of environmental attributes that is encoded with little effort and is minimally influenced by the intentions of a person. The present research seeks to extend our understanding of exactly what types of information receive obligatory "counts."

The presentation stimuli in the two experiments were instances from conceptual categories; the category names were the superordinates whose implicit or covert elicitations concerned us. This choice of materials was made because of our assumption that when the representation of a cate-

This research was supported by a grant from the National Institute of Mental Health, MH-33140-01.

Thanks are due to James Arbuckle and Robert Weisberg for their assistance in this project and to an anonymous reviewer for his helpful comments.

Walter Chromiak is now at Dickinson College.

Requests for reprints should be sent to Lynn Hasher, Department of Psychology, Weiss Hall, Temple University, Philadelphia, Pennsylvania 19122.

gory instance is accessed, the representation of its superordinate is also momentarily activated. Several lines of evidence justify this assumption. College students have been shown to spontaneously encode this information in a variety of memory tasks when not explicitly instructed to do so (e.g., Jacoby, 1974; Nelson, Fehling, & Moore-Glascock, 1979). Conceptual categories are also an aspect of experience to which even children as young as 2 and 3 yr. old are known to be sensitive (Faulkender, Wright, & Waldron, 1974).

Given that superordinates are elicited, our central question concerns whether adults keep frequency counts for superordinate level information. The two experiments reported here suggest that people do tag occurrence rates of covertly referenced category names when only their basic level exemplars appear.

Experiment 1

In this experiment, college students saw a list of categorized words. Critical categories were represented by 0, 3, 6, or 9 exemplars, each of which occurred only once. Students were informed only after they had seen the list that when cued by the category name, their task was to judge the number of exemplars that they had seen. The category names were shown for the first time at test, and frequency decisions were made at one of four different pacing rates (2, 4, 5, or 10 sec per decision).

The primary question of this study was whether or not adults store a representation of occurrence rates for category names when they only see exemplars. A second question concerned the time at which the representation of frequency (or, more conservatively, the information that mediates the frequency judgments) is encoded. If stored as each exemplar is presented, then test pacing should have little influence on either the size or accuracy of the judgments. If frequency is a secondary attribute, inferred at test (perhaps by subjects generating instances to the superordinate and counting or guestimating), then the rate of testing should influence the accuracy of those judgments.

Method

Materials. Nine words were chosen from each of 12 categories in the Battig and Montague (1969) norms. The instances were the 9 with the highest frequency, except in those cases in which a word was judged to have more than one meaning. The unambiguous replacement was then the next most frequent instance of the category. These items, together with one high frequency instance from each of 10 other Battig-Montague categories, were then used to construct the lists that were presented to subjects.

The stimulus lists consisted of 64 words, 10 of which were unrelated buffer items (instances from other categories in the norms). Half of the buffer items were placed at the beginning and half at the end of a list. The 54 categorized instances were the critical materials in the study. These came from 9 different categories with 3 categories represented at each of the critical frequency levels; that is, 3 categories each contributed 3 instances, 3 other categories each contributed 6 instances, and 3 other categories each contributed 9 instances, for a total of 54 items. That left 3 other categories that could be used on the judgment test as the zero-presented items. The actual category exemplars that occurred were selected randomly for each subject, except, of course, in the 3 nine-item categories.

Four unique lists were generated by counterbalancing categories with the instance-frequency variable. Thus, across subjects each category was represented equally often at each of the four instance levels, 0, 3, 6, and 9. This was done to eliminate the influence of unknown aspects of particular categories on memory for the number of experimental instances.

The category instances were placed in the list in such a way that a third of the items from each number of instances level would occur in each third of the list. Also, we attempted to maximize the distance between members of the same category. This ensured that subjects could not be aware of the categorized nature of the list until they had seen a substantial number of items (from 10 to 15, depending on the arrangement of words).

Procedure. Subjects were tested on an individual basis with the stimuli presented one at a time on the cathode ray tube (CRT) of a PDP-11/40 minicomputer. Each word remained on the screen for 3 sec, during which time the subject was required to name it aloud. Subjects were told that there would be a memory test. Instructions about the test were given immediately after the last list item was presented.

On the test, the names of the categories were presented on the CRT at one of four experimental test rates. Within a critical interval, subjects were required to read the category name aloud and estimate the number of exemplars that had appeared on the preceding word list. Subjects were fully informed about the test rate. Any responses given after the next category name appeared on the screen were discounted.

Subjects. Subjects were 95 university undergraduates who were fulfilling an option for an introductory psychology course. Fifteen of the original

Table 1
Judgment Size (Means) and Accuracy (Mean Deviation) as a Function of the Number of Category Exemplars

No. exemplars	Mean judgments (Test rate)					Mean deviations (Test rate)				
	2	4	5	10	*M*	2	4	5	10	*M*
0	.55	.35	.40	.10	.35	.56	.35	.40	.12	.36
3	3.37	2.93	2.62	2.60	2.88	1.37	1.17	1.22	1.27	1.26
6	4.62	3.85	4.33	4.24	4.26	2.53	2.25	2.15	2.20	2.28
9	5.91	5.08	5.20	5.58	5.44	3.71	3.92	3.38	3.81	3.70
M	3.61	3.05	3.14	3.13	—	2.04	1.92	1.90	1.85	—

subjects were dropped from this pool due to an inability to follow instructions (e.g., naming the words as they were presented), to failure to keep pace either at presentation or at test, or to experimenter error.[1]

Results

Judgments of occurrence. A 4 (test rate) × 4 (number of exemplars) analysis of variance with repeated measures on the second factor was performed separately on two measures of frequency knowledge—mean and median judgments. Since the results from the two analyses were identical, only those for the mean estimates will be presented. All results reported as significant are at or beyond the .05 level.

As Table 1 indicates, judgments of the number of instances of a category increased as the number of exemplars increased, $F(3, 228) = 334.47$, $MS_e = 1.09$. In fact, each increase in actual occurrence rate resulted in significantly higher judgments (Newman-Keuls test). Adults are then sensitive to categorical-level frequency information when only exemplars occur.

The data further suggest that the information that mediates this decision might be stored prior to the time of the test. Test rate had minimal impact on performance ($F < 1$). The Test Rate × Number of Instances interaction was also not significant ($F = 1.08$). Thus, subjects who were given only 2 sec to read the category label aloud and to decide on and then produce a response performed no differently from those subjects who were given 10 sec to do the same tasks. This was the case even in the present situation in which the information needed to make the judgment was stored incidentally during stimulus presentation.

Accuracy of judgments. Two measures of the accuracy of judgments were calculated. One was the absolute number correct. The other was an unsigned deviation score based on the difference between actual and estimated numbers of instances. Analyses of variance were performed on both measures with identical results; only those for deviation scores will be presented (see Table 1).

As the number of exemplars from a given category increased, accuracy in recalling that information declined, $F(3, 228) = 168.32$, $MS_e = 1.05$. What is particularly important, however, is that the testing rate had no influence on accuracy scores (Fs < 1). Subjects who were forced to decide at a 2-sec rate were no less accurate than those who had up to 10 sec to make their decisions.

The pattern of judgments and of accuracy correspond to a Weber function, and so are in keeping with those reported elsewhere (e.g., Hasher & Chromiak, 1977) for judgments of explicitly presented items.

[1] Of the 15 subjects who were replaced, 6 were subjects in the 2-sec test-rate condition who failed to keep pace with the test. These subjects missed giving answers for from 1 to 5 of the 12 categories, with a mean miss rate of 2.5 categories. This differential drop-out rate would be serious if it could be shown that the eliminated subjects were less sensitive to frequency differences than were the retained subjects. To test for this, estimates given by the eliminated subjects to categories they responded to were compared to those given by the 20 subjects who were able to keep pace with the test. Analysis of variance showed a main effect for frequency with no tendency toward differences between the two groups of subjects in their sensitivity to frequency, $F(1, 24) = 1.49$, $MS_e = 7.40$. Thus, failure to keep pace seems to be attributable to factors other than differences in frequency sensitivity.

Discussion

Despite the incidental nature of acquisition of the information underlying performance on this frequency judgment task, and the fact that the categorized nature of the materials was not initially obvious, subjects' judgments were sensitive to actual differences in the numbers of category exemplars. Further, neither the value of those judgments nor their accuracy varied with test rate. Whether subjects were required to respond in less than 2 sec or less than 10, their judgments were the same. (Informal measurement showed that it takes subjects a minimum of .5 sec just to say the name of the category aloud.) It appears unlikely that subjects in the fastest pacing condition were using the category name, shown for the first time at test, to generate and count instances of the category that had occurred in the experimental context. Rather, it seems more plausible to assume that some internal representation of the category name had been tagged with frequency information during the presentation of exemplars, and that this representation was available to the subject to assign a judgment of magnitude on the test. Whether subjects tested at the slower rates engaged in generate-and-count strategies is unclear. What is certain is that if such a strategy were used, it did not result in performance that was superior to the direct estimation that we presumed to have occurred at the fastest rate in this study.

Evidence from the present study suggests that frequency tags are kept for superordinates of category instances even when those superordinates are referenced only implicitly. Further, the findings suggest that this information is stored without the deliberate intention of the subjects. The next experiment raised the question of the similarity of this mechanism to those that tabulate the frequency with which individual events occur.

Experiment 2

Whatever the knowledge is that underlies people's abilities to judge the frequency with which events occur, it does not seem to be aided by instructions to intentionally process that information (e.g., Howell, 1973). In fact, uninformed subjects ranging in age from kindergarten to late adulthood give judgments of frequency that are as sensitive to small differences as are those given by forewarned college students (Attig & Hasher, 1980; Hasher & Chromiak, 1977).

The present experiment compared the impact of intentional frequency instructions with incidental instructions when the task was to judge the frequency of implicitly referenced events, again, categories when only exemplars occurred. Since the literature on instructions raises the possibility of finding no effect, performance conditions were biased to maximize the usefulness of intentional instructions. We assumed that at least under optimal conditions, people will use instructions about a forthcoming frequency judgment test to try to keep a running tabulation during list presentation. To facilitate this, we varied the ease with which subjects could actually count exemplars during list presentation. In the easy condition, all of the exemplars of a single category appeared in succession (blocked presentation). In the more difficult condition, a random arrangement of exemplars was used. The deliberate invocation of a counting strategy induced by intentional instructions should increase the accuracy of judgments. Informed subjects should do particularly well under the blocked presentation sequence.

Method

Design. As in Experiment 1, each subject saw a list of exemplars of familiar categories. These varied in their number of instances (0, 3, 6, and 9). Half of the subjects were instructed about the nature of the frequency judgment task before presentation of the list, and half of the subjects were not so informed. Within each of these two instructional conditions, two basic presentation patterns were used: In the blocked presentation, all instances of each category occurred in series before the presentation of any instances of another category; in the random presentation, exemplars from different categories were intermingled. The design was than a 2 (instructions) × 2 (list structure) × 4 (number of exemplars) factorial with repeated measures on the last factor.

Materials. The pool of words from Experiment 1 was used again, with each presentation list containing 64 words, 54 of them members of categories and 10 serving as buffers.

The four lists from Experiment 1 were used for the random condition. Blocked lists were created by having the words from each category appear consecutively. The list was divided into thirds with one category at each frequency level occurring in each third. The order of the three frequency levels was different in each third of the list. These different orders were then rotated through each segment of the list to produce three different list orders. Across all subjects each category occurred in each third of a list. Finally, to ensure that each category also occurred at each of the four levels of the frequency variable, list order was combined with frequency level, yielding 12 lists.

Procedure. Subjects were tested individually with materials presented on the CRT of a PDP-11/40 minicomputer. The procedure was identical to that used in Experiment 1, with the exception that half of the subjects were now fully informed about the nature of the test prior to seeing the list. All subjects were given explicit test instructions after the last item was presented. Words were presented at a 3-sec rate, and judgments were paced at a 4-sec rate.

A total of 24 subjects were assigned to each cell of the design. In the blocked condition, each list was used twice. In the random condition, each list was used eight times.

Subjects. Subjects were 111 undergraduate students fulfilling a course option. The data from 96 of these subjects were used. The remaining subjects were eliminated because of inability to keep pace with stimulus presentation, with the testing procedure, because of failure to follow instructions, or because of experimenter error.

Results

Judgments of frequency. Analyses of variance were conducted on both mean and median estimates. These yielded similar results, and only those based on the mean estimates are presented (see Table 2).

Results showed significant main effects for frequency, $F(3, 276) = 482.25$, $MS_e = 1.40$, instructions, $F(1, 92) = 7.97$, $MS_e = 3.95$, and their interaction, $F(3, 276) = 5.85$. Newman-Keuls tests (at .01) revealed that both instructed and uninstructed subjects gave progressively higher estimates with each increase in the actual number of exemplars. The significant interaction stems from the instructed subjects giving higher estimates (5.34 and 7.06, respectively) for Category Sizes 6 and 9 than is the case for the uninstructed subjects (4.66 and 5.84).

No other effects approached significance, including those involving the order of presentation ($Fs < 1.29$). Viewing the instances of a category in consecutive order provided no particular advantage over viewing them interspersed with instances of other categories, and this was true even for instructed subjects.

Accuracy of judgments. To assess accuracy, the deviation measure reported for Experiment 1 is discussed here. The pattern of results seen for estimates was maintained here (in Table 2). Instructions, $F(1, 92) = 7.43$, $MS_e = 1.49$, frequency, $F(3, 276) = 174.78$, $MS_e = .77$, and their interaction, $F(3, 276) = 10.98$, were all significant. Newman-Keuls tests showed that accuracy declined with each increase in actual number of exemplars. The instructed subjects were significantly more accurate than the uninstructed subjects only for the nine-instance categories.

Again, the order in which category instances were presented had no impact on performance. Both the blocked and random conditions resulted in equally accurate per-

Table 2
Judgment Size (Means) and Accuracy (Mean Deviation) as a Function of Number of Category Exemplars, Instructions, and Presentation Order

	Mean judgments				Mean deviations			
	Instructed		Uninstructed		Instructed		Uninstructed	
No. exemplars	Blocked	Random	Blocked	Random	Blocked	Random	Blocked	Random
0	.44	.26	.18	.25	.44	.26	.18	.25
3	3.29	3.32	3.19	2.93	1.13	1.32	1.14	1.15
6	5.58	5.11	4.67	4.65	1.69	1.56	1.67	2.10
9	6.71	7.40	5.79	5.88	~~1.44~~	~~1.43~~	~~1.65~~	~~1.82~~
M	4.01	4.02	3.46	3.43	~~1.18~~	~~1.14~~	~~1.16~~	~~1.33~~

9 2.51 2.57 3.60 3.80
x̄ 1.44 1.43 1.65 1.82

formance, even though a counting strategy should have been far easier to employ in the former condition.

Discrimination coefficients. At this point, the interpretation of the differences between the instructed and uninstructed subjects is unclear. If explicit instructions led to the use of a counting strategy, then the absence of an Instruction × Item Sequence effect is puzzling. Counting seems far more difficult to do (and indeed proved impossible) when the instances of a single category span the entire length of the list than when they all occur in sequence.

It is possible that the instructional effects reflect differences in the criteria used by subjects to assign a value to a subjective magnitude rather than to differences in sensitivity to occurrence rates. To answer this question, a measure first introduced by Flexser and Bower (1975) was employed. This measure is the correlation between actual and estimated frequency, and is calculated for each subject. As Flexser and Bower pointed out, what results is a measure of a subject's ability to distinguish one frequency from another, rather than a measure of how subjects assign values to subjective magnitudes.

Discrimination coefficients were calculated for every subject in each of the four experimental conditions. The mean correlations for subjects in the blocked-instructed, blocked-uninstructed, random-instructed, and random-uninstructed conditions were .83, .83, .83, and .80, respectively (all significant at .05). These discrimination coefficients are strikingly high and are virtually identical to those reported elsewhere for overtly referenced events (Flexser & Bower, 1975; Hasher & Zacks, 1979). A 2 (list type) × 2 (instructions) analysis of variance failed to detect any differences ($Fs < 1$).

Although the Flexser and Bower (1975) correlation is very sensitive to violations in the rank ordering assigned by subjects to items of different true frequencies, it is rather insensitive to differences in sensitivity among subjects, all of whom may have given the same or similar rank ordering but different raw estimates. To assess this possible difference in sensitivity, a measure that simulates a forced-choice score was calculated. The measure was based on the Estes (1976) scanning model, and was calculated for each subject by summing together the probabilities of choosing the more frequent of two categories of different sizes (based on the actual frequency judgments assigned to the categories) had there actually been a two-alternative forced-choice procedure. Values representing the probability of choosing the more frequent category were obtained for each of the six possible pairs of category-instance occurrence rates (i.e., 0–3, 0–6, 0–9, 3–6, 3–9, 6–9). The mean of these values was taken for each subject, and represents the probability of choosing the more frequent category given all possible alternatives in a forced-choice procedure. For any given pair of actual frequency levels, the derived probability of choosing the larger frequency in a forced-choice procedure would be

$$P_{ij} = \sum (\Pi_{is}\Pi_{jt}) + .5 \sum (\Pi_{is}\Pi_{js}),$$

where P_{ij} is the probability that a stimulus presented i times will be chosen as more frequent than one presented j times; Π_{is} is the probability that a stimulus presented i times will be perceived as having been presented s times; Π_{jt} is the probability that a stimulus presented j times will be perceived as having been presented t times; Π_{js} represents the probability that a stimulus presented j times will be perceived as having been presented s times. The variables i and j take on the values of any two different actual category sizes, and variables s and t take on the values of the estimates given for those two category sizes, respectively, with the constraint that $0 \leq t < s \leq$ the largest judgment given. The second half of the equation represents a correction for guessing, and occurs whenever a subject assigns identical estimates to two different categories of different actual sizes.

The mean values for subjects in the blocked-instructed, blocked-uninstructed, random-instructed, and random-uninstructed conditions were .89, .88, .91, and .87, respectively. The values suggest that subjects would have shown excellent performance in a forced-choice task. An analysis of variance revealed no significant differences among the conditions.

Thus, two analyses show that neither instructions nor item order has an influence on a person's ability to make discriminations among categories that differ in the number of instances by which they are represented.

Discussion

A basic result of this experiment is the replication of the first: Adults have knowledge of the frequency of occurrence of higher order concepts even when their experience is with the subordinate instances of those concepts. However, the final conclusion one draws depends on the appropriate interpretation of the instructional effects found here. Is the frequency sensitivity of instructed subjects actually greater than that of uninstructed subjects? Disregarding for the moment the directly contradictory findings from the discrimination coefficient analysis, how might this be so?

The most obvious advantage of instructions about a forthcoming frequency test would result if subjects used this information during instance presentation to keep a running tabulation of category size. Two sets of facts argue against such a view. First, a counting strategy should be far easier to deploy under a blocked presentation sequence than under a random sequence. However, neither the list sequence nor the Sequence × Instruction interaction even bordered on significance for any of the dependent measures.

Second, the results of a follow-up experiment appear to directly contradict the counting hypothesis. Six university students were given one of the random lists used in Experiment 2. They were informed of the categorical nature of the words in the list and were asked only to report verbally, as each word was presented, a running tally of the size of its category. For all six subjects, the task proved to be impossible well before a third of the list had elapsed. This was before any category's full size could be known. In fact, at this point in the list, a maximum of three items from a particular category had been presented. In postexperimental interviews, subjects reported that the best they could do was to estimate the size of most if not all of the categories.[2] It is, of course, doubtful that subjects in the random condition of the main experiment were any more able or likely to count than were the subjects in our follow-up experiment. Given the equivalent performance of random and blocked instructed subjects in the main experiment, the present finding raises the question of whether even the blocked condition subjects effectively used a counting strategy.

Thus the difference found in the experiment between instructed and uninstructed subjects may well be tied to differences in the criteria used by subjects when they assigned numbers to their subjective magnitudes. Instructed subjects were more likely to give larger numbers than were uninstructed subjects. The knowledge that larger numbers were appropriate might have stemmed from at least two sources: (a) if some of the instructed subjects were able to keep a relatively accurate count of one of the larger categories throughout the entire list or (b) if subjects abandoned the strategy of counting after reaching a count of three or four and then realized that there were still more instances of that category. In either case, the result would be the impression that correct estimates could range into values more extreme than four. Thus the larger (and so more accurate) numbers given by the instructed subjects at the highest frequency level would be due to a criterion difference rather than to increased sensitivity to frequency. This conclusion is buttressed by the findings from the discrimination coefficient analysis and the simulated forced-choice analysis, which are based on measures devised to eliminate

[2] In other unpublished work, subjects were given lists containing repeated instances of the same words distributed in a long list and were informed of the frequency estimation task that would follow the end of the list. Of the 24 subjects, only 8 reported specific employment of a counting strategy. Most of the 8 subjects who tried counting spontaneously said that it was useful only for the beginning part of the list.

the contribution of response bias differences. Instructional differences were not seen with these measures.

General Discussion

These results extend previous findings that demonstrate sensitivity to frequency of occurrence of both explicit (cf. Hasher & Zacks, 1979) and implicit (cf. Johnson, Taylor, & Raye, 1977) repetitions of individual events themselves. Further, sensitivity to differences in the frequency of implicitly referenced superordinates appears to be equivalent whether subjects are forewarned or not that this is the test-relevant aspect of their task. This equivalence is also found under presentation sequences which differ in the ease with which frequency differences are apparent. These findings, together with the suggestion that the information mediating the frequency judgment is encoded at presentation, suggest that superordinate frequency is automatically encoded. This encoding would then be expected to extend to events outside of the laboratory, and indeed there is some empirical evidence that in everyday life, people mark superordinate frequency when exposed to individual exemplars: Adults show knowledge of the category size of various professions (Lichtenstein et al., 1978).

One of the major cognitive functions of automatically encoded information is to guide the retrieval process (Hasher & Zacks, 1979). Recent evidence suggests the importance of a word's superordinate in long-term retention (Nelson et al., 1979). The automatic accumulation of frequency information of both events and their superordinates has important implications for social behavior as well. Indeed, this information may underlie the establishment of a network of expectations and tacit assumptions that are of fundamental importance in guiding behavior in public places (Goffman, 1959), and memory for that information (Schank & Abelson, 1977). Such information may also lead to the creation and utilization of assumptions about people grouped according to a variety of criteria (e.g., age, religion, style of dress).

References

Attig, M., & Hasher, L. The processing of frequency of occurrence information by adults. *Journal of Gerontology*, 1980, *35*, 66–69.

Attneave, F. Psychological probability as a function of experienced frequency. *Journal of Experimental Psychology*, 1953, *46*, 81–86.

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*(3, Pt. 2).

Carroll, J. B. Measurement properties of subjective magnitude estimates of word frequency. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 722–729.

Estes, W. K. The cognitive side of probability learning. *Psychological Review*, 1976, *83*, 37–64.

Faulkender, P. J., Wright, J. C., & Waldron, A. Generalized habituation of concept stimuli in toddlers. *Child Development*, 1974, *45*, 1002–1010.

Flexser, A. J., & Bower, G. H. Further evidence regarding instructional effects on frequency judgments. *Bulletin of the Psychonomic Society*, 1975, *6*, 321–324.

Goffman, E. *The presentation of self in everyday life*. Garden City, N.Y.: Doubleday, 1959.

Gude, C., & Zechmeister, E. G. Frequency judgments for the "gist" of sentences. *American Journal of Psychology*, 1975, *88*, 385–396.

Hasher, L., & Chromiak, W. The processing of frequency information: An automatic mechanism? *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 173–184.

Hasher, L., & Zacks, R. T. Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 1979, *108*, 356–388.

Hintzman, D. L., & Stern, L. D. Contextual variability and memory for frequency. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, *4*, 539–549.

Howell, W. C. Storage of events and event frequency: A comparison of two paradigms in memory. *Journal of Experimental Psychology*, 1973, *98*, 260–263.

Jacoby, L. L. The role of mental contiguity in memory: Registration and retrieval effects. *Journal of Verbal Learning and Verbal Behavior*, 1974, *13*, 483–496.

Johnson, M. K., Taylor, T. H., & Raye, C. L. Fact and fantasy: The effects of internally generated events on the apparent frequency of externally generated events. *Memory & Cognition*, 1977, *5*, 116–122.

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, *4*, 551–578.

Lindsay, P. H., & Norman, D. A. *Human information processing*. New York: Academic Press, 1977.

Nelson, T. O., Fehling, M. R., & Moore-Glascock, J. The nature of semantic savings for items forgotten from long-term memory. *Journal of Experimental Psychology: General*, 1979, *108*, 225–250.

Schank, R. C., & Abelson, R. P. *Scripts, plans, goals, and understanding*. Hillsdale, N.J.: Erlbaum, 1977.

Shapiro, B. J. The subjective estimate of relative word frequency. *Journal of Verbal Learning and Verbal Behavior*, 1969, *8*, 248–251.

Underwood, B. J. Recognition memory. In H. H. Kendler & J. T. Spence (Eds.), *Essays in neo-behaviorism*. New York: Appleton-Century-Crofts, 1971.

Received September 20, 1979
Revision received February 11, 1980 ■