Primacy, recency, and the availability heuristic

CAROL L. CURT and EUGENE B. ZECHMEISTER
Loyola University, Chicago, Illinois

Primacy and recency effects in recall were used to assess an availability heuristic for judging category-frequency information. By "over-representing" the less frequently occurring category in the early and late positions of a list, relative-frequency information for these positions was biased. This manipulation significantly affected frequency estimates, but only when recall was not requested prior to the frequency-judgment task. Factors other than differential recall of category instances are likely involved in judgments of category frequency and the possible role of "automatic" frequency encoding is discussed.

People use an availability heuristic for judging category frequency when they assess the ease with which instances of a category can be brought to mind (Tversky & Kahneman, 1973). For example, a decision as to whether there are more foreign-made cars or more American-made cars on the nation's highways might be made by recalling instances of each car type that have been seen. If instances of one category are more easily recalled, then this is evidence of greater frequency of that category. Although such a strategy is perfectly reasonable when availability corresponds directly to true frequency of occurrence, errors in frequency estimation can arise when it does not (see Tversky & Kahneman, 1973).

The well-known serial position effect in free recall suggests an interesting test of the availability heuristic. Probability of recall for the first and last items in a list can be twice that of items in the middle of the list (Murdock, 1962). In the present experiment, relative frequency information was "biased" by "over-representing" the less frequently occurring category in the early and late portions of a list. Specifically, the frequency of category exemplars in these sections of a list was disproportionate (i.e., exaggerated) relative to the overall frequency of these items in the list. By biasing input in this way, it was expected that subjects, if relying on an availability heuristic, would be misled when judging category frequency.

It is possible, however, that people are sensitive to changes in relative frequency within a list and would not be deceived by differential availability due to primacy and recency effects. Recent analyses of the frequency judgment process have implicated "automatic" mechanisms for encoding frequency (Hasher & Chromiak, 1977; Hasher & Zacks, 1979). Automatic encoding of frequency information presumably occurs for both

individual event frequency and category frequency (Alba, Chromiak, Hasher, & Attig, 1980). It is important, therefore, to understand better when so-called automatic processes are involved in frequency estimation and when apparently nonautomatic, and potentially error producing, strategies, such as availability, are employed. The present experiments were designed with this goal.

METHOD

Lists

Eight lists were constructed. Each list contained 29 names of well-known movie stars (e.g., John Wayne, Katherine Hepburn). Four of these lists contained 15 male names and 14 female names; the other four lists consisted of 15 female names and 14 male names. Each list was formed from the same pool of 30 names, and, therefore, each set of four lists was identical with respect to particular names. Experimental and control lists were defined by the relative frequency with which male and female names appeared in Positions 1-6 and Positions 26-29 of the lists. Specifically, in the experimental lists, the less frequently occurring sex was "over-represented" in the first and last portions of the lists. In these lists, there were four instances of the less frequent sex in Positions 1-6 and three instances of the less frequent sex in Positions 26-29. In addition, instances of the less frequent sex were always the first and last items in the experimental list. In the control lists, male and female names were equally represented in Positions 1-6 (3 each) and Positions 26-29 (2 each). For half the control lists, an instance of the less frequent sex appeared in the first position of the list and an instance of the more frequent sex was the last item. The reverse was true for the other half of the control lists. Names were assigned randomly to positions in the list under the restrictions already outlined and with the additional restriction that no more than 3 names of one sex appeared consecutively. Two lists of each type (experimental and control) had more male than female names, and two of each type had more female than male names. These pairs of lists differed only in the randomization of specific names within the list. Lists were tape recorded for auditory presentation. Rate of presentation was approximately 3 sec/name.

Procedure

Two experiments were conducted. In both experiments, the subjects were informed initially that they were participating in a memory experiment that involved names of famous personalities, but that the exact nature of the memory test could not be revealed. In Experiment 1, immediately after presentation of the list of names, a brief questionnaire, containing three filler questions (e.g., "Was the name of any person presented twice?")

The authors would like to thank Lynn Hasher and John J. Shaughnessy for their comments on an earlier draft of this article. Requests for reprints should be sent to Eugene B. Zechmeister, Department of Psychology, Loyola University, 6525 North Sheridan Road, Chicago, IL 60626.

and two target questions, was presented. One of the target questions (Question 3, or Q3) asked subjects to indicate which of the following statements was correct: (a) "There were more female than male names," or (b) "There were more male than female names." The order of statements (a) and (b) was reversed in approximately half the questionnaires. The other target question (Q4) requested subjects to estimate the exact number of male and female names.

Experiment 2 was identical to Experiment 1, except that recall of the names was tested immediately after the list had been read and before the questionnaire had been presented to the subjects. Five minutes were allowed for free recall.

Subjects

In Experiment 1, 169 subjects were tested; in Experiment 2, a total of 84 subjects participated. All subjects were undergraduate students from a large introductory psychology class who earned class credit for participation. Testing was done in small groups.

RESULTS

Measures of frequency estimation were based on the number of subjects in each group who correctly identified whether there were more male (M) or female (F) names (Q3) and on the actual estimates of category frequency (Q4). Unexpectedly, 24 subjects in Experiment 1 and 12 subjects in Experiment 2 gave contradictory responses to the two target questions. For example, a subject who indicated that there were more M than F names (Q3) then estimated the actual frequency of F names to be greater than that of M names. Approximately an equal number of control and experimental subjects in each experiment provided inconsistent information. Analyses were carried out both with and without these inconsistent subjects. Since the pattern of results was similar and the conclusions reached on the basis of statistical tests were the same for both sets of analyses, only the results for consistent subjects are reported.

Table 1 summarizes the frequency-judgment data obtained from the experimental and control subjects in the two experiments. In Experiment 1, the subjects hearing the biased lists were more likely than the control subjects to say, mistakenly, that the less frequent category was more frequent [$\chi^2(1) = 5.79, p < .05$]. Differ-

ences between estimates given to M and F names also differed reliably between the two groups [$t(143) = 2.53, p < .01$]. Experimental subjects gave higher frequency estimates to the less frequent category of names, whereas control subjects correctly gave higher frequency estimates to the more frequent category.

In Experiment 2, when name recall was requested before asking for frequency judgments, the proportion of subjects judging category frequency correctly did not differ between the experimental and the control subjects ($\chi^2 < 1$). Differences in the estimates of actual frequency for the two categories also were similar for the two groups ($t < 1$). The pattern of results seen in Experiment 1, therefore, apparently was eliminated by requiring recall of the names before asking for frequency judgments. Although experimental subjects gave slightly lower overall frequency estimates than did control subjects, this difference was not reliable [$t(70) = 1.07, p > .05$].

Analysis of the recall data from Experiment 2 showed that distinct primacy and recency effects were obtained. Mean proportion correct for early (1-6), middle (7-25), and late (26-29) serial positions was .54, .39, and .48, respectively, for control subjects, and .61, .48, and .58, for experimental subjects.

For each of the list types in both experimental and control conditions of Experiment 2, individual subjects were classified according to whether recall of M or F names was greater and whether M or F names were judged to be more frequent. Twenty subjects were not considered in this comparison, because they either recalled the same number of M and F names (17) or estimated the actual frequency of M and F names to be the same (3). For the remaining 52 subjects, in both conditions there was a clear pattern of agreement between relative difference in recall of M and F names and category-frequency judgments. Considering only the direction of the differences in recall and frequency estimates for M and F names, 79% of the control subjects (19/24) and 71% of the experimental subjects (20/28) showed agreement. These results indicate a positive relationship between relative differences in recall of category instances and relative-frequency judgments; however, this relationship was somewhat poorer for experimental than for control subjects.

DISCUSSION

In Experiment 1, control and experimental subjects differed significantly in their judgments of category frequency. Experimental subjects were more likely to say that the less frequent category of names was more frequent. These results can be said to be expected, given the assumption that the subjects used an availability heuristic in judging category frequency and on the basis of the fact that instances of the less frequent category had a high probability of being recalled due to their "over-representation" in the early and late positions of the experimental lists. Differences between control and experimental subjects were not found in Experiment 2 when recall was tested before asking for frequency judgments. However, an examination of individual

Table 1
Frequency Judgments of Control and Experimental Subjects

Subjects	n	Proportion Correct (Q3)	Mean Frequency Estimates	
			>	<
Experiment 1				
Control	72	.58	13.51	12.49
Experimental	73	.38	11.79	13.14
Experiment 2				
Control	35	.45	13.61	13.77
Experimental	37	.44	12.54	12.77

Note—> = mean frequency estimate for names of greater frequency; < = mean frequency estimate for names of lesser frequency. Q3 = were there more male or more female names?

recall and frequency scores in Experiment 2 showed a relationship between relative differences in category availability and relative differences in category-frequency judgments.

The finding that prior recall of category instances affects frequency judgments was also reported by Beyth-Marom and Fischhoff (1977). They asked Hebrew-speaking subjects to estimate the frequency of names of countries and kibbutzim (Israeli settlements) that began with a certain (Hebrew) letter. Estimates were made both before and after attempted recall of category instances. Measures of availability included the latency to produce the first instance of a category and the number of instances produced in the first 5 sec. These availability measures successfully predicted some, but not all, frequency estimates. Furthermore, the subjects produced more accurate estimates of category frequency after attempting recall of instances from the category. Beyth-Marom and Fischhoff suggested that "at least some of the bias induced by relying on the availability heuristic can be eliminated by using the heuristic properly, that is, by working hard to see how many instances are in memory" (p. 237).

It can be suggested that, without the requirement to recall all names within the list, the subjects in Experiment 1 made frequency estimates on the basis of information most available, namely, that which was presented in the early or late portions of the list. Thus, because they did not make an effort to determine exactly how many instances of each type were in memory, the experimental subjects made more errors in frequency estimation than did the control subjects. An alternative possibility is that experimental subjects estimated frequency on the basis of an "impression" that they gained when first presented the list of names. That is, during list presentation, the experimental subjects may have developed an erroneous impression regarding differential frequency. Such an impression may have led them to judge frequency without relying on attempted retrieval of category instances.

In the present study, a recall test given prior to a frequency-judgment test apparently eliminated the bias produced by the primacy and recency manipulations, but, unlike the results of Beyth-Marom and Fischhoff (1977), it did not improve frequency estimation. A comparison of the results of Experiment 1 with those of Experiment 2 indicates that the control subjects in Experiment 2 were somewhat less accurate overall in judging category frequency. Both the control and the experimental subjects in Experiment 2 were correct about 45% of the time, whereas the control subjects in Experiment 1 were correct on 58% of their judgments. It must be emphasized, however, that the frequency difference in the present task was minimal (14 vs. 15), and, in fact, statistical analysis showed that frequency-judgment performance was not significantly different from chance (.50) in the control conditions of either Experiment 1 or Experiment 2.

Alba et al. (1980) suggested that category-frequency information, like that of individual event frequency, is encoded automatically. They based this conclusion on the results of experiments requiring college students to judge the frequency of categories that were represented in a long list of items by various numbers of category exemplars (0, 3, 6, and 9). The category names were shown to subjects for the first time at test when subjects made frequency judgments for each of the categories. Judgments of category frequency increased significantly as the number of exemplars increased. Moreover, since the rate of presentation of test names was varied (Experiment 1), and since this variable did not affect frequency judgments, Alba et al. concluded that information about category frequency was encoded at the time of study. These results were obtained despite the fact that subjects were not forewarned of a frequency test. In the second experiment, when subjects were either told of an up-

coming frequency-judgment test or not told, the researchers also obtained results that they interpreted as being consistent with an automatic encoding process.

In the present study, the subjects, especially those in the experimental condition of Experiment 1, appear not to have made frequency judgments on the basis of an "automatic" process. However, given the nature of the frequency-discrimination task employed, it is perhaps not a fair test of an automatic frequency-encoding process as described by Hasher and Zacks (1979). For example, it might be expected that automatic category-frequency encoding is most likely to be observed when a category name is reliably elicited by category exemplars. In other words, if each time an exemplar is presented the category name is elicited, then one need talk only about a frequency "counter" for implicit events. There is evidence that subjects can accurately estimate the frequency of implicit events, such as the gist of a sentence (Gude & Zechmeister, 1975) or the internally generated image of an item (Raye, Johnson, & Taylor, 1980). Judgments of category frequency may be considerably less accurate (and not automatic) when the category name is not reliably elicited by each exemplar. Because Alba et al. (1980) used only exemplars that had a strong normative relation with the category name, the category name likely had a high probability of occurring implicitly when an exemplar was presented. When category names are less reliably elicited, as perhaps when names of male and female movie stars are presented, category-frequency judgments will possibly depend on less accurate and less automatic processes. Moreover, this may be especially true when the frequency-discrimination task is very difficult, as it was in the present study. Therefore, it is perhaps not unreasonable in this situation to expect subjects to use an availability strategy. On the other hand, it can be pointed out that the difference in mean frequency judgments for M and F names in the control group of Experiment 1 was almost exactly the same as the actual frequency difference in the list, namely, 1.0. Future experiments will help decide to what extent this difference was arrived at "automatically."

REFERENCES

- ALBA, J. W., CHROMIAK, W., HASHER, L., & ATTIG, M. S. (1980). Automatic encoding of category size information. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 370-378.
- BEYTH-MAROM, R., & FISCHHOFF, B. (1977). Direct measures of availability and judgments of category frequency. *Bulletin of the Psychonomic Society*, 9, 236-238.
- GUDE, C., & ZECHMEISTER, E. B. (1975). Frequency judgments for the "gist" of sentences. *American Journal of Psychology*, 88, 385-396.
- 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.
- MURDOCK, B. B., JR. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, 64, 482-488.
- RAYE, C. L., JOHNSON, M. K., & TAYLOR, T. H. (1980). Is there something special about memory for internally generated information? *Memory & Cognition*, 8, 141-148.
- TVERSKY, A., & KAHNEMAN, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.

(Manuscript received for publication December 27, 1983.)