Journal of Experimental Psychology:
Learning, Memory, and Cognition
1986, Vol. 12, No. 2, 232–240

Frequency Discrimination: Assessing Global-Level and Element-Level Units in Memory

Howard S. Hock and Lawrence Malcus
Florida Atlantic University

Lynn Hasher
Temple University

Subjects' knowledge of how often various events occur was used to assess the retention of memory units for word-like strings of letters. A series of strings was presented at one of three exposure durations. Within the series, the frequencies of occurrence of different strings and of the letters composing the strings were varied orthogonally. At relatively long exposure durations, subjects could discriminate the frequency of occurrence for both strings and their constituent letters. The formation of global-level (string) memory units was indicated by judgments of string frequency being unaffected by either the frequencies of their component letters or experimental conditions (brief exposures) that prohibited accurate judgment of letter frequency. Although judgments of letter frequency were sometimes biased by the frequency of the strings containing the letters, the success with which the judgments discriminated different levels of letter frequency did not depend on the activation of string-level memory units. Furthermore, subjects' frequency judgments for letters were not predictable from their recall of the strings containing the letters. These results, which could not be explained by Tversky and Kahneman's (1973) "availability heuristic," provided evidence for the formation of element-level (letter) memory units. A converging experiment established that element-level frequency information could be abstracted from words as well as nonwords, and further, that this information was stored in long-term memory.

Of central concern to both perceptual and cognitive theories of visual processing (e.g., Hochberg, 1981; Neisser, 1967) is the issue of whether the functional units are elements, subsets of elements, or the entire visual array. For example, the functional units in the identification of printed words could be individual letters, orthographically regular combinations of letters, or the entire word. Regardless of the size of functional units during the identification of printed strings of letters, our concern in this article was to determine whether informational units of different size are stored in memory.

It is well established that people are sensitive to information about the frequency with which events occur (Hintzman & Block, 1971; Underwood, 1969). This information appears to be processed with little effort (Hasher & Zacks, 1979) or intention (Howell, 1973). Our experiments were designed to capitalize on this sensitivity to occurrence-rate information as a way of identifying memory units for words and word-like items. We did this in our first experiment by varying the frequency of occurrence of strings of letters orthogonally to the frequency of occurrence of the individual letters composing the strings. After presenting the strings at varying exposure durations, we asked subjects to judge either letter or string frequency. Our results provided converging evidence for the retention of both global-level (string) and element-level (letter) memory units. Retrieval of memory units at both levels influenced subjects' judgments of letter frequency. String frequency judgments appeared to be influenced only by string-level memory units.

Previous researchers have been concerned with frequency judgments for items that were physically identical for each repetition in a list as well as information abstracted from the items in a list. Gude and Zechmeister (1975) and Burnett and Stevenson (1979) have compared frequency judgments for sentences that were literally identical on each repetition with sentences that were literally different but kept the same meaning on each repetition. Jacoby (1972) and Rowe (1973a, 1973b) have similarly compared frequency judgments for words that were literally identical on each repetition with frequency judgments for words that were literally identical but varied in meaning (homonyms) and words that were literally different but the same in meaning (synonyms). Our approach is somewhat different. We are interested in frequency judgments for information abstracted from the items in the list, but the abstraction of interest is across the list rather than within the individual items composing the list. From this point of view, letter-level memory units, coded for frequency of occurrence, could constitute an abstract description of the compositional characteristics of the list. By analogy with category acquisition research, the stimulus list presented at some point in time could be thought of as a particular information category, with the global-level (string) units corresponding to the exemplars of the category and the element-level (letter) units corresponding to an abstract, featural description of the category.

The initial purpose of Experiment 1 was to determine if, at relatively long presentation durations, subjects' judgments would discriminate different levels of occurrence frequency for both

This research was supported by Grant #MDA903-82-C-0317 from the Army Research Institute and Grant #MH33140 from the National Institute of Mental Health. The authors thank Stephen Reed and John Jonides for careful reading of an earlier version of the manuscript, Gerald O'Rourke for conducting Experiment 2, and Clay Cavedo for analyzing the data.

Correspondence concerning this article should be addressed to Howard S. Hock, Department of Psychology, Florida Atlantic University, Boca Raton, Florida 33431.

the strings and the letters composing the strings. Correlational data suggest that people have reliable knowledge about the frequency with which both words (Carroll, 1971) and individual letters (Attneave, 1953) occur in natural language. It was important to demonstrate that under relatively ideal circumstances, subjects could discriminate frequency of occurrence at both the element (letter) and global (string) levels. By varying letter and string frequency orthogonally, we could determine whether frequency judgments at the element level were derived from stored frequency information at the global level and vice versa.

Experiment 1 also manipulated the presentation duration for the strings, the purpose being to determine the levels of memory representation, if any, that are sacrificed when processing constraints are increased (here, by reducing the presentation duration). If, for example, brief exposure durations eliminated subjects' ability to judge the frequency of letters while leaving intact their ability to judge the frequency of strings, it would provide evidence that judgments of frequency for strings were not derived from stored frequency information for their constituent letters.

The final purpose of the first experiment was to determine whether subjects use a strategy in which letter-frequency judgments are based on the activation of string-level memory units. According to one such strategy, subjects' estimates of letter frequency would increase if they activate the memory representation for a string containing the letter. This strategy, which is similar to the "availability heuristic" proposed by Tversky and Kahneman (1973), could be the basis for accurate judgments of letter frequency. The latter have argued that estimates of the frequency of various events depend on the activation in memory of specific instantiations or associates of the events. The activation of strings could be responsible for the accurate judgment of letter frequency because in our experiments, as in natural language, the reason some letters are higher in frequency than others is that high-frequency letters occur in more different strings than do low-frequency letters. As a result, there is more opportunity for subjects to activate memory representations of strings containing high-frequency letters compared with strings containing low-frequency letters, and further, to use this as the basis for judgments of the frequency of occurrence of constituent letters.

Experiment 1 was designed to control for the possibility that subjects would be biased to judge letters as high in frequency simply because they occurred in "more available," high-frequency strings. The orthogonal manipulation of letter and string frequency prevented such a bias from being the basis for subjects' discrimination between high- and low-frequency letters. Subjects' free-recall protocols were also used to investigate the role of string activation in letter-frequency judgments. In analyzing the recall data we assumed that the likelihood of a string being recalled was directly related to the likelihood that its memory representation was activated while subjects were making judgments of letter frequency.

Experiment 1

Method

Subjects. Seventy-two students in undergraduate psychology classes at Florida Atlantic University voluntarily participated in this experiment without pay.

Table 1
The Consonants and Strings Assigned to the Four Experimental Conditions Used in Experiments 1 and 1A

	String frequency			
	High (6)		Low (3)	
Consonant frequency	Consonants	Strings	Consonants	Strings
High (12)	*N,P,W*	*AWUP* *WENY* *NIPO*	*S,B,M*	*SUMY* *EMSA* *BOIM* *MUBY* *ASEB* *OBIS*
Low (6)	*X,H,G,C,R,F*	*ROCY* *AGIF* *HUXE*	*Z,T,K,L,V,D*	*ZAKY* *TIDU* *OLEV* *UZIL* *KEDO* *VATY*

Design. The experiment was conducted in three phases. Phase 1 involved presenting a sequence of 72 letter strings at one of three exposure durations. In Phase 2, independent groups of subjects judged the frequency of occurrence for either the strings or the letters composing them. Only subjects judging letter frequency during Phase 2 participated in Phase 3, which involved recalling the strings presented during Phase 1. The orthogonal combination of the three exposure durations and the two frequency judgment conditions (letter vs. string) generated six between-subject experimental conditions. Twelve subjects were randomly assigned to each condition. The stimuli, which are presented in Table 1, were designed such that the orthogonal combination of two variables, high- versus low-frequency letters and high- versus low-frequency strings, produced four within-subject conditions.

Stimuli. The initial step in generating the stimuli was to assign three consonants to each of the two high-letter-frequency conditions and six consonants to each of the two low-letter-frequency conditions. The average frequency of usage in English (Mayzner & Tresselt, 1965) was virtually identical for the four sets of consonants. The five vowels, plus *Y*, were then combined with the consonants to produce the 18 orthographically regular, four-letter strings presented in Table 1.

In the high-letter-frequency, high-string-frequency condition, each string was presented a total of six times. Because each consonant in this condition appeared in two different strings, each was presented a total of 12 times. In the low-letter-frequency, high-string-frequency condition, each string was again presented six times, but now each consonant appeared in only one string. As a result, each was presented a total of six times. The same logic was applied in the two remaining stimulus conditions. Following this procedure, 18 strings were constructed. Six were presented six times each and 12 were presented three times each, producing a total of 72 stimuli. As can be seen in Table 1, consonants that appeared in two different strings were presented in different positions in each string and in combination with different vowels and consonants. Each vowel, including *Y*, appeared equally often; two vowels were always combined with two consonants. The strings were all typed in upper-case, Letter Gothic type font.

None of the strings used in the experiment were English words. However, most could be converted to words by changing the identity of a single letter, leaving position unchanged (e.g., *WENY* could become *WENT*). The strings assigned to the four experimental conditions were matched in that two thirds of the items in each condition could be converted to English words by changing one letter.

Procedure. The 72 strings were presented, in random order, at one of three presentation rates. The fastest rate, which was achieved by holding

Table 2
Judgments of Letter Frequency: The Proportion of "12" Responses

Exposure duration (s)	Frequency of letters in high-frequency strings				Frequency of letters in low-frequency strings				Letter frequency for all strings			
	12	6	12 − 6	*M*	12	6	12 − 6	*M*	12	6	12 − 6	*M*
					Experiment 1							
4.2	.70	.26	.44	.48	.44	.22	.22	.33	.57	.24	.33	.41
1.2	.36	.17	.19	.27	.39	.24	.15	.32	.38	.21	.17	.30
0.2	.42	.53	−.11	.48	.40	.25	.15	.33	.41	.39	.02	.40
M	.49	.32	.17	.41	.41	.24	.17	.33	.45	.28	.17	.37
					Experiment 1A							
0.2	.48	.47	.01	.48	.44	.42	.02	.43	.46	.45	.01	.46

Note. The proportion of "12" responses refers to how often subjects judge a letter to have occurred 12 rather than 6 times.

the advance button on the Kodak Carousel projector in a depressed position, resulted in an exposure duration of approximately 0.2 s per string and a duration of approximately 0.8 s between exposures. The slower two presentation rates resulted in exposure durations of approximately 1.2 and 4.2 s per string, the duration between exposures remaining at approximately 0.8 s. The projected width of each string was approximately 15 cm. Each letter was approximately 2.2 cm wide. Because the experiment was conducted in classrooms, the visual angle intercepted by each string varied from subject to subject but was never greater than 3.0°. Prior to the presentation of the strings, each group of subjects was instructed to try to remember the information presented on the screen. There were no directions concerning whether subjects should attend to individual letters or strings and no indication that we would subsequently assess frequency of occurrence information.

Following the presentation of the strings, subjects in alternate seats were assigned to two groups. One group made decisions about consonant frequency, the other group about string frequency. Booklets were prepared with one consonant (or string) per page, each booklet comprising a different random order. Subjects judging string frequency were required to circle either the "3" or "6" typed above each string. Subjects judging letter frequency were required to circle either the "6" or "12" typed above each consonant. Following these responses, subjects in the letter (consonant) judgment condition were instructed to recall as many strings as possible.

Results

Letter-frequency judgments. The overall percentages of letter-frequency judgments that were correct were 67%, 59%, and 52%, in the 4.2-, 1.2-, and 0.2-s conditions, respectively. These data indicated that subjects could discriminate letter frequency for the 4.2-s and 1.2-s exposures (chance was 50% correct), but frequency discrimination following the 0.2-s exposures was equivocal. Average performance in the latter condition was close to chance (52% correct), with subjects performing below chance (45%) for letters from high-frequency strings and above chance (58%) for letters from low-frequency strings.

For purposes of analysis, the dependent variable was the proportion of subjects' responses for which letters were judged to have occurred 12 times (half the letters were presented 12 times, half 6 times). This measure allowed us to determine whether letter-frequency judgments (a) discriminated between high- and low-frequency letters (as indicated by positive 12–6 difference scores), and (b) were biased by the frequency of the strings containing the letters (as indicated by the mean proportion of "12" judgments at each level of string frequency). These results are summarized in Table 2. In addition to showing the previously described effect of exposure duration on discrimination accuracy, the data in Table 2 indicate a general bias to judge letters as low in frequency.

An analysis of variance (ANOVA) on the proportion of "12" responses in subjects' letter judgments indicated that there was a significant interaction between the effects of letter frequency and exposure duration, $F(2, 33) = 6.20$, $p < .01$, $MS_e = .045$. Tests of simple effects indicated that the effect of letter frequency was significant for the 4.2-s and 1.2-s exposure durations, $F(1, 33) = 28.32$, $F(1, 33) = 24.08$, $p < .01$, $MS_e = .045$, respectively, but was not significant for the 0.2-s exposure duration, $F(1, 33) < 1.0$, $MS_e = .045$. High- and low-frequency letters were not differentiated more accurately when they were presented in high-frequency compared with low-frequency strings. That is, the interaction between letter frequency and string frequency was not significant, $F(1, 33) < 1.0$, $MS_e = .086$. The three-way interaction between exposure duration, letter frequency, and string frequency also was not significant, $F(2, 33) = 2.00$, $p > .05$, $MS_e = .086$.

With regard to response bias, the main effect of string frequency, $F(1, 33) = 5.63$, $p < .05$, $MS_e = .042$, and the interaction between string frequency and exposure duration, $F(2, 33) = 3.63$, $p < .05$, $MS_e = .042$, were significant. The latter two effects were obtained because frequency judgments were biased to be relatively high for letters from high-frequency strings, at least for the 4.2-s and 0.2-s exposure durations. This bias, which was smaller and unreliable in the replication reported in Experiment 1A, did not directly contribute to the discrimination of letter frequency because string frequency and letter frequency were varied orthogonally in both experiments.

A further analysis, in which the letters used in the experiment replaced subjects as the random variable in the ANOVA, indicated that the results were generalizable over the letters used in the experiment.[1] The interaction between letter frequency and ex-

[1] Because the experiment was designed with different numbers of letters assigned to the high- and low-frequency conditions, we performed a least-

Table 3
Judgments of String Frequency: The Proportion of "6" Responses

Exposure duration (s)	Frequency of strings with high-frequency letters				Frequency of strings with low-frequency letters				String frequency for all letters			
	6	3	6 − 3	*M*	6	3	6 − 3	*M*	12	6	6 − 3	*M*
					Experiment 1							
4.2	.75	.26	.49	.51	.81	.14	.67	.48	.78	.20	.58	.49
1.2	.81	.29	.52	.55	.81	.38	.43	.60	.81	.34	.47	.58
0.2	.67	.43	.24	.55	.84	.36	.48	.60	.76	.40	.36	.58
M	.74	.33	.41	.54	.82	.29	.53	.56	.78	.31	.47	.55
					Experiment 1A							
0.2	.67	.37	.30	.52	.79	.46	.33	.63	.73	.42	.31	.58

Note. The proportion of "6" responses refers to how often subjects judge a letter to have occurred 6 rather than 3 times.

posure duration was again significant, $F(2, 28) = 8.21, p < .005$, $MS_e = .011$. Tests of simple effects again indicated that the effect of letter frequency was significant at the 4.2-s and 1.2-s exposure durations, $F(1, 28) = 39.56$, $F(1, 28) = 11.24, p < .005, MS_e = .011$, respectively, but not the 0.2-s exposure duration, $F(1, 42) < 1.0, MS_e = 0.11$. Once again, the interaction between letter frequency and string frequency was not significant, $F(1, 42) < 1.0$, $MS_e = .020$, but the main effect of string frequency, $F(1, 28) = 4.39, p < .05, MS_e = .020$, and the interaction between exposure duration and string frequency, $F(2, 28) = 6.38, p < .01, MS_e = .011$, were significant. The one difference from the analysis in which subjects was the random variable was that the three-way interaction between exposure duration, letter frequency, and string frequency was now significant, $F(2, 28) = 5.50, p < .05$, $MS_e = .011$. This interaction, which was not reliable over the full set of participating subjects, reflected the relatively small effect of exposure duration on frequency discriminability for letters from low-frequency strings. This insensitivity to exposure duration may have been the result of performance being relatively poor (at "floor") for the letters from low-frequency strings. Our clearest evidence for successful letter-level frequency discrimination (and its elimination at brief exposure durations) was obtained for letters from high-frequency strings.

The above analyses indicated that subjects could discriminate high-frequency from low-frequency letters for 1.2- and 4.2-s exposure durations. It remained possible, however, that their judgments of letter frequency were based on the activation of global-level memory units (i.e., strings) containing the letters rather than the retrieval of letter-level frequency information. To evaluate this possibility, letter-frequency judgments were compared at the 1.2-s and 4.2-s exposure durations for high-frequency consonants embedded in low-frequency strings and low-frequency consonants embedded in high-frequency strings (these data are part of the full data set presented in Table 2). If letter-frequency judgments were based on the activation of strings containing either high- or low-frequency consonants, differences in string frequency would have favored the activation of strings containing low-frequency letters. Nonetheless, high-frequency letters were judged as higher in frequency than low-frequency letters for both the 1.2-s exposure duration, $t(11) = 3.74, p < .005$, and for the 4.2-s exposure duration, $t(11) = 2.57, p < .05$.

squares analysis of these data. A different subset of letters was assigned to the four stimulus conditions generated by the orthogonal combination of letter frequency and string frequency. The latter were therefore treated as between factors in the analysis. Because each letter was presented at all three exposure durations, exposure duration was treated as a within factor in the analysis. The same approach was taken for the other item analyses reported in this experiment.

String-frequency judgments. The overall percentages of string frequency judgments that were correct were 80%, 79%, and 68%, in the 4.2-, 1.2-, and 0.2-s conditions, respectively. The proportion of responses for which the strings were judged to have occurred six times (half the strings had been presented six times, half three times) are presented in Table 3. It can be seen from the 6–3 difference scores that subjects' judgments discriminated between the two levels of string frequency at all three exposure durations. It can also be seen from the mean proportion of "6" responses in Table 3 that differences in letter frequency introduced little response bias into subjects' judgments of string frequency.

An ANOVA was performed in which frequency judgments were contrasted for high- and low-frequency strings, which in turn were composed of high- and low-frequency letters. The analysis indicated that the effect of string frequency on the proportion of "6" responses in subjects' string judgments was significant, $F(1, 33) = 107.07, p < .001, MS_e = .074$. This effect was obtained for all three exposure durations; the interaction between exposure duration and string frequency was not significant, $F(2, 33) = 1.96, p > .05, MS_e = .074$. Although string-frequency judgments were slightly higher for the low-letter-frequency than the high-letter-frequency condition, the main effect of letter frequency was not significant, $F(1, 33) < 1.0, MS_e = .035$. Two interactions were marginally significant. The three-way interaction between string frequency, letter frequency, and exposure duration, $F(2, 33) = 3.30, MS_e = .026$, barely reached significance at the .05 level. The interaction between string frequency and letter frequency, $F(1, 33) = 4.09, MS_e = .026$, fell just short of significance at the .05 level. These interactions may have been due to uncontrolled characteristics of some of our items (e.g., certain items may have been more likely than others to remind subjects of familiar words). This was suggested by the results of an additional ANOVA in which items replaced subjects as the random variable.

Table 4
Probability of a Letter Being Present in at Least One String Recalled by a Subject (Experiment 1)

Exposure duration (s)	Frequency of letters in high-frequency strings				Frequency of letters in low-frequency strings				Letter frequency for all strings			
	12	6	12 − 6	*M*	12	6	12 − 6	*M*	12	6	12 − 6	*M*
4.2	.81	.53	.28	.67	.64	.40	.24	.52	.73	.47	.26	.60
1.2	.50	.47	.03	.49	.58	.14	.44	.36	.54	.31	.23	.43
0.2	.61	.31	.30	.46	.14	0	.14	.07	.38	.16	.22	.27
M	.64	.44	.20	.54	.45	.18	.27	.32	.55	.31	.24	.43

The results of this item analysis again indicated that the effect of string frequency on the proportion of "6" responses was significant, $F(1, 14) = 79.55$, $p < .001$, $MS_e = .033$, and the interaction between string frequency and exposure duration was not significant, $F(2, 28) = 2.09$, $p > .05$, $MS_e = .026$. However, neither the interaction between string and letter frequency, $F(1, 14) = 1.22$, $p > .05$, $MS_e = .033$, nor the three-way interaction between string frequency, letter frequency, and exposure duration, $F(2, 28) = 1.22$, $p > .05$, $MS_e = .026$, was significantly generalizable over the items used in the experiment. The interaction effects obtained with subjects serving as the random variable were thus limited to a relatively small number of strings in the stimulus list.

The results obtained for the 0.2-s presentation duration indicated that string frequency could be discriminated under conditions that prohibited the retention of letter frequency information (subjects' frequency judgments following the 0.2-s exposures did not discriminate between the two levels of letter frequency). This evidence that frequency judgments for strings were not based on the frequency of occurrence of their constituent letters was supplemented by the following. At each exposure duration, string-frequency judgments were compared for high-frequency strings composed of low-frequency letters and low-frequency strings composed of high-frequency letters. If string-frequency judgments were based on the frequency of occurrence of the letters composing each string, differences in letter frequency would have favored judging the low-frequency strings as high in frequency (the relevant data are part of the full data set presented in Table 3). Nonetheless, high-frequency strings were judged as higher in frequency than low-frequency strings at the 0.2-s exposure duration, $t(11) = 6.87$, $p < .001$, the 1.2-s exposure duration, $t(11) = 7.65$, $p < .001$, and the 4.2-s exposure duration, $t(11) = 4.32$, $p < .005$.

Free recall. As indicated earlier, subjects who judged letter frequency were subsequently asked to recall as many strings as possible. Our reason for doing this was to determine whether subjects' letter-frequency judgments were derived from activated memory representations for previously seen strings. For each letter we computed the proportion of subjects for whom the letter appeared in at least one correctly recalled string. As can be seen from the mean values in Table 4, we obtained the expected difference in probability of recall between high- and low-frequency letters; high-frequency letters appeared more often in correctly recalled strings than did low-frequency letters. This difference was expected because high-frequency letters occurred in more strings than low-frequency letters. What was important, however, was that this difference was virtually identical, when averaged over string frequency, for each of the three exposure durations.[2] An ANOVA, with the letters used in the experiment serving as the random variable, indicated that strings with high-frequency letters were recalled significantly more often than strings with low-frequency letters, $F(1, 14) = 104.57$, $p < .001$, $MS_e = .007$. The effect of exposure duration on string recall was significant, $F(2, 28) = 27.24$, $p < .001$, $MS_e = .017$, but neither the interaction between letter frequency and exposure duration, $F(2, 28) < 1.0$, $MS_e = .017$, nor the interaction between letter frequency and string frequency, $F(1, 14) = 2.13$, $p > .05$, $MS_e = .007$, was significant. Finally, the three-way interaction between exposure duration, letter frequency, and string frequency was significant, $F(2, 28) = 5.44$, $p < .02$, $MS_e = .017$. There was no obvious explanation for this interaction. It is worth noting, however, that the data pattern leading to the interaction did not match the pattern obtained for judgments of letter frequency.

The results obtained from this analysis of subjects' recall protocols stand in contrast with the significant interaction between letter frequency and exposure duration that was obtained when we assessed subjects' judgments of frequency of occurrence for individual letters. Subjects' frequency judgments differentiated among the high- and low-frequency letters for the 4.2-s and 1.2-s, but not for the 0.2-s exposure duration. This interaction was not obtained for the recall data. Subjects recalled more strings with high-frequency letters than strings with low-frequency letters, even at the briefest exposure duration. If subjects' frequency judgments for letters were based on whether or not they could recall a letter string containing the letter being judged, they would have been as accurate discriminating letter frequency for the 0.2-s exposure duration as they were for the 1.2-s and 4.2-s exposure durations. The case was particularly clear for letters from high-frequency strings. For 0.2-s exposures, subjects were clearly unable to judge the frequency of occurrence for these letters (see Table 2). However, it can be seen from the recall analysis, summarized in Table 4, that this combination of conditions (i.e., high-frequency strings, 0.2-s exposures) provided the greatest potential for subjects to base discriminative letter-frequency judgments on the activation of memory representations for previously seen strings. Yet, successful letter-frequency discrimination was not obtained.

[2] A separate analysis was performed for strings that were listed in subjects' recall but did not correspond with strings that were previously presented. High- and low-frequency letters appeared equally often in these incorrectly recalled letter strings.

An alternative version of the strategy discussed above is one in which frequency judgments are based on whether or not subjects can recall more than one string containing the letter being judged. Accordingly, the likelihood of a subject judging that a letter occurred 12 rather than 6 times would increase when the subject recalled more than one string containing the letter. As above, the potential utility of this strategy stems from high-frequency letters occurring in more different strings than low-frequency letters. Multiple recall was therefore more likely for the high- than the low-frequency letters.

In parallel with the preceding analysis, we computed the proportion of subjects for whom each letter appeared in at least two correctly recalled strings. In contrast with the preceding analysis, a letter appeared in more than one correctly recalled string very infrequently. The mean probabilities of multiple-letter recall, averaging across high- and low-frequency letters, were .02, .06, and .17 for the brief, intermediate, and long presentation durations, respectively. These probabilities were too low to permit the comparison of high- and low-frequency letters that was the basis for the preceding analysis. Instead, we deleted subjects' frequency judgments for letters that subsequently appeared in more than one of their correctly recalled strings. The results for judgments of letter frequency remained the same. Thus, there was no support for the hypothesis that subjects' frequency judgments for letters were based on the number of strings they could recall that contained the letter being judged.

Experiment 1A

The evidence obtained when the exposure duration was 0.2 s showed that subjects could accurately judge string frequency under conditions that prohibited accurate discrimination of letter frequency. The purpose of this experiment was to replicate that finding.

Method

As indicated above, the exposure duration was 0.2 s per string (the time between exposures remained at approximately 0.8 s). The design and procedure were identical to Experiment 1. A group of 32 subjects voluntarily participated in the experiment. All were students in an undergraduate psychology class at Florida Atlantic University. Half judged letter frequency and half judged string frequency.

Results

Letter-frequency judgments. The overall percentage of letter-frequency judgments that were correct was 52%, which was effectively at chance. The proportions of responses for which the letters were judged to have occurred 12 times are presented in Table 2. As in the brief exposure condition of Experiment 1, subjects' judgments did not differentiate between letters presented 12 times and letters presented 6 times (their proportion of correct frequency judgments was at chance). This was the case for letters from both high- and low-frequency strings. An ANOVA on the proportion of "12" responses in subjects' letter judgments indicated that the effect of letter frequency, $F(1, 15) < 1.0$, $MS_e = .042$, the effect of string frequency, $F(1, 15) < 1.0$, $MS_e = .090$, and the interaction between letter frequency and string frequency, $F(1, 15) < 1.0$, $MS_e = .055$, were not significant. Identical results were obtained when the letters used in the experiment replaced subjects as the random variable in the ANOVA. The effects of letter frequency, string frequency, and the interaction between letter frequency and string frequency were not significant, $F(1, 14) < 1.0$, $MS_e = .024$ (for all of the above).

String-frequency judgments. The overall percentage of string-frequency judgments that were correct was 66%. The proportions of responses for which the strings were judged to have occurred six times are presented in Table 3. As in Experiment 1, subjects' responses differentiated between the strings presented six times and the strings presented three times. An ANOVA indicated that the effect of string frequency on the proportion of "6" responses in subjects' string judgments was significant, $F(1, 15) = 27.33$, $p < .001$, $MS_e = .058$. Although string frequency judgments were significantly higher for strings with low-frequency compared to high-frequency letters, $F(1, 15) = 5.06$, $p < .05$, $MS_e = .034$, the extent to which subjects' judgments discriminated between the two levels of string frequency was not affected by the frequency of the letters composing the strings. That is, the interaction between letter frequency and string frequency was not significant, $F(1, 15) < 1.0$, $MS_e = .023$. Similar results were obtained when items replaced subjects as the random variable in the ANOVA. The effect of string frequency was significant, $F(1, 14) = 11.42$, $p < .005$, $MS_e = .035$, and the interaction between string frequency and letter frequency was not significant, $F(1, 14) < 1.0$, $MS_e = .035$. In contrast to the analysis with subjects serving as the random variable, the main effect of letter frequency was not significant, $F(1, 14) = 1.17$, $p > .05$, $MS_e = .035$. The final analysis contrasted frequency judgments for high-frequency strings composed of low-frequency letters with low-frequency strings composed of high-frequency letters. Judgments were significantly higher for the high-frequency strings, $t(15) = 7.09$, $p < .001$.

Discussion

The results closely replicated those obtained in the comparable 0.2-s exposure duration of Experiment 1; subjects' judgments discriminated string frequency under conditions that prohibited the discrimination of letter frequency.[3] Subjects could not have

[3] A potential objection to this conclusion was that subjects may have had knowledge of letter frequency for briefly presented strings, but it was more difficult to discriminate 12 from 6 events, as required for individual letters, than it was to discriminate 6 from 3 events, as was required for strings. An additional experiment was conducted to evaluate this possibility. The stimuli in this experiment were individual letters, half of which were presented 12 times each and half presented 6 times each. The exposure duration was 0.2 s. The results of the experiment indicated that the difference in the proportion of "12" responses between the high-frequency letters (.65) and low-frequency letters (.32) was comparable to the difference in the proportion of "6" responses between the high- and low-frequency strings in the 0.2-s exposure conditions of Experiments 1 and 1A. We could therefore rule out the possibility that subjects failed to discriminate letter frequency for 0.2-s exposures in Experiments 1 and 1A because of the relative difficulty inherent in discriminating between 12 and 6 events. Although several explanations are possible, it seems most likely to us that differences in frequency judgments for strings and letters were due primarily to the relative ease of retaining information for one string relative to retaining information for four independent,

based their judgments of string frequency on the frequency of occurrence of their component letters because they apparently had no knowledge of the letter frequencies.

One aspect of the data that is difficult to explain is that strings tended to be judged higher in frequency when they were composed of low-frequency compared with high-frequency letters. This difference, which was not significant in Experiment 1, did not directly contribute to the discrimination of letter frequency because string frequency and letter frequency were varied orthogonally in both experiments. In addition, the direction of the difference was opposite what would be expected if the presence of high-frequency letters in a string biased subjects to judge the string as high in frequency. Because it was not reliable when subjects was the random variable in the analysis, it may be that the effect of letter frequency on string frequency judgments was due to uncontrolled differences in familiarity among some of the strings.

Experiment 2

The purpose of this experiment was to replicate and generalize the finding that letter-level frequency information can be abstracted from sequences of strings containing the letters. The replication extended the results of Experiment 1 in the following ways: (a) Subjects' task during the initial presentation of the strings involved making a lexical (word/nonword) decision for each item rather than remembering the information presented on the screen, (b) there were four rather than two frequency levels for the target letters, (c) instead of immediate testing, a 30-min delay was introduced between the initial presentation of the strings and the letter-frequency test, and (d) a partial-correlation procedure was used to show that subjects' ability to estimate letter frequency did not depend on their frequency estimates being based on the number of strings that they could recall that contained the letter being judged.

Method

Subjects. Sixteen undergraduate students in psychology classes at Florida Atlantic University participated in the experiment, for which they received class credit.

Stimuli. Sixteen consonants were selected as target letters for the experiment. For one stimulus set, eight of the target letters appeared only in common, four-letter words and the other eight appeared only in pronounceable four-letter nonwords. Each set of eight letters was further subdivided into four subsets of two letters, with each subset assigned to one of four frequency levels (4, 8, 16, and 32). The average frequencies of usage (i.e., the background frequencies) were similar for each pair of letters (based on the Mayzner & Tresselt, 1965, norms). For the second stimulus set, the eight target letters appearing in words and the eight target letters appearing in nonwords were switched. The frequency levels to which the letters were assigned were also changed.

Every string contained either one or two target letters. A target letter never appeared more than once in any string, and each string appeared only once in each stimulus set. One stimulus set had 100 words and 100 nonwords. The other stimulus set had 40 more filler items (evenly divided between words and nonwords), which were strings composed entirely of nontarget letters. The strings belonging to each stimulus set were presented in two random orders (one order was the reverse of the other), resulting in four different stimulus sequences. Each subject was presented one of these sequences. Letter repetitions were separated in each stimulus sequence by a minimum of two intervening strings. Preceding each experimental sequence were 40 randomly ordered practice strings (half words, half nonwords). None of the practice strings was presented more than once, and none included any of the target letters.

Procedure. The stimuli were displayed on an Electrohome black and white television monitor that was controlled by a Data General Eclipse computer. Each string was presented inside a small rectangular box that always remained on the screen. The box intercepted a visual angle of 1.1° vertically and 2.8° horizontally. Each string intercepted a visual angle of 0.7° vertically and 2.4° horizontally. The exposure duration for each string was 1 s. The interstimulus interval was 1 s, except for the occasional trials for which the subject required more than 1 s to respond. Then, a 1-s delay was introduced between the subject's response and the presentation of the next stimulus. Subjects were provided with a response box and instructed to press the button marked "yes" if the string presented was a word. They were to press the button marked "no" if the string presented was not a word. Subjects were told to respond as quickly as possible, but to keep their errors to a minimum (incorrect responses were signalled by a brief flash of the rectangular box on the screen). They received no instructions suggesting that there would be any sort of memory test. When asked at the conclusion of the experiment, none of the subjects indicated that they expected to receive a memory test. Intervening between the lexical-decision task and the letter-frequency test was a 30-min interval during which subjects participated in a choice-reaction-time task involving judgments of visual direction.

During the letter-frequency test, subjects were presented with one of two random orders of all 26 alphabet letters (one order was the reverse of the other). Accompanying each letter was a number. This number was at the midpoint of the range of frequency values for all the target and nontarget letters appearing in the stimulus sequence. For two of the stimulus sequences, the range of letter frequencies was 1 to 90. For the longer two stimulus sequences the range was 1 to 130. Subjects were required to adjust the number appearing on the monitor upward or downward, using the same response buttons as in the preceding lexical-decision task, to reflect their estimate of how often each letter appeared over the entire sequence of strings. They were given all the time they needed for each response before the next letter was presented.

Results

Mean reaction times for subjects in the lexical-decision task were 632 ms for "yes" responses and 680 msec for "no" responses. The advantage in processing time for "yes" responses is typical of the lexical-decision paradigm.

Correlations between actual letter frequency and estimated letter frequency were computed for each subject. One correlation was computed for the eight target letters assigned to the words and another for the eight target letters assigned to the nonwords. The mean correlations were .41 for the words and .40 for the nonwords. The overall mean was significantly greater than zero, $t(15) = 5.78$, $p < .001$ (of the 32 correlation coefficients that were computed, 28 were positive).[4]

simultaneously occurring letters. Regardless of why frequency judgments failed to discriminate between the two levels of letter frequency (for letters embedded in strings) under conditions for which they successfully discriminated between the two levels of string frequency, we could conclude that subjects' frequency judgments for strings were not derived from the frequencies of their constituent letters.

[4] Correlation coefficients were used as descriptive statistics in this experiment. The computation of mean correlation coefficients and statistical analyses of individual correlation coefficients were based on Fischer's r to z transformation.

As in Experiment 1, recall protocols were used to evaluate the likelihood that subjects' letter-frequency judgments were based on the retrieval of global-level memory representations of the strings containing the letters. We adopted the criterion that three of the four letters in the string had to be correct before the string was considered to be correctly recalled and then counted the frequency with which each target letter appeared in a subject's correct-recall protocols. These recall frequencies were correlated with both the actual and estimated letter frequencies, and a partial correlation coefficient was computed between the actual and estimated frequencies. Using this procedure, the relationship between actual and estimated frequency was determined, with the effects of recall frequency held constant (McNemar, 1962).

Partial correlation coefficients for each subject were computed separately for the eight target letters assigned to the word and the eight target letters assigned to the nonwords. The mean correlation coefficients, .46 for the words and .40 for the nonwords, were not statistically different, $t(15) < 1.0$. However, the overall mean, $r = .43$, was significantly greater than zero, $t(15) = 4.75$, $p < .001$ (of the 32 partial correlation coefficients that were computed, 27 were positive).

Discussion

The results of this experiment extended the previously reported evidence for the abstraction of letter-level frequency information. Subjects' estimates discriminated between the frequencies of letters as they appeared in both words and pronounceable nonwords during a lexical-decision task. The partial correlation procedure showed that subjects' success at letter-frequency estimation could not be attributed to the derivation of their estimates from the recall of the strings containing the letters. Furthermore, the introduction of a 30-min delay between the presentation of the strings and the letter-frequency test showed that the letter-frequency information was stored in long-term memory. This result paralleled Warren and Mitchell's (1980) evidence for the retention of string-level frequency information over a 20-min delay. Substantial losses in string-level frequency information have been reported for 1-week delays (Underwood, Zimmerman, & Freund, 1971). Whether or not extended delays would result in similar losses in abstracted letter-level frequency information remains to be determined.

General Discussion

Our basic evidence for the storage of element-level units was that high-frequency letters were judged as high in frequency significantly more often than low-frequency letters (for exposure durations of 1.2 and 4.2 s). We then considered the counterargument that frequency information was not associated with letter-level memory units and that subjects based their letter judgments on the activation of global-level memory units containing the letters. We noted first that the stimulus list was designed such that letter and string frequency were varied orthogonally, so that string frequency had no predictive value for judgments of letter frequency. Furthermore, high-frequency letters were judged as higher in frequency than low-frequency letters (for the 1.2- and 4.2-s durations) even when the high-frequency letters appeared in low-frequency strings and the low-frequency letters appeared in high-frequency strings. Because differences in string frequency would have favored the activation of global-level units with low-frequency letters, this comparison provided evidence against the argument that element-level frequency judgments depended on the activation of global-level memory units. The final evidence for this conclusion came from an analysis of subjects' recall protocols. As expected from the structure of the stimulus list, subjects recalled more strings with high-frequency than low-frequency letters (there were more of the former to recall). However, this was the case even at the 0.2-s exposure duration. Despite this indication of greater availability of global-level memory units containing high-frequency letters compared with low-frequency letters, subjects still failed to discriminate between high- and low-frequency letters following 0.2-s exposures of the strings. It could be concluded, at least for the longer exposure durations of Experiment 1, that subjects abstracted letter-level frequency information that characterized the componential structure of the stimulus list. Further evidence indicating that this information was stored in long-term memory was reported in Experiment 2.

The letter judgment data provided some indication that Tversky and Kahneman's (1973) "availability heuristic" was operative in our experiments. That is, there was some bias for subjects to judge letters as relatively high in frequency simply because they occurred in more available high-frequency strings (the main effect of string frequency on letter judgements was significant in Experiment 1 but fell short of significance in Experiment 1A). However, the orthogonal manipulation of letter and string frequency prevented this bias from being the basis for subjects' discrimination between high- and low-frequency letters.

Experiments involving judgments of string frequency generally do not control for the frequency of occurrence of the letters composing the strings. The above evidence for letter-frequency discrimination suggests the possibility that frequency judgments for strings might be based, not on their frequency of occurrence, but on the frequency of occurrence of their constituent letters. Although the use of such a strategy is possible, the stimulus list was designed so that string frequency could not be predicted from letter frequency (they were varied orthogonally). Furthermore, high-frequency strings were judged as higher in frequency than low-frequency strings even when the former were composed of low-frequency letters and the latter of high-frequency letters. The opposite result would have been obtained if judgments of string frequency were based on the frequencies of the strings' constituent letters. Finally, string-frequency judgments discriminated between high- and low-frequency strings under experimental conditions (the 0.2-s exposure duration) for which letter-frequency judgments did not discriminate between high- and low-frequency letters. Under these conditions, it was impossible for string frequency to be judged on the basis of stored information involving letter frequency because the latter information was not available (as measured by judgments of letter frequency).

To summarize, our experimental results provided evidence for the independent formation of both global-level (strings) and element-level (letters) memory units. Although this conclusion refers to the sort of information subjects remember, the results are potentially informative concerning the way in which this information is retrieved. That is, the tendency for string-frequency information to bias letter-frequency judgments suggests that both types of information may be retrieved together, even when the task calls only for the judgment of letter frequency. (The case for joint retrieval when subjects judge string frequency remains uncertain because these judgments appeared to be influenced only by string-level information.) The joint retrieval of string-level and letter-level memory units might suggest, but does not demand, that they are stored together. One possibility is that representations involving global-level (string) memory units include individual element-level units (the spelling for the string), with each of the elements tagged with regard to its frequency of occurrence in the list. Whether one can distinguish between this and other representational formats (e.g., separate storage of global-level and element-level units) is questionable. This indeterminacy is analogous to the indeterminacy regarding the representational format for item-specific, exemplar information (global-level units in our study) and category-level, summary information (element-level units in our study) in category-acquisition tasks (Medin, Dewey, & Murphy, 1983). In the absence of experimental paradigms that definitively distinguish among alternative representational formats, our further research involving the formation of memory units has replicated the results reported in this article and focused on the task characteristics that influence the retention of informational units of different size. For example, we are finding that the phonemic processing of strings of letters facilitates the formation of memory units for the letters (Malcus, Hock, Cavedo, & Smith, 1985).

Evidence that subjects can accurately judge the frequency of occurrence of component elements of larger order units is not exclusive to this article. Jacoby (1972) showed that subjects can judge the frequency of occurrence of words embedded in grammatical sentences. Investigators testing feature-frequency models of concept formation have shown that subjects can judge frequency of occurrence for the component parts of schematic faces (Kellogg, 1981). In work in progress, we are obtaining evidence that subjects presented with a sequence of dot patterns can accurately judge how often individual locations have been occupied by a dot. In other work in progress, we are finding that subjects can discriminate differences in the frequency with which various spatial relations (e.g., inside, below) occur across a series of scenes involving different objects. However, obtaining evidence for successful frequency discrimination does not definitively indicate that subjects have stored frequency information regarding elements and relations between elements. It may be necessary to show, as in the present article, that subjects' frequency judgments for the componential characteristics of a series of stimuli are not based on the activation of global-level memory units corresponding to the particular stimuli presented in the series.

The frequency-judgment procedure seems to be particularly useful for assessing memory units, especially when the vocabulary of elementary units is limited, as is the case for alphabet letters. For example, when frequency judgments are used to assess memory for the constituent letters in strings, the full set of alphabet letters can be presented during the acquisition phase of the experiment. In contrast, recognition and recall procedures would have to exclude a significant proportion of the alphabet from the acquisition phase because they must be used to detect false recognition responses or recall intrusions during the subsequent memory test. Furthermore, the frequency-judgment procedure eliminates the need for production (as in recall paradigms) and also eliminates the problem of performance depending on the perceptual similarity of previously seen and new distractor stimuli (as in recognition paradigms).

References

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

Burnett, S. A., & Stevenson, J. H. (1979). Instructional effects on frequency judgments for sentence meanings. *American Journal of Psychology, 92*, 711–721.

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

Gude, C., & Zechmeister, E. B. (1975). Frequency judgments for the 'gist' of sentences. *American Journal of Psychology, 88*, 385–396.

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

Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology, 88*, 297–306.

Hochberg, J. (1981). Levels of organization. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, NJ: Erlbaum.

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

Jacoby, L. L. (1972). Context effects on frequency judgments of words and sentences. *Journal of Experimental Psychology, 94*, 255–260.

Kellogg, R. T. (1981). Feature frequency in concept learning: What is counted? *Memory & Cognition, 9*, 157–163.

Malcus, L., Hock, H. S., Cavedo, L. C., & Smith, L. B. (1985, March). *Orthographic processing of strings affects frequency discrimination for component letters*. Paper presented at the meeting of the Eastern Psychological Association, Boston, MA.

Mayzner, M. S., & Tresselt, M. E. (1965). Tables of single letter and diagram frequency counts for various word-length and letter-position combinations. *Psychonomic Monograph Supplements, 1*(No. 2), 13–78.

McNemar, Q. (1962). *Psychological statistics*. New York: Wiley.

Medin, D. L., Dewey, G. I., & Murphy, T. D. (1983). Relationships between item and category learning: Evidence that abstraction is not automatic. *Journal of Experimental Psychology: Learning, Memory and Cognition, 9*, 607–625.

Neisser, U. (1967). *Cognitive psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Rowe, E. J. (1973a). Context effects in judgment of frequency. *Bulletin of the Psychonomic Society, 2*, 231–232.

Rowe, E. J. (1973b). Frequency judgments and recognition of homonyms. *Journal of Verbal Learning and Verbal Behavior, 12*, 440–447.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207–232.

Underwood, B. J. (1969). Attributes of memory. *Psychological Review, 76*, 559–573.

Underwood, B. J., Zimmerman, J., & Freund, J. S. (1971). Retention of frequency information with observations on recognition and recall. *Journal of Experimental Psychology, 87*, 149–162.

Warren, L. R., & Mitchell, S. A. (1980). Age differences in judging the frequency of events. *Developmental Psychology, 16*, 116–120.

Received October 15, 1984
Revision received July 1, 1985 ■