

# Frequency Judgments: The Problem of Defining a Perceptual Event

Marcia K. Johnson  
Princeton University

Mary A. Peterson, Evan Chua Yap, and Patricia M. Rose  
State University of New York at Stony Brook

In four experiments the conditions under which frequency judgments reflect the relative frequency of complex perceptual events were explored. Subjects viewed a series of  $4 \times 4$  grids each containing seven items, which were letters and numbers in one of four typefaces. Later judgments of the relative frequency with which particular letters appeared in particular typefaces were unaffected by a warning about an upcoming frequency judgment task, but were affected by both the time available for processing the stimuli and the nature of the cover task subjects engaged in while viewing the grids. Frequency judgments were poor when exposure durations were less than 2 s and when the cover task directed subjects' attention merely to the locations of the items within the grids. Frequency judgments improved when the cover task directed subjects' attention to the identity of the stimuli, especially to the conjunction of letter and typeface. The results suggest that frequency estimation of complex stimuli may be possible only for stimuli that have been processed as phenomenal objects.

Largely prompted by Hasher and Zacks's (1979, 1984) proposal that certain attributes of events are automatically encoded, there has been considerable recent interest in people's ability to judge the relative frequency of events (e.g., Barsalou & Ross, 1986; Begg, Maxwell, Mitterer, & Harris, 1986; Birnbaum, Taylor, Johnson, & Raye, 1987; Fisk & Schneider, 1984; Greene, 1984, 1986; Hanson & Hirst, 1988; Kausler, Lichty, & Hakami, 1984; Maki & Ostby, 1987; Naveh-Benjamin & Jonides, 1986; Williams & Durso, 1986). The available studies collectively highlight a number of important aspects of frequency judgments that are not yet well understood. One of the most central is the relation between encoding processes and sensitivity to differential frequency. The present experiments were concerned with this question. Specifically, we were interested in determining the minimal processing conditions necessary for discrimination of relative frequency.

It has been established that sensitivity to frequency is affected by the type of orienting task subjects engage in (e.g., Hanson & Hirst, 1988; Jonides & Naveh-Benjamin, 1987; Maki & Ostby, 1987; Rowe, 1974). For example, Jonides and Naveh-Benjamin found that instructions to produce semantic

associations to words led to more accurate frequency estimates for the words than instructions to produce acoustic associates. However, as the authors note, frequency estimates were reasonably accurate, even under the conditions yielding the poorest relative performance (Jonides & Naveh-Benjamin, 1987, p. 239). Two exceptions to this robustness of frequency estimates were reported by Fisk and Schneider (1984). First, subjects showed no sensitivity to the differential frequencies of foveally presented words in conditions in which their attention was focused elsewhere in the visual field, that is, on target digits in the corners of the array (Fisk & Schneider, 1984, Experiment 1). Second, subjects showed no sensitivity to the differential frequencies of words that were presented as distractors in a highly practiced category search task in which subjects searched among these distractors for six vehicle items after hundreds of trials of practice (Fisk & Schneider, 1984, Experiment 2). These exceptions show that some types of encoding processes do not support accurate memory for frequency of events.

Hasher and Zacks (1979) suggested that a minimum requirement for frequency judgments is attention to the stimulus (p. 359), and Zacks, Hasher, and Hock (1986) further proposed that conscious attention may be required. Fisk and Schneider's (1984) conditions do not meet these "boundary conditions" for frequency judgments, namely, conscious or attended processing (Zacks et al., 1986).

It is not entirely clear, however, what it is about a stimulus that must be attended to or consciously processed in order to create memory records that support reliable judgments of frequency. For example, Hock, Malcus, and Hasher (1986) suggested that when subjects process letter strings (e.g., AWUP), global-level (string) and element level (letter) memory units are independently formed and perhaps tagged for frequency. The implication is that subjects can attend to one aspect of a stimulus and yet store information relevant to judging the frequency of another aspect of the stimulus (also

---

This research was supported by National Science Foundation Grant BNS-8510633 to Marcia K. Johnson and by a Public Health Service BRSG award to Mary A. Peterson. The article was completed while Marcia Johnson was a fellow at the Center for Advanced Study in the Behavioral Sciences, Stanford, California, with financial support provided by the John D. and Catherine T. MacArthur Foundation and Princeton University.

We would like to thank Stanley Thompson for his help in making the stimulus materials and testing pilot subjects and Atsuko Watanabe for her help in analyzing the results of Experiment 4. Mary Peterson is now at the University of Arizona, Tucson.

Correspondence concerning this article should be addressed to Marcia K. Johnson, Department of Psychology, Princeton University, Princeton, New Jersey 08544.

see Barsalou & Ross 1986; Hanson & Hirst, 1988). Further complicating the picture are findings suggesting that if an item's membership in a particular category (e.g., the category "red things") is not made salient at the time the stimulus (e.g., ketchup) is attended to, subjects may not later be able to judge the relative frequency of categories (Barsalou & Ross, 1986; Freund & Hasher, 1986). Thus, overall, it remains to be specified which aspects of a stimulus (e.g., its "identity," its subordinate elements, or its superordinate properties) must be attended to in order to support accurate frequency judgments of that aspect and which must only be activated. To clarify this issue, we must better specify either the meaning of "stimulus" or the meaning of "attention."

### Defining Perceptual Events

The ability to judge the relative frequency of stimuli or events is likely to depend on the role that various types of cognitive processes play in defining the events that are initially experienced and encoded. That is, it is probably possible to attend to a physical stimulus, especially a complex physical stimulus, for purposes that do not require that all aspects of that stimulus be identified or experienced. In most previous work on frequency judgments, simple events were used; for example, words were presented individually, affording little ambiguity about what constituted an event. In contrast, stimuli in the world often occur in a crowded perceptual field in which events are defined in terms of the purposes of the perceiver. In addition, relevant events outside the laboratory are likely to involve the conjunction of more simple attributes, and the particular attributes that are conjoined are likely to depend on the perceiver's purpose. Therefore, any general theoretical statements about frequency judgments should take into account the purposes of the viewer, which may affect both the attention paid to the stimulus and the nature of the effective stimulus. In the present experiments, we began to specify those conditions that do and do not later yield successful judgments of the relative frequency of complex events in a crowded visual field. This information, in turn, should imply something about the nature of the memory representations that are the basis of frequency judgments.

The idea that different purposes of the viewer recruit different encoding processes with different consequences for frequency judgments is consistent with Johnson's (1983) Multiple Entry, Modular Memory System (MEM). According to MEM, encoding is accomplished by at least three distinguishable processing subsystems: the sensory, perceptual, and reflection subsystems. The sensory system is critical for such things as detecting the location of stimuli, resolving figure-ground relations, developing visual scanning habits, and other sensory-motor skills. The sensory system deals with perceptual information that has cue value and can enter into learned relations but is not itself ordinarily the object of perception. The objects of perception are usually conjunctions of information, such as size, shape, and location, into specific objects and events that are organized with respect to each other and that produce patterns that later can be recognized as familiar. That is, the perceptual system deals with the information that typically constitutes our phenomenal perceptual experience.

The reflection system deals with information generated by the subject, such as that created during planning, organizing, or relating a present event to a past event. In other words, compared with sensory and perceptual processes, reflection depends less on immediate perceptual data.

According to this model, records are produced in all subsystems in an obligatory fashion as inevitable consequences of the corresponding processes. Memory tasks differ in the extent to which they draw on sensory, perceptual, and reflection records. For example, certain tasks measuring increased perceptual skills as a function of practice (e.g., Cohen & Squire, 1980; Lewicki, 1986) may draw primarily on sensory records, whereas tasks measuring object recognition are more likely to involve the perceptual (and perhaps the reflection) subsystems (Johnson, 1983, in press). Assuming there is some commonality between processes underlying recognition and those underlying frequency judgments (e.g., Underwood, 1983, p. 106), a straightforward extension of the MEM model is that judgments of the frequency with which particular objects have been experienced should, like recognition judgments, draw on records created by perceptual subsystem processes. Reflective records may contribute as well (as they do in recognition), but sensory records alone should be a poor basis for frequency judgments concerning particular objects.

To examine the conjecture that perceptual processing is necessary for frequency judgments, we had subjects engage in different tasks while viewing a series of briefly exposed  $4 \times 4$  grids that contained seven letters and numbers in various typefaces (as shown in Figure 1). Later, we asked subjects to rank order how often letters appeared in particular typefaces. Three cover tasks were designed to vary in the involvement of the proposed critical perceptual subsystem processes. In one cover task, the detecting-locations task, subjects were directed to identify which cells were filled. Location can be specified on the basis of processing that does not depend on a viewer's experiencing particular objects in the locations specified (e.g., figure-ground information). Thus, this task was expected to draw primarily on the sensory system. In a second cover task, the detecting-letter-locations task, subjects were directed to identify which cells contained letters as opposed to numbers. This task was expected to draw on the perceptual system to the extent that it requires some categorization of the grid items. During this categorization process, letter information should be experienced, and perhaps typeface information as well (depending on whether or not processing typeface is necessary for processing letter shape). In a third task, the looking-at-objects task, subjects looked at the grids without a cover task (Experiment 1) or with a cover task that specifically directed their attention to the fact that letters would occur in different typefaces (Experiment 4). The looking-at-objects cover task was expected to involve the kind of conjunction of perceptual attributes that results in phenomenal perception of complex objects or events. Subjects in the looking-at-objects condition should, therefore, be most likely to integrate letter and typeface information into percepts of particular, complex objects, either spontaneously as a consequence of normal looking habits (Experiment 1), or as a consequence of the cover task demands (Experiment 4). If the accuracy of later judgments about the frequency of occur-

rence of letters in particular typefaces depends on the nature of the records established at input, the looking-at-objects condition would be expected to yield the most accurate frequency judgments and the detecting-locations task the least accurate judgments. The detecting-letter-locations condition provided an index of the degree to which letter and typeface information is incidentally conjoined when the task demands require only categorization of letters.

### Exposure Duration

In addition to manipulating the nature of the cover task, we also investigated the effects of the duration of stimulus exposure in Experiments 1–3. We assumed that the processes that produce perceptual records take more time than the processes that produce sensory records (although both types of processes may begin at the same time). There is growing evidence that object or event perception is a result of complex processing that takes both resources and time (Mewhort, Campbell, Marchetti, & Campbell, 1981; Treisman & Gelade, 1980). For example, following Mewhort et al.'s (1981) model, information contained in an early, preidentification "feature" buffer might support the detecting-locations task, whereas detecting the location of letters (or looking at particular letters in particular typefaces) might require information in a later postidentification "character" buffer. Although MEM is not a stage model (the input to the perceptual subsystem is not the output of the sensory subsystem), it seems reasonable to assume that the longer the exposure duration, the more likely it is that featural information has been identified as specific letters in specific typefaces before other tasks (e.g., marking the locations) or other stimuli (due to uncontrolled eye movements, for example) interfere with perceptual processing.

### The Role of Intention

In Experiments 2 and 3 we also investigated the effects of warning the subjects about the upcoming frequency task. On the basis of arguments such as those made by Postman (1964) and Greene (1986), we did not expect a warning to help subjects who viewed the grids in either the detecting-locations task or the detecting-letter-locations tasks: Unembellished intention instructions that do not specifically suggest to subjects a type of processing relevant to the task may have little effect on frequency judgments (Greene, 1986). In addition, these tasks are quite demanding and we did not expect subjects to be able to engage in much additional processing beyond that demanded by the cover tasks.

### Experiment 1

In Experiment 1, we used the first version of the looking-at-objects condition. In this condition, observers who were warned about the upcoming frequency test viewed the stimuli without a cover task. We expected that the individual letters would be experienced as phenomenal objects in this condition and, therefore, that viewers would later be able to judge their relative frequency of occurrence. This condition is similar to

conditions in which reliable frequency judgments of simple items have been demonstrated (e.g., Hasher & Zacks, 1984). Nevertheless, it was not obvious that subjects would be able to accurately judge the relative frequency of various conjunctions of letter and typeface, especially when the visual field is crowded with different items occupying 7 of a possible 16 locations. Accordingly, we examined whether accurate relative frequency judgments regarding complex objects could be made under these conditions. In addition, we explored whether some minimum amount of processing time is necessary by using three conditions of exposure duration: 50 ms, 1 s, and 2 s.

### Method

**Subjects.** The participants in this experiment were 46 students in an undergraduate research methodology course at the State University of New York at Stony Brook. All had vision that was normal or corrected to normal and were naive as to the purpose of the experiment.

**Stimulus materials and apparatus.** The presentation stimuli were 128  $4 \times 4$  grids containing letters and numbers in one of four typefaces: Computer, Boldface, Old English, or Script (see Figure 1A). One typeface was used per grid, and there were 32 grids in each typeface. In each grid, seven randomly chosen cells contained letters or numbers. Of the 128 grids, 32 contained only letters, and 95 contained six letters and one digit (because of an error in making up the stimuli, one grid contained six digits and one letter). The digits were randomly selected from single digits, excluding 0. Sixteen letters (C, E, G, H, J, K, L, M, O, P, Q, S, U, V, Y, and Z) appeared either

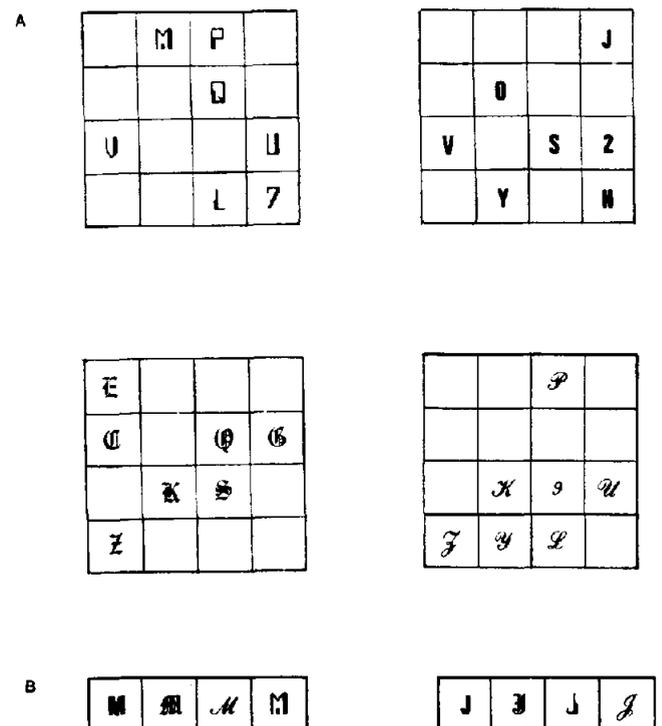


Figure 1. Four examples of the stimulus grids (A) and two examples of the test items (B) used in Experiments 1–4.

2, 8, 16, or 24 times in each of the four typefaces, so that four letters appeared with each frequency in each typeface. We used a larger frequency of exposure range (2 to 24 times) than is typical because we expected that, under some conditions, frequency discrimination would not be above chance. Thus, to increase the sensitivity of our design, we increased the range of exposure frequency. The greater the frequency differential that is not discriminated, the more persuasive null results should be. Particular letters were assigned to cells in the grid randomly, with the constraint that letters of various frequency of usage in English were evenly distributed across both presentation frequency and the grid rows and columns. Once a letter was assigned to a grid position, the various versions of the letter could appear in that position only.

The test stimuli were 16  $1 \times 4$  grids, each containing the four versions of each of the aforementioned letters (see Figure 1B). The order of the four different typefaces in each test stimulus was randomized with the constraint that letters that had occurred 2, 8, 16, and 24 times in the presentation stimuli occurred with equal probability in each location of the test grids.

The stimuli were presented as slides via a Kodak carousel projector. An electronic shutter mounted on the projector lens and controlled by a Gerbrands timer (Model 300-6T) controlled the duration of exposure.

*Procedure.* Subjects were tested in three groups; each group was tested by a different experimenter who was naive as to the purpose of the experiment. One group saw the test stimuli for 50 ms ( $n = 18$ ), another for 1 s ( $n = 16$ ), and the third for 2 s ( $n = 12$ ). With only seven stimuli in each grid, and no masking stimulus, even the shortest stimulus duration (50 ms) was sufficient to produce a brief impression of a matrix of letters. Such exposures typically produce correct immediate reports of a substantial number of letters when they are tested individually by location cuing (Sperling, 1960; Coltheart, 1980).

Subjects in all groups were shown a sample grid and were told that they would be viewing a series of  $4 \times 4$  grids containing symbols in different typefaces. Next, in order to acquaint subjects with the different typefaces, we showed four single exposures of the letter *A*, one exposure in each typeface, followed by four typeface versions of the letter *R* (neither *A* nor *R* was used as a stimulus in the experiment). Subjects were told that each symbol would have four versions like the ones they had just seen. They were told to pay attention to the typeface in which each symbol appeared in order to form a rough idea of which version of a symbol appeared more often than other versions of the same symbol. Subjects were told that after seeing 128 different grids, they would be asked to rank the four versions of a given symbol according to how frequently each had occurred. They were told that they would see each grid for a short duration only, and they could increase the probability that they would see all the symbols by keeping their eyes fixed on the center of the screen.

Subjects viewed the 128 presentation stimuli in random order. Regardless of exposure duration, the interstimulus interval (ISI) was 2 s. Immediately after the last slide, subjects were given data sheets on which 16 empty  $1 \times 4$  grids were drawn. Next, they were shown the test stimuli (in random order) and were asked to use the numbers 1 through 4 in the empty boxes of each grid to rank order the relative frequency with which each version of the test letter had appeared during the experimental trials: 1 represented the least frequent and 4 the most frequent. We had subjects rank order relative frequency rather than make frequency estimates because this was expected to provide a more sensitive test of subjects' ability to discriminate relative frequency, especially given the difficulty of the cover tasks to be introduced in the following experiments.<sup>1</sup> Subjects were told that they had seen each typeface equally often across the 128 grids and that for each letter, any typeface could have been the most frequent. Thus, rather than making a judgment based on how often they thought they saw a typeface overall, they were to make a separate

judgment for each letter. Subjects were asked to trust their first impressions and to do as well as they could, despite the difficulty of the task.

## Results

As shown in Figure 2, when subjects viewed these stimuli with no cover task and with a warning about the upcoming frequency test, their mean frequency rankings increased as presentation frequency increased,  $F(3, 129) = 14.41$ ,  $MS_e = 1.33$ ,  $p < .001$ , and there was a marginal Frequency X Exposure Duration interaction,  $F(6, 129) = 1.99$ ,  $MS_e = .18$ ,  $p < .07$ . Given that this interaction was predicted, subsequent planned analyses were conducted.

A one-factor analysis of variance (ANOVA) was computed within each exposure duration, and where the main effect of presentation frequency was significant, follow-up comparisons of individual pairs of frequency values were made in order to further specify the nature of the discrimination reflected in the main effect. Each comparison was made at the .05 alpha level. A correction for multiple comparisons seemed unnecessary because the comparisons were planned in advance and implied by the design. Furthermore, because we were looking for the conditions in which frequency judgments would not be possible as well as those in which they would, it seemed appropriate to guard against Type II errors as well as Type I errors (Keppel, 1973, pp. 153–155). This basic format was used for subsequent analyses in all four experiments.

The results of the subsequent analyses were as follows: In the 50-ms condition, letters shown 16 or 24 times were ranked higher than those shown 2 times ( $ps < .05$ ). Although the 1-s condition looked very much like the 50-ms condition (see Figure 2), the main effect of frequency was not significant in the 1-s condition ( $p > .10$ ). In the 2-s condition, letters shown 24 times were ranked higher than those shown 2, 8, or 16 times, and those shown 16 times were ranked higher than those shown 2 times ( $p < .01$ ).

## Discussion

The results of Experiment 1 show that frequency judgments can be obtained with complex objects in a crowded visual field when viewers have been warned about the frequency test. Further, these results suggest that exposure duration influences memory for frequency information: Although there was some evidence that subjects could discriminate 2 from 16 or more presentations after 50-ms exposures, the effect was not significant in the 1-s condition. It is notable that even with intention instructions, 2-s exposures (or approximately 285 ms/item) were required for more systematic frequency discrimination under these conditions. Overall, performance

<sup>1</sup> Given that subjects rank ordered the stimuli from 1 to 4 in each test set, the overall mean for each subject had to be 2.5. Therefore, exposure duration (Experiments 1–3), warning (Experiments 2 and 3), and cover task (Experiment 4) could not result in main effects; the potential interactions of these variables with presentation frequency were of primary interest in these experiments.

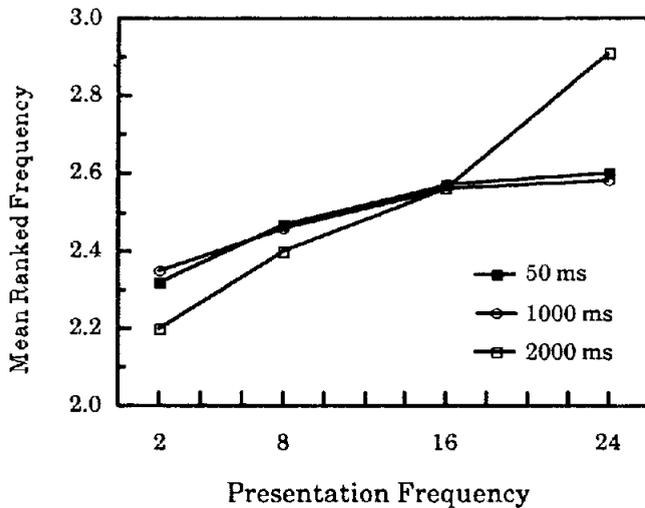


Figure 2. Mean ranked frequency as a function of presentation frequency and exposure duration (all conditions warned) in Experiment 1.

on the frequency test suggests that a 2-s exposure condition is near the boundary for frequency judgments using these stimuli.

### Experiment 2

According to the MEM model, increasing exposure time should be effective only if it results in the creation of perceptual, rather than merely sensory, records. Consequently, in Experiment 2 we varied exposure time and introduced the detecting-locations cover task in order to reduce the chances that subjects would process the targets as particular objects. Subjects were told that we were interested in memory for location; immediately after viewing each grid, they indicated which cell locations had been filled by marking a blank grid. In this task, subjects had to attend to the nominal stimuli (unlike Fisk & Schneider, 1984, Experiment 1, in which the subject's attention was focused elsewhere in the visual field), but the task could be accomplished without object identification. The detecting-locations task was relatively difficult, especially at short exposure durations, so there was little incentive to engage in unnecessary processing. Thus, we were able to examine frequency judgments under conditions in which viewers had to attend to the symbols in the grids but did not have to process those symbols as phenomenal objects. If accurate frequency judgments require perceptual records, sensitivity to relative frequency should be poor. Of course, if object processing is an automatic consequence of simply attending to stimuli (or if object processing is not necessary for frequency judgments), then the detecting-locations task may not reduce sensitivity to relative frequency. In addition, we were interested in whether subjects who were performing the locations task could elect to perform processing appropriate for frequency performance, if warned that it would be of consequence. Accordingly, we examined the effect of intention by varying the presence or absence of a warning about an upcoming frequency test.

### Method

**Subjects.** There were 70 subjects from the same research methodology class as the subjects in Experiment 1 (none had participated in Experiment 1). All had vision that was normal or corrected to normal and were naive as to the purpose of the experiment.

**Stimulus and apparatus.** The presentation and test stimuli and the apparatus were the same as those used in Experiment 1. In addition, booklets containing 128 blank grids were used for recording the filled grid locations after each exposure.

**Procedure.** There were four groups of subjects: Two groups were warned about the frequency rating task and two were not. One of the warned groups viewed the slides for 50 ms ( $n = 14$ ); the other for 2 s ( $n = 16$ ). Likewise, one of the no-warning groups viewed the slides for 50 ms ( $n = 21$ ); the other for 2 s ( $n = 19$ ).

All groups were told that we were interested in testing memory for locations. Subjects were shown the sample grid used in Experiment 1; "locations" were defined as squares in the  $4 \times 4$  grids. For each grid, subjects were asked to note all the locations that contained symbols and, after each stimulus presentation, to mark the blank response grids with slashes to indicate the filled locations.

All subjects were shown the four versions of letters *A* and *R* shown in Experiment 1. Those in the no-warning condition were told that they need not report which particular letter or digit they saw, or whether they saw a letter or a digit at all, but only the filled locations. Subjects in the warning condition were told that we were asking two questions in addition to the location question: First, Could they tell us which version of a symbol shown in various typefaces occurred more often relative to the other versions of the same symbol? and second, Could they acquire this information even though they had another task to do at the same time (i.e., note locations in order to fill in the blank grids)? Accordingly, when presented with a grid they were told to look to see which squares were filled with symbols so that they could mark those locations on their blank grids, and they were also asked to notice the typeface in which the symbols occurred. They were required, however, to indicate only location when they marked the blank grids.

All groups participated in four practice trials, and feedback was given regarding performance on the location task. During practice and experimental trials, the ISI was 6 s, in order to allow time to fill in the blank grids. Otherwise, the testing procedure was the same as in Experiment 1. Four experimenters, naive as to the purpose of the experiment, conducted the four groups.

### Results

The data are shown in Figure 3, collapsed across the warning variable, which had no effect. As is clear from Figure 3, with the detecting-locations cover task used in Experiment 2, we found no effects of frequency ( $F < 1$ ); nor were there any effects of duration. A subsequent replication of this experiment with 64 subjects and different experimenters also failed to produce systematic increases in ranked frequency with increases in exposure frequency.

### Discussion

These results suggest that processes necessary for later frequency discrimination are not engaged by a task that merely requires detecting the locations of symbols. This was true regardless of exposure duration and regardless of the presence or absence of a warning. The fact that the warning was not

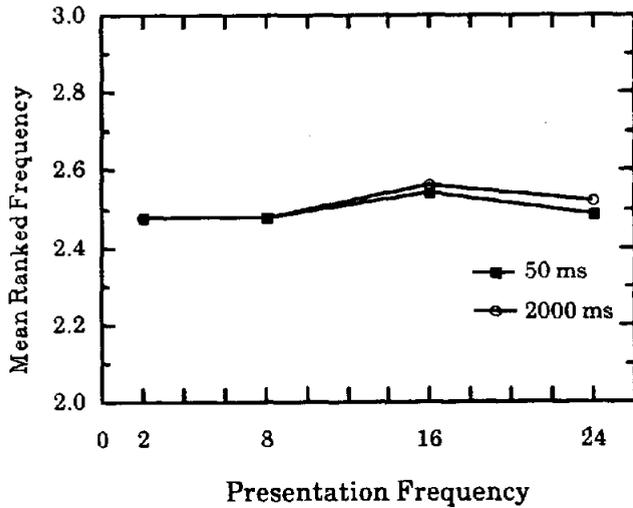


Figure 3. Mean ranked frequency as a function of presentation frequency and exposure duration (collapsed across warning) for the detecting-locations cover task conditions in Experiment 2.

effective might imply that phenomenal object status is necessary but not sufficient for better-than-chance frequency judgments. Alternatively, even with the warning, the items might not have become phenomenal objects because the demands of the location task were too great. If subjects first attended to locations (because this was the primary task) there may not have been enough time left for fully processing object identity (the outcome of Experiment 1 indicates 2 s is near the boundary even with no cover task).

### Experiment 3

The detecting-locations cover task used in Experiment 2 was designed to draw primarily on the sensory subsystem in MEM. To perform this task, subjects did not need to identify the stimuli in the grid locations, which, according to MEM, would require the involvement of the perceptual subsystem. These results suggest that frequency judgments (at least those requiring the conjunction of typeface and letter information) must be based on perceptual records. Of course, there were other differences between Experiments 1 and 2 besides the nature of the processing subjects engaged in while looking at the slides. One difference is that longer ISIs were used in Experiment 2. A second difference is that subjects were engaged in an activity during the ISI in Experiment 2 but not in Experiment 1. Accordingly, in Experiment 3 we introduced a cover task that specifically directed viewers' attention to the letters as particular objects: We asked subjects to note the locations of the letters only (detecting-letter-locations task). Better-than-chance frequency judgments with the detecting-letter-locations cover task would be consistent with the general idea that perceptual subsystem processes are required for frequency judgments. Thus, we were able to examine whether some of the difference between outcomes of Experiments 1 and 2 was due to the nature of stimulus processing and the resultant memory records and not merely the presence of a

cover task. Note that this cover task was expected to be more difficult than the cover task used in Experiment 2, inasmuch as subjects had to segregate letters from numbers. Therefore, if frequency judgments obtained in Experiment 3 were better than those obtained in Experiment 2, the improvement could not be attributed to an easier cover task.

### Method

The subjects were 63 undergraduates at the State University of New York at Stony Brook who had enrolled in the research methodology course for the following semester. All were naive as to the purpose of the experiment, and all had vision that was normal or corrected to normal.

We continued to explore the effects of warning and exposure duration in Experiment 3. Four groups were tested: 50-ms exposure and no warning ( $n = 13$ ); 50-ms exposure with warning ( $n = 17$ ), 2-s exposure and no warning ( $n = 15$ ), and 2-s exposure with warning ( $n = 18$ ).

The stimuli and apparatus were the same as those used in Experiments 1 and 2. The 128 slides were divided in half. Half the subjects saw slides 1-64 first, the other half saw slides 65-128 first. The procedure was generally the same as that used in Experiment 2, except that subjects were asked to note the locations of the letters only: They were told to leave blank on their response grid any cell corresponding either to a blank cell or to a cell that contained a digit during the stimulus presentation. Four different experimenters conducted the groups; all were naive as to the experimental purpose.

### Results

Using the detecting-letter-locations cover task, which specifically directed subjects' attention to the letters, we obtained a main effect of frequency,  $F(3, 177) = 13.81$ ,  $MS_e = 1.07$ ,  $p < .001$ , and a Frequency  $\times$  Duration interaction,  $F(3, 177) = 3.18$ ,  $MS_e = .25$ ,  $p < .025$ , both of which are shown in Figure 4. The data are collapsed across warning, which had no effect. Subsequent analyses showed that the ranks assigned in the

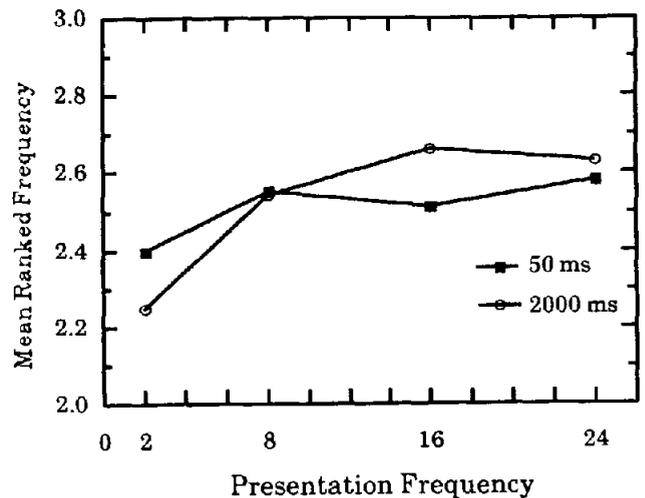


Figure 4. Mean ranked frequency as a function of presentation frequency and exposure duration (collapsed across warning) for detecting-letter-locations cover task conditions in Experiment 3.

50-ms condition did not vary with presentation frequency ( $p > .10$ ). In the 2-s condition, however, subjects ranked the letters shown 2 times as occurring significantly less frequently than those shown 8, 16, and 24 times (all  $ps < .001$ ). Differences among 8, 16, and 24 item presentations were not significant.

### Discussion

These results suggest that at least for subjects viewing the presentation stimuli for 2 s, the absence of any knowledge of relative frequency found in Experiment 2 was not due merely to the presence of a cover task or to the presence of activity between slide presentations, but to the nature of the cover task as well. The memory representations produced by the detecting-letter-locations cover task used in Experiment 3 were sufficient for differentiating two occurrences of specific letter–typeface conjunctions from eight or more occurrences. Once again, a warning about the upcoming frequency judgment task had no effect; frequency judgments were above chance regardless of the presence or absence of a warning.

Taken together, the results of Experiments 1–3 indicate that a 2-s exposure duration is sufficient for some information to be accessible to subjects as a basis of frequency judgments, provided subjects have given attention to the items to be tested as specific perceptual objects.

### Experiment 4

Comparing the results across Experiments 1–3, it appears that as subjects devote more attention to the items as objects (especially as complex objects), frequency judgments improve. That is, the looking-at-objects condition of Experiment 1 seemed to produce more differentiated frequency information than the detecting-letters condition of Experiment 3. One explanation is that subjects in the looking-at-objects condition were more likely to integrate letter and typeface information as part of their perceptual experience. However, given the differences between Experiment 1 and Experiments 2 and 3 (in length of ISI and in whether the ISI was filled or not), these conclusions remain suggestive. Therefore, in Experiment 4, frequency judgments obtained after subjects viewed the stimuli in the three different cover tasks were compared directly using a 6-s filled ISI for all conditions. Because of the results of Experiments 2 and 3, no warning was used in this experiment and the grid was presented for 2 s in all conditions.

In addition, the looking-at-objects cover task was changed for the following reasons: First, in Experiment 1, we assumed that subjects in the looking-at-objects condition would spontaneously process object information because that is the most natural thing to do when looking at slides. However, inasmuch as no cover task was used, we had little control over the type of processing subjects actually engaged in, although subjects were warned about the frequency test. Therefore, in Experiment 4, we omitted the warning and introduced a cover task that should have induced the critical object processing. Subjects were asked to judge the legibility of the stimuli, which should specifically direct their attention to the conjunction of

letter and typeface. Second, subjects recorded their legibility judgments, using a 5-point scale, during the interval between slides. These changes (increasing the specificity in the looking-at-objects instructions, equating the ISI, and including a cover task in all conditions) permitted more confidence about any results regarding the relation between particular types of processing and frequency judgment performance.

### Method

*Subjects.* The subjects were 120 undergraduate students enrolled in an introductory psychology course at the State University of New York at Stony Brook. They participated in the experiment in order to fulfill a course requirement. All subjects had vision that was normal or corrected to normal, and all were naive as to the purposes of the experiment.

*Stimulus materials and apparatus.* The stimuli were the 128 4 × 4 grids used in Experiments 1–3. In this experiment, 32 grids contained only letters, and 96 contained six letters and one digit. Presentation order was counterbalanced as in Experiment 3. A John Bell card in an Apple IIe microcomputer controlled both the exposure duration and the ISI.

*Procedure.* The subjects were tested in groups ranging in size from 3 to 9 students. Groups were assigned randomly to the three cover task conditions, looking at objects ( $n = 39$ ), detecting letter locations ( $n = 42$ ), and detecting locations ( $n = 39$ ). All subjects were shown a sample 4 × 4 grid that contained only digits. In addition, they all saw four presentations of the letter *A*, one in each of the different typefaces. Subjects in the looking-at-objects group were told that we were interested in their impressions of how legible the symbols on each slide were—that is, their immediate judgment of how easily they could identify the symbols while the slide was on. Subjects were asked to use the letters *a* through *e* as follows to indicate how easily they could identify the symbols: (*a*) easily, (*b*) fairly easily, (*c*) somewhat difficult, (*d*) difficult, and (*e*) very difficult. Subjects were told that because the particular combination of other symbols and other locations might be a factor in determining legibility, they were to make separate judgments for each slide.

Subjects in the detecting-locations and detecting-letter-locations groups were given the instructions used in Experiments 2 and 3, respectively.

All subjects were told that they could increase their chances of seeing all the filled locations if they looked at the center of the grid. They were instructed to look at the slides for the entire 2 s they appeared on the screen and then to mark their packets (or make their legibility ratings) during the 6 s between each slide.

*Data analyses.* In addition to analyzing frequency judgments, we compared subjects' performance in the detection cover tasks in order to examine whether differences in performance on the frequency task could be attributed to differences in capacity allocated to the cover tasks. That is, it is possible that subjects found the process of isolating letters from numbers in the detecting-letter-locations task to be so difficult that they abandoned the cover task. If so, it would not be surprising that their frequency judgments look more like those of the subjects who participated in no cover task (Experiment 1) than like those of the subjects in the detecting-locations task (Experiment 2).

### Results

*Frequency ranking data.* As shown in Figure 5, frequency ratings increased as a function of presentation frequency,  $F(3, 351) = 9.28$ ,  $p < .001$ . In addition, there was a significant interaction between presentation frequency and cover task,

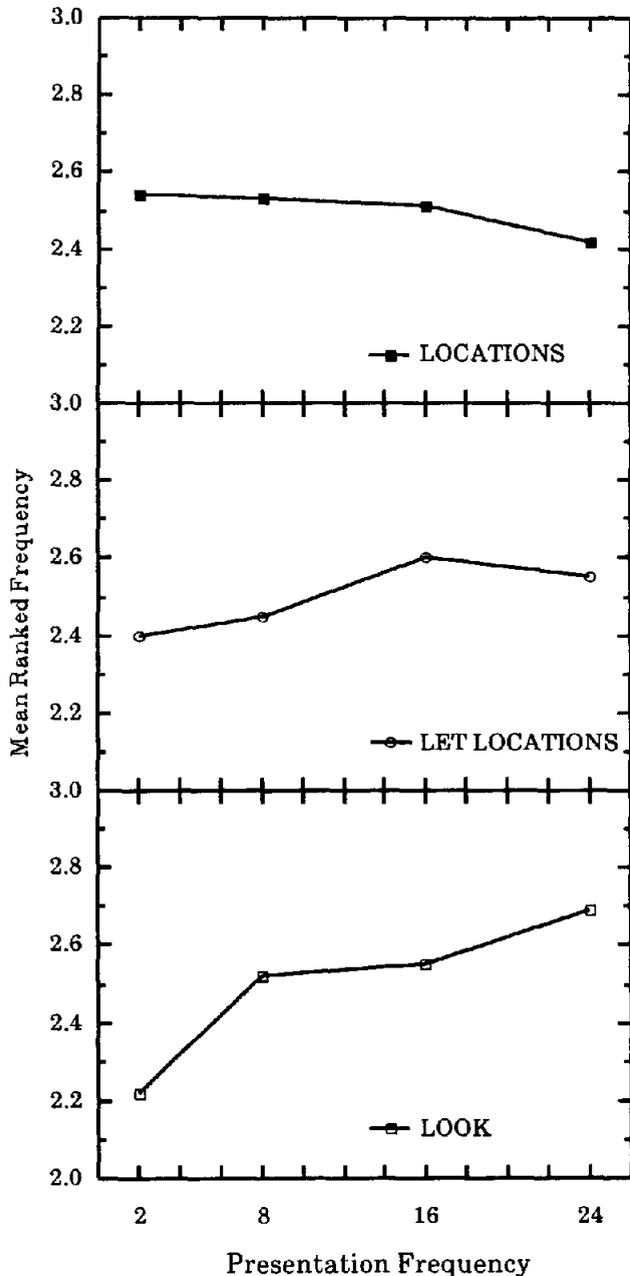


Figure 5. Mean ranked frequency as a function of presentation frequency for three cover task conditions: looking at objects (LOOK), detecting letter locations (LET LOCATIONS), and detecting locations (LOCATIONS) in Experiment 4.

$F(6, 351) = 7.88, p < .001$ . Subsequent analyses showed that frequency ratings increased with presentation frequency in the looking-at-objects condition,  $F(3, 114) = 22.09, p < .001$ , and the detecting-letter-locations condition,  $F(3, 123) = 3.72, p < .01$ , but not in the detecting-locations condition,  $F(3, 114) = 1.70, p > .17$ .

Subjects in the detecting-letter-locations condition ranked letters that had occurred 2 times as significantly less frequent than those that had occurred 16 times ( $p < .001$ ) or 24 times

( $p < .05$ ) and letters that had occurred 8 times as significantly less frequent than those that occurred 16 times ( $p < .01$ ). Subjects in the looking-at-objects condition performed even better, ranking letters that occurred 2 times as significantly less frequent than those that occurred 8, 16, or 24 times ( $ps < .001$ ), and also ranking letters occurring 8 and 16 times as significantly less frequent than those occurring 24 times ( $p < .01$  and  $p < .025$ , respectively).

*Cover task performance.* As a manipulation check, we examined performance on the cover tasks. The subjects in the looking-at-objects condition rated the legibility of the typefaces as follows: Boldface (1.2), Computer (1.6), Script (2.2), and Old English (2.6). In the detecting-locations cover task, for each subject the hit rate ( $H$ ) was the mean proportion of target cells that were marked, and the false alarm rate ( $FA$ ) was the mean proportion of nontarget cells that were marked. For the detecting-letter-locations condition, nontarget cells included cells filled with digits. We calculated  $(H - FA)/(H + FA)$  to compare performance in the two conditions. By this index, the detecting-letter-locations group (.85) did not differ from the detecting-locations group (.83),  $t(78) < 1$ , indicating that the superior performance of the detecting-letter-locations group in the frequency test was not because subjects failed to perform the detection task.

Discussion

When the different viewing purposes represented in Experiments 1-3 were compared directly in Experiment 4, we obtained evidence that frequency judgments about letter-typeface combinations improved with increases in the likelihood that the letter-typeface combinations would be seen as phenomenal objects. As predicted, memory for presentation frequency was best in the looking-at-objects condition and worst in the detecting-locations task. The detecting-letter-locations condition fell in between the other two conditions. This replicates the general pattern observed across Experiments 1-3.

General Discussion

Frequency and Event Processing

The present experiments explore the nature of the processing required for frequency judgments of complex objects or events. We set out to examine the minimum processing that might support frequency judgments of attended objects that required the conjunction of at least two characteristics. For example, consider the nature of the processing that occurs when we walk down the street engaged in an engrossing conversation. We are concerned with the location of objects in order to avoid walking into them, but we are less concerned with their identity. We are generally successful in avoiding objects and other people, thus indicating that we have processed stimuli-occupying locations, but we have little conscious sense of what these objects were. Under these conditions, would we later be able to judge the relative frequency of men and women in red and blue sweaters? The present results suggest it is unlikely: In order for frequency judgments to

accurately reflect presentation frequency, the objects to be judged must have been experienced as phenomenal objects during the initial presentation. The general idea that attention to the stimulus is a precondition for frequency judgments is present in other investigators' discussions of frequency and related issues (e.g., Hasher & Zacks, 1979; Hintzman, 1986a; Zacks et al., 1986). Here we have attempted to make this idea more specific, to tie it to experimental manipulations, and to relate it to a particular theoretical framework (Johnson, 1983).

The three tasks investigated here—looking at objects, detecting letter locations, and detecting locations—were expected to vary in the likelihood that the subjects would process the target stimuli as specific complex objects (i.e., as particular letters in particular typefaces), with the first task requiring the most and the last task requiring the least attention to the target stimuli as specific objects. These tasks were designed to differ in the relative involvement of processing subsystems as postulated in MEM (Johnson, 1983). Filled locations can be specified on the basis of sensory information that does not involve identifying particular objects, which may be necessary to establish a perceptual record. Specifying whether a letter (as opposed to a digit) occupies a particular location involves the perceptual subsystem because it requires the phenomenal identification of objects. Note that identifying the objects as letters does not necessarily require the phenomenal conjunction of letter and typeface information, but this conjunction should be more likely to occur if subjects are looking for objects than if they are only identifying filled locations. Finally, rating legibility involves perceptual subsystem processes that are directed at producing a conjunction of letter and typeface information. In this latter condition, then, the phenomenal objects or events are more likely to be letters in specific typefaces than in the detecting-letter-locations task. If judgments of relative frequency are based on records of phenomenal perceptual experience of those events, then the better the match between original experience and the object or events to be later judged, the better should be the judgment. The superior performance on the frequency task in the looking-at-objects condition compared with the detecting-letter-locations condition (along with the chance performance in the detecting-locations condition) suggests that this is the case.

In addition to requiring some minimal type of processing, frequency judgments seem to require some minimal processing time: Even the looking-at-objects task required 2 s per grid (approximately 285 ms/item) for systematic relative frequency judgments to emerge, even under intentional instructions (Experiment 1).

The exact nature of the processing required to support frequency judgments remains to be specified precisely. For example, the detecting-letter-locations task did not direct viewer's attention to the typeface of the letters. Nevertheless, frequency judgments about letters in particular typefaces were better than in the detecting-locations task. In processing shape information to identify letters, shape information that specifies typeface may also be processed to some extent, perhaps because it is structurally relevant to the letter shape. In any event, a reasonable conclusion at this point is that attention or visual processing directed at stimuli merely for the purpose of identifying their locations is not sufficient for accurate

judgments about the relative frequency of complex events in those locations. Rather, accurate relative frequency judgments appear to require that complex target events have been experienced as phenomenal objects of perception.

Although we cannot say precisely what is required for an object or an event to be a phenomenal object of perception, we can begin to consider variables that are likely to be critical. For example, the amount of time and the type of processing required for a target event to achieve phenomenal status may depend on the nature of the event and its familiarity. Simple, familiar stimuli (e.g., letters and words) should be easy to perceive; complex stimuli that are the conjunction of various features (e.g., letters in a particular typeface) should be harder to perceive, requiring a greater number of processes or more time (Treisman & Gelade, 1980). Presumably, estimating the frequency of even more complex events (e.g., how often have you observed an Oedipus complex?) would require even more complex processing and more processing time, and may involve reflective as well as perceptual processes.

A related point is that when a subject is asked to judge the relative frequency of events, better judgments should result when subjects perceive the initial event in the same terms as those in which the question is asked (e.g., when a subject perceives a symbol as a particular letter in a particular typeface). That is, frequency estimation is a judgment about "events" or "objects," and "events" or "objects" are defined in terms of phenomenal experience. This is an extension of the encoding specificity principle (Tulving, 1983): An event must be encoded in terms relevant to the later frequency judgment task (also see Hintzman, 1986b.) Thus, it is not a question of whether frequency judgments are a sensitive measure of long-term memory modification (Fisk & Schneider, 1984; Hock et al., 1986), but rather whether the form of the frequency estimation task is compatible with the form of the memory representation resulting from the type of processing engaged in during presentation. Evidence consistent with this claim is that the match between encoding and test should be better for such combinations as carrots, corn, peas, and broccoli (at encoding) and vegetables (at test) than for such combinations as ketchup, blood, apple, and stop sign (at encoding) and red things (at test) (Alba, Chromiak, Hasher, & Attig, 1980; Barsalou & Ross, 1986; Freund & Hasher, 1986), because "vegetables" is a more probable encoding category than "red things."

This proposal is not necessarily at odds with recent work demonstrating frequency judgments for elements within objects—specifically for letters within strings (Hock et al., 1986). We agree with Hock et al. that nominal stimuli may be represented in more than one way (or at more than one level) in memory. However, this does not mean that while subjects attend to or process one aspect of a stimulus (e.g., the strings), another unattended aspect is simultaneously or automatically being counted (e.g., the letters). This may be the case (especially for familiar elements), but current evidence is not sufficiently compelling. Hock et al. (1986) found the strongest support for this interpretation in a comparison between performance with long and short exposures. They found that with long exposures, subjects could differentiate both higher from lower frequency strings and higher from lower frequency

letters, even though string and letter frequency were varied orthogonally. With brief exposures (.2 s), however, string frequency estimates were good, but letter frequency estimates were not. The differential performance at long and short exposures provides the most convincing evidence that the judgments about the letters were not derived from the memory representations of the strings. However, it is possible that the types of processing engaged in at short and long exposures were quite different. For example, with long exposures, letter-by-letter processing may have occurred, and this processing may support both letter and string judgments. With short exposures, however, overall shape or contour processing may have occurred and this may have been sufficient to support string judgments, but not letter judgments. Thus, inasmuch as the type of processing engaged in at short and long exposures cannot be specified, Hock et al.'s (1986) results do not imply that elements not initially experienced as phenomenal objects can support frequency judgments. (For a related discussion of the Hock et al. conclusions, see Hintzman, 1986b.)

Of course, whether subordinate elements and/or superordinate categories (or associates) are a reliable part of either the unconscious or the phenomenal consequences of a nominal stimulus should change with experience (e.g., Barsalou & Ross, 1986). In any case, a fundamental problem is the relative contributions of unconscious and conscious activation of information to creating memory records that will support frequency judgments.

*Intention.* In Experiments 2 and 3, a warning about the upcoming frequency test did not improve performance, even though there was considerable room for improvement. This result is consistent with the idea that intention per se has little effect on memory (Postman, 1964), without instructions concerning the proper strategy (Greene, 1984, 1986). Warnings may have effects under other circumstances, however: For example, warning effects have been reported when simple or well-integrated stimuli (e.g., letters and words) are presented briefly (Fisk & Schneider, 1984; Naveh-Benjamin & Jonides, 1986). Although this may be true for simple stimuli such as Fisk and Schneider or Naveh-Benjamin and Jonides used, it does not appear to hold for more complex stimuli such as we used. With complex stimuli, longer exposure durations may be necessary for a warning to improve performance. Even the longest exposure duration used here (2 s for seven items) did not leave much time for extended perceptual processing or for embellishing individual occurrences of target stimuli through reflective processes (cf. Maki & Ostby, 1987).

In summary, the present results are consistent both with the general idea that different memory tasks draw on different types of information or records in memory (e.g., Cohen & Squire, 1980; Johnson, 1983; Moscovitch, Winocur, & McLachlan, 1986; Roediger & Blaxton, 1987; Tulving, 1985) and with the more specific proposal, derived from MEM (Johnson, 1983), that perceptual records may be a minimum requirement for frequency judgments of complex events. Further specification of the processes that define perceptual events, of the nature of the information on which frequency judgments can be based, and of the conditions that affect sensitivity to relative frequency provides a potentially useful focus for future research.

## References

- Alba, J. W., Chromiak, W., Hasher, L., & Attig, M. S. (1980). Automatic encoding of category size information. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 370-378.
- Barsalou, L. W., & Ross, B. H. (1986). The roles of automatic and strategic processing in sensitivity to superordinate and property frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 116-134.
- Begg, I., Maxwell, D., Mitterer, J. O., & Harris, G. (1986). Estimates of frequency: Attribute or attribution? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 496-508.
- Birnbaum, I. M., Taylor, T. H., Johnson, M. K., & Raye, C. L. (1987). Automatic encoding of event frequency? The case of alcohol intoxication. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 251-258.
- Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, 210, 207-210.
- Coltheart, M. (1980). Iconic memory and visual persistence. *Perception and Psychophysics*, 27, 183-228.
- Fisk, A. D., & Schneider, W. (1984). Memory as a function of attention, level of processing, and automatization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 181-197.
- Freund, J. S., & Hasher, L. (1986, November). *Judgments of category size: Now you have them, now you don't*. Paper presented at the annual meeting of the Psychonomic Society, New Orleans.
- Greene, R. L. (1984). Incidental learning of event frequency. *Memory & Cognition*, 12, 90-95.
- Greene, R. L. (1986). Effects of intentionality and strategy on memory for frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 489-495.
- Hanson, C., & Hirst, W. (1988). Frequency encoding of token and type information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 289-297.
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 356-388.
- Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist*, 39, 1372-1388.
- Hintzman, D. L. (1986a). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, 93, 411-428.
- Hintzman, D. L. (1986b). *Judgments of frequency and recognition memory in a multiple-trace memory model* (Tech. Rep. No. 86-11). Eugene, OR: University of Oregon, Cognitive Science Program.
- Hock, H. S., Malcus, L., & Hasher, L. (1986). Frequency discrimination: Assessing global-level and element-level units in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 232-240.
- Johnson, M. K. (1983). A multiple-entry, modular memory system. In G. H. Bower (Ed.), *The Psychology of learning and motivation: Advances in research and theory* (Vol. 17, pp. 81-123). New York: Academic Press.
- Johnson, M. K. (in press). Functional forms of human memory. In J. L. McGaugh, N. M. Weinberger, & G. Lynch (Eds.), *Brain organization and memory: Cells, systems and circuits*. New York: Oxford University Press.
- Jonides, J., & Naveh-Benjamin, M. (1987). Estimating frequency of occurrence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 230-240.
- Kausler, D. H., Lichty, W., & Hakami, M. K. (1984). Frequency

- judgments for distractor items in a short-term memory task: Instructional variation and adult age differences. *Journal of Verbal Learning and Verbal Behavior*, 23, 660-668.
- Keppel, G. P. (1973). *Design and analysis: A researcher's handbook*. Englewood Cliffs, NJ: Prentice-Hall.
- Lewicki, P. (1986). *Nonconscious social information processing*. Orlando, FL: Academic Press.
- Maki, R. H., & Ostby, R. S. (1987). Effects of level of processing and rehearsal on frequency judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 151-163.
- Mewhort, D. J. K., Campbell, A. J., Marchetti, F. M., & Campbell, J. I. D. (1981). Identification, localization, and "iconic" memory: An evaluation of the Bar-Probe Task. *Memory & Cognition*, 9, 50-67.
- Moscovitch, M., Winocur, G., & McLachlan, D. (1986). Memory as assessed by recognition and reading time in normal and memory-impaired people with Alzheimer's disease and other neurological disorders. *Journal of Experimental Psychology: General*, 115, 331-347.
- Naveh-Benjamin, M., & Jonides, J. (1986). On the automaticity of frequency coding: Effects of competing task load, encoding strategy, and intention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 378-386.
- Postman, L. (1964). Short-term memory and incidental learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 145-201). New York: Academic Press.
- Roediger, H. L. III, & Blaxton, T. A. (1987). Retrieval modes produce dissociations in memory for surface information. In D. S. Gorfein, & R. R. Hoffman (Eds.), *Memory and cognitive processes: The Ebbinghaus centennial conference* (pp. 349-379). Hillsdale, NJ: Erlbaum.
- Rowe, E. J. (1974). Depth of processing in a frequency judgment task. *Journal of Verbal Learning and Verbal Behavior*, 13, 638-643.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monograph: General and Applied*, 74 (Whole No. 498).
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97-136.
- Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.
- Tulving, E. (1985). How many memory systems are there? *American Psychologist*, 40, 385-398.
- Underwood, B. J. (1983). *Attributes of memory*. Glenview, IL: Scott, Foresman.
- Williams, K. W., & Durso, F. T. (1986). Judging category frequency: Automaticity or availability? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 387-396.
- Zacks, R. T., Hasher, L., & Hock, H. S. (1986). Inevitability and automaticity: A response to Fisk. *American Psychologist*, 41, 216-218.

Received June 15, 1987

Revision received March 28, 1988

Accepted March 31, 1988 ■

### Squire Appointed Editor of *Behavioral Neuroscience*, 1990-1995

The Publications and Communications Board of the American Psychological Association announces the appointment of Larry R. Squire, Veterans Administration Medical Center and University of California, San Diego, as editor of *Behavioral Neuroscience* for a 6-year term beginning in 1990. As of January 1, 1989, manuscripts should be directed to

Larry R. Squire  
 Veterans Administration Medical Center (116 A)  
 3350 La Jolla Village Drive  
 San Diego, California 92161