Bulletin of the Psychonomic Society
1985, 23 (4), 387-390

The effects of rehearsal on frequency coding

MOSHE NAVEH-BENJAMIN
Ben-Gurion University of the Negev, Israel

and

JOHN JONIDES
University of Michigan, Ann Arbor, Michigan

It has been claimed (e.g., Hasher & Zacks, 1979, 1984) that encoding of frequency information is a result of an automatic process. This claim has been supported by experimentation that has included tests of several criteria suggested for the identification of automatic processes. The present experiment extends these tests to a case in which frequency judgments were obtained in a rehearsal task with massed presentation. In this task, frequency coding was very sensitive to strategy manipulation. In fact, in one condition there was no reliable evidence of any frequency coding. These results are at odds with the claim that frequency coding is automatic.

In recent years, there has been a growing interest in the study of automatic processes. One frequently cited example of such processes is the encoding of frequency information. Hasher and Zacks (1979, 1984) and Zacks, Hasher, and Sanft (1982) have argued that people automatically code how often an event has occurred. In order to test whether frequency information is encoded automatically, Hasher, Zacks, and colleagues have examined a number of criteria which should be satisfied by a process if it is automatic. Their results have shown that encoding of frequency information is not affected by age (Attig & Hasher, 1980; Hasher & Chromiak, 1977; Hasher & Zacks, 1979); does not benefit from practice (Hasher & Chromiak, 1977; Zacks, Hasher, & Sanft, 1982); is not influenced by competing task demands (Zacks, Hasher, & Sanft, 1982); does not show individual differences (Zacks, Hasher, & Sanft, 1982); and is not enhanced by intention to code frequency (Flexser & Bower, 1975; Zacks, Hasher, & Sanft, 1982).

The major assumption underlying all of these criteria is that automatic processes should function at a constant level under all circumstances. Thus, if frequency coding is such a process, we would expect that manipulation of coding strategy would not affect the fidelity of frequency judgments.

The purpose of the present experiment was to test whether variation in encoding strategy has an effect on frequency judgments. To render this a powerful test, we intentionally confounded two encoding strategies. The first one was the level of processing used by the subjects while encoding the information: In one condition subjects created a mental image of each pair of words presented; in the second condition subjects simply repeated the words aloud. The second encoding strategy was the intention to learn the information: In the condition in which subjects used mental imagery, they were also instructed to remember the words in order to be prepared for a later memory test; in the condition in which they only repeated the words, they were not led to expect a later memory test. If automatic processes function at a constant level under all circumstances, as Hasher and Zacks (1979) have claimed, we should expect no difference between the frequency judgments of the two groups.

This variation in encoding strategy was implemented in a task that involved variation in the number of times that subjects rehearsed the words whose frequency was later tested. The basic task was a modification of the classic Brown-Peterson short-term memory task (Glenberg, Smith, & Green, 1977). On each trial, subjects were given a set of three two-digit numbers to hold in memory. During the retention interval, they were given a pair of words. One group was told to create an image linking the words in the pair in preparation for an upcoming memory test for the words, and to elaborate this image as they rehearsed the words aloud. The other group was told simply to repeat the words aloud and was told nothing about a later memory test. For both groups, the number of overt rehearsals was varied from 1 to 10. After a series of such trials, we queried subjects from both groups about the number of times each word pair had been repeated—in short, a test of frequency for the word pairs. If frequency is automatically coded, there should have been no difference in performance of the groups on the frequency test. But there most definitely was.

METHOD

Subjects

Forty-two students at the University of Michigan served as paid volunteers in a session of approximately 75-min duration.

We would like to thank Pat Wilson for her help in data collection and analysis. This work was supported in part by NSF grant BNS-79-13757, and in part by a grant from AFOSR to John Jonides.

Requests for reprints 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.

Design

The experiment included two major independent variables. The first, varied within subjects, was the number of rehearsals required on each trial (1, 2, 4, 6, 8, or 10). The second, varied between subjects, was the instruction under which rehearsal occurred. In one condition, subjects were told simply to repeat the words; in the other, they were told to create images that would link the words and to elaborate these images in preparation for a later memory test. The dependent variables were the proportions correct in recall of numbers, and the frequency estimations for the words.

Apparatus

Stimuli were displayed on the face of a Hewlett-Packard terminal which was connected to a PDP-11/34 computer. The terminal was located in a lighted soundproof booth. A small microphone was placed in front of the subjects.

Stimuli

There were three types of stimuli employed: numbers and words in the learning phase and frequency test words.

Numbers. From the subjects' point of view, the main experimental task was to memorize and recall sets of three two-digit numbers that were selected randomly for each subject with the following constraints: All stimuli were drawn from the numbers 12-98 inclusive; no numbers with two identical digits were used (e.g., 55); no two numbers in a set were chosen from the same decade; no two numbers in a set were integer multiples of one another.

Words. Throughout each interval during which the numbers were to be retained, subjects were required to recite aloud (rehearse) a pair of words. The words chosen for rehearsal were all one-syllable words of three to five letters in length and of greater frequency than 20 occurrences per million (Kucera & Frances, 1967). In addition, they were paired so that the words in a pair neither rhymed nor had any obvious semantic relationship. Otherwise the selection and pairing of words was accomplished randomly.

Frequency test words. The test of frequency judgments was administered after all trials were completed. A list of 70 pairs of words was selected that included the 60 rehearsed pairs from the first part of the experiment plus 10 completely new word pairs. These 10 new pairs had the same characteristics as the other words. Subjects were asked to provide a value between 0 and 10 that corresponded to their estimate of the number of times that each pair of words had been rehearsed in the first part of the experiment. They were told that some of the pairs did not appear in the first part of the experiment, and these should be assigned a value of 0.

Trial structure. There were 80 trials in the experiment. Of these, the first 10 were considered practice and were not analyzed. The next 5 trials after the practice trials and the last 5 trials of the experiment were also excluded from analysis in order to minimize the effects of primacy and recency in the later frequency test. Thus, there remained 60 test trials, 10 for each value of rehearsal interval. Subjects rehearsed items 1, 2, 4, 6, 8, or 10 times on different trials. For each trial, the following sequence of events occurred:

First the word "ready" appeared for 2 sec, centered on the face of the terminal. Following this, three two-digit numbers appeared for 4 sec; they were horizontally arrayed and centered on the screen. The numbers were then replaced by two words that remained in view for 0.6 sec. Following presentation of the words, subjects heard a set of tones. The tones appeared at a rate of one every 1.2 sec. There was (were) 1, 2, 4, 6, 8, or 10 tone(s) presented. Each length of tone sequence appeared 10 times, and the various lengths were distributed randomly in a session in order to prevent subjects from predicting the length of rehearsal interval. Following the tones, the word "recall" appeared on the screen. Subjects then typed their recall of the three numbers and pressed the "return" key at their leisure, initiating a 5-sec interval before the next "ready" signal appeared.

Procedure

Subjects were run individually. The subjects in the incidental-maintenance rehearsal condition were told that they were in an experiment on short-term memory for digits and that they would have to repeat words aloud as a distractor activity between presentation and recall of the digits. They were not told about any test of the words. The subjects in the intentional-elaborative rehearsal condition were told about a later unspecified memory test for the words, and they were urged to rehearse the words elaboratively (using an interactive mental image for each pair) in order to remember them. They were cautioned, however, to keep their recall of the numbers at a reasonable level. Both groups were told to rehearse each word pair into a microphone at a pace set by the tones. Following the completion of the experimental trials, subjects were engaged in 2 min of interpolated counting activity after which the frequency estimation task was administered. Subjects had as much time as they needed to complete the test.

RESULTS

The reported results included only 40 of the 42 subjects. Two subjects were excluded from further analysis (one in the incidental condition who suspected the final frequency test; the other in the intentional condition who did not expect any memory test).

Number Recall

Table 1 presents the percentage of numbers recalled as a function of rehearsal condition and number of rehearsals. A strict criterion was used in which a response was scored as correct only if all three two-digit numbers were recalled in their correct order of presentation. The analysis showed that recall performance was significantly higher for the incidental-maintenance rehearsal condition [$F(1,38) = 10.45$; $p < .01$]. This is not surprising because subjects in the intentional-elaborative rehearsal condition were presumably devoting more effort to learning the words than were subjects in the incidental-maintenance rehearsal condition. Recall performance also dropped significantly with increasing number of rehearsals [$F(5,190) = 3.18$; $p < .01$]. This effect has been found in some other rehearsal studies (e.g., Naveh-Benjamin & Jonides, 1984a, 1984b). The interaction of rehearsal condition and number of rehearsals proved to be significant [$F(5,190) = 2.31$; $p < .05$], because, in the incidental-maintenance

Table 1
Mean and Standard Deviation of Percentage of Numbers Recalled as a Function of Actual Frequency and Rehearsal Condition

	Actual Frequency (Number of Rehearsals)											
	1		2		4		6		8		10	
Rehearsal Condition	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Incidental-Maintenance	63.5	36	62.3	31	63.9	38	53.4	30	57.4	31	61.4	32
Intentional-Elaborative	43.5	28	39.0	30	28.3	25	33.9	26	31.7	32	28.9	28

condition, the decrease in number performance across rehearsal intervals was rather small [$F(5,95) = 1.54$; n.s.].

Frequency Estimations

Several measures of performance were calculated and compared across conditions (see Naveh-Benjamin & Jonides, in press, for a detailed discussion of these measures). The first measure is the slope of the function relating estimated to actual 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 conditions. The values for these averaged slopes were 0.18 for the intentional-elaborative rehearsal condition and 0.01 for the incidental-maintenance rehearsal condition (standard deviations were 0.12 and 0.09 respectively). Analysis of these slopes revealed a significant effect of condition [$t(38) = 4.79$; $p < .01$]. This result indicates that subjects in the intentional-elaborative rehearsal condition could better discriminate among the actual frequencies of rehearsals. Furthermore, the analysis revealed that the slope in the incidental-maintenance condition was not different from zero [$t(19) = 0.81$], a fact which implies that the subjects in this condition could not reliably discriminate among any of the actual frequencies of presentation.

It is also possible to examine how the different rehearsal conditions affected the absolute level of frequency estimates. The estimated frequencies for each of the two rehearsal groups are shown in Table 2 as a function of actual frequency. The results show that the absolute level of judged frequency was higher in the intentional-elaborative rehearsal group. A two-way analysis of variance revealed a marginally significant effect of rehearsal condition [$F(1,38) = 2.95$; $p < .10$]. The effect of actual frequency and the interaction of actual frequency and rehearsal condition were both significant [$F(5,190) = 7.22$; $p < .01$, and $F(5,190) = 3.72$; $p < .01$, respectively]. Because high estimates for small frequencies imply overestimation and low estimates for large frequencies imply underestimation, the significant interaction reflects the superiority of performance of the intentional-elaborative rehearsal group over the incidental-maintenance rehearsal group.

The final measure of performance was the mean variability of estimates (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 to yield one mean standard deviation per subject. The resulting means of these values were 1.95 for the intentional-elaborative rehearsal group and 2.23 for the incidental-maintenance rehearsal group (standard deviations of these means were 0.43 and 0.58 respectively). These results indicate that the dispersion of the judgments was larger in the incidental-maintenance group. This trend was confirmed by the analysis which showed the difference in standard deviations to be marginally significant [$t(38) = 1.71$; $p < .10$]. (One reason for obtaining only a marginally significant effect could be that some subjects in the incidental-maintenance rehearsal condition tended to give only low estimates, a fact which reduced the range of their judgments' dispersion).

DISCUSSION

The results of the above experiment are quite clear. All three measures of performance on the frequency judgment task show the superiority of the intentional-elaborative rehearsal condition over the incidental-maintenance rehearsal condition. Subjects' estimations in the intentional-elaborative rehearsal condition showed better discriminability among the presented frequencies, higher absolute level of judged frequencies, and less dispersion of the frequency judgments than those in the incidental-maintenance rehearsal condition. Such results contradict one of the major criteria suggested by Hasher and Zacks (1979) for automatic processes, and they extend the conditions under which it has been shown that strategy manipulations such as intention to learn or level of processing affect frequency judgments (Fisk & Schneider, 1984; Greene, 1984; Naveh-Benjamin & Jonides, in press). The present experiment reveals an effect of coding strategy under conditions in which repetition is produced by rehearsal and in which presentation is massed rather than spaced as in previous studies.

What could lead to performance differences between the conditions? There are two obvious variables: intention to learn and level of processing. Some studies have found no effect of intention (e.g., Flexser & Bower, 1975; Zacks, Hasher & Sanft, 1982); however, there are recent reports showing that under some conditions there can be an effect of intention (fast presentation rate: Naveh-Benjamin & Jonides, in press; use of true incidental learning: Greene, 1984).

There are demonstrations showing that the levels-of-processing variable has an effect on judgments (e.g., Rose & Rowe, 1976). However, Hasher and Zacks (1984) have claimed that this effect is actually mediated by extra covert rehearsals given to the stimuli under study, rehearsals which were later confused with actual presentation. We find this claim somewhat troublesome since it could make tests of the automaticity of frequency coding impossible: any obtained effects on frequency judgments of strategy manipulations, practice, or competing load could be claimed to be mediated by extra covert rehearsals. How is this tested? In addition, how could extra covert rehearsals account for the effect of elaborative rehearsal on all three reported measures of frequency judgments? For example, we would expect the effect of extra covert rehearsals produced in this condition to increase the variability of the judgments rather than to decrease it, as actually happened.

Note that in general, the fidelity of the judgments obtained in the present experiment is quite low in comparison to studies that have used stimu-

Table 2
Mean and Standard Deviation of Estimated Frequencies as a Function of Actual Frequency and Rehearsal Condition

	Actual Frequency (Number of Rehearsals)											
	1		2		4		6		8		10	
Rehearsal Condition	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Incidental-Maintenance	2.9	1.6	3.1	1.4	3.2	1.7	3.1	1.8	3.4	1.8	3.1	1.3
Intentional-Elaborative	2.8	1.0	3.1	0.8	3.6	0.9	3.9	0.9	4.3	1.1	4.4	1.0

lus repetition and spaced presentation (e.g., Greene, 1984; Naveh-Benjamin & Jonides, in press). For example, the mean coefficients of discrimination (the computed correlation between estimated and actual frequency, which is the measure of fidelity used in some previous studies) in the current study were 0.03 for the incidental-maintenance condition, and 0.22 for the elaborative-intentional condition. These are much lower than values reported in the past (e.g., 0.75 obtained by Flexser & Bower, 1975; 0.85 for the incidental condition obtained by Rose & Rowe, 1976; and a range between 0.41 and 0.71 obtained by Greene, 1984).

Even more striking is the result for the incidental-maintenance rehearsal condition. Here the slope of the estimated versus actual frequency function did not reliably differ from zero, so we cannot reject the hypothesis that subjects in this condition did not encode any information that enabled them to discriminate among frequencies of occurrence. Such results are in clear contradiction to a major assumption concerning automatic processes: that they function at a constant level under all circumstances. At the very least, this assumption seems to be inappropriate as a description of frequency coding. In fact, it may be too stringent an assumption for automaticity assessment in general.

One possible explanation of the poor judgmental performance in this experiment, and in the incidental-maintenance rehearsal condition in particular, is that it is due to the use of massed presentation (i.e., rehearsal). The most prevalent theories of the spacing effect attribute the disadvantage of massed presentation to a voluntary change in the direction of processing effort, to involuntary habituation to the presented stimuli (Hintzman, 1974), or to a blend of these two mechanisms (the effort theory, suggested by Johnston & Uhl, 1976). Although there are some differences between these hypotheses, they both assert that decreased performance under massed repetition is due to some lack of attention paid to further appearances of a stimulus. Of course, if frequency information were coded automatically, with minimal resources required from a limited-capacity attentional mechanism, it should not be affected heavily by the spacing of presentation, regardless of which hypothesis provides the better account of the spacing effect (see also Rose & Rowe, 1976, for a reported effect of spacing on frequency judgments).

Note also that the lack of decrease in number performance in the incidental-maintenance condition, which might reflect changes in allocation strategies of subjects, does not help the automaticity position because a major characteristic of automatic processes is their insensitivity to competing task demands.

The results of the present experiment do not support the claim of automaticity for frequency coding. We do not stand alone in failing to support this claim. There have been reports showing effects of intention on frequency coding (Greene, 1984; Fisk & Schneider, 1984; Naveh-Benjamin & Jonides, in press). Also, there are apparently effects of other strategy manipulations and of competing task demands on judgments of frequency (Naveh-Benjamin & Jonides, in press). To supplement these recent reports, Greene (1984) has raised questions about some of the other criteria used to test for automaticity in frequency coding. Upon careful examination, the automaticity position seems weak.

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 & 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.

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

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

Hasher, L., & Chromiak, W. (1977). The processing of frequency information: An automatic mechanism? *Journal of Verbal Learning & 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. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola symposium* (pp. 77-100). Potomac, MD: Erlbaum.

Johnston, W. A., & Uhl, C. N. (1976). The contributions of encoding effort and variability to the spacing effect on free recall. *Journal of Experimental Psychology: Human Learning & Memory*, **2**, 153-160.

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

Naveh-Benjamin, M. & Jonides, J. (1984a). Cognitive load and maintenance rehearsal. *Journal of Verbal Learning & Verbal Behavior*, **23**, 494-507.

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

Naveh-Benjamin, M., & Jonides, J. (in press). On the automaticity of frequency coding: Effects of competing task load, encoding strategy, and intention. *Journal of Experimental Psychology: Learning, Memory, & Cognition*.

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

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

(Manuscript received for publication June 3, 1985.)