Truly incidental encoding of frequency information

LYNN HASHER
Temple University
ROSE T. ZACKS
Michigan State University
KAREN C. ROSE
Temple University
HENRIANNE SANFT
Carnegie Mellon University

Four experiments demonstrated that adults can reliably remember frequency of occurrence information about items they have been exposed to under truly incidental memory conditions. Subjects neither knew that the ultimate test task would concern item frequency nor that they had any reason to remember the items. This was accomplished by presenting items under the guise of one of three cover tasks: anagram solving, sentence completion, and picture naming in a Stroop-like task. In addition, one experiment found that subjects who were prewarned for either a nonspecific memory test or a frequency test were no better able to judge frequency than were subjects operating under truly incidental conditions.

Hasher and Zacks (1979, 1984) have argued that frequency of occurrence information is encoded into memory automatically (see also Hintzman & Stern, 1978). To make this argument, Hasher and Zacks stipulate that a number of performance criteria must be jointly satisfied (1984, p. 1373). Two of these are that (a) automatically encoded information is stored in memory in the absence of conscious intentions to do so, and (b) the quality of this encoding is uninfluenced by the voluntary goals or strategies of an individual (e.g., Hasher & Zacks, 1979, 1984; Posner & Snyder, 1975).

Most experiments on intentionality effects vary the frequency with which critical items occur and compare the frequency-test performance of subjects given one of two types of instructions: instructions that prewarn of an unspecified memory test (here called an "incidental" or "nonspecific" condition) or instructions that explicitly inform subjects of the nature of the memory test (an "intentional"

AMERICAN JOURNAL OF PSYCHOLOGY
Spring 1987, Vol. 100, No. 1, pp. 69–91

condition). In no case is there evidence that giving specific instructions for a forthcoming frequency test boosts performance over that obtained with incidental instructions for either a judgment or a forced-choice discrimination test (e.g., Attig & Hasher, 1980; Flexser & Bower, 1975; Hasher & Chromiak, 1977; Hasher & Zacks, 1979, Experiment 2; Kausler & Puckett, 1980; Rose & Rowe, 1976). Other experiments have contrasted specific and correct instructions about the frequency test with specific but misleading instructions about another memory test such as free recall (Howell, 1973; Zacks, Hasher, & Sanft, 1982). These studies also report no effect of instructions on memory for frequency of occurrence.

Such methods are, however, open to criticism (see Greene, 1984; Kausler, Lichty, & Hakami, 1984). In particular, concern centers on the validity of assessing the impact of specific intentional strategies on frequency encoding by comparing intentional instructions with incidental/nonspecific and/or with misleading instructions. The criticism rests on two considerations: (a) instructions given to both the incidental/nonspecific and misleading groups warn of a future memory test and so are likely to activate some conscious encoding strategies; and (b) intentionally instructed subjects may not have any preexisting strategies appropriate for encoding frequency, and so they may rely on rehearsal strategies that are similar to those used when other (more familiar) sorts of tests are expected. Thus, the incidental/nonspecific condition may not achieve the experimenter's goal of studying the encoding of frequency in the absence of intentional strategies; also, the different types of instructions may fail to provide strong contrasts among different encoding strategies (see Postman, 1982, for a related argument).

These objections suggest that a test of the intention and instruction criteria for automaticity will require the use of truly incidental instructions: instructions that not only do not inform subjects that frequency knowledge will be tested for but that also hide from subjects the fact that *any* memory test will be administered. Of course, work on subjects' knowledge of frequency differentials in such naturally occurring events as syllables, single letters, and pairs of letters (see Zechmeister & Nyberg, 1982, pp. 149–150) suggests that frequency information may well be coded under truly incidental conditions. Further, three recent papers (Greene, 1984; Kausler et al., 1984; Marmurek, 1983) report confirming evidence of memory for frequency under truly incidental conditions. In all three, a paradigm developed by Glenberg, Smith, and Green (1977) to study maintenance rehearsal was used to present the target words. This paradigm uses a variant of the Brown-Peterson short-term memory task in which the ostensive to-be-remembered items are sequences of four or five digits. The distractor activity that fills the retention interval for each digit sequence is overt repetition of words. Because subjects do not intend to remember the distractor words, any knowledge of the frequency with which they occurred (i.e., of the number of different distractor trials in which they appeared) can be considered to have been obtained under truly incidental conditions. The finding that subjects can reliably discriminate differences in frequency under truly incidental conditions is in clear agreement with the automaticity view.

However, the general question of whether frequency is encoded under truly incidental conditions is too important for a conclusion to be reached on the basis of the results of a single paradigm. Thus, one objective of the present research was to provide demonstrations of the truly incidental coding of frequency information under a variety of different cover tasks. The four experiments reported here used truly incidental conditions in which the experimental subjects were not aware, until the time of the final test, that they were participating in a memory experiment. From a subject's point of view, the entire purpose of each experiment was to measure his or her performance on an orienting task carried out on a series of items. Across the four experiments three different orienting tasks were used: solving anagrams in Experiment 1; judging whether a word fit meaningfully into a sentence frame in Experiment 2; and naming pictures in Experiments 3 and 4. Thus the experiments permit us to determine whether frequency is encoded in the absence of any conscious intention to do so, as well as in the absence of any intention to remember the target items.

According to the automaticity view, not only should frequency be coded under truly incidental conditions, but in addition, performance should be no better when subjects are warned of a memory test, whether or not that warning includes information about the exact nature of the test. Two papers using the Glenberg et al. (1977) paradigm provide a test of this automaticity criterion. In one of them (Kausler et al., 1984, Experiment 1; see also Experiment 2), there were no significant differences among the three instructional conditions. However, in the other (Greene, 1984), performance on the frequency test was poorer in the truly incidental condition than in the two prewarned conditions, one giving subjects only general memory test instructions and the other explicitly informing subjects of the upcoming frequency test. The latter two conditions did not differ from each other (Experiment 2).

Not surprisingly, there are any number of procedural differences between the two sets of studies (e.g., in the specific instructions and in the actual testing procedures) that might account for the discrepancy in findings. In our view the crucial difference probably lies in the differential rehearsal patterns initiated by critical differences in the distractor task procedures used in the two investigations. These, in turn, probably resulted in differential rates of reality-monitoring confusions (Johnson & Raye, 1981), resulting in differences in frequency performance. We discuss this issue in some detail in the general discussion.

In any event, the conflicting data regarding the impact of the instructional effects on frequency discrimination led us to take another look at the effect of instructions on the processing of frequency. This was done in Experiment 4, in which we compared frequency test performance for subjects under two intentional conditions with that of subjects under truly incidental conditions.

EXPERIMENT 1

This experiment assessed the encoding of frequency under truly incidental instructions using a cover task in which subjects believed that they were being timed to solve a series of anagrams. No mention was made of a memory test until the anagram series was completed. Subjects were then asked to estimate the frequency with which words had been presented. Critical words appeared 0, 1, 2, and 3 times. In this experiment, as in the subsequent three, subjects had no reason to attend deliberately to the frequency with which words occurred, nor did they have any reason to prepare for a memory test.[1]

METHOD

Subjects and design

The 32 subjects were introductory psychology students, who participated to fulfill a course requirement. The experiment was a 2 × 4 within-subjects design in which two levels of anagram difficulty (easy and hard)[2] were crossed factorially with four levels of frequency of occurrence (0, 1, 2, and 3). In this and all subsequent experiments, subjects were tested individually.

Materials

Each subject was exposed to anagrams based on 18 different words selected from a pool of 24 words. The words were ones whose letters could not be rescrambled to make any other English word (i.e., single-solution anagrams; Olson & Schwartz, 1967).

The normative data of Mayzner and Tresselt (1958) were used to generate easy and hard anagrams. They found that when the letter orders of anagram and word were maximally similar, the solution was quickly obtained (easy letter orders); by contrast, when the orders of anagram and word were maximally dissimilar, the solution process took longer (hard letter orders). Three unique easy letter orders and three unique hard letter orders were chosen. Words occurring 2 and 3 times were shown in a different order at each presentation. For words occurring 1 and 2 times, the particular letter orders used were selected randomly from the three available at each difficulty level.

Throughout this series of experiments, multiple lists were constructed from the basic set of materials. These functioned to independently counterbalance the items across the difficulty and the frequency levels.[3] Also, the presentation orders used in this research always spaced the repetitions of items receiving multiple presentations. Items representing different conditions (e.g., easy vs. hard anagrams) were evenly distributed across input positions. Otherwise, presentation order was random. In Experiment 1, there were eight unique lists consisting of 36 anagrams with each list given to 4 subjects.

Procedure

Each anagram was typed on a separate 3 × 5-in. (7.6 × 12.7-cm) card. Subjects turned to the next anagram as soon as the current one was solved or at the end of 2 min if the anagram remained unsolved. After 90 s of unsuccessful work on an anagram, subjects were given the first letter of the solution word to use as a hint. Subjects reported their solutions aloud, and solution times were recorded. Following the final anagram, subjects were unexpectedly given a frequency judgment test that consisted of a listing of 24 words, 6 at each frequency of occurrence (0–3). For each of the (actually presented) frequency levels, three of the words had been easy anagrams and three had been hard. The words were randomly arranged down the left-hand side of a page, with the constraint that one word from each of the eight conditions occurred in each third of the series. Subjects were instructed to write a number next to each word that indicated how often it had appeared on the list. They were informed that some words had never been presented and that these should be given a zero. The test was unpaced.

RESULTS AND DISCUSSION

An alpha level of .05 is used throughout this paper. Two measures of performance on the anagram-solving task were used: solution probabilities and solution times. Words that occurred more than once were always solved after the first presentation, but on their first occurrence hard anagrams were slightly less likely to be solved ($p = .94$) than easy anagrams ($p = .99$), $t(31) = 5.00$. The solution times are shown in Table 1. The solution times for hard anagrams were longer than those for easy ones, $F(1, 31) = 32.39$, $MS_e = 216.67$ (confirming the validity of the anagram difficulty manipulation). Also, solution times declined with practice, $F(2, 62) = 41.14$, $MS_e = 161.60$, and the decline was greater for hard than for easy anagrams, $F(2, 62) = 3.36$, $MS_e = 159.58$.

Mean frequency judgments for presented items are also shown in Table 1. Because of few nonzero frequency judgments to nonpresented items (e.g., less than 5% in Experiment 1), these items are not included in the analyses of the frequency data for any of the experiments that included nonpresented items on the frequency judgment test (Experiments 1, 3, and 4).[4] As is apparent in the table, solution difficulty did not affect the frequency judgments ($F < 1$). Only the effect of number of presentations was significant, $F(2, 62) = 168.82$, $MS_e = .26$. Mean judged frequency increased regularly as a function of increased actual frequency, and Newman-Keuls tests showed that each increase was significant ($p < .05$). Additionally, for each subject, Pearson product-moment correlations were computed between the judged frequency of individual items and their actual frequency (nonpresented items were not included). The means of these correlations, calculated separately for easy and hard anagrams, were .70 and .73, respectively.

The group data describe fully the performance of individual subjects: For 29 of the 32 subjects, the mean judgment for the once-presented items was lower than that for the twice-presented items which in turn was lower than that for items presented three times. Two of the remaining 3 subjects had a tie between two adjacent frequencies, and only 1 subject had a reversal where a higher frequency received a lower mean judgment. Furthermore, of the 64 individual correlations between judged and true frequency, 63 were positive and 52 were significant at $p < .05$.

These data join with those using the distractor paradigm (Greene,

Table 1. Experiment 1. Mean solution times (s) and mean frequency judgments at each difficulty and presentation frequency level

	Frequency level			
Difficulty	1	2	3	*M*
Solution times				
Easy	18.05	8.22	3.70	9.99
Hard	36.01	20.13	10.10	22.08
M	27.03	14.18	6.90	—
Frequency judgments				
Easy	1.12	2.16	2.75	2.01
Hard	1.08	1.98	2.78	1.95
M	1.10	2.07	2.77	—

1984; Kausler et al., 1984; Marmurek, 1983) to confirm that frequency information is encoded under truly incidental conditions. When subjects have no particular reason to code items for frequency and, indeed, no particular reason to be prepared for a memory test, they are nonetheless able to discriminate differences reliably in frequency of occurrence.

EXPERIMENT 2

This experiment permits an assessment of whether frequency is encoded under truly incidental conditions in a sentence completion task. In this cover task, subjects were shown a target word together with a sentence frame that omitted a single word. Subjects decided whether the target word completed the frame to make a meaningful sentence. In some cases, only a very few words could complete the sentence (easy decisions); in others, a broad range of words could complete the frame (hard decisions; cf. Tyler, Hertel, McCallum, & Ellis, 1979). Target words occurred 1, 2, and 3 times in the series. The instructions informed subjects that the experiment was concerned with "sentence completion tasks"; only when the orienting task was completed was the memory task mentioned.

METHOD

Subjects and design

Twenty-four undergraduate students, who were each paid $2.00, participated in the experiment. The design was a 2 (Easy vs. Hard Decisions) × 2 (Yes vs. No Answers) × 3 (Frequencies 1, 2, and 3) within-subjects factorial.

Materials

Thirty-six common words were chosen from the Thorndike-Lorge (1944) norms. For each of them, six sentences were written to serve as meaningful frames (i.e., the correct response to the question of whether the word fits the frame is *yes*); three were in an easy format and three in a hard format. Most often, hard format sentences were created by removing a concrete referent having some preexisting associative strength to the target word. For example, for the target word *doctor*, one of the easy/yes frames was "The emergency room was chaotic because there was no ____ available." In the hard/yes version of this frame, the word *emergency* was deleted. In addition to the six meaningful frames for each target word, there were 108 frames used in the no condition; these could not be meaningfully completed by the target words. Ratings of a separate group of 15 subjects indicated that although our yes condition materials provided a valid manipulation of decision difficulty, our no condition materials did not (see Zacks et al., 1983, p. 750, for details). However, because we are not concerned in this paper with encoding difficulty effects (see Note 2), we will report the data from both the yes and the no conditions.

The initial pool of 36 words was divided among the four sets of 9 words each. These were rotated among the four basic experimental conditions (Easy vs. Hard × Yes vs. No). Words were also rotated across frequency levels (1, 2, and 3 presentations), and the result was 12 different lists, each of which was given to 2 subjects. With two buffer items at the beginning and two at the end of the list, each list's total presentation series consisted of 76 items. When target words appeared more than once, each repetition was paired with a different sentence frame.

Procedure

The apparatus consisted of an AIM 65 microcomputer interfaced with a slide projector and a response box containing two operative response buttons, the left one for *no* answers and the right one for *yes* answers. On each presentation trial, a blank slide was presented for 1,500 ms, followed by a target word presented for 1,500 ms, followed by a sentence frame. The subject terminated the sequence by pressing the appropriate response button. The computer controlled the slide projector and recorded the subject's responses as well as the response latencies.

Instructions for the encoding task asked subjects to decide whether each target word could complete the sentence frame with which it was paired. Both speed and accuracy of response were emphasized. Subjects were also informed that some of the target words would occur more than once. An unpaced and unexpected frequency judgment test immediately followed the encoding task. For the frequency test, subjects were given a sheet of paper listing the 36 target words in random order. Next to each word was a blank line on which subjects were to write their best estimate of how many times that word had occurred in the preceding series.

RESULTS AND DISCUSSION

Mean decision times on the encoding task are shown in Table 2. (On approximately 5% of the trials, subjects disagreed with our determination that the target word did or did not fit the sentence frame; these trials are included in the reported means and in the analyses.) Analysis of the decision times revealed three significant effects: *No* responses took longer than *yes* responses, $F(1, 23) = 19.19$, $MS_e = 236{,}511$; response times were longer in the easy than in the hard condition, $F(1, 23) = 19.54$, $MS_e = 83{,}482$; and there was a significant interaction between these two variables, $F(1, 23) = 7.26$, $MS_e = 141{,}756$, such that the easy-hard difference was limited to the no condition. (The trend toward slower decision times in the easy condition was unexpected and probably occurred because the easy sentence frames were on average 1.6 words longer than the hard ones; see Zacks et al., 1983, Experiment 2.)

Table 2 also presents the frequency judgment data. Here the only significant effect is for the number of presentations, $F(2, 46) = 79.59$, $MS_e = .75$. Each increase in actual frequency resulted in a significant increase in mean judged frequency ($p < .05$, Newman-Keuls tests). The subjects' sensitivity to frequency is also apparent in correlational analyses of their judgments. The means of the individual subject correlations between judged and true frequency ranged from .53 to .55, depending on the subset of items included in the computations. For example, for all the items together, the mean correlation was .53; for the *no* items only, it was .53; and for the *yes* items only, it was .55.

Further, the group results are fully descriptive of the data of individual subjects.

1. The overwhelming majority of the subjects (23 of 24) showed the following pattern in their mean frequency judgments: Items presented once had a lower mean judgment than items presented twice, which in turn had a lower mean judgment than items presented three times. The one subject who deviated from this pattern had a tie between two adjacent frequencies.

2. Individually, almost all subjects had significant positive correlations between judged and true frequency. For correlations that included all items, 24 of 24 were positive correlations of which 23 were significant.

As in Experiment 1 then, subjects stored information permitting frequency judgments even though they had no specific reason to do

Table 2. Experiment 2. Mean decision times (ms) and mean frequency judgments for *yes* and *no* conditions at each difficulty and presentation frequency level

	Frequency level							
	Yes conditions				*No* conditions			
Difficulty	1	2	3	*M*	1	2	3	*M*
Decision times								
Easy	2,793	2,929	2,960	2,954	3,360	3,254	3,358	3,324
Hard	2,860	2,846	3,062	2,923	3,124	3,034	3,007	3,055
M	2,917	2,888	3,011	—	3,242	3,144	3,183	—
Frequency judgments								
Easy	1.26	2.14	2.94	2.11	1.33	2.11	2.89	2.11
Hard	1.31	1.96	2.86	2.04	1.25	2.04	2.78	2.02
M	1.29	2.05	2.90	—	1.29	2.08	2.84	—

so. Nor did they have any particular reason to try deliberately to remember items. The next two experiments report similar findings from yet another truly incidental paradigm.

EXPERIMENT 3

Here, and in the fourth experiment, a picture-word version of the Stroop interference task served as the orienting activity. Subjects had to name pictures of familiar objects that in some cases had interfering labels superimposed on them (e.g., a picture of a cigar with the word *pipe*) and in other cases did not (e.g., the picture of the cigar alone). The presence of interfering labels makes picture naming more difficult and slows it down (e.g., Lupker, 1979). Instructions informed subjects that the experiment was concerned with the speed and accuracy of picture naming. Critical pictures occurred 0, 1, 2, and 4 times in the series, and as in the previous studies, no mention of a memory test was made until after the presentation series was completed.

METHOD

Design and subjects

Experiment 3 had a 2 (Labeled vs. Unlabeled Pictures) × 4 (Frequency Levels 0, 1, 2, and 4) design with the label factor manipulated between subjects and the frequency factor within subjects. Twenty volunteer undergraduate subjects participated in each group.

Materials

The Snodgrass and Vanderwart (1980) norms were used to select 36 line drawings that met two criteria: (a) Each had received at least a 90% agreement level for the labels assigned, and (b) taken together, the pictures represented a broad range of semantic categories. Interfering labels for the pictures were generated by 10 undergraduates, none of whom took part in the previous experiments. Each saw the pictures one at a time and was asked to generate the first word thought of (other than the name of the picture) from the same semantic category as the picture. For each picture, the modal response was chosen to serve as the interfering label in the experimental task. In each label condition, there were four unique lists, each containing nine pictures at frequency levels 0, 1, 2, and 4. The remaining nine pictures were reserved for use as nonpresented items on the frequency judgment task. The presentation series contained 72 pictures, the 1st and last 3 of which were buffer pictures. For pictures presented more than once, the labels were always the same.

Procedure

Experiment 3 used the same equipment as Experiment 2 except that a voice-activated relay replaced the yes/no response box. The computer timed subjects from the onset of each slide to their naming of the picture. There was a 1,000-ms delay between the naming of a picture and the presentation of the next one. Subjects were instructed to name each picture as quickly and as accurately as they could. They were warned not to correct out loud any errors they might make but instead to wait for the next slide. Before the experiment proper began, subjects were given practice at responding into the microphone by having them read a series of 20 slides of first names. Subjects were not prewarned of a memory test.

RESULTS AND DISCUSSION

In accordance with the Stroop phenomenon, naming times were longer for the group with interfering labels than for the no-label group (see Table 3), $F(1, 38) = 8.18$, $MS_e = 80{,}530$. Also, naming latencies declined over presentations, $F(3, 114) = 47.98$, $MS_e = 7{,}433$, but the decline was greater in the label than in the no-label condition, $F(3, 114) = 3.67$, $MS_e = 7{,}433$.

In the analysis of the frequency judgments, only the effect of number of presentations was significant, $F(2, 76) = 301.32$, $MS_e = 21.18$. Newman-Keuls tests showed that the differences in mean judgments were all significant ($p < .05$). The mean correlation between judged and actual frequency was .73 for the label condition and .76 for the no-label condition.

The consistency of the frequency results across subjects is again striking: For all 40 subjects of this experiment, the mean judgment for items presented four times was higher than that for items presented two times, and in turn the mean judgment of the twice-presented items was greater than that for items presented once. Additionally, each of the 40 individual-subject correlations of judged and true

Table 3. Experiment 3. Mean picture-naming times (ms) and mean frequency judgments for no-label and label conditions at each presentation frequency level

	Frequency level				
Label condition	1	2	3	4	*M*
Naming times					
No label	1,636	1,507	1,484	1,482	1,527
Label	1,814	1,673	1,584	1,551	1,655
M	1,725	1,590	1,534	1,517	—
Frequency judgments					
No label	1.38	2.66	—	4.17	2.75
Label	1.31	2.58	—	4.13	2.67
M	1.35	2.62	—	4.15	—

frequency was positive, and in all but one case, the correlation was significant at $p < .05$.

Thus, three experiments, each using a different cover task, demonstrate that during exposure to items, subjects store information that permits reliable frequency judgments, even though they neither had the intention to encode frequency nor the intention to remember the items themselves. These findings, together with data from the truly incidental conditions in the distractor task paradigm (Greene, 1984; Kausler et al., 1984; Marmurek, 1983) provide solid support for the assumption that frequency information is encoded without any intention to do so, thus meeting at least one of the explicit criteria for determining the automaticity of encoding of an attribute (Hasher & Zacks, 1979, 1984). A second criterion stipulates that the addition of intention to encode an attribute that is otherwise encoded automatically will not improve performance. This criterion is considered in the fourth and final study in this series.

EXPERIMENT 4

In this experiment we again assess the automaticity criterion of incidental encoding. In addition, we compare frequency judgments made by subjects operating under truly incidental instructions with those made by subjects given two kinds of intentional instructions. As in Experiment 3, a picture-word Stroop task was used.

METHOD

Design and subjects

The design of this experiment was a 3 (Label Condition) × 4 (Frequencies 0, 1, 2, and 4) × 3 (Instructions) factorial. Only the last variable was tested between subjects. The three label conditions were no labels, labels unrelated to the picture they were superimposed on, and related labels.[5] The three instructional conditions included a truly incidental group informed only about the Stroop task, a standard intentional group told about the frequency task and asked to treat both the frequency and the Stroop tasks as equally important, and a frequency emphasis intentional group told that the frequency task was more important than the Stroop task. There were 16 undergraduates in each of the three groups.

Materials

The 36 critical pictures of Experiment 4 were line drawings taken from the Peabody Picture Vocabulary Test. The related labels for these pictures were selected by searching through association norms and by asking 8 subjects to give free associations (nouns) to a list of nouns that named the experimental pictures. Labels that were unrelated to any of the pictures were chosen. In each of the four unique lists, three pictures represented each of the combinations of Label Condition × Frequency 0, 1, 2, or 4. The presentation order for each list contained 80 pictures that included five buffers at the beginning and end and seven fillers spread throughout. Repeated pictures had the same label on all presentations.

Procedure

Experiment 4 used the same equipment and procedures as Experiment 3 with the exception of the instructions in the intentional conditions. Subjects in the standard intentional group were asked to "try to work at remembering how often each picture occurs while doing the naming task." Subjects in the frequency-emphasis intentional group were told to put primary emphasis on preparing for the frequency task and to treat the picture-naming task as secondary. On the frequency judgment task, subjects were given a list of the picture names (including those of the nine pictures that were in the zero-presentations condition) and were asked to judge the frequency with which each occurred.

RESULTS AND DISCUSSION

Mean naming latencies for each instructional condition are presented in Table 4. Three significant effects in these data are those for the main effect of label condition, $F(3, 135) = 24.86$, $MS_e = 12{,}478$; the main effect for presentation number, $F(3, 135) = 205.14$, $MS_e = 5{,}654$; and their interaction, $F(6, 270) = 4.41$, $MS_e = 4{,}190$. These results are similar to those of Experiment 3.

Table 4. Experiment 4. Mean picture-naming times (ms) and mean frequency judgments at each instruction condition and presentation frequency level

	Frequency level				
Instruction condition	1	2	3	4	*M*
Naming times					
Truly incidental	1,495	1,349	1,311	1,295	1,362
Standard intentional	1,580	1,426	1,403	1,389	1,449
Frequency-emphasis	1,549	1,412	1,361	1,356	1,419
M	1,541	1,396	1,358	1,346	—
Frequency judgments					
Truly incidental	1.23	2.49	—	3.84	2.52
Standard intentional	1.26	2.67	—	4.12	2.69
Frequency-emphasis	1.25	2.83	—	4.25	2.78
M	1.25	2.66	—	4.07	—

Note. The data are collapsed across the label conditions.

1. Naming times were faster in the no-label condition (M = 1,364 ms) than in the conditions with interfering labels (M = 1,436 and 1,431 ms) for the unrelated-label and the related-label conditions, respectively.

2. Naming time decreased over presentations.

3. The decrease was slightly greater in the conditions with labels than in the no-label condition.

There was in addition a significant effect of instructional condition, $F(2, 45) = 3.63$, $MS_e = 104{,}051$: Subjects in the standard intentional condition took longer to name the pictures than did subjects in the truly incidental condition (Newman-Keuls test), suggesting that subjects in at least this intentional condition were trying to follow instructions and prepare for the frequency test. Naming times for the frequency emphasis group were longer than times for the truly incidental group, suggesting that these subjects, too, were responding to the instructions, but the increased time was not significantly different from that taken by the truly incidental group.

Also shown in Table 4 are the mean frequency judgments (collapsed across the label conditions). The two significant effects in the judgment data are those for number of presentations, $F(2, 90) = 578.73$, $MS_e = 4.46$, and for the Instruction × Label Condition interaction, $F(4, 90) = 2.71$, $MS_e = 2.20$. The differences that produced the latter effect are small and not readily interpretable because the order of the three label conditions was different in each of the instructional groups. Moreover, this effect does not include the frequency variable, so it is not indicative of any differences in ability to discriminate frequency of occurrence as a function of cover task instructions. More important is the fact that neither the Frequency × Instruction interaction, $F(4, 90) = 1.09$, $MS_e = 4.46$, nor the Frequency × Instruction × Label Condition interaction $F(8, 180) = 1.85$, $MS_e = 1.88$, was significant. The main effect of instructional condition was also nonsignificant, $F(2, 45) = 1.44$, $MS_e = 15.56$. Further analysis of the frequency main effect by Newman-Keuls tests showed that (as in Experiments 1–3) the differences between successive frequency levels were significant ($p < .05$). The sensitivity of the subjects to frequency of occurrence in this paradigm is also demonstrated by the correlations between judged and actual frequency: The means of the individual-subject correlations were .82, .80, and .81 for the truly incidental, standard intentional, and frequency emphasis conditions, respectively.

Finally, we note that again the group data apply to each of the subjects. For all 48 subjects in this experiment, the mean judgment of items presented four times was higher than that for items presented two times, which was in turn higher than that for items presented one time. Indeed across the four experiments reported here, only 4 of 144 subjects failed to conform completely to this pattern. Additionally, all 48 subjects of Experiment 4 individually had a significant positive correlation between judged and true frequency.

Two conclusions are clearcut: (a) Frequency of occurrence is stored under trully incidental conditions, and (b) the incidental sensitivity that has been found here (Experiments 1–4) and elsewhere (Greene, 1984; Kausler et al., 1984; Marmurek, 1983) is *not* dependent upon subjects' deliberate attempts to remember either items or their frequencies. Intentional strategies may slow down the processing of items (as they did for one group of subjects in this experiment), but they do not necessarily result in enhanced sensitivity to frequency information. Thus the data are in agreement with those reported by Kausler et al. rather than those reported by Greene.

GENERAL DISCUSSION

In the course of performing an orienting task in which subjects have no conscious intention to remember a series of items, they nonetheless store information that subsequently permits them to discriminate differences in the frequency with which the items occurred. In addition, we found that a specific intention to remember frequency of occurrence is of no benefit (Experiment 4). These laboratory data are in clear agreement with the many demonstrations that people can reliably discriminate the frequencies of such naturally occurring events as single letters, letter pairs, surnames, occupations, and diseases (see Zechmeister & Nyberg, 1982, pp. 149–150, for a review). This knowledge, too, was presumably acquired under truly incidental conditions; that is, under conditions in which people had no particular reason to attend to frequency or to prepare for a memory test.

The acquisition of knowledge in the absence of intentional strategies is what is expected of an encoding process that operates automatically (Hasher & Zacks, 1979, 1984; Posner & Snyder, 1975). Further, according to Hasher and Zacks (1979, 1984), an automatic process should be insensitive to the presence of other demands on capacity, to the age of the subject, to individual differences in ability or mood, and to the effects of practice (see Hasher & Zacks, 1979, 1984). When these other dimensions or criteria of automaticity have been investigated in relation to encoding of event frequency, the typical finding has been one supporting the automaticity view: The process of encoding frequency information appears to be largely impervious to the vicissitudes of competing demands, old age, and depressed mood; it is also largely impervious to benefits typically associated with practice, with explicit preknowledge of what will be tested, with superior intellectual ability, and with the greater sophistication about memory of older compared with younger children (see Hasher & Zacks, 1984, for a recent review).

Still, there are some findings on memory for frequency that are seemingly in contradiction with one or another of the automaticity criteria (e.g., Fisk & Schneider, 1984; Greene, 1984; Hockley, 1984).[6] Greene, in particular, has presented a broad challenge to the automaticity view. He argues that the evidence "is not very strong" (p. 94) that frequency encoding meets any of the Hasher-Zacks (1979) criteria for automaticity. Whether this challenge must be accepted as valid cannot be completely decided on objective grounds—there are no hard and fast rules for deciding what is strong versus weak evidence. Another review of the literature on memory for frequency (Hasher & Zacks, 1984) suggests that Greene may have ignored some supportive findings while overemphasizing the small number of contradictory findings. Here we address a few of the specific claims made by Greene.

Intentionality effects

First, consider Greene's claim that encoding of frequency is affected by intentionality. Greene bases this claim on his findings of differences in performance between prewarned and truly incidental subjects, findings that contradict those reported by Kausler et al. (1984). We first address the discrepancy between the two sets of findings and then return to the general issue of intentionality effects.

As mentioned in the introduction, both sets of studies used the Glenberg et al. (1977) distractor procedure. In the Kausler et al. (1984) experiments, the 6-s retention interval on each trial was filled with the successive exposure of three words at a 2-s rate. Subjects said each word aloud once during an exposure. In the Greene experiments, a single word was presented on each distractor trial, and subjects repeated it aloud twice per second for the duration of the 4- or 12-s retention interval. Thus, each occurrence of a distractor word was associated with 8 or 24 massed rehearsals of the word. To show how these differences in distractor task procedures might account for the discrepancy in instructional effects, we rely first on a recent analysis of maintenance rehearsal (Naveh-Benjamin & Jonides, 1984a, 1984b) and, second, on an analysis of the likelihood of confusions occurring between memories of rehearsals of items and memories of actual presentations of items (Johnson & Raye, 1981).

The Naveh-Benjamin and Jonides (1984a, 1984b) analysis suggests that maintenance rehearsal, at least as operationalized in the Glenberg et al. paradigm, is not a unitary phenomenon. According to Naveh-Benjamin and Jonides, generating a series of consecutive rehearsals of an item requires effortful processing only for the first few rehearsals. If subjects are not expecting a memory test on the rehearsed items, then later rehearsals require only automatic processing. However, effortful processing on later rehearsals can be reinstated by warning subjects of a future memory test on the distractor items and asking them to elaboratively rehearse the distractor words (Naveh-Benjamin & Jonides, 1984b).

As applied to the two competing sets of distractor procedures, this analysis of rehearsal is informative. The Kausler et al. (1984) procedure did not allow for the development of automatized rehearsal under any of their instructional conditions. The Greene (1984) procedure allowed for the automatization of later rehearsals in the truly incidental condition but not in the two prewarned conditions; prewarned subjects would have continued with effortful rehearsals through an item's exposure. These differential rehearsal patterns between the truly incidental and the two intentional conditions are a likely source of the differences in frequency judgments reported by Greene, as the following argument will show.

The literature on "reality monitoring" demonstrates that our ability to discriminate between memories that have self-generated sources (e.g., rehearsals) and ones that have external sources is far from perfect (see Johnson & Raye, 1981, for a review). In particular, judgments of *presentation* frequency are influenced by self-generated repetitions or rehearsals; that is, increases in the number of self-generations of an item are associated with increases in the judged frequency of environmental occurrences of that item (Johnson, Taylor, & Raye, 1977; Raye, Johnson, & Taylor, 1980). Although no research has directly addressed the issue of how automatized and nonautomatized rehearsals differ in their effects on judgments of presentation frequency, it stands to reason that automatized rehearsals will have a smaller impact. This would especially be true if, as claimed by Naveh-Benjamin and Jonides (1984a, 1984b), it is primarily the nonautomated rehearsals that result in the creation of long-term memory traces. This line of argument suggests that Greene's procedures, which promoted automatized rehearsal in the truly incidental condition but not in the two prewarned ones, complicate interpretation of the obtained instructional-group differences in frequency judgments. By contrast, the Kausler et al. (1984) procedures do not seem to be characterized by a similar source of complication.

However, even if we conclude that Greene's instructional results should be taken at face value, their impact is weakened by a consideration of the total set of results on truly incidental encoding of frequency of occurrence. Consider for example the fact that in all cases (Greene, 1984; Kausler et al., 1984; Marmurek, 1983; the four experiments of this paper) subjects in truly incidental conditions have demonstrated the ability to discriminate frequency differences. In the present research, for example, we found that whether subjects are solving anagrams, deciding if a word completes a sentence frame, or naming pictures with or without superimposed labels, they are able to discriminate differences in frequency on a subsequent and completely unexpected memory test. In addition, the evidence from studies that have directly compared the performance of subjects given truly incidental instructions with those given intentional instructions is decidedly mixed. Of five experiments (two in each of the Greene and Kausler et al. papers, one in the present one), three failed to find differences. When these findings are coupled with others that show in adult subjects reliable, incidentally acquired knowledge of the relative frequency of naturally occurring events, there seems considerable support for the view that frequency of occurrence information is stored without deliberate intention. By deliberate intention, we refer first and foremost to intention to store frequency information. In addition, frequency information is stored whether or not subjects intend to remember target items.

Noninterference effects

Greene (1984) also criticized the fit between existing data and the Hasher and Zacks (1979, 1984) "noninterference" criterion which states that automatic encoding is not hindered by other simultaneous mental activity. Here, Greene apparently disagrees with us, as well as with Kausler et al. (1984), as to the kinds of results that do or do not support this criterion. The nature of the disagreement is demonstrated by our differing conclusions about the results of Experiment 3 by Zacks et al. (1982). In that study, subjects were told, prior to list presentation, to expect one of three kinds of tests on the items: free-recall, frequency, or both. Half the subjects in each of the three instructional conditions subsequently received a frequency discrimination test and half received a free-recall test.

We and others assume that subjects try to prepare for whatever memory tests are specified by an experimenter. This would include frequency tests, for which preparation is largely wasted effort because the information that permits these judgments is clearly stored (as the present research demonstrates) without the intention of subjects and shows minimal benefit from the deliberate strategies subjects adopt. Thus instructions to prepare for both frequency and recall tests will result in competition for processing capacity induced by subjects' attempts to prepare for the two tests. This competition for capacity should affect free-recall performance (which depends on effortful strategies) but not frequency performance (which depends on automatic processing). The findings (Zacks et al., 1982, Experiment 3) are in agreement with these expectations: Instructions had no impact on the frequency test. By contrast, instructions had a substantial impact on the free-recall test: Subjects who expected both memory tests did more poorly than subjects who expected only a free-recall test. An identical analysis of this situation was presented by Kausler et al. (Experiment 1, 1984), with findings directly comparable to those reported by Zacks et al. (1982).

Greene claims, however, that according to the Hasher-Zacks noninterference criterion, we and presumably Kausler et al. should have expected a completely symmetric pattern of results on the two memory tasks: "Not only must frequency encoding be unimpaired by other activities, but also performance on these other tasks must not suffer as a result of frequency encoding" (p. 94). This seems to us (and to Kausler et al.) to be an incorrect assessment of the situation, possibly based on the mistaken assumption that if subjects try to adopt a strategy for coding frequency, this strategy should use no capacity.

Reduced capacity effects

Greene (1984) also contends that only a single weak study (Hasher & Zacks, 1979, Experiment 3) supports the criterion that frequency encoding would be unaffected by reduced capacity. The study referred to by Greene examined encoding of frequency information by subjects whose processing capacity was presumed to have been reduced by depression. Depressed and nondepressed individuals did not differ on a frequency judgment task. We partially agree with Greene's evaluation of this study. The sample size was small and the depressed subjects were only mildly depressed (mean score of 15.0 on the Beck Depression Inventory). However, we disagree that it is the only relevant study. For example, in the last few years Weingartner and his colleagues have published several papers (e.g., Roy-Byrne, Weingartner, Bierer, Thompson, & Post, 1986) presenting frequency and recall data for clinically depressed subjects. These data are fully consistent with the hypothesis that (at least severe) depression is associated with reduced cognitive capacity and that the reduced capacity is detrimental to effortful but not to automatic encoding. A similar approach to the cognitive deficits seen in elderly persons also received some support (Attig & Hasher, 1980; Kausler & Puckett, 1980) as does an approach to the problem of learning disabled children (Goldstein & Dundon, 1987; Goldstein, Hasher & Stein, 1983).

CONCLUSION

Greene (1984) also examined frequency encoding in relation to developmental trends and practice effects, claiming for the former that the data are mixed and for the latter that more extended practice periods need to be studied (cf. Hockley, 1984). As demonstrated above, some of the sting could probably also be taken out of these claims. We want, however, to conclude on a different note: Rarely does a body of literature in cognitive psychology contain entirely uniform findings. When data on an issue are somewhat mixed, one can focus on the deviant findings or one can look for a message in the preponderance of results. Obviously, we prefer to do the latter when considering the manner in which frequency of occurrence is encoded. There can be no dispute about the fact that the majority of the findings show people to be remarkably sensitive to the frequency with which events occur. Furthermore, as the present series of studies confirms, this sensitivity occurs whether or not subjects intend to remember either the target information itself or the frequency of targets. And this sensitivity is relatively unaffected by usually powerful variables (at least in the memory literature) such as individual differences in ability, age, or mood.

There also can be no doubt about the fact that evidence of frequency sensitivity does not just emerge from averaging across subjects; the evidence is apparent in the data of individual subjects as well. Whatever criticisms may be made of the notion of automatic encoding and whatever revisions the larger theoretical framework may ultimately face, the sensitivity of individuals to frequency of occurrence is an indisputable fact, found in all reported experiments.

From our perspective, the importance of such findings is partly related to how they contrast with much of the memory research literature of the last 20 years. That literature has focused on the acquisition and retention of laboratory-learned information, and has in particular emphasized intentional learning and memory tasks (e.g., free recall) that benefit from conscious strategies. By contrast, we suggest that much of what we learn of our environments in the course of our everyday activities involves acquisition mechanisms more similar to automatic encoding than to intentional strategies. Further, with respect to frequency, we have argued that a good deal of what people know and expect about their environments is based on sensitivity to this attribute (Hasher & Zacks, 1984).

Notes

The research that is reported here was supported by National Institute of Mental Health Grant MH33140, Army Research Institute Grant MDA903–82–C–0317, and National Institute on Aging Grant AGO4306, as well as by Biomedical Research Grants from Temple and Michigan State Universities.

Requests for offprints should be sent to Lynn Hasher at her present address, Department of Psychology, Duke University, Durham, NC 27706. Received for publication October 8, 1985; revision received January 15, 1986.

1. To prevent information from being disseminated through the large pool of subjects available to us, we tested all subjects in the briefest possible time span.

2. These experiments were run simultaneously with a series of experiments concerned with effort effects on free recall. In the latter, we were pursuing the suggestion that the greater the effort expended during an orienting or cover task, the greater the incidental retention of the items encountered in the cover task (e.g., Tyler, Hertel, McCallum, & Ellis, 1979). However, because our effort manipulations failed to influence free recall (see Zacks, Hasher, Sanft, & Rose, 1983), the finding that they also did not influence frequency encoding is neither surprising nor particularly interesting. We include a description of each of our effort manipulations to complete our descriptions of the materials and design of the four experiments.

3. We consider the counterbalancing of items across frequency levels a critical aspect in the design of studies assessing knowledge of frequency. This is because Underwood and Freund (1970) showed that there are Item × Frequency × Subject interactions that affect judgments.

4. Across all experiments using zero-presented items, the false positives were as rare as in this experiment. Mean judged frequency for these items never rose above .20 and were more commonly below .10.

5. The goal here was to have three levels of naming difficulty, but as will be seen, the naming times for the unrelated- and related-labels conditions did not differ, and this suggests that related labels did not produce more interference than unrelated ones. The lack of a difference between the related- and unrelated-label conditions was probably due to our failure to control the nature of the relation between picture and word in the related-label condition. Lupker (1979) has shown that only certain kinds of related labels (e.g., ones from the same semantic category) produce more interference than unrelated labels.

6. For a discussion of the Fisk and Schneider data (1984), see Zacks, Hasher, and Hock (1986).

References

Attig, M. S., & Hasher, L. (1980). The processing of frequency of 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 on 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 and Verbal Behavior, 16,* 339–352.

Goldstein, D., & Dundon, W. D. (1987). Affect and cognition in learning disabilities. In S. J. Ceci (Ed.), *Handbook of cognitive, social and neuropsychological aspects of learning disabilities* (Vol. 2, pp. 233–249). Hillsdale, NJ: Erlbaum.

Goldstein, D., Hasher, L., & Stein, D. K. (1983). The processing of occurrence rate and item information by children of different ages and abilities. *American Journal of Psychology, 96,* 229–241.

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

Hockley, W. E. (1984). Retrieval of item frequency information in a continuous memory task. *Memory & Cognition, 12,* 229–242.

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

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88,* 67–85.

Johnson, M. K., Taylor, T. H., & Raye, C. L. (1977). Fact and fantasy: The effects of internally generated events on the apparent frequency of externally generated events. *Memory & Cognition, 5,* 116–122.

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.

Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory & Cognition, 7,* 485–495.

Marmurek, H. H. C. (1983). Multiple processes in the automatic encoding of frequency information. *Canadian Journal of Psychology, 37,* 512–520.

Mayzner, M. S., & Tresselt, M. E. (1958). Anagram solution times: A function of letter order and word frequency. *Journal of Experimental Psychology, 56,* 376–379.

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

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

Olson, R., & Schwartz, R. (1967). Single and multiple solution five-letter words. *Psychonomic Monograph Supplements, 2,* 105–152.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Postman, L. (1982). An examination of practice effects in recognition. *Memory & Cognition, 10,* 333–340.

Raye, C. L., Johnson, M. K., & Taylor, T. T. (1980). Is there something special about generated information? *Memory & Cognition, 8,* 141–148.

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.

Roy-Byrne, P. P., Weingartner, H., Bierer, L. M., Thompson, K., & Post, R. M. (1986). Effortful and automatic cognitive processes in depression. *Archives of General Psychiatry, 43,* 265–267.

Snodgrass, J. G., & Vanderwart, M. A. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174–215.

Thorndike, E. L., & Lorge, I. (1944). *The teacher's wordbook of 30,000 words.* New York: Columbia University Teacher's College, Bureau of Publications.

Tyler, S. W., Hertel, P. T., McCallum, M. C., & Ellis, H. C. (1979). Cognitive effort and memory. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 607–617.

Underwood, B. J., & Freund, J. S. (1970). Relative frequency judgments and verbal discrimination learning. *Journal of Experimental Psychology, 83,* 279–285.

Zacks, R. T., Hasher, L., & Hock, H. S. (1986). Inevitability and automaticity: A reply to Fisk. *American Psychologist, 41,* 216–218.

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.

Zacks, R. T., Hasher, L., Sanft, H., & Rose, K. C. (1983). Encoding effort and recall: A cautionary note. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 747–756.

Zechmeister, E. B., & Nyberg, S. E. (1982). *Human memory: An introduction to research and theory.* Monterey, CA: Brooks/Cole.