

Interactions among gender and task variables in retention of verbal materials

YUNG CHE KIM

Keimyung University, Daegu, South Korea

and

MELVIN H. MARX

University of Missouri, Columbia, Missouri

Data were collected from 684 South Korean 11th-graders in 36 groups of 19 each. They first completed three- or four-letter Korean words, each with one letter provided, using associated-meaning cues. The words occurred one, two, or three times in the list. Half of the students performed (made the actual response), and the other half observed the performance. Then 0, 1, or 7 days later, they were given an unexpected free-recall test. The final cued-retrieval test was administered 0, 1, or 7 days thereafter. The students were instructed to form two Korean words, again using associated-meaning cues, from each of 24 six- to eight-letter anagrams. In each letter string, one word had been used in the initial word-completion task and one had not. The major results were that females showed not only the usual superiority in retaining verbal materials, but also greater resistance to the inhibiting effects of a single prior recall on subsequent cued retrieval (Marx, 1982; Marx & Kim, 1984—a prior report of part of this experiment). Female superiority in free recall was due largely to less forgetting of the more frequently completed words over the longer test intervals. Male performers retrieved more words than male observers. Overall, performers recalled more words than observers, mainly the words that had been completed two or three times. The various interactions obtained confirm and extend those previously reported for gender and task variables.

We report here some interrelated effects of gender and frequency of occurrence of target words, test interval, and performance or observation in acquisition on free recall and subsequent cued retrieval. In the experiment, the free-recall test was manipulated in order to influence subsequent cued retrieval. Because the primary questions addressed by the experiment were entirely different from the present ones, the present data were not included in the previous report of the experiment (Marx & Kim, 1984).

Manipulation of the performance/observation variable was intended to replicate certain earlier results. Briefly, these have shown superior acquisition under performance (subject makes response) compared with observation (subject watches performer and receives equivalent feedback, if provided) conditions (Marx, Homer, & Marx, 1980; Marx, Marx, & Homer, 1980, 1981) and an interaction with gender, such that performing males, but not females, have shown both a suggestion of more repetition of correct responses (Marx & Witter, 1972)

This research was supported in part by a Research Career Award from the National Institute of Mental Health to the second author. The first author was responsible for the collection and analysis of data in Korea. M. H. Marx's mailing address is: Department of Psychology, University of Missouri, Columbia, Missouri 65211. Yung Che Kim's mailing address is: Department of Psychology, Keimyung University, Daegu, South Korea.

and a clearer indication of increased repetition of errors (Marx et al., 1981; Marx, Witter, & Farby, 1973).

METHOD

Methodological details of the experiment have been reported (Marx & Kim, 1984), and only the highlights are given here.

Subjects and Design

A total of 684 Korean 11th-graders in two large senior high schools served as the subjects. They were tested in regular male or female classes, which were assigned randomly to treatment groups. The basic design was a mixed-factorial analysis of variance (ANOVA) with number of occurrences of target words (3) as the within-subjects variable and gender (2) x performance or observation in word completion (2) x free-recall test interval (3) x cued-retrieval test interval (3) as between-subjects variables.

Materials and Procedure

Word completion. The students were given booklets with 48 incomplete three- and four-letter Korean words, for each of which only one letter was provided. The words were to be completed with the help of cues (associated meanings). Five minutes were allowed for this self-paced task. A total of 24 words were used, 8 of them appearing one time, 8 two times, and 8 three times. Different cues were used for the repetitions.

The performance/observation variable was manipulated essentially as had been done in earlier research (e.g., Marx et al., 1981). The students worked in same-sex pairs. One student actually made the responses (filled in the words) as a performer, and the other student, the observer, watched.

Free recall. An unexpected free-recall test was administered

immediately (0) or after delay intervals of 1 or 7 days. All students, now working individually, were requested to write down as many of the words that had been presented previously as they could remember. They were allowed 2 min in which to do so.

Cued retrieval. This test was also administered immediately (0) or after 1 or 7 days. All students, again working individually and pacing themselves, were allowed 5 min in which to solve as many two-word anagrams as they could. Twenty-four six- to eight-letter strings were presented in booklets. Each letter string could be broken down into two three- or four-letter Korean words. Cues (associated meanings) were provided with the letter strings for each of the target words. One word in each pair was the same as one of the earlier words presented for completion; the cue for it was one that had previously been used.

RESULTS

Results are presented separately for performers alone and for performers and observers.

Free Recall

Performers. All of the main effects showed reliable differences. Free-recall scores were a positive function of the number of target words actually completed (means of 1.00, 3.41, and 4.96 for one, two, and three occurrences, respectively, hereafter labeled F1, F2, and F3) [$F(2,672) = 1,044.39, MSe = 1.31, p < .01$] and a negative function of test interval (means of 3.56, 3.24, and 2.57 for delays of 0, 1, and 7 days, respectively) [$F(2,336) = 44.05, MSe = 2.01, p < .05$]. Overall, females recalled more words than males (means of 3.42 and 2.82) [$F(1,336) = 46.00, MSe = 2.01, p < .01$].

ANOVA of performers' errors in free recall indicated that there had been no differences in this respect for gender (male mean = 1.79, female = 1.94) [$F(1,336) = 1.97, MSe = 2.39, p > .05$] or test interval (1.90, 1.85, and 1.85 errors for 0-, 1-, and 7-day intervals, respectively) [$F(2,336) = 2.17, MSe = 2.39, p > .05$].

Obviously there were no surprises in these main effects, but the interactions were more interesting. Word frequency interacted reliably with test interval [$F(4,672) = 5.16, MSe = 1.31, p < .05$]. This interaction was due to a relatively greater decrement after 7 days for words that occurred two or three times (mean scores of 1.21, 3.93, and 5.55 for one, two, and three occurrences, respectively, for the 0-day interval, corresponding means of 1.04, 3.58, and 5.10 for the 1-day interval, and .75, 2.73, and 4.22 for the 7-day interval). There was also a reliable interaction between gender and test interval [$F(2,336) = 5.73, MSe = 2.01, p < .05$]. Female superiority in recall was due mainly to less forgetting over the delayed test and particularly after the 1-day delay, when a slight increment in the number of mean words recalled actually occurred.

These relationships are shown in Figure 1, which depicts the reliable triple interaction of gender, test interval, and word frequency [$F(4,672) = 2.55, MSe = 1.31, p < .05$]. Female superiority is evident for all frequencies and test intervals. The female increment in free recall from 0- to 1-day test interval is apparent for F2 and F3 words (and exactly the same means occurred

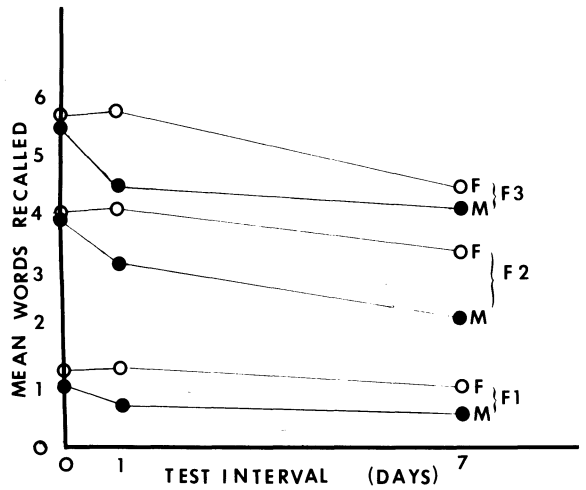


Figure 1. Free recall as a function of gender (M, F), word frequency (F1, F2, F3), and test interval.

for F1 words). It is also apparent that the interaction between word frequency and test interval was due to the relatively greater loss in recall scores for F2 and F3 words.

Performers and observers. Overall, performers recalled more words than observers (means of 3.12 and 2.84) [$F(1,680) = 14.44, MSe = 2.78, p < .01$], and the performance of females exceeded that of males at each word frequency [$F(1,680) = 64.17, MSe = 2.78, p < .01$]. Moreover, as Figure 2 suggests, there was a reliable interaction between the performance/observation and word-frequency variables [$F(2,1360) = 3.99, MSe = 1.44, p < .05$]. It is evident that advantages in F2 and F3 words were responsible for the performer superiority in recall.

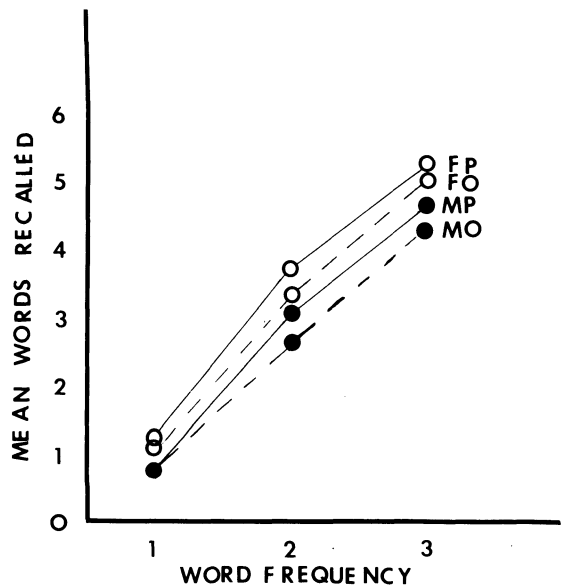


Figure 2. Free recall as a function of gender (M, F), performance (P) or observation (O), and word frequency.

Although there was no reliable interaction between gender and word frequency for performers alone, a reliable interaction was obtained when both performers and observers were analyzed [$F(2,1360) = 3.69$, $MSe = 1.44$, $p < .05$]. This effect can be attributed to the greater recall of females after two and three word occurrences, as indicated in Figure 2.

Cued Retrieval

Performers. Data for the two test intervals, both showing reliable negative functions, and word frequency, showing a reliable inverted curve (highest scores for F2 words), have been reported previously (Marx & Kim, 1984). That report showed a surprisingly strong inhibitory effect of a single recall on F1 words even after a full 14-day delay (cf. also Marx, 1982). Resistance to this inhibitory effect of recall was observed in both F2 and F3 words, and especially for the short delay intervals. Now, with the inclusion of the gender variable, it seems that there was greater resistance to this kind of inhibition of retrieval in the female subjects. Figure 3 shows this result. Both sexes showed greater resistance (higher percentage of relatively more recalled-word retrievals) for the short delay interval, but there was a consistent female advantage in this measure for both F2 and F3 words. Except for the F2 short delay condition [$\chi^2(1) = 2.39$, $p > .05$], the gender differences shown in Figure 3 were reliable [$\chi^2(1) = 6.45$, $p < .05$, for the F2 long delay, $\chi^2(1) = 7.76$, $p < .01$, for the F3 short delay, and $\chi^2(1) = 5.44$, $p < .05$, for the F3 long delay].

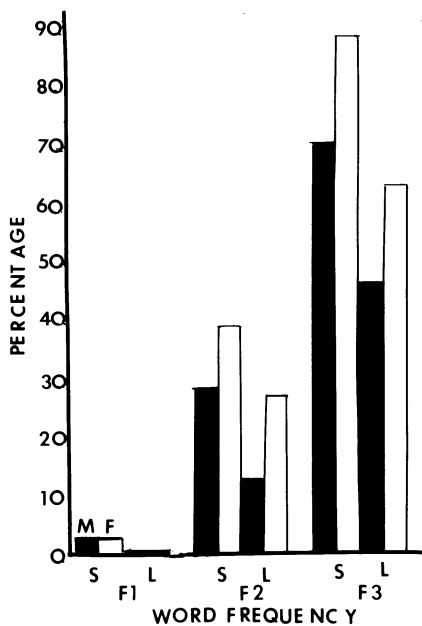


Figure 3. Percentage of male (black bar) and female (white bar) subjects who retrieved more recalled than nonrecalled words for the three word frequencies (F1, F2, F3) after short (S) and long (L) test intervals.

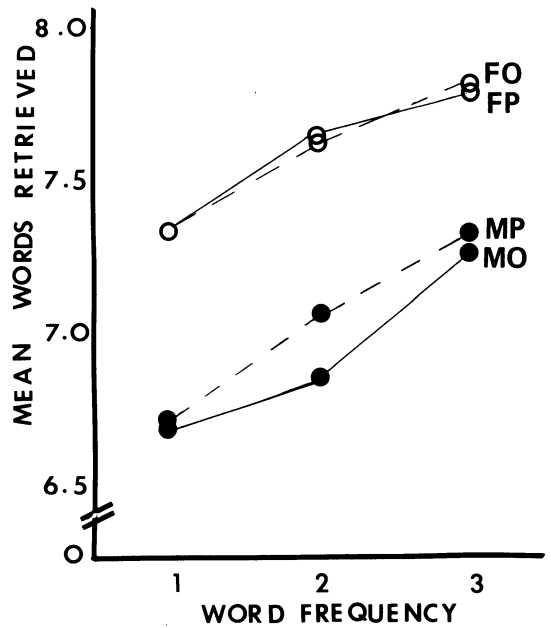


Figure 4. Mean number of words retrieved by male(M)/female(F) performers(P)/observers(O) for the three word frequencies.

Performers and observers. There were reliable main effects of gender (overall male score = 6.98; overall female score = 7.58) [$F(1,680) = 65.79$, $MSe = 2.82$, $p < .01$] and word frequency (overall means of 7.01, 7.30, and 7.54 for one, two, and three occurrences, respectively) [$F(2,1360) = 127.65$, $MSe = .37$, $p < .01$].

Gender interacted reliably with word frequency [$F(2,1360) = 3.83$, $MSe = .37$, $p < .05$], reflecting the reduced female advantage for F3 words (female means of 7.34, 7.63, and 7.78 for one, two, and three word occurrences, respectively, compared with male means of 6.69, 6.96, and 7.29).

The main effect of performance/observation was not reliable [$F(1,680) = .66$, $MSe = 2.82$, $p > .05$]. There was, however, an interesting triple interaction among performance/observation, gender, and word frequency [$F(2,1360) = 21.92$, $MSe = .37$, $p < .01$], shown in Figure 4. It indicates that male performers retrieved more words (especially F2 words) than male observers. Overall, however, females were obviously superior to males, both as performers and as observers, and did not differ in this respect.

DISCUSSION

The overall superiority of females in learning and remembering verbal materials has been amply demonstrated (Maccoby & Jacklin, 1974), but, beyond this generalization, more specific relationships need to be delineated. The present results invite some such suggestions.

As is evident in Figure 1, there was an increasing female advantage in free recall over time—even a slight increment in mean score for females from 0 to 1 day, contrasting with the

consistent male decrement. This result complements the earlier reports (Marx, Marx, & Homer, 1980, 1981) that female superiority in acquisition is an increasing function of trials. It might well be fruitful for researchers who have appropriate subject samples to look for such interactions between gender and trials or time variables.

The consistently greater resistance of female subjects to the inhibitory effects of recall for F2 and F3 words (Figure 3) is an additional reflection of their superiority in verbal learning and memory.

The interactions of gender with performance/observation are generally consistent with previous reports. For example, the clear indication in Figure 4 that males are superior retrievers after performance supports an earlier suggestion (Marx & Witter, 1972). Apparently, males tend to repeat their own responses more than do females, even when they are errors (Marx et al., 1981, Figure 1; Marx et al., 1973).

The interaction between performance/observation and word frequency (Figure 2) in free recall indicates the dependence of the performer advantage on the higher word frequencies. Whether responses are really strengthened more by frequency of occurrence, whether "reinforced" or not, in this kind of incidental learning situation (no instruction to remember) cannot be determined from the present data. It is also possible, of course, that the observers paid less attention to the word repetitions. But the fact that they were equally proficient at recalling single responses suggests no overall lessening of attention. However interpreted, these results are consistent with the recent reports of performer superiority in acquisition (Marx, Homer, & Marx, 1980; Marx, Marx, & Homer, 1980, 1981) under various conditions of intentional learning of verbal materials.

The interaction between word frequency and test interval is clearly attributable to the markedly reduced level of recall of F1 words at the 0-day interval. Eventual closing of the gap between word frequencies would be anticipated if longer test delays were used, in somewhat the same manner as retention and transfer scores have been shown to converge over long time intervals (Bunch, 1941).

The main effect of interval is itself of some interest. Nearly a century ago, Ebbinghaus (1885/1913) reported a positive relationship between amount of practice and retention for nonsense syllables, and, more recently (only half a century ago), a similarly uniform relationship was observed for verbal materials (Krueger, 1929) and finger mazes (Krueger, 1930). In still more recent years, however, relating long-term temporal intervals to

frequency and similar variables seems to have been of small interest to learning-memory researchers and theorists.

REFERENCES

- BUNCH, M. E. (1941). A comparison of retention and transfer of training from similar material after relatively long intervals of time. *Journal of Comparative Psychology*, **32**, 217-231.
- EBBINGHAUS, H. (1913). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Teachers College, Columbia University. (Original work published 1885)
- KRUEGER, W. C. F. (1929). The effect of overlearning on retention. *Journal of Experimental Psychology*, **12**, 71-78.
- KRUEGER, W. C. F. (1930). Further studies in overlearning. *Journal of Experimental Psychology*, **13**, 152-163.
- MACCOBY, E. E., & JACKLIN, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.
- MARX, M. H. (1982). Effects of frequency of prior incidental occurrence and recall of target words on anagram solution. *Bulletin of the Psychonomic Society*, **19**, 253-255.
- MARX, M. H., HOMER, A. L., & MARX, K. (1980). Verbal discrimination learning and retention as a function of task and performance or observation. *Bulletin of the Psychonomic Society*, **15**, 167-170.
- MARX, M. H., & KIM, Y. C. (1984). Inhibition of learned-response availability: Reduction of cued retrieval by prior recall of target words. *Bulletin of the Psychonomic Society*, **22**, 29-32.
- MARX, M. H., MARX, K., & HOMER, A. L. (1980). Verbal discrimination learning and retention as a function of performance or observation and ease of conceptualization of task materials. *Bulletin of the Psychonomic Society*, **16**, 135-136.
- MARX, M. H., MARX, K., & HOMER, A. L. (1981). Interactions among performance, task, and gender variables in verbal discrimination learning. *Bulletin of the Psychonomic Society*, **18**, 9-11.
- MARX, M. H., & WITTER, D. W. (1972). Repetition of correct responses and errors as a function of performance with reward or information. *Journal of Experimental Psychology*, **92**, 53-58.
- MARX, M. H., WITTER, D. W., & FARBRY, J. (1973). Greater repetition of errors under performance compared to observation in multiple-choice human learning. *Perceptual and Motor Skills*, **37**, 949-950.

(Manuscript received for publication November 22, 1983.)