

Enhancement of frequency judgments by response choice

MELVIN H. MARX

University of Missouri, Columbia, Missouri

Four classes of mixed fifth- and sixth-grade students selected the more positive word of 36 word pairs, in each of which a negative-word foil was present with the positive word, and simply copied 36 equally positive words when the words were presented singly. All subjects were then given a forced-choice frequency-judgment test in which they selected the more frequent of the two matched positive words, one in each pair having been shown with a negative-word foil and one having been shown singly and copied. All positive words tested together had actually been shown equally often (one, three, or five times). In both experimental groups, reliably more students selected more of the words that had been paired with the foils, and that presumably had received some degree of "reinforcement," but there were no differences between the two experimental groups. These results are seen as somewhat unusual in two respects. First, they indicate the sensitivity of the noneffortful frequency-judgment function to certain orthodox experimental manipulations. Second, they suggest some role for a strength factor in the interpretation of frequency processing.

This experiment was planned as an initial test of the proposition that frequency judgment can be influenced by more variables than have hitherto been shown to be effective. As the comprehensive literature review by Hasher and Zacks (1979) made clear, none of the usual determinants of recall, such as age, practice, intention, emotional state, and the like, seem to exercise any influence on the presumably automatic and noneffortful function of frequency judgment.

The particular variable with which this experiment was concerned was the presumed automatic effect of reward (Marx, 1981). It was hypothesized that frequency judgments might be subject to the same kind of enhancement in retention as previously had been shown to occur for instrumental behaviors in successfully completed tasks (Bottenberg, Marx, & Pavur, 1976; Marx, Pavur, & Seymour, 1977).

A "cover" task (choosing the more "positive" of two words) was used instead of the customary memory task. There are unavoidable ambiguities involved in the use of a memory task, including uncontrolled amounts and types of rehearsal or other processing, especially if the instructions are kept deliberately vague. These potential contaminants are best avoided in frequency-judgment research by using a different kind of task.

The specific purpose of the present experiment was

to determine if there is enhancement of frequency judgments of events in which the subject makes a decision and is almost invariably "correct" because of the simplicity of the problem. A secondary purpose was to determine if a presumably stronger form of such "reinforcement," manipulated by a slight change in the task instructions, would produce a greater degree of enhancement than would a weaker form.

METHOD

Subjects

Students in four mixed fifth- and sixth-grade classes in a public school served as subjects. The students were accustomed to testing because of the nature of the school (parents had been required to consent to experimental testing of their children as a condition of the children's admission to the school) but had engaged in no previous frequency-judgment experimentation. All students in attendance on the days of the preliminary and test phases participated. There were a total of 97 participants in the test phase.

Experimental Design

In the preliminary phase, the students were asked to rate a large number of common English words in terms of the words' positive or negative character. During the experiment proper, the same students were shown selected pairs of these words and were asked to choose the more positive word of each pair. They were also shown single "positive" words, which they were instructed simply to copy. Without advance notice, a forced-choice frequency-judgment test, in which the positive member of each of the pairs of words was pitted against equally positive and equally frequent words that had been presented singly and copied, was then administered. Two lists of equivalently rated positive words were counterbalanced in the single/double format over classes to control for adventitious variations among the words.

The students in the first two classes tested were instructed simply to select the more positive word when two words were

This research was supported in part by a Research Career Award from the National Institute of Mental Health. I thank Dorothy Lyon, manager of the psychology laboratory at Florida International University, for the preparation of the slides, and Patricia Frost, principal of West Laboratory School, Coral Gables, Florida, for making the subjects available.

shown (the weak experimental group). The students in the last two classes were instructed to select the word that they thought the class had earlier rated as more positive. This instruction was designed to produce a more explicit type of "reinforcement" function (the strong experimental group). Because of the clear contrast between positive target words and negative foils, the task of selecting the more positive word was a very easy one, and predominantly "correct" responses were anticipated in both groups.

Procedure

In the preliminary phase, the teacher in each of the four classes passed out copies of a questionnaire in which a total of 78 English words appeared. Half of the words had been judged by the experimenter to be clearly positive (e.g., "money, beauty, rainbow") and half to be clearly negative (e.g., "death, crime, war"). The students rated each of the words on a 6-point scale: highly positive, positive, slightly positive, slightly negative, negative, and highly negative. This work was self-paced.

The 78 words were then scaled on the basis of the 52 fifth-grade students' ratings. The 2 highest (most positively) ranked words ("friend" and "mother") were paired, and 1 of them was placed randomly in List A and the other in List B. Similar random divisions were made for the next 11 pairs of positively ranked words, so that Lists A and B each contained 12 highly positive words and were essentially equivalent in this respect. A set of 12 negative foils was then selected from the pool of clearly negative words (e.g., "itch, shame, fear") for use in the double-word test items. Right-left order of these items was determined randomly.

In the experimental phase, conducted 10 weeks after the preliminary phase, the first class tested was given List A words paired with the negative foils and List B words in the single-word format. The opposite arrangement of Lists A and B was used for the second class tested. Both of these classes were instructed that, when two words were shown, by slide projector, their task was "to decide which of the two is the more positive—that is, the one that you like better—and to write that word on the answer sheet."

The third and fourth classes tested were treated in the same way with respect to the counterbalancing of Lists A and B in single- and double-word format, but were instructed that, when two words were shown, their task was "to decide which of the two was judged most positive—that is, most preferred—in the earlier work we did" and that "we want to see how many of the class decisions each of you can guess correctly."

The frequency variable was manipulated over one, three, and five occurrences. The 12 positive target words were assigned randomly to the three frequencies, 4 words at each level. Negative-word foils were similarly placed in the same one, three, and five frequencies and were then assigned randomly to the positive words in the double-word format, with the restriction that no 2 words appeared together more than once.

The 72 items were presented in a random order, with the restriction that no more than two successive appearances were permitted of either single- or double-word items. The presentation was approximately 8 sec per item, with a somewhat longer period used for the first few items. A 30-sec rest period was interpolated after the first 36 items. The students were told that, if necessary, they could abbreviate the words to keep pace.

The frequency-judgment test immediately followed the word-selection exercise. Each of the original 12 pairs of positive words was shown on a test slide. One such word in each case had been presented singly and simply had been copied, and the other had been a member of a pair of words and had been written on the answer sheet. The students were instructed to select, without regard to degree of positiveness, the word that had appeared more often in the word exercise just completed, and to guess when necessary. Thus, choices between items of

equal frequency were forced, in order to determine the degree of selection of previously chosen words compared with that of words that merely had been copied.

The 12 negative foils were presented in pairs, providing six more test items. These words were arranged in terms of contrasting frequency, such that there were two 1:3, two 1:5, and two 3:5 items. Two of these appeared as the first two test slides and two as the final two slides, with the remaining two randomly interpolated among the positive-word slides. These items were used primarily to provide relatively easy frequency choices and thereby to make the test more acceptable—that is, to make it seem more legitimate—as well as to relieve the tedium of so many difficult forced choices between equal frequencies. Secondly, they were intended to be scored for correctness as a check on the accuracy of ordinary frequency judgments by the participants.

RESULTS

Word Selections

Spot checks were made of the word selections. Answer sheets filled in by two males and two females were scored from each class. A very high degree of compliance with the instructions was indicated. Fourteen of the 16 papers scored were letter perfect. One paper showed a single omission, and the other paper had three omissions and one duplication. On the whole, these errors were held to be relatively inconsequential, in light of the total task. There seemed to be little point in scoring all of the papers, not only because of the very low error rate expected on the basis of the spot checks, but also because there did not seem to be any very productive way to include these small errors in the analysis of the data. Moreover, any failures to select the correct (positive) word would reduce the actual frequency of that word (in re response, of course, and not stimulation) and thus would tend to bias against the hypothesis. Substantial compliance with instructions was therefore assumed, and all scoring was done in terms of objective (stimulus) frequencies.

Positive Word Choices

Because of the large discrepancy in *n* (58 subjects in the first two classes, 39 in the second two classes) and, more importantly, because of questions concerning homogeneity and restrictions of variance in the test scores, analysis of variance was not applied, and statistical analysis was done by the less sensitive, but also less demanding, chi-square test.

The major dependent variable was the proportion of subjects in each group that selected the positive word from the double-word format in the frequency-judgment test. Because of the need to control possible differences in attractiveness among positive words, both of the classes in each group were scored together.

In the weak experimental group, 32 of the 58 subjects selected more words from the double-word format, 17 showed the opposite result, and there were 9 ties. If a 50-50 chance distribution is assumed and the ties are

divided evenly between the two conditions, this difference yields $\chi^2(1) = 3.88$, reliable, but barely so, at the .05 level of confidence.

In the strong experimental group, 21 of the 39 subjects selected more of the words from the double-word format and 7 selected more single words, with 11 ties. Applying the same statistical technique to these data, we obtain $\chi^2(1) = 5.02$, which is comfortably reliable at the .05 level.

If all of the subjects are combined, yielding 53 who selected more double-format words and 24 who showed the opposite tendency (with 20 ties, equally distributed), $\chi^2(1) = 9.90$ is obtained, reliable at the .01 level.

A direct comparison of the strong and weak experimental groups yields $\chi^2 < 1.00$, indicating no difference in proportions of double-word selections. No gender difference appeared, with approximately the same proportions in both sexes (28 females favoring double-format words, compared with 13 of the opposite tendency, and 5 ties; comparable male results were 25, 11, and 15, respectively).

Negative Word Foils

Frequency judgments of the differentially occurring negative foils revealed the usual high accuracy. The two experimental groups were essentially equal in this measure, the weak group averaging 4.9 correct judgments (of a possible 6) and the strong group averaging 5.0. The two sexes were similarly equivalent, females averaging 5.0 correct judgments and males 4.9. This mean percentage correct of slightly more than 80% is quite consistent with typical published results. Choices on the two 5:1 comparisons were 88% correct, with corresponding mean scores of 87% for the 3:1 comparisons and 72% for the 5:3 comparisons.

DISCUSSION

The major significance of the present results is probably the clear suggestion of a break in what has seemed to be a kind of monolithic insensitivity of the frequency-judgment process to a large number of common variables that strongly influence other cognitive functions, such as recall. The implication is that differential frequency judgments may be obtained as a result of some such manipulations, at least, perhaps, if a forced-choice test technique is used. An obvious next step is to see if the more customary technique of letting subjects make their own estimates yields similar differences. In any event, this potential chink in the massive wall that has seemed to close frequency processing off from so many experimental manipulations merits exploration in further research.

Apart from this general implication, how are the present results to be explained? Processing theorists may immediately point to the presumably greater elaboration involved in the decision making induced by the double-word items. But if this difference is used to account for the differential results obtained, it is difficult to see why such added processing, deliberately induced in other situations, such as those with instructions to count and opportunities to practice, has not been effective in those situations.

The rationale underlying the present design was that re-

warded responses are differentially encoded, and that such enhancement should produce behavioral differences in retention (Marx, 1981). Obviously, this is also at present a tenuous interpretation of the differences obtained, especially in light of the failure of the strong/weak reward operation to make a difference. Nevertheless, it should be noted that an indeterminate amount of self-reinforcement was probably present in both the weak and the strong groups. A more effective manipulation of the reward variable is needed to test this hypothesis.

The results have some implications for the major theoretical interpretations of frequency processing (Hintzman, 1976). Certainly, it is difficult to see how the multiple-trace hypothesis can be used in any way to explain the obtained difference, because of the equivalence in this respect between the experimental and control conditions. If a difference is assumed to have occurred in encoding between the two conditions, in such a way as to favor the experimental condition, then some form of the propositional encoding hypothesis may be entertained. However, it is not clear why the propositions encoded in the experimental condition, relating to choice and, typically, a kind of success, should produce generally higher frequency estimates. Finally, the results do obviously offer at least a modicum of support for some form of the strength construct, at least as a kind of terminal factor. Of course, a strength concept is ideally suited to the forced-comparison nature of the task. Nevertheless, the differences obtained in the experiment suggest some role for a strength factor, for which little evidence could be found by Hintzman (1976) in the interpretation of frequency processing.

More directly, the results suggest the important role that choice can play in behavior.

Methodologically, the statistically reliable differences produced by the combination of equal objective frequencies and forced-choice responses suggest that this method may be a useful complement to the more orthodox procedures. Unfortunately, time restrictions precluded the use of open-ended or other questions at the end of the experiment; as a result, what strategies, if any, were used, and what other cognitive reactions there were to the forced-choice test, cannot be determined. (It should be mentioned that, in one instance, both of the positive words were included in an answer, suggesting that the use of some kind of confidence rating might yield useful supplementary data.) In any case, the utility of this technique, especially in combination with an appropriate cover task, is suggested by the results obtained.

REFERENCES

- BOTTENBERG, R. A., MARX, M. H., & PAVUR, E. J., JR. (1976). Differential recall of problem names and clues as a function of problem solution or nonsolution. *Bulletin of the Psychonomic Society*, 7, 445-448.
- HASHER, L., & ZACKS, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 356-388.
- HINTZMAN, D. L. (1976). Repetition and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 10, pp. 47-91). New York: Academic Press.
- MARX, M. H. (1981). Habit activation in human learning. In G. d'Ydewalle (Ed.), *Cognition and memory: Essays in honor of J. Nuttin* (pp. 87-122). Hillsdale, N.J.: Erlbaum.
- MARX, M. H., PAVUR, E. J., JR., SEYMOUR, G. E. (1977). Differential recall of problems, clues, and solutions from completed and uncompleted tasks. *Bulletin of the Psychonomic Society*, 9, 322-324.