Journal of Experimental Psychology:
Learning, Memory, and Cognition
1986, Vol. 12, No. 3, 378–386

On the Automaticity of Frequency Coding: Effects of Competing Task Load, Encoding Strategy, and Intention

Moshe Naveh-Benjamin
Ben-Gurion University of the Negev, Israel

John Jonides
University of Michigan

Two criteria of central importance to the identification of an automatic process are insensitivity to competing task loads and to strategy manipulation. Past experimentation using these criteria has led to the view that judgments of frequency of occurrence are a result of an automatic process. Experiment 1 is a reexamination of the effect of competing load. This experiment's results make clear that judgments of frequency are markedly affected by the level of difficulty of a secondary task. Results from Experiment 2 suggest that although intention to code frequency may not have an effect on frequency judgments under all conditions, there are strategic effects on coding nonetheless. Finally, Experiment 3 is included to show that under some conditions, in fact, one can find an effect of intention as well. Taken together, the results that are reported question the view that frequency judgments are the result of an automatic process.

There has been a growing interest in the past several years in identifying automatic processes in cognition and specifying details of their operation. A case in point is the encoding and judging of frequency of occurrence of events. Hasher and Zacks and their colleagues, as well as others, have amassed evidence about a number of criteria that they proposed as tests for automaticity of frequency coding (Hasher & Zacks, 1984). Briefly, the evidence is this:

(a) Sensitivity to the frequency of events is a characteristic that is demonstrable in young children and shows little change with age (Attig & Hasher, 1980; Hasher & Chromiak, 1977; Hasher & Zacks, 1979; Kausler & Puckett, 1980).

(b) Encoding of frequency does not benefit from practice in tasks that require such encoding (Hasher & Chromiak, 1977; Zacks, Hasher, & Sanft, 1982).

(c) Individual differences in cognitive functioning do not affect the accuracy of reports about frequency of occurrence (Hasher & Zacks, 1979; Zacks et al., 1982).

(d) Intention to code frequency does not enhance its memorability (Flexser & Bower, 1975; Howell, 1973; Zacks et al., 1982). Similarly, variations in effortful encoding strategies or instructional manipulations will not affect the efficiency of registering frequency information.

(e) Encoding of frequency information is not influenced by competing task demands (Zacks et al., 1982).

These criteria have led to the view that the encoding of frequency information is effortless, inflexible, and inaccessible—in short, that it is automatic (Hasher & Zacks, 1984). Perhaps it is. However, in order to be confident about this conclusion, one must be sure that the operational definitions of the various criteria are properly reflective of the underlying characteristic in question. Greene (1984) has recently called some of these criteria into question, showing that as evidence for automaticity, some may be ambiguous. Accordingly, there is a need to examine more carefully the evidence that has been used to support the automaticity argument.

The most central evidence bearing on automaticity comes from studies of the effects of competing task demands and strategy manipulations. Both theoretically and empirically, various investigators have argued that insensitivity to these variables is the hallmark of automaticity (Jonides, 1981; Naveh-Benjamin & Jonides, 1984; Posner & Snyder, 1975; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977).

To begin, let us concentrate on the criterion of competing task demands. If an automatic process is assumed to require minimal mental capacity for its operation, performance of a task simultaneously with an automatic process should not affect its operation. In a study reported by Zacks et al. (1982), competing demand for capacity has been shown not to affect performance on frequency judgments. In this study, three groups of subjects received different types of instructions. One group was informed about a forthcoming frequency test before being exposed to the test stimuli. A second group was informed about a forthcoming recall test on the material, and a third group was informed about both an upcoming frequency test and a recall test. The group instructed about both tests was included as a way to place competing demands on subjects. Preparation for a recall test is assumed to require considerable capacity, especially if subjects are using mnemonics. Zacks et al. (1982) claimed that the competing demands induced by instructing subjects to prepare for both the frequency and recall tests should have had no impact on frequency judgments if these are mediated by an automatic process. However, competing de-

We thank G. Logan and F. Durso for their helpful comments on an earlier draft of this article. We would also like to thank Joanne Alexander and Dorit Lavi-Tubin for their help in data collection and analysis. This work was supported in part by a grant from the Air Force Office of Scientific Research to John Jonides.

Correspondence concerning this article should be addressed to Moshe Naveh-Benjamin, Department of Behavioral Sciences, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva, 84120, Israel.

mands were predicted to reduce performance for the free recall task, which was hypothesized to be effortful. The results showed that whereas preparing for both tasks interfered with free recall, it had no impact on frequency judgments.

However, the results of this experiment are problematic for two reasons: First, if frequency judgments themselves are mediated by accessing information about a memory trace that also serves as the basis for recall (Hintzman & Block, 1971; Howell, 1973), one may not particularly expect that preparation for a free recall test should impede frequency judgment performance. Second, there is no direct way of scaling the amount of cognitive demand induced by the different kinds of test preparation.

There is another report of an experiment measuring the effect of competing task demands on frequency judgments whose results do not support the automaticity contention. In this experiment, by Kausler, Wright, and Hakami (1981), judgments about frequency were obtained under two conditions. In one, subjects were asked to prepare for just an upcoming frequency judgment task. In the other, subjects had to perform an ancillary case-monitoring task in addition to preparing for frequency judgments. Results showed that the group that performed both tasks was less accurate in its frequency estimations, a result that is inconsistent with an automaticity hypothesis. However, the results of this study cannot be taken as definite evidence about automaticity because it could be argued that the reduced performance in the dual task was due to an effect of "structural" interference by the competing task. It could be that performing the ancillary case-monitoring task involves the same mental mechanisms (e.g., rehearsal of recent information) as frequency coding. If so, the frequency judgments would be less accurate because of such specific interference, not because the same pool of general processing capacity was being taxed by the case-monitoring task (see Kahneman, 1973; Posner & Boies, 1971).

Another experiment that is relevant to the issue of competing demands was reported by Fisk and Schneider (1984). Although their results also show reduced performance in judging frequency with heavy competing loads, the effect is not easily interpretable because the conditions with the different loads also differed in instructions about the frequency test.

In short, there is still need for a test of the effect of competing demands on frequency judgments. This test should include a primary task that requires attention to frequency of occurrence for various events and a secondary task that varies in difficulty. Effects of secondary task load on performance of the primary task would establish an effect of competing demands. Furthermore, by varying secondary task difficulty and finding an effect of this variation, one could be confident that the interference of the secondary task on the primary task was due to more than structural interference. Thus, an experiment of this sort would be relevant to the issue of whether frequency coding meets the criterion of insensitivity to competing load, which is one hallmark of an automatic process.

The first experiment meets this need. Subjects were required to engage in a digit-counting task while preparing for a test of frequency estimation. The load of the secondary task was varied by varying the difficulty of the counting task. In the easiest load condition, subjects had to count forward by 1s. In a more difficult load condition, subjects had to count backward by 3s, and in the most difficult condition, they had to count backward by 11s.

Before proceeding, let us comment on a feature of all the experiments to be reported concerning treatment of the data. There has been a remarkable lack of uniformity in analysis of frequency judgments in past work. Early experimenters were concerned largely with analysis of the deviation of judged from actual frequencies (e.g., Rose & Rowe, 1976; Rowe, 1974). Later, Flexser and Bower (1975) introduced the use of a correlation between judged and actual frequency as a measure of performance. Their rationale was that a correlation would assess the discriminability of the judgments rather than some absolute deviation that the judgments may have had from the actual frequencies (a kind of constant error). Since then, investigators have used a mixture of these different measures with little discussion of the information provided by each (e.g. Greene, 1984).

Some reflection about the value of each measure reveals that there are three important pieces of information that one should extract from the relationship between judged and actual frequency (assuming, as is usually the case, that the relationship is linear or nearly so). One is the slope of the function relating estimates to actual frequencies. This slope is an indicator of sensitivity to variations in frequency. As the slope becomes closer and closer to unity, one can conclude that subjects are more and more sensitive to variations in presented frequency.

A second measure of interest is the deviation of each judgment from the mean judgment for a given presentation frequency. Larger deviations obviously imply less consistency in judgments and therefore less reliability.

Notice that, contrary to the recommendation of Flexser and Bower (1975), coefficients of discrimination (correlations of estimated against actual frequency) are not sufficient to distinguish between the above two measures of interest. Correlation coefficients are sensitive to both the slope of the line relating the two variables of interest and the variability of the data points. This is most easily seen by examining the equation of the Pearson r: $r = b(Sx/Sy)$. If the variable x is fixed, as it is in typical studies of frequency (because frequency is treated as an independent variable), then this equation makes clear that r is directly related to the slope and inversely related to variability in y. In short, r is ambiguous by itself about the first two measures that we propose.

A third measure of interest is the absolute magnitude of the judgments compared to the actual frequencies. This reveals whether subjects are sensitive to the range of frequencies that were actually presented and whether their judgments of frequency tend to underestimate or overestimate the actual frequencies. Of course, it is of value to compare these absolute magnitudes between various experimental conditions as well as comparing them to the objectively presented frequencies.

Given this analysis, it seems most appropriate to consider results of a frequency judgment experiment in a way that allows all three measures of performance to be examined. The analyses presented below are consistent with this recommendation.

Experiment 1

Method

Subjects. The subjects were 30 undergraduate students at Ben-Gurion University of the Negev who were paid for their participation.

Materials and apparatus. The pool of words used to construct the stimulus list for the study phase of the experiment consisted of 42 two-syllable high-frequency Hebrew words. Seven of these words appeared once in the list, 7 twice, and so on, with the last 7 appearing six times. In addition to these 147 stimuli, 5 additional words were added to the beginning of the list and 5 to the end to absorb primacy and recency effects, respectively. In addition, 4 words were added to the beginning of the list for practice. Altogether, then, there were 161 words in the stimulus list for the study phase of the experiment. These words were presented sequentially via slide projector. There were several orders of these slides, all of which were random with the constraints that repetitions of a word were separated by at least 4 other words and that the words that were repeated appeared about equally in the two halves of the list.

Procedure. There were three groups of 10 subjects each. Each subject was tested individually. Before the study phase of the experiment, subjects were told that they would have to perform two tasks during the presentation of the words. The first was to pay attention to the words in order to prepare for the upcoming frequency judgment test. To assure that subjects actually paid at least minimal attention to the words, they were asked to read each word aloud when it appeared. The second task was to count once for each slide appearance, starting with a specific number. Subjects in one group were told to count forward by 1s starting with the number 1,146. Subjects in the second group were told to count backwards by 3s starting with 1,250, and subjects in the third group were told to count backwards by 11s starting with 1,825. Subjects in all groups were told to pay attention equally to the word and the counting tasks. After the practice trials, there were almost no errors in the counting task, and the errors were evenly distributed across conditions. (Mean number of errors was 1%, 2%, and 3.5%, respectively.)

In order to be sure that the digit-counting tasks were ordered in difficulty as described, a preliminary experiment was performed. Each of 24 subjects was asked to count as fast as possible for 1 min forward by 1s (+1 condition), backward by 3s (−3 condition), or backward by 11s (−11 condition). The order of the three counting tasks was counterbalanced. The mean number of digits counted was 28.6, 18.1, and 10.9 for the respective tasks. There were significant differences in the number of digits that subjects were able to count, $F(2, 21) = 47.6$, $p < .01$. Subjects were on the average able to count significantly more digits in the +1 condition than in the −3 condition, $F(1, 21) = 32.1$, $p < .01$, and they were able to count significantly more digits in the −3 condition than in the −11 condition, $F(1, 21) = 49.3$, $p < .01$. These results assured us that the difficulty of the counting conditions was as hypothesized.

During the study phase of the main experiment, the words were presented at a rate of one word per 4 s in one of the several random orders. Following the last word of the list, the experimenter engaged subjects in approximately 2 min of conversation before the second phase of the experiment began.

The second phase of the experiment consisted of a frequency test for which there were 84 words, the 42 words presented in the study phase plus 42 distractor words that met the same criteria as the test words but that had not been presented during the study phase. These 84 words were randomly intermixed. Subjects were given a sheet with these words, and they were asked to give a value of zero to six for each word, a value that represented the judged frequency of appearance of each word during the study phase of the experiment. Subjects were told explicitly that half of the words on the frequency test had not appeared during the study phase.

Table 1
Mean Slopes and Standard Deviations of Frequency Judgments as a Function of Load Condition

Load condition	Slope *M*	Slope *SD*	Standard deviation *M*	Standard deviation *SD*
+1	0.51	0.18	1.13	0.22
−3	0.33	0.14	1.45	0.24
−11	0.21	0.13	1.40	0.26

Results and Discussion

Several measures of performance were compared across conditions. The first measure is the slope of the regression line fit to the estimates. This measures the discriminability among the presented frequencies. A higher slope is an indication of greater discriminability among the actual frequencies of presentation. A linear regression analysis was performed on each subject's judged estimates, regressing estimated on actual frequencies, and the resulting slopes were compared across load conditions. The values for these averaged slopes can be seen in Table 1. Analysis of these slopes revealed a significant effect of load, $F(2, 27) = 15.13$, $p < .001$.

It is also possible to examine how the different load conditions affected the absolute level of frequency estimates. Each subject's estimates were averaged across each of the actual frequencies to yield mean absolute frequency estimates that were then averaged across subjects in each load condition. The mean estimated frequencies for each of the load groups are shown in Figure 1 as a function of actual frequency. The results show that the heavier the competing task load, the lower the absolute level of judged frequencies. This trend was confirmed by an analysis of variance (ANOVA), $F(2, 27) = 8.22$, $p < .01$ ($MS_e = 0.95$).

The third measure of performance provides information about the degree to which the frequency judgments dispersed around the mean judgments (a measure of the reliability of the judgments). For each subject, a standard deviation was calculated of the estimates given for each actual frequency. These standard deviations were then averaged across the actual frequencies within each subject to yield one mean standard deviation per subject. Table 1 presents the resulting means of these values for the different load conditions. It shows that dispersion was the largest for the two heaviest competing load conditions. This was confirmed by an ANOVA, $F(2, 27) = 4.90$, $p < .05$.

Altogether, the results are quite clear: The heavier the load of the competing task, the worse the performance on the frequency judgment task by any of the measures used. The encoding of frequency information is affected by competing task demands in line with Kausler et al.'s (1981) study. This result is not consistent with the claim that frequency coding is automatic.

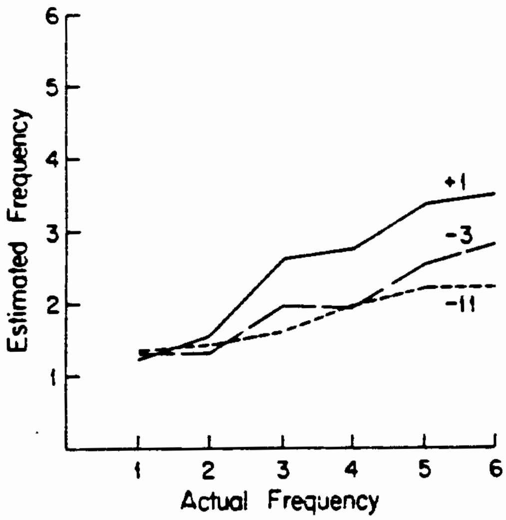

Figure 1 Mean absolute level of frequency estimates for each load condition.

Experiment 2

The second experiment focused on the second major criterion of automatic processes. According to this criterion, intention to code frequency, or manipulation of strategy used in coding frequency, should have no effect on estimations.

Consider first the effect of intention to code frequency on later judgments. Previous research that has compared intentional with incidental instructions has been mixed. Some investigators have found no effect at all of this manipulation (e.g., Flexser & Bower, 1975; Howell, 1973), whereas others (e.g., Greene, 1984) have found such an effect. There are some problems with the studies that have not found an effect of intentional instructions on frequency judgments, a result which has been taken as support for the automaticity of frequency coding. These problems are summarized by Greene (1984). To mention just the major one, Greene suggests that incidental instructions have been inappropriately operationalized. He claims that the appropriate way to use incidental instructions is by involving subjects in a task in which the learning of the items themselves will be incidental (pure incidental condition), in contrast to previous studies that manipulated only the intent to learn the frequency attribute itself. When comparing intentional with pure incidental instructions, Greene (1984) found an advantage of intentional over incidental instructions.

Proponents of the automaticity position could claim that we have to be cautious in accepting Greene's results, however. The way in which he operationalized an incidental instruction has problems of its own. He did so by using the digit-recall paradigm of maintenance rehearsal (Glenberg, Smith, & Green, 1977; Rundus, 1977). In this paradigm, subjects were presented words for a later frequency test as rehearsal distractors in a Brown-Peterson task. In the incidental condition, they were encouraged to forget the words as quickly as possible because their alleged task was to remember the digits that were presented on each trial. The problem with this paradigm is that the incidental condition is actually an intentional forgetting condition. This might be too extreme a way of having subjects pay no attention to the frequency of words. (Kausler, Lichty, & Hakami, 1984, using a similar procedure obtained no effect of intentionality, although a ceiling effect in the digit task might have been troublesome.) We prefer a more neutral incidental condition but one that is also truly incidental.

In addition to intention, there is also a need to examine other strategic manipulations to test their effect on frequency judgments. The one on which we focused is the level of processing applied to the stimulus words, because there have been some attempts in the past to study the effects of this variable on frequency judgments (e.g., Greene, 1984; Rowe, 1974; Rowe & Rose, 1977). Although these experiments have shown some effects of level of processing, the measures of frequency performance have been incomplete, as we argued above.

To address these issues, in Experiment 2 we manipulated level of processing in conjunction with intention. As such, the experiment is a replication and extension of Greene (1984). According to the automaticity view of frequency coding, neither intention to code frequency nor level of processing should affect the fidelity of the frequency judgments.

Method

Design. The experiment consisted of two phases. During the first phase, subjects were sequentially presented a list of words and asked to write down associates for each. One variable in the experiment was the kind of associate that was to be generated: Half of the subjects were to provide as many different acoustic associations (rhymes) as they could for each word presented, whereas the other half were to provide as many semantic associates as they could during the interval when each word was presented. The other variable was the learning paradigm: Half of the subjects were told nothing of the later frequency test, (they were led to believe that we simply needed to create a new set of association norms for another purpose), whereas the other half were told that the words of the first phase would be presented again and on that occasion they would have to judge how frequently each had appeared in the first phase. These two variables resulted in four groups of subjects altogether because the design was completely between subjects. The second phase of the experiment consisted of a frequency test given to all subjects.

Subjects. One hundred and nineteen undergraduate students at the University of Michigan (77) and Ben-Gurion University (42) participated as part of a course requirement. Students from each university were distributed approximately equally among experimental conditions.

Materials. There were 48 one- and two-syllable common English or Hebrew nouns used as stimuli (frequency greater than 20 per million in Kučera & Frances, 1967). Eight words appeared in the stimulus list at each of six frequencies, from one to six, so that there were 168 tokens in the list. In addition, 5 words appeared at the beginning as practice, 5 words appeared after the practice words to absorb any primacy effect, and 5 words appeared at the list's end to overcome any recency effect. The order of appearance of the 168 critical words was randomized with the same constraints as in Experiment 1.

There were 48 additional words with the same characteristics as those of the stimulus list. These were intermixed randomly with the 48 critical words of the stimulus list to create a list of 96 words for the frequency test.

Procedure. Subjects were tested in eight groups with 10 to 23 subjects on the frequency test.

Results and Discussion

The reported results include only 117 of the 119 subjects. Two subjects in the incidental learning groups were excluded from the analysis because they suspected the final frequency test.

The same performance measures used in Experiment 1 were calculated for the present data. The first measure was the mean slope of the regression lines fit to each subject's data. The mean values for these slopes for each condition are presented in Table 2. The analysis revealed a significant effect of levels of processing, $F(1, 113) = 7.29$, $p < .01$. There was no reliable effect of learning instructions, $F(1, 113) = 0.13$, nor was there a significant interaction of the two variables, $F(1, 113) = 0.35$.

The comparison of the above variables for the absolute level of frequency estimates showed the same trend as the slopes. These results can be seen in Figure 2. A three-way ANOVA revealed a significant effect of levels of processing, $F(1, 113) = 16.68$, $p < .01$ ($MS_e = 0.79$), and neither an effect of intention 0.54 and 0.59, respectively ($MS_e = 0.79$). The effects of actual frequency and the interaction of actual frequency and levels of processing were both significant, $F(5, 565) = 120.7$, $p < .01$, and $F(5, 565) = 4.04$, $p < .01$, respectively ($MS_e = 0.16$).

As for the dispersion of the frequency judgments, the average standard deviations for each subject (which were obtained by averaging the mean of the standard deviations for each actual frequency) were calculated. Table 3 presents the values for these deviations for each group. The results show that there is a trend for semantic processing to lead to more reliable judgments. However, ANOVAs reveal that neither of the variables, nor the interaction between them, had a significant effect, $F(1, 113) = 0.61$ for levels of processing, $F(1, 113) = 1.39$ for learning instructions, and $F(1, 113) = 0.02$ for the interaction.

The trend for all three measures (which was significant for two) opposes the contention that frequency coding is automatic because level of processing had an effect on judgments. This replicates the effect of levels of processing reported by Greene (1984) under somewhat different instructions.

Recently Hasher and Zacks (1984) have claimed that previ-it took subjects to generate associations. In a follow-up experiment, we had subjects engage in the same task but on an ad lib basis. This experiment's results showed that it took much longer to generate acoustic than semantic associates. If this time difference is another index of covert rehearsal activity, one would predict better frequency judgment performance in the acoustic condition. Quite the contrary, we once again found frequency judgments in the semantic condition to be superior. For all these reasons, we conclude that the effect of levels of processing is due to just that, not to covert rehearsal.

The results for the intention variable are more problematic. On the one hand they agree with many of the previous studies that have shown no effect of intention on the coding of frequency information. But on the other hand, they are at odds with the results reported by Greene (1984). We believe that the effect reported by Greene (1984) could hinge on the unique procedure he used, in which an effect might be attributed to the intentional forgetting characteristic of the incidental condition.

There is another feature of our experiment that characterizes

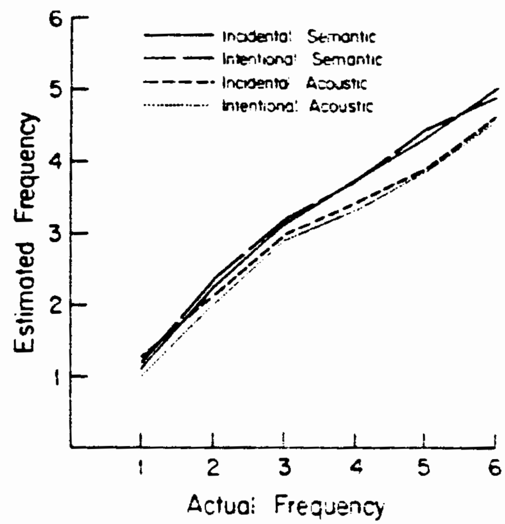

Figure 2. Mean absolute level of frequency estimates as a function of levels of processing and learning instructions.

all failures to demonstrate an effect of intention. Our experiment and others have included fairly long presentation times of 3 to 5 s for the stimulus items. The only study of intention to use short times (0.6 s) was by Fisk and Schneider (1984).[1] As they discussed, it may be that modification of long-term memory traces for word frequency occurs only in the first few seconds of processing. After this, encoding processes may be inflexible, so that further attempts to modify frequency codes will produce little or no change. This would account for failures to find an effect of intention at long presentation times and also for the large effect reported by Fisk and Schneider (1984) for short presentation times.

This account of the effect of time on frequency coding may also be relevant to Experiment 1. Perhaps the effect of competing demand in that experiment is actually mediated by variation in the time available for processing frequency. Heavy demand effectively reduces processing time from the nominal presentation time of 4 s to something shorter. According to this hypothesis, a time less than 4 s would not be sufficient to achieve optimal frequency coding.

Experiment 3 was designed to test this hypothesis directly. It included variation in intentionality instruction, but it also included variation in presentation time of the words. If the hypothesis is correct, short presentation times (less than approximately 4 s) will lead to an effect of intention. At a longer presentation time, however, there should be no such effect (e.g., Hasher & Zacks, 1979).

Experiment 3

Method

Subjects. Sixty-six undergraduate students at Ben-Gurion University of the Negev were paid for their participation.

Materials and apparatus. The materials and apparatus were identical to that of Experiment 1 except for the use of a timer and shutter that was attached to the slide projector to control exposure times. Exposure times were 1, 2, or 4 s, and the interslide interval was approximately 1 s.

Procedure. Subjects were tested in groups of two to four. The instructions in each group forewarned subjects about the exposure times designated for that group. Subjects in the intentional condition were told to expect a frequency test at the end, whereas subjects in the incidental group were told to pay attention to the words in order to answer some questions about them later (this sort of incidental instruction allowed us to more nearly replicate the typical incidental conditions in previous experiments). One third of the subjects saw each word for 1 s, one third for 2 s, and one third for 4 s.

Following list presentation, subjects were engaged in conversation for about 2 min, after which they were given the same frequency test as in Experiment 1.

Table 3
Mean Standard Deviations of Judgments as a Function of Level of Processing and Learning Instructions

	Level of processing			
	Semantic		Acoustic	
Learning instructions	*M*	*SD*	*M*	*SD*
Incidental	1.05	0.22	1.12	0.15
Intentional	1.16	0.24	1.21	0.27

Table 4
Mean Slope Values as a Function of Learning Instructions and Exposure Time

	Exposure time					
	1s		2s		4s	
Learning instructions	*M*	*SD*	*M*	*SD*	*M*	*SD*
Incidental	0.40	0.17	0.45	0.12	0.54	0.13
Intentional	0.59	0.19	0.58	0.13	0.56	0.11

Results and Discussion

The reported results include only 60 of the 66 subjects. Six subjects, two in each of the incidental learning groups, were excluded from the analysis because they suspected the final frequency test. Altogether, therefore, the following analysis includes six groups of 10 subjects each.

In order to test the hypothesis that exposure duration modulates the effect of intention in the specific manner mentioned earlier, three planned comparisons were performed for each measure of performance. These analyses involved comparison of intentional to incidental conditions at each of the three levels of exposure time.

The first measure was the mean slope of the line regressing estimated against actual frequencies. Each subject's data were regressed individually producing a slope estimate for that subject. These were then averaged across subjects in each condition. The data are presented in Table 4 by learning instructions and exposure times. The slope estimates for the intentional groups were significantly higher than those for the incidental groups for the 1-s and 2-s exposure time conditions but not for the 4-s condition. Planned comparisons for each exposure time reveal: $t(18) = 2.23$, $p < .05$, for the shortest exposure time; $t(18) = 2.20$, $p < .05$, for the intermediate exposure time; and $t(18) = 0.58$, *ns*, for the longest exposure time.

The second performance measure on which comparisons were made was that of the mean estimates themselves. These are shown in Figure 3, in which we plot mean estimates for each learning instruction condition and exposure time. Comparing conditions at each exposure time by a two-way ANOVA revealed the following: For the shortest exposure time there was no effect of instruction condition, $F(1, 18) = 1.19$ ($MS_e = 0.70$); there was a significant effect of actual frequency, $F(5, 90) = 10.09$, $p < .01$ ($MS_e = 0.25$), and a significant effect of the interaction of instruction condition and actual frequency, $F(5, 90) = 2.52$,

[1] Hintzman (1970), using general memory instructions, showed a weak interaction between frequency and duration of presentation for intervals between 2 and 6 s.

... reflects the superiority of performance of the intentional over the incidental instructions. For the intermediate exposure time, again there was no effect of instruction condition, $F(1, 18) = 0.68$ ($MS_e = 0.60$), a significant effect of actual frequency, $F(5, 90) = 12.71$, $p < .01$ ($MS_e = 0.21$), and a significant interaction of instruction condition and actual frequency, $F(5, 90) = 2.43$, $p < .05$ ($MS_e = 0.21$). For the longest exposure time there was a significant effect of actual frequency, $F(5, 90) = 11.1$, $p < .01$ ($MS_e = 0.31$), but neither an effect of instruction, $F(1, 18) = 0.64$ ($MS_e = 0.83$), nor of the interaction of the two variables, $F(5, 90) = 0.58$ ($MS_e = 0.31$).

The final measure of performance was the mean variability of estimates. For each subject, a standard deviation was calculated of the estimates given for each actual frequency. These standard deviations were then averaged across the actual frequencies for each subject to yield one mean standard deviation per subject. The resulting means of these values are listed in Table 5 by intention condition and exposure time. Planned comparisons for each exposure time showed significant differences in favor of the intentional condition for the shortest exposure time, $t(18) = 2.24$, $p < .05$, and no differences for intermediate and long exposure times, $t(18) = 0.90$ and 0.39, respectively. The trend here is the same as with the other measures: Although there is no effect of intention for the longest exposure time, there is such an effect for the shortest exposure time.

We have replicated previous failures to find an effect of intention at the 4-s exposure time. By contrast, the shorter exposure

Table 5
Mean Standard Deviations of Judgments as a Function of Learning Instructions and Exposure Time

	Exposure time					
	1s		2s		4s	
Learning instructions	*M*	*SD*	*M*	*SD*	*M*	*SD*
Incidental	1.41	0.28	1.20	0.26	1.16	0.24
Intentional	1.13	0.23	1.08	0.31	1.22	0.30

... conditions.

General Discussion

The results of all three experiments seem quite clear. The first experiment showed that encoding of frequency information is influenced by competing task demands. Increasing the load of a competing task causes a decline in the accuracy of frequency judgments. These results are at odds with one of the main criteria suggested by Hasher and Zacks (1979, 1984) and by Zacks et al. (1982) to test whether frequency coding is based on the operation of automatic processes.

The results of the second experiment showed that a specific strategy manipulation (i.e., levels of processing) affected the encoding of frequency information, which is also at odds with a major criterion of automatic processing.

There are two conclusions that can be drawn from the results of Experiment 3. The first is that the effect of load in Experiment 1 was not due simply to a decrease in the time left for processing of frequency information after the load tasks had been completed. We conclude this because presentation time had relatively little effect on frequency judgments with intentional instructions in Experiment 3 (recall that Experiment 1 was also run under intentional instructions). Instead, the results of Experiment 1 are probably more directly attributable to either decreased momentary capacity devoted to frequency coding or to less effective coding strategies with heavier competing demands. In either case, the results remain at odds with a major criterion for automaticity suggested by Hasher and Zacks (1979).

The second conclusion to be derived from Experiment 3 is that the suggestion proposed by Fisk and Schneider (1984) has some merit: Automatic encoding of event frequency may be an "early asymptote" controlled process. That is, modification of long-term memory for frequency information may occur only in the first few seconds of processing, before it reaches an asymptotic level. Before reaching this level, intention to code frequency has an impact on later judgments. This effect could be mediated by either the amount of effort expended, by the type of strategy used, or by both.

One point about the suggested measures of frequency judgments. Although it turned out that in our experiments all three measures converged in the same direction, this might not have been the case. One can conceive of manipulations that will affect only the absolute level of performance, for example, but not the discriminability of the judgments (reflected in the slopes of the functions). Changing the instructions concerning the range of frequencies to be used in the test phase might have this effect. The convergence of all three measures used in our work is, thus, a robust indication of the differences we obtained.

Altogether, the picture that emerges from this work is consistent with the picture that has been developing from the research of Fisk and Schneider (1984) and Greene (1984). Two major criteria for automaticity are not satisfied for judgments of frequency: They are neither insensitive to competing task demand nor insensitive to manipulation of strategy.

Is frequency coding automatic, then? There are three answers: The first is to argue that frequency coding is probably not an automatic process because it violates some of the conditions claimed to be necessary criteria for automaticity (Greene, 1984). This interpretation should be viewed with caution for two reasons. First, frequency estimation is a task that has been subjected to tests of several criteria for automaticity, and it may satisfy some of these criteria (Hasher & Zacks, 1984). Second, there has been no convincing argument that each of the criteria frequently mentioned as a condition for automaticity is really a necessary condition. Rather, the most compelling demonstrations of automatic processing are those that have relied on strength in number. Until there is good theoretical reason to rely on one or another criterion, it is imprudent to reject the conclusion that a process is automatic simply because it is not consistent with some of the currently popular criteria. A principled argument for the choice of critical criteria is necessary first (see Jonides, Naveh-Benjamin, & Palmer, 1985, for a discussion of this issue).

The second answer hinges on whether automatic and effortful processes are as distinct as once thought. Perhaps there is a continuum linking completely automatic and nonautomatic processes. The differences among points on this continuum may be only a matter of degree. If so, then our results show that frequency coding is not at the extremely automatic end of this continuum, as previously argued. Where it is located is a question that cannot be addressed without a principled understanding of how the criteria suggested for automatic processes can be used to order the processes on the continuum when some criteria are satisfied and others are not (for a related discussion, see Zacks, Hasher, Alba, Sanft, & Rose, 1984).

The third answer to whether frequency judgments are mediated by an automatic mechanism is to question whether the question is appropriate. It should be obvious that there are many processes involved both in coding frequency information at the time of presentation and in retrieving information at the time of testing to produce estimates. For example, there may be contributions of direct and indirect mechanisms of coding that accumulate counts of frequency (Underwood, 1969) and that rely on other aspects of memory traces (Hintzman & Block, 1971; Howell, 1973). Although some of these processes may dominate under certain conditions, all may participate to some extent. Some of these processes may be automatic and some not. Our research involved manipulation of variables that had their most direct effect on the coding process (e.g., having to count at the same time that the words were presented). This research indicates that some aspect of coding is not automatic. Perhaps the most illuminating research at this point would be to further partition all the processes involved in generating frequency judgments to assess their characteristics. In short, a finer grained analysis of the frequency judgment task should be the order of the day.

References

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

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 Psychonomics Society, 6,* 321–324.

Glenberg, A., Smith, S. M., & Green, C. (1977). Type I rehearsal: Maintenance and more. *Journal of Verbal Learning and Verbal Behavior, 16,* 339–352.

Greene, R. L. (1984). Incidental learning of event frequency. *Memory and Cognition, 12,* 90–95.

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

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. (1970). Effects of repetition and exposure duration on memory. *Journal of Experimental Psychology, 83,* 435–444.

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

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

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye movement. In J. Long & A. D. Baddeley (Eds.), *Attention and Performance IX* (pp. 187–204). Hillsdale, NJ: Erlbaum.

Jonides, J., Naveh-Benjamin, M., & Palmer, J. (1985). Assessing automaticity. *Acta Psychologica, 60,* 157–171.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kausler, D. H., Lichty, W., & Hakami, M. K. (1984). Frequency judgments for distractor items in a short-term memory task: Instructional variation and adult age differences. *Journal of Verbal Learning and Verbal Behavior, 23,* 660–668.

Kausler, D. H., & Puckett, J. M. (1980). Frequency judgments and correlated cognitive abilities in young and elderly adults. *Journal of Gerontology, 35,* 376–382.

Kausler, D. H., Wright, R. E., & Hakami, M. K. (1981). Variation in task complexity and adult age differences in frequency-of-occurrence judgments. *Bulletin of the Psychonomics Society, 18,* 195–197.

Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Naveh-Benjamin, M., & Jonides, J. (1984). Maintenance rehearsal: A two-component analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 369–385.

Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review, 78*, 391–408.

Posner, M. I., & Snyder, C. R. (1975). Attention and cognitive control. In R. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55–86). Hillsdale, NJ: Erlbaum.

Rose, R. J., & Rowe, E. J. (1976). Effects of orienting tasks 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.

Rundus, D. (1977). Maintenance rehearsal and single-level processing. *Journal of Verbal Learning and Verbal Behavior, 16*, 665–681.

Schneider, W., & Shiffrin, R. N. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*, 1–66.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127–190.

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

Zacks, R. T., Hasher, L., Alba, J. W., Sanft, H., & Rose, K. C. (1984). Is temporal order encoded automatically? *Memory & Cognition, 12*, 387–394.

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 February 18, 1985
Revision received August 19, 1985 ■