Journal of Experimental Psychology:
Learning, Memory, and Cognition
1992, Vol. 18, No. 2, 368-378

0278-7393/92/$3.00

Direct Coding for Frequency of Occurrence

John Jonides
University of Michigan

Caren M. Jones
Stanford University

Using an interference paradigm, we demonstrate that there is a place for a direct coding mechanism in a comprehensive theory of frequency coding. Ss were presented words whose frequency was judged later. Under one set of instructions, these words were coded in terms of numerical associates; under another set of instructions, the coding was governed by nonnumerical associates. The condition using numerical associates resulted in frequency estimations that were of lesser quality than those produced in the control condition. This effect, moreover, was a function of the encoding of the target words, not just their retrieval.

People are quite good at remembering how often events occur. Ask about the relative numbers of many kinds of events, and you are likely to get answers that reflect the actual relative frequencies of the events with great fidelity. In one study, for example, Shedler, Jonides, and Manis (1985) showed that, when queried about the relative number of restaurants in various fast-food chains, people's estimates bore a close relation to the actual number of extant outlets.

Models of frequency judgments have focused on the coding processes involved in retaining information about the frequency of occurrence of events. One class of theory has become quite popular as an account of the representation of such frequency information. This class subscribes to the assumption that information about the frequency of occurrence of an event is derived from aspects of the memory for an event other than its frequency per se, such as the number of traces of that event that are stored when it occurs repeatedly (Hintzman, 1988; Howell, 1973).

One line of evidence for such indirect coding models comes from experiments in which variables thought to have effects on memory traces for events are also shown to have similar effects on frequency judgments. One example of such a variable is instruction about the level of processing that should be applied to items in a list. Ample documentation exists that semantic coding of list items leads to better recall than does coding that is tied to some physical aspect of the stimuli (Craik & Lockhart, 1972). It has also been shown that semantic coding (compared with more superficial coding) results in frequency estimations with greater fidelity to actual frequency (Fisk & Schneider, 1984; Greene, 1984; Jonides & Naveh-Benjamin, 1987; Naveh-Benjamin & Jonides, 1985, 1986; Rose, 1980; Rose & Rowe, 1976; Rowe, 1974; Rowe & Rose, 1977). Thus, there is a similarity between the effects of semantic coding on both memory for an event and on frequency judgments for that event, leading to the claim that frequency is represented by some aspect of the memory code, such as the multiplicity of traces stored. Examples of this sort strengthen the case for a model proposing indirect coding of frequency by means of some aspect of the memory traces of the events other than a direct coding of frequency per se.

An alternative to indirect coding models rests on the assumption that frequency is a separately and *directly* coded attribute of memory traces (e.g., Underwood, 1969). According to models of this class, the registration of a trace about some event has, as one of its components, a frequency tag that is incremented with successive appearances of the same event (see Hintzman & Stern, 1978; Howell, 1973; Underwood, 1969, for descriptions of such a model).

A version of a direct coding model that is not particularly attractive is based on the notion that the direct code is updated using explicit counting. Although such a mechanism may have some usefulness in certain limited circumstances, counting is not a skill that can be applied generally and with precision to the range of tasks for which humans and other animals show impressive frequency coding ability (see Fisk & Schneider, 1984; Flexser & Bower, 1975; Greene, 1984; Jonides & Naveh-Benjamin, 1987; Naveh-Benjamin & Jonides, 1986; Zacks, Hasher, & Sanft, 1982, for relevant discussions). On a task like those presented here, subjects given explicit counting instructions along with paper and pencil aids performed no better and no worse than those without such instructions and aids (see Greene, 1986; Jones, 1990).

By contrast, there is reason to take seriously a version of direct coding that is more subtle. According to this version, frequency information about an event is accumulated automatically on successive appearances of that event. Evidence on the automatic updating of frequency information is somewhat mixed. However, even though recent studies cast doubt on a *wholly* automatic updating process, there is still reason to believe that at least *some* contribution to the coding of frequency may be mediated automatically (e.g., Birnbaum, Taylor, Johnson, & Raye, 1987; Hasher & Zacks, 1979, 1984; Naveh-Benjamin & Jonides, 1986; Sanders, Gonzalez, Murphy, Liddle, & Vitina, 1987). If such automatic coding is still a viable possibility, then an automatic mechanism based on direct coding of frequency may also be viable.

Tests of a direct coding theory have been based on a compelling empirical logic: If a code for frequency is stored in the form of numerical information, then it ought to be

---

We thank David Bryant, Robert Greene, Robert Rose, and William Whitlow for their incisive comments on an earlier version of this article.

Correspondence concerning this article should be addressed to John Jonides, Department of Psychology, University of Michigan, 330 Packard Road, Ann Arbor, Michigan 48104.

possible to interfere with this numerical information by simultaneous presentation of other numerical data that must be stored in memory. This is the rationale that has guided tests by Whitlow and Skaar (1979), Hintzman (1982), and Hintzman, Nozawa, and Irmscher (1982).

Whitlow and Skaar (1979) tested whether spatial numerosity (the number of occurrences of an event within a frame) and total frequency are mutually exclusive, noninterfering aspects of memory. If they are, then manipulation of either should not affect the storage or recall of the other. Whitlow and Skaar attempted to influence frequency judgments by varying the correlation of numerosity and total frequency of an event. Hence, for half of the subjects, a stimulus with a high numerosity also had a high total frequency, whereas for the remaining subjects, a stimulus with a high numerosity had a low total frequency. If frequency and spatial numerosity are noninterfering aspects of memory, then subjects should have been able to choose the more frequent stimulus of a test pair regardless of the relative numerosities.

Letter strings were presented in one of two lists. In one list, numerosity was positively correlated with total frequency, whereas in the other list the correlation was negative. Subjects were presented with only one of the lists and subsequently were asked to make relative frequency judgments of letter pairs.

It was found that subjects showed a tendency to choose the letter with the highest numerosity as being the more frequent regardless of actual total frequency. In addition, accuracy on the positively correlated list was generally higher than on the negatively correlated list. These two results suggest that there may have been interference of the numerosity information on internal codes of the information about total frequency of occurrence. This evidence for the intrusion of a numerical associate on frequency information suggests that frequency data may be stored in a numerical format. This is exactly the storage format implied by a direct coding mechanism.

A similar rationale was used by Hintzman (1982): Distinct aspects of memory should not intrude on each other if they are separately coded. That is, variations in one aspect of memory should not influence the recall of another aspect. Hintzman attempted to show that frequency information is stored in a form different from that of other numerical associates by varying both the spatial numerosity and the presentation frequency of stimulus words. He found no evidence for any intrusion of numerosity on frequency estimates.

In another attempt to show that frequency information is stored in a unique manner, Hintzman et al. (1982) used the same experimental logic: Distinct aspects of memory should not intrude on one another. In this case, the experimenters examined the difference between presentation frequency and the magnitude of digit associates. Picture–digit pairs were used as stimuli, with both digits and stimulus frequencies in the range of 1 to 5. Subjects were asked to learn the digit associate for each picture. One half of the subjects were then given the expected associate recall test, whereas the rest were given an unexpected absolute frequency judgment test before the associate test.

Frequency judgments were fairly accurate and were not affected by the digit associate, but the accuracy of digit recall was reduced when frequency estimates were made before the digit recall test. The fidelity of the frequency estimates indicates that the encoding of the digit associates had no effect on the encoding of frequency. In addition, the lack of intrusion of the stored digit associates on the frequency estimates during frequency recall indicates that frequency information was stored in a manner different from that of other numerical associates.

These studies lead to no solid conclusion about the plausibility of a direct coding mechanism. Whitlow and Skaar (1979) found some interference between numerosity and frequency, but Hintzman (1982) and Hintzman et al. (1982) found no evidence of interference in the encoding of frequency by other digit information (although there was some evidence of interference at the time of retrieval).

The present article reports the results of three experiments that continue to explore a direct coding mechanism, again making use of the rationale that other digit-related information about an event might interfere with a frequency code for that event if that code is represented in the form of a count.

There are two reasons to continue investigation of a direct code given past results. The first is that some influence of direct coding seems apparent in the data of Whitlow and Skaar (1979). The second is that the interference techniques used by Hintzman et al. (1982) may not have been wholly effective. There is the possibility in those studies that the frequency and digit codes were sufficiently different in internal form that there was not adequate opportunity for the two codes to interfere mutually. The frequency code was generated internally by subjects, whereas the digit code was given by the experimenter. This difference in origin of the codes may have rendered them sufficiently different that they were separable in memory and therefore did not maximally interfere with one another.

With that problem in mind, we designed a technique that would allow both the frequency and the digit codes to be generated internally by subjects, thereby overcoming any differences in origin. Of course, it is straightforward to have subjects generate a frequency code internally: They are assumed to do so automatically according to the most plausible version of a direct coding model. To include a potentially interfering numerical code that was also generated internally, we used as stimuli words that themselves have numerical values associated with them. These words (as well as other words that do not have a characteristic numerical value associated with them) were presented at several frequencies. Subjects were instructed to generate associates to each word as the words were presented, and they were asked to estimate the frequency of presentation of each word after presentation of them all. We then examined whether there was interference on the frequency estimates by the numerical associates of the words.

Experiment 1

The purpose of Experiment 1 was to determine if we could establish an interference effect on frequency judgments by introducing sources of numerical information into the experimental setting. To this end, we chose words that have characteristic numerical associates, and we varied the frequency of presentation of these words. In addition, we included instructions to subjects that exhorted them to process the words with either numerical or word associates in mind. Would the characteristic numerical associates of the words influence the frequency judgments for these words especially under conditions in which the numerical associates were made salient? This was the issue at hand.

Method

Subjects. The subjects were 44 undergraduates of the University of Michigan who participated as part of a course requirement.

Materials. The pool of critical words used in constructing the stimulus list consisted of 12 words previously found to have naturally and strongly associated numerical values (couple, few, lone, quartet, alone, solo, triplicate, quadrangle, pair, wheels, triplets, twins).

These words were generated in a two-stage preliminary procedure. First, 24 subjects were asked to list as many words as possible that they considered to have a natural numerical value. The 104 most frequently generated words were given to a different set of 24 subjects who were asked to indicate the numerical value associated with each word. Of course, not every subject assigned the same numerical associate to each of the words in this pool. Those words having associates within the range of digits 1 through 4 and with variances in their values of less than 0.44 qualified to be critical words in the main experiment.

There were at least four satisfactory words at each of the numerical values. However, for the values 3 and 4, three of the four satisfactory words began with the prefixes "tri" or "qua," respectively. Such repetition seemed to draw attention to the fact that the words had numerical values. To minimize this problem, three words at each associate value were randomly selected with the constraint that not all of the words having a particular value could begin with the same prefix. These words became the 12 critical words for the experiment proper.

The 3 words each from Associate Categories 1, 2, 3, and 4 were assigned to presentation frequencies 1 through 4 as shown in Table 1. Table 1 shows that an orthogonal combination of frequency and association value was approximated. Because there were only 3 critical words for each association, there was not an opportunity to complete the orthogonal assignment. In addition to the critical words, 6 filler words were assigned to each frequency. These words had no known associated numerical values as mutually judged by the authors. The mean linguistic frequencies of the filler and the critical words were equated. An additional 10 filler words were used with presentation frequency of 1 at the beginning and end of the presentation sequence to absorb primacy and recency effects.

All of the words at their various frequencies were photographed one to a slide. They were then arranged in random order for presentation with two constraints. First, the 10 filler words used to overcome primacy and recency effects were placed such that 5 began the sequence and 5 ended it. Second, repetitions of a word were separated by at least three slides of other words. Altogether, 100 stimuli were presented: 9 words of the 4 frequencies and 10 additional words of frequency 1.

Procedure. Four groups of subjects were tested. Subjects in all groups began their session by viewing the 100 slides containing the stimulus words at a rate of 5 s per word. During this presentation phase, subjects in all four groups were given a booklet consisting of 20 pages with five blank lines on each page. Two of the groups (the word-encoding conditions: $n = 11$, $n = 14$) were instructed to write down on each line the first word that came to mind when they saw a stimulus slide. The other two groups (the number-encoding conditions: $n = 10$, $n = 9$) were instructed to write down for each stimulus word the first *number* that came to mind when they viewed that word. Therefore, subjects in all groups were engaged in an associate-generation task as they viewed the stimulus slides. All of the subjects were asked to use the same associate on each appearance of a repeated word. The subjects were led to believe that they were engaged in a task to generate a new set of associative norms. This cover instruction seemed effective in focusing subjects' attention on the task of generating associates as we wished.

Table 1
Approximation of Orthogonal Combination of Presentation Frequency and Associated Numerical Values for the Critical Words of Experiment 1

Associated numerical values	Presentation frequency			
	1	2	3	4
1	+	+	+	—
2	+	+	—	+
3	—	+	+	+
4	+	—	+	+

Note. A cross indicates that a word of this type was used.

Immediately after the slide presentation, the associate booklets were collected. One group of subjects in each encoding condition was then given an unexpected written test of recall for the associates they had generated. For this task, they were asked to indicate the associate that they had written in their answer booklet for each stimulus word. For this group of subjects, the associate test was immediately followed by an unexpected test of the frequencies with which each word had been presented: Subjects were given each word (other than the 10 primacy and recency words) and were asked to estimate its frequency of presentation. The second group of subjects in each encoding condition received the frequency test first and then the test of associate recall.

Results and Discussion

Effects of instruction. Figure 1 plots the frequency estimates as a function of actual frequency for both the word-encoding condition (Figure 1a) and the number-encoding condition (Figure 1b). Both panels make it clear that subjects were able to estimate the frequency of appearance of the words with considerable fidelity. This was confirmed by an omnibus analysis of variance (ANOVA) including the factors of frequency, order of testing on recall versus frequency, encoding condition, and work type. This analysis showed a very reliable increase of estimates with actual frequency, $F(3, 120) = 283.63$, $p < .001$, $MS_e = 0.137$.

We found no reliable effect of the order in which subjects were tested for frequency versus associate recall, $F(1, 40) = 2.28$, $p > .10$, nor did we find any significant interactions of test order with any other variable ($p > .10$). Our failure to find any reliable influence of order led us to collapse over this variable in further analyses.

The fidelity of the estimates differs between the two types of words. This is shown in two ways. First, the estimates for the filler words are closer to the true frequencies in absolute magnitude than are the estimates for the critical words. Second, the slopes of the estimates for the filler words are steeper than the slopes for the critical words.

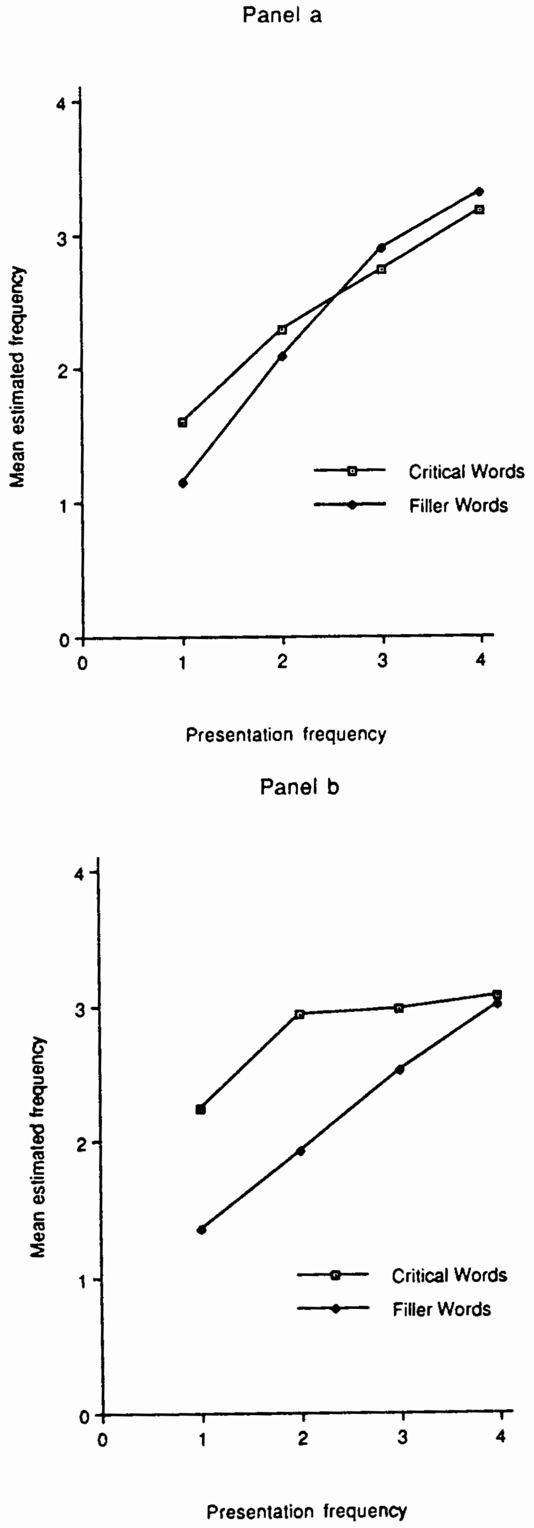

Figure 1. Estimated versus actual frequency for the critical and filler words of Experiment 1 for (a) the word-encoding condition and (b) the number-encoding condition.

The effect of word type on absolute magnitude was confirmed by the results of a three-way ANOVA on the estimates. The main effect of word type in this analysis was reliable, $F(1, 42) = 41.12, p <.001, MS_e = 0.255$, and readers can see from Figure 1 that this is because the estimates for the filler words were nearer the true frequencies. The slope difference was also confirmed by statistical analysis. Slopes were estimated for each subject's functions, and they were compared for critical words ($M = 40, SD = .25$) and filler words ($M = .66, SD = .18$). As shown in the figure, the comparison confirmed that the slopes of the estimates for the filler words were steeper, $F(1, 86) = 30.49, p < .001, MS_e = 0.048$.

A highly significant Word Type × Encoding Condition interaction, $F(1, 42) = 22.72, p < .001, MS_e = 0.255$, indicated that encoding condition had the most influence on the judgments for the critical words. Consider now the data for these critical words only. These are the data of central interest in that they offer the possibility of revealing an interference effect. A comparison of Figure 1a and 1b suggests just such an interference effect. It appears that when subjects judged the frequency of the critical words after they had produced numerical associates to those words, their judgments were compromised compared with the condition in which they generated word associates. This is confirmed by a two-way ANOVA, including the factors of encoding condition and frequency. This analysis showed a main effect of encoding condition, $F(1, 42) = 11.22, p < .002, MS_e = 0.512$, as well a Frequency × Condition interaction, $F(3, 126) = 6.97, p < .001, MS_e = 0.204$.

Experiment 1 was designed to find an effect of number encoding by including the critical words for which a natural numerical associate within the range of the actual frequencies was present. By virtue of the design, however, it was also possible to examine whether such an interfering effect of encoding might appear for the filler words as well. Indeed it does: A two-way analysis on the filler words comparable to the one for the critical words shows both a main effect of encoding condition, $F(1, 42) = 5.83, p < .05, MS_e = 0.259$, and a Frequency × Condition interaction, $F(3, 126) = 8.35, p < .001, MS_e = 0.090$.

Effect of numerical magnitude. These analyses on both the critical and filler words show, then, that the production of numerical associates reliably interfered with the later production of frequency judgments. Is this interference effect due to the *magnitudes* of the numbers involved, or is it due merely to the activation of a number *system* for both frequency coding and numerical association? We studied this issue by examining whether the numerical magnitude of the value associated with each critical word had a systematic relationship to the frequency estimate given. Recall from Table 1 that the critical words were assigned frequencies of appearance such that the natural associates of the words were systematically related to their frequencies. In some cases, there was a match between associate and frequency, and in some cases there was not. When there was not, furthermore, there were various values of difference (numerically from 1 to 3) between frequency and associate.

Our analysis required the definition of a metric that would relate frequency to associated value. We chose to define this "distance value" metric as the absolute value of the natural numerical associate minus the presentation frequency for that word. The larger this distance value, the greater the difference in magnitude between a word's frequency and its natural numerical associate. Having determined this for each word, we then examined the accuracy of the frequency estimates as a function of distance value. Accuracy in this case was simply the absolute value of the estimated frequency for a word minus its actual presentation frequency. These data are shown in Figure 2 in which accuracy is plotted against distance value. Note that there are four words with distance value of 0, five words with distance value of 1, two words with distance value of 2, and one word with distance value of 3 that contribute to these data.

One thing that is reiterated by Figure 2 is the difference in accuracy between encoding conditions: The number-encoding condition yields frequency estimates that deviate more from the actual frequencies than does the word-encoding condition. This is borne out by a two-way ANOVA including the factors of encoding condition and distance on these data. The ANOVA shows a significant main effect of encoding condition, $F(1, 42) = 30.00$, $p < .001$, $MS_e = 0.295$. Notice also that the functions have essentially zero slope in each condition; that is, the actual value of the distance between associate and frequency does not affect the amount of over- or underestimation, $F(3, 126) = .09$, $p > .88$. This leads to the conclusion that simple activation of a numerical system by generating associates causes the interference in frequency judgments. It is not the particular value of the associate in question that matters.

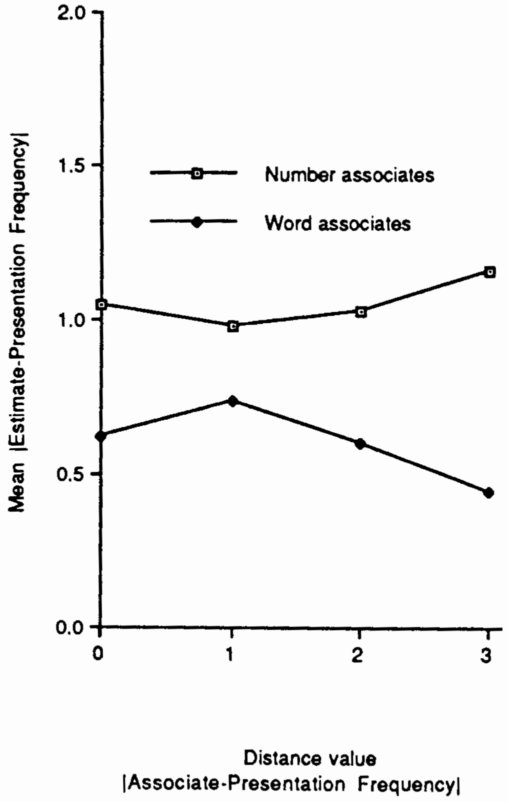

Figure 2. Accuracy of frequency estimates as a function of distance value for the critical words of Experiment 1. (Data are plotted for the conditions in which numerical and word associates were required.)

Number of associates generated. One possible basis for the general effect of numerical associates (shown in Figures 1 and 2) that must be investigated is that past research (e.g., Jonides & Naveh-Benjamin, 1987) showed that the greater the number of associates generated for a given word during presentation, the more accurate the frequency estimate for that word. To be sure, this effect is quite small for frequencies in the range used in the present experiment. Therefore, it is unlikely a priori that there would be much influence of number of associates on frequency estimates. Nonetheless, if the word-encoding condition resulted in the generation of more associates for each repeated stimulus word than the numerical condition did, this could provide an explanation for the differences in the frequency estimates given in the two encoding conditions. Recall that subjects were instructed to repeat their associate on repetition of a word, but in some cases they did not do so. To test for the possibility that the number of associates may have compromised the frequency results, the mean number of associates given for each repeated stimulus word was calculated.

Using mean number of associates generated as the dependent measure, an ANOVA was performed with encoding condition, presentation frequency (2, 3, or 4), and word type as the variables. This analysis showed no main effect of number of associates as a function of encoding condition, $F(1, 42) = 2.65$, $p > .10$, $MS_e = 0.879$. So it is unlikely that number of associates could account for the reliable effect of encoding condition on frequency estimates that we found.

This analysis did reveal two reliable effects. First, more associates were given for the filler words than for the critical words, $F(1, 42) = 4.67$, $p < .05$, $MS_e = 0.114$. This is not surprising because there were twice as many filler words as critical words and, hence, there were twice as many opportunities for extra associates to be generated. Second, the critical words in the word associate condition were given more associates than any other word type and encoding condition combination, $F(1, 42) = 26.72$, $p < .001$, $MS_e = 0.114$. This second result gives us another opportunity to test for an effect of number of associates on frequency judgments: If the accuracy of the frequency judgments was determined by the number of associates generated, then the most accurate frequency estimates should be for the critical words in the word-encoding condition because this was the condition that led to the most associates. A comparison of Figure 1a and 1b shows that this was not the case, however. In fact, the most accurate estimates were given for filler words in the word-encoding condition, the category of stimulus words given the second *fewest* associates.

Experiment 2

Experiment 1 demonstrated an interfering effect of numerical associates. The implication of this effect is that there is some contribution of a numerical processing mechanism to the generation of frequency judgments, and the task of generating numerical associates produces an interference with this process. However, what is the locus of the process? Two possibilities suggest themselves. One is that a numerical code created at the time of stimulus presentation has a role in the eventual judgment of frequency. This is what others have called a direct or propositional code for frequency (e.g., Hintzman, 1982; Hintzman et al., 1982). The other possibility is that the interference occurs only at the time that subjects produce the judgments and not at the time of encoding. This possibility is implied by the work of Hintzman et al. (1982) in which they documented a similar effect on the recall of experimenter-assigned digit associates. Experiment 2 was designed to discriminate whether the numerical interference effect was lodged only at retrieval or whether encoding processes were implicated as well.

Four groups of subjects were used. Two groups replicated Experiment 1 in that they generated and recalled either word or number associates. The other two groups generated associates of one type during the encoding of the stimuli and then were asked to generate the other type of associates immediately before making frequency estimates. If the interference occurs during encoding, then the type of associates generated immediately before the frequency test should not influence the estimates; only the type of associates generated at encoding should be influential. Interference only at retrieval, however, would imply that the type of associates generated during encoding would not affect the frequency estimates, whereas the associates generated at the time of judgment would impact the estimates.

Method

Subjects. The subjects were 65 University of Michigan students who were paid for their participation. None of the subjects had participated in Experiment 1.

Materials. The stimuli were identical to those used in Experiment 1.

Procedure. The procedure was identical to that of Experiment 1 except for a variation in testing. All subjects received a type of associate task followed by the frequency judgment test. For the associate retrieval task, one half of the subjects in each encoding condition ($n = 16$, $n = 16$) were asked to recall the associates they had generated during the slide presentation. The remaining subjects were asked to generate the alternative type of associates to the stimulus words: Subjects in the number-encoding condition generated word associates ($n = 18$), whereas those in the word-encoding condition generated numerical associates ($n = 15$).

Results and Discussion

The frequency estimates for the word and number-encoding conditions are shown in Figures 3 and 4, respectively. Notice that the slopes of the estimate functions are somewhat steeper overall in the word-encoding condition and that the estimates do not differ as a function of retrieval condition. It appears, then, that the interference effect is located at the encoding stage. These observations were confirmed by an omnibus ANOVA. A marginal effect of encoding condition was found $F(1, 16) = 3.05$, $p < .09$, $MS_e = 1.01$, but there was no main effect of retrieval condition ($p > .88$). A highly significant Word Type × Encoding Condition interaction, $F(1, 61) = 19.65$, $p < .001$, $MS_e = 0.339$, indicated that encoding condition had the most influence on the estimates for the critical words just as would be expected. These critical words are of primary interest.

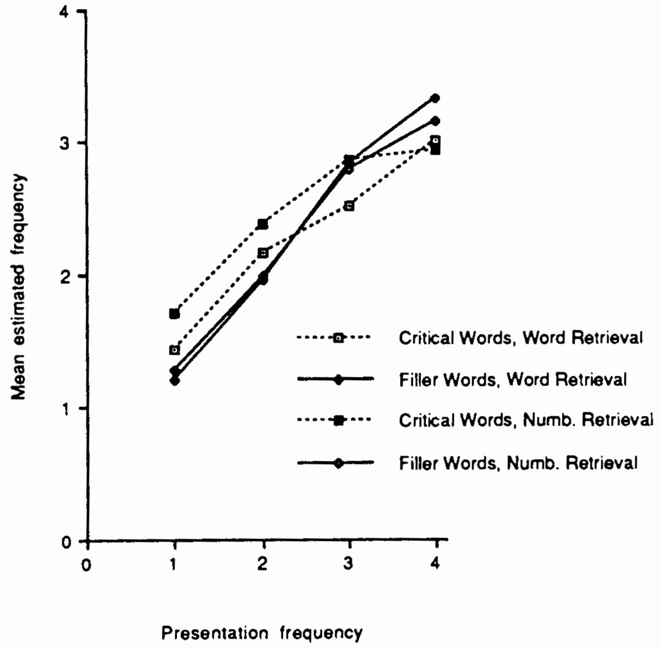

Figure 3. Estimated versus actual frequency for the critical and filler words in the word-encoding conditions of Experiment 2. (There were two conditions of retrieval: generation of either number or word associates to the stimuli.)

Consider the data for the critical words only. Notice two features of these data. First, the absolute magnitudes of the estimates differ as a function of encoding condition but not as a function of retrieval condition. Second, the slopes of the

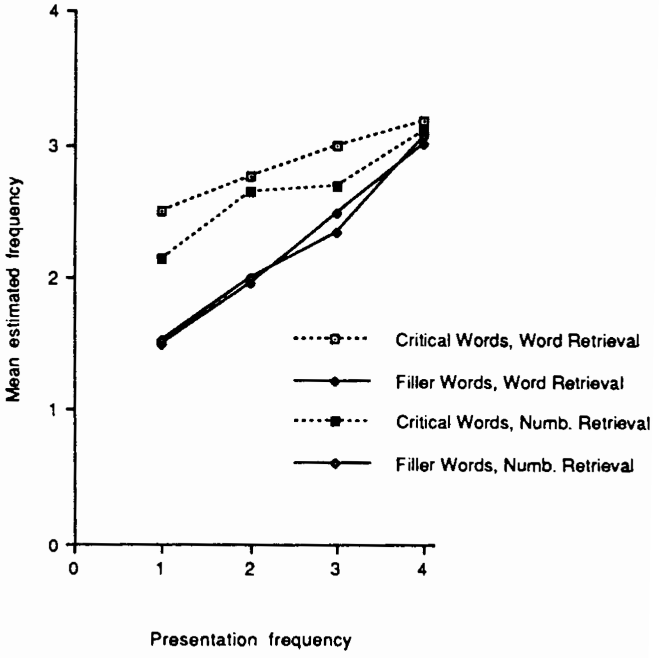

Figure 4. Estimated versus actual frequency for the critical and filler words in the number-encoding conditions of Experiment 2. (There were two conditions of retrieval: generation of either number or word associates to the stimuli.)

functions, shown in Table 2, also differ with encoding condition but not retrieval condition. Both of these effects were confirmed by ANOVA on the frequency judgments including factors of encoding condition, retrieval condition, and frequency: There was a main effect of encoding condition, $F(1, 61) = 12.30$, $p < .001$, $MS_e = 0.763$, but there was no effect of retrieval task ($p > .9$). An interaction between encoding condition and presentation frequency was also found, $F(3, 183) = 5.93$, $p < .001$, $MS_e = 0.222$. Of course, there was a main effect of presentation frequency as well, $F(3, 183) = 63.81$, $p < .0001$, $MS_e = 0.222$.

To examine the interaction, we calculated estimated slopes for each subject's frequency function, and we analyzed these slopes for reliable differences. The analysis revealed a main effect of encoding condition, $F(1, 61) = 12.60$, $p < .001$, $MS_e = 0.056$, which confirms the steeper slopes of the word-encoding conditions. No other reliable effects were found.

Now consider the data for the filler words only. As was expected, the encoding condition had a much attenuated effect for these words because the words have no natural numerical values as associates. However, a three-way analysis parallel to the one conducted for the critical words did reveal an interaction between encoding condition and presentation frequency, $F(3, 183) = 11.02$, $p < .0001$, $MS_e = 0.119$, indicating that the fidelity of the frequency estimates was indeed reduced by number encoding. Aside from the normal effect of presentation frequency, $F(3, 183) = 318.55$, $MS_e = 0.119$, no other effects were found to be reliable.

The interaction was examined by calculating slopes as for the critical words. The slopes are shown in Table 2. Again a main effect of encoding condition was found, $F(1, 61) = 14.54$, $p < .001$, $MS_e = 0.038$, confirming the observation that numerically encoding the stimuli decreased the slope of the frequency estimates. No other significant effects were revealed.

As in Experiment 1, we were able to examine the effect of numerical distance between each word's associate and its frequency. Figure 5 plots accuracy of judgments against distance value in a manner parallel to that of Figure 2. As before, we found that the number-encoding condition resulted in poorer estimates than did the word-encoding condition. It also appears that there might be an effect of the magnitude of the distance value on accuracy. Notice, however, that the effect is entirely due to the distance value of 3, and recall that there is only one stimulus word with that distance value. It is entirely possible that this effect is due idiosyncratically to the word at this value (quartet). We have no way of assessing this at present. Ignoring this value, there does not appear to be an effect of distance value in general. In fact, this is shown by an ANOVA on the data in Figure 5 including the factors of encoding and retrieval conditions and distance. This analysis shows a reliable effect of encoding condition only, $F(1, 61) = 6.47$, $p < .02$, $MS_e = 0.163$. No other significant effects were found. This confirms that the condition of retrieval does not influence the quality of frequency estimates. These results, then, are in concordance with those of Experiment 1, and they expand on those results by showing that the locus of interference effects seems to be at the stage in which the material's frequency is encoded.

Table 2
Slopes of Frequency Estimate Functions for Experiment 2

Variable	*M*	*SD*
Critical words		
Word-encoding condition		
Word-retrieval task	.49	.18
Number-retrieval task	.55	.19
Number-encoding condition		
Word-retrieval task	.16	.29
Number-retrieval task	.36	.10
Filler words		
Word-encoding condition		
Word-retrieval task	.79	.12
Number-retrieval task	.66	.18
Number-encoding condition		
Word-retrieval task	.63	.18
Number-retrieval task	.49	.13

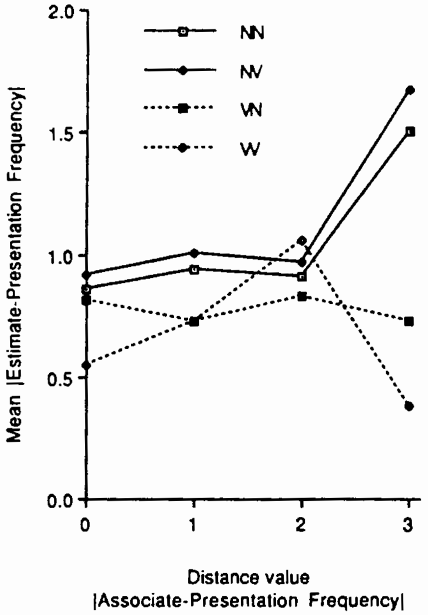

Figure 5. Accuracy of frequency estimates as a function of distance value for the critical words of Experiment 2. (Data are plotted for the four conditions of encoding versus retrieval defined by whether number [N] or word [V] associates were involved.)

As in Experiment 1, the number of associates generated did not differ significantly between the encoding conditions. A three-way ANOVA on presentation frequency (2, 3, or 4), encoding condition, and stimulus word type revealed main effects of stimulus word type, $F(1, 63) = 6.02$, $p < .02$, $MS_e = .084$, and frequency, $F(2, 126) = 8.97$, $p < .001$, $MS_e = 0.132$, but no effect of encoding condition ($p > .15$). An interaction between word type and encoding condition, $F(1, 63) = 13.12$, $p < .001$, $MS_e = 0.084$, reflected the subjects' difficulty in settling on a particular numerical associate to a filler word and in generating a single word associate for a critical word. A marginal Word Type × Frequency interaction, $F(2, 126) = 3.05$, $p = .05$, $MS_e = 0.046$, reflected the finding that the number of associates generated was more influenced by the frequency of the filler words than of the critical words.

Experiment 3

Experiment 1 demonstrated an interfering effect of numerical associates on frequency judgments. Experiment 2 replicated the findings of Experiment 1 and further showed that the major locus of that interference is at encoding rather than at retrieval. However, there were several limitations of those studies that deserve attention. First, the order of stimulus presentation and the assignment of stimuli to presentation frequencies were not randomized for each subject in Experiments 1 and 2. Also a more precise matching of the critical words and the filler words in terms of word length and linguistic frequency is needed to rule out the possibility that the effect on the critical words of number encoding is due to some unique idiosyncracy of these words having nothing to do with their numerical meaning. Finally, it is possible that the poorer frequency estimates in the conditions with numerical associates was a function simply of poorer memory overall in these conditions. If this were so, an interpretation based on the numerical meaning of these words would be compromised. Experiment 3 addresses all of these concerns.

Method

Subjects. The subjects were 28 University of Michigan and 33 Stanford University undergraduates who either participated as part of a course requirement or were paid. None of the subjects had participated in either of the previous experiments.

Materials. The stimuli were nearly identical to those used in Experiments 1 and 2. Twelve of the original filler words were randomly selected and removed. They were replaced by 12 new filler words that were matched to particular critical words on length and linguistic frequency. The matched filler words were linked to the corresponding critical words so that they also matched on presentation frequency. For each subject, the unmatched filler words were randomly assigned presentation frequencies of 1 through 4 such that 5 words appeared with each frequency. The critical words were randomly assigned presentation frequencies 1 through 4 with the constraints that (a) no word be assigned a frequency that equaled its associated value (e.g., *twins* was never assigned the frequency value 2), and (b) no words with the same associated values be assigned the same presentation frequency. The matched filler words, of course, were presented as many times as their linked critical words were shown.

Procedure. The procedure was similar to that of Experiment 1. However, each subject viewed a different random ordering of the stimuli. As in the previous experiments, at least three other words intervened between repetitions of a word.

As in Experiments 1 and 2, one half of the subjects generated word associates to the stimuli ($n = 31$), whereas the remaining subjects generated number associates ($n = 30$). After the presentation, one half of the subjects in each associate condition (word condition, $n = 16$; number condition, $n = 15$) were given an unexpected test of frequency estimation, and the remaining subjects were given an unexpected free-recall test for the words that had been used in the slide presentation (word condition, $n = 15$; number condition, $n = 15$).

Results and Discussion

The free-recall data indicated that the stimuli were just as well remembered in the number associate condition as in the word associate condition. Subjects in each condition recalled the same percentage of words overall (word condition: 40.7%, $SD = 10.0$; number condition: 41.3%, $SD = 9.8$), $t(27) = -.16$). In addition, the same number of the words recalled were actually critical words (word condition: 37.8%, $SD = 8.2$; number condition: 37.6%, $SD = 8.3$), $t(27) = .04$. Therefore, any effect of associate condition on frequency judgments cannot be attributed to overall memory differences between the two groups of subjects.

In Figure 6, the frequency estimates for the word-encoding condition are shown (Figure 6a), and the estimates for the number-encoding condition are plotted (Figure 6b). As in Experiments 1 and 2, subjects were clearly sensitive to presentation frequency. This was confirmed by an omnibus ANOVA including the factors of frequency, word type, and encoding condition. Frequency estimates increased reliably with actual frequency, $F(3, 87) = 197.45$, $p < .001$, $MS_e = 0.318$.

The frequency estimates given for the unmatched filler and matched filler words did not differ from each other. The estimates for the critical words, however, were less accurate than those for the filler words. This is evidenced by the slopes of the estimates for the filler words being closer to 1.0 (perfect performance) than are the slopes of the estimates for the critical words. Also the estimates for the filler words are closer in absolute magnitude to the true frequencies.

Two analyses confirmed these observations. First, the omnibus ANOVA revealed a main effect of word type, $F(2, 58) = 19.06$, $p < .001$, $MS_e = 0.179$, indicating that the critical words were given higher estimates overall and a Frequency × Word Type interaction, $F(6, 174) = 7.72$, $p < .0001$, $MS_e = .248$, suggesting that the slope of the critical word estimates was flatter than that of the unmatched filler and matched filler word estimates. Second, the slopes for each subject's critical word (word condition: $M = .53$, $SD = .30$; number condition: $M = .36$, $SD = .21$), unmatched filler word (word condition: $M = .74$, $SD = .24$; number condition: $M = .70$, $SD = .21$), and matched filler word (word condition: $M = .84$, $SD = .29$; number condition: $M = .75$, $SD = .27$) frequency functions were calculated. These slopes were subjected to a two-way ANOVA with word type and encoding condition as the independent variables. The analysis indicated a main effect of word type, $F(2, 58) = 28.1$, $p < .001$, $MS_e = 0.037$, and Tukey's method of multiple comparisons confirmed that the slope for the critical words differed from both the unmatched filler and matched filler words, which were not different from each other.

Therefore, we have evidence that the interference effect demonstrated in Experiments 1 and 2 was not simply an idiosyncratic result of the particular word frequencies and presentation order. Both the presentation order and the assignment of words to frequencies were random for each subject in this experiment. We also now have evidence that the differences between the filler words and the critical words found in Experiments 1 and 2 were not simply caused by some extraneous variable such as word length or linguistic frequency. The matched filler words were carefully matched (word by word) to the critical words, and the estimates for those fillers do not differ from the estimates for the original set of unmatched filler words.

As in Experiments 1 and 2, we examined the effect of numerical distance between a critical word's inherent associative value and its presentation frequency. Figure 7 plots accuracy of frequency estimates against distance value as in Figures 2 and 5. Note that, for this experiment, there were no distance zero words. Instead, there were six words with distance 1, four words with distance 2, and two words with distance 3. This change in distance values was made to increase the power of the distance analysis. The particular words at each distance were different for each subject. An ANOVA on the data in Figure 7 with factors of encoding condition and distance indicated that there were no reliable effects of condition or distance. A marginal Distance × Condition interaction was found, $F(2, 58) = 3.26$, $p < .06$, $MS_e = 0.216$.

A closer examination of the estimates revealed that subjects are not merely substituting the associate value for the presentation frequency when making a frequency estimate. The probability of reporting the associate value rather than the frequency actually decreased with increasing distance. That is, for distances 1, 2, and 3, the probabilities of reporting the associate value as the frequency were .278, .133, and .067, respectively, for the numerical associate condition, linear trend not significant: $F(1, 2) = 21.4$, $p < .14$, $MS_e = 0.001$;

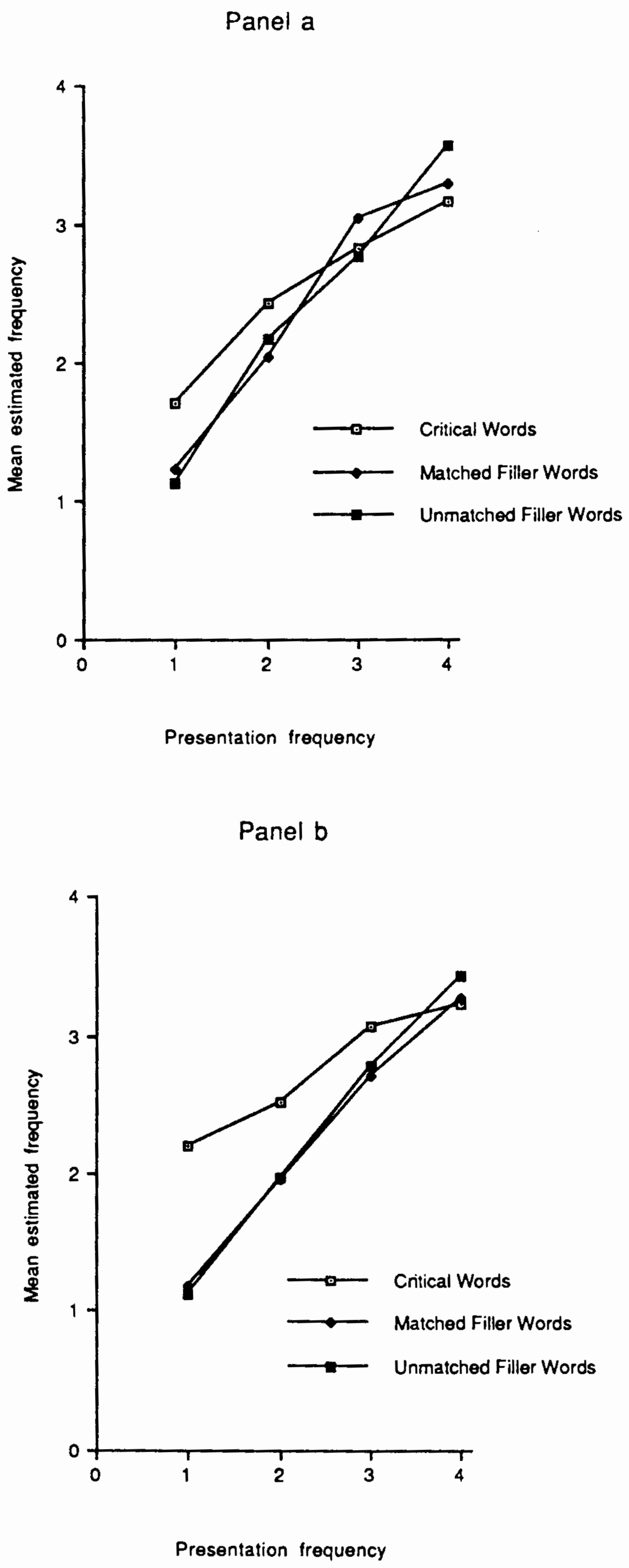

Figure 6. Estimated versus actual frequency for the critical and filler words of Experiment 3 for (a) the word-encoding condition and (b) the number-encoding condition.

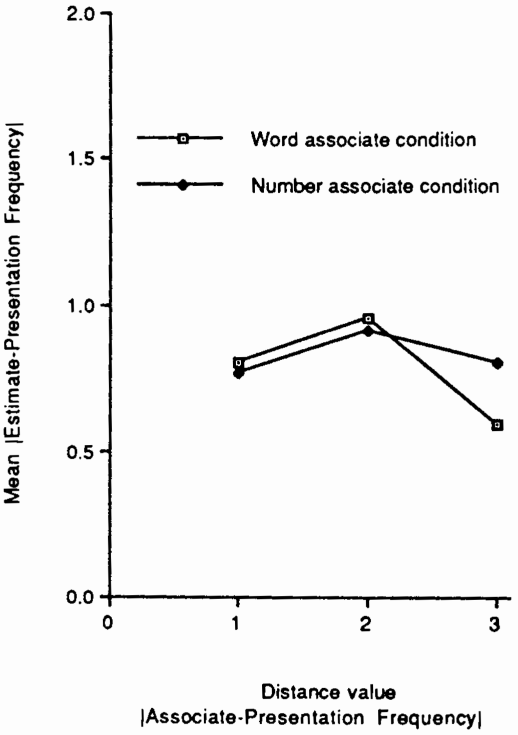

Figure 7. Accuracy of frequency estimates as a function of distance value for the critical words of Experiment 3. (Data are plotted for both the number- and the word-encoding conditions.)

and .208, .109, and 0, respectively, for the word associate condition, linear trend: $F(1, 2) = 1297.92$, $p < .02$, $MS_e = 0.00002$. The increase in the amount of error in the numerical associate condition as distance increased was not caused by a substitution of the associate value for the presentation frequency in the estimate.

The number of associates generated did not differ between encoding conditions and thus cannot account for the interference effect. A three-way ANOVA on frequency, encoding condition, and word type revealed a main effect of frequency, $F(2, 54) = 3.62$, $p < .05$, $MS_e = 0.048$.[1] A Word Type × Condition interaction showed that critical words were given more associates in the word associate condition, whereas filler words were given more associates in the number associate condition, $F(1, 27) = 31.25$, $p < .0001$, $MS_e = 0.070$. If number of associates generated were responsible for the difference in frequency estimates, then the critical words should have been given more accurate estimates in the word associate condition. Looking back at Figure 6 shows that this was not the case. In addition, a three-way interaction with frequency, word type, and condition, $F(2, 54) = 7.81$, $p < .01$, $MS_e = 0.027$, revealed that, for filler words, the word associate condition led to number of associates generated decreasing with frequency but the number condition led to number of associates increasing with frequency. For the critical words, the word associate condition led to number of associates generated increasing quite a bit with frequency, but in the number associate condition number of associates increased only slightly with frequency.

General Discussion

Experiment 1 provided evidence of an interference effect between a task requiring the generation of numerical associates and one requiring the generation of frequency estimates. Our analyses suggest that this effect is not due simply to the sheer number of associates generated; rather, it is apparently due to the numerical nature of these associates. Experiment 2 replicated this effect and showed that the major (perhaps exclusive) locus of this effect was the encoding of stimuli. Little interference seemed to occur when the numerical associates were generated only at the time of testing. Experiment 3 replicated and extended Experiments 1 and 2, demonstrating that the effect of type of associate at the time of encoding is not a general memory effect but rather is specific to frequency judgments.

It might be argued that the interference effect we have demonstrated was not caused by the numerical nature of the critical words but rather by their close similarity to one another.[2] Such an account of the data is consistent with Hintzman's (1988) MINERVA 2 model of memory. According to that model, stimuli that are more similar to each other than to the other stimuli will be given higher frequency estimates. At first glance, the data presented here seem to be explained merely by such similarity effects. However, data collected by Jones and Heit (1991) indicate that the amount of frequency overestimation associated with semantically similar stimuli in a task much like the present one is approximately constant across presentation frequencies 0 to 7. The interference effect we demonstrated does not involve constant overestimation across even the frequencies 1 to 4. So it seems quite unlikely that the similarity of the critical words could completely explain the numerical interference effect.

Our favored account of these effects is based on the proposal that a direct code for frequency is constructed and stored at the time of an item's appearance in a given context (see Howell, 1973; Underwood, 1969, for discussions of similar proposals). The nature of this direct code is such that its creation requires the activation of a numerical coding system, just the one that is activated by the natural associates of some of the words we used in our experiments. Because there is a common system activated, interference is caused between the two codes.

Why has evidence for such an interference effect not been generated previously? In fact, it has. The data collected and cited by Whitlow and Skaar (1979) showed a systematic interaction of frequency estimates with other numerical information in their task as we argued earlier. Their interpretation of these data does not make this interaction obvious, but an effect does appear to be present. We suspect that the failure of Hintzman et al. (1982) to find an effect is due to the fact that, in their tasks, the generation of a frequency code and the generation of an associative code were not the responsibility of the same system. The associative code was generated by the experimenter and merely stored by the subject, whereas the frequency code was internally generated by the subject.

Suppose, then, that a direct code does play a role in frequency coding. What is its role? One thing is clear: A direct code is not the only one used in the estimation of frequency. Ample evidence now exists to implicate an indirect coding mechanism of some sort. The likeliest possibility for this mechanism is one based on the multiplicity of traces that is formed with repeated presentation of an item (see, e.g., Hintzman, 1988, for a review of evidence about the plausibility of such a mechanism). Therefore, if a direct coding model is to have any currency, it must be one in which a direct code and an indirect code are combined during frequency estimation.

Our present data do not permit unambiguous inferences about the nature of the integration process between the presumed direct and indirect codes. In fact, several possibilities are open. One is that a direct code serves as the basis of judgment when there is a subjective feeling of confidence in its accuracy; but when this fails, an indirect code acts as the default source of information about frequency. This would be consistent with a regular feature of the data from both Experiments 1 and 2: The greatest interference of the number-encoding instruction was on the lowest frequencies. If readers consult Figures 1, 3, 4, and 6, they will see that the difference in frequency estimates between the numerical and word associate conditions was greatest for frequencies 1 and 2 and

---

[1] The data from 2 subjects were not included in this analysis. For 1 subject, the associate booklet was lost. For the second subject, the data file containing information about that subject's presentation order was lost, rendering the generated associates impossible to analyze.

[2] This point was brought to our attention by David Bryant.

least for frequencies 3 and 4. If a direct code were the first line of information regarding frequency, it seems reasonable to suppose that it would have greater subjective validity the lower the frequency. At lower frequencies, the subject has done less updating of the code, and so there might be less opportunity for misremembering the current value of frequency assigned to an item.

An alternative model combines direct and indirect codes: It may be that the two sources are separately evaluated, and their outputs are combined in a weighted manner to arrive at an overall impression of frequency. This could also be consistent with the amount of interference caused by numerical associates as a function of frequency. The relative weight assigned to a direct code might be greater at lower frequencies, and so the interference caused by a competing numerical code could be greater.

Yet another possible model is based on the principle of anchoring and adjustment (Tversky & Kahneman, 1974). According to this model, one source of information about frequency would serve as the basis for creating an anchoring value, and this value would be adjusted by the other source of information. For example, the direct code might provide a "ballpark" estimate of frequency for an item, and this anchor would then be finely tuned on the basis of the multiplicity of traces stored for that item. Jonides, Jones, and Smith (1991) reported evidence that implicates an anchoring and adjustment mechanism in the production of frequency judgments, making this model a viable option.

At present, there is insufficient evidence to discriminate among alternative models that combine direct and indirect codes for frequency, both the ones speculatively cited here and others. However, there *is* sufficient evidence to advocate the position that both direct and indirect representations are involved in the encoding of frequency information.

References

Birnbaum, I. M., Taylor, T. H., Johnson, M. K., & Raye, C. L. (1987). Is event frequency coded automatically? The case of alcohol intoxication. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 251–258.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.

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.

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

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.

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. (1982). Are presentation frequency and spatial numerosity distinct attributes of memory? *Bulletin of the Psychonomic Society, 20*(4), 196–198.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95,* 528–551.

Hintzman, D. L., Nozawa, G., & Irmscher, M. (1982). Frequency as a nonpropositional attribute of memory. *Journal of Verbal Learning and Verbal Behavior, 21,* 127–141.

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

Howell, W. C. (1973). Representations of frequency in memory. *Psychological Bulletin, 80,* 44–53.

Jones, C. M. (1990). [Counting strategy for coding frequency information.] Unpublished raw data.

Jones, C. M., & Heit, E. (1991). *Effects of similarity on frequency judgments: A multiple-trace account.* Manuscript submitted for publication.

Jonides, J., Jones, C. M., & Smith, R. W. (1991). *Effects of range instructions on the judgment of frequency.* Manuscript in preparation.

Jonides, J., & Naveh-Benjamin, M. (1987). Estimating frequency of occurrence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 230–240.

Naveh-Benjamin, M., & Jonides, J. (1985). The effects of rehearsal on frequency coding. *Bulletin of the Psychonomic Society, 23*(4), 387–390.

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.

Rose, R. J. (1980). Encoding variability, levels of processing, and the effects of spacing of repetitions upon judgments of frequency. *Memory & Cognition, 8,* 84–93.

Rose, R. J., & Rowe, E. J. (1976). Effects of orienting task and spacing of repetitions on frequency judgments. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 142–152.

Rowe, E. J. (1974). Depth of processing in a frequency judgment task. *Journal of Verbal Learning and Verbal Behavior, 13,* 638–643.

Rowe, E. J., & Rose, R. J. (1977). Effects of orienting task, spacing of repetitions, and list context on judgments of frequency. *Memory & Cognition, 5,* 505–512.

Sanders, R. E., Gonzalez, E. G., Murphy, M. D., Liddle, C. L., & Vitina, J. R. (1987). Frequency of occurrence and the criteria for automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 241–250.

Shedler, J. K., Jonides, J., & Manis, M. (1985, November). *Availability: Plausible but questionable.* Paper presented at the 26th annual meeting of the Psychonomic Society, Boston, MA.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

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

Whitlow, J. W., Jr., & Skaar, E. (1979). The role of numerosity in judgments of overall frequency. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 409–421.

Zacks, R. T., Hasher, L., & Sanft, H. (1982). Automatic encoding of event frequency: Further findings. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 106–116.

Received October 19, 1989
Revision received August 21, 1991
Accepted August 22, 1991 ■