

# Recall as a function of method of presentation and individual differences in test anxiety

JOHN H. MUELLER, MICHAEL CARLOMUSTO, and MATTHEW MARLER  
*University of Missouri, Columbia, Missouri 65201*

Subjects differing in test anxiety had four free recall trials with a list of words that could be organized by associates or rhymes. The words were presented in a simultaneous display, with related items either grouped together spatially or separated. Anxiety deficits in organization previously observed with single-item presentation were reduced, but low-anxiety subjects still recalled a greater number of words. Blocking related items led to superior performance overall, but spatial arrangement did not interact with anxiety level.

Recent research has found that high-anxiety subjects typically recall fewer words on a free recall test, and their recall generally involves less clustering of related words (e.g., Mueller, in press). One interpretation of this deficit follows from the observation that high-anxiety subjects utilize fewer cues during learning (cf. Kausler & Trapp, 1960). Applied to free recall, this analysis suggests that word features which might help organize the list are not employed to benefit recall (Mueller, 1976).

This experiment was concerned with the generality of the deficit, that is, which word features are not used by high-anxiety subjects. Memory theorists currently find heuristic value in conceptualizing the encoding of words in terms of a hierarchy of features differing in the "level" of processing required for encoding (Craik & Lockhart, 1972), with the assumption that the durability of the memory trace will be directly related to the depth of processing. Examination of earlier findings on the anxiety deficit suggested that high-anxiety subjects did not process material as deeply as low-anxiety subjects (Mueller, 1976). In other words, compared to low-anxiety subjects, high-anxiety subjects fail to utilize all of the available semantic (deep) content, but may use a comparable proportion of the physical (shallow) attributes.

Tests of the depth analysis of the high-anxiety deficit in free recall have used word lists that can be classified in alternative ways (Mueller, in press). For example, given a list with words like KING, RING, BELL, subjects might cluster by associates or rhymes, or perhaps by both. This allows a contrast of deep and shallow encoding by the same subject for the same words, as well as a comparison across subjects differing in anxiety level. Experiments with such lists (Mueller, 1977, in press) have demonstrated a high-anxiety deficit for both deep

and shallow bases of organization, in apparent contradiction of the levels analysis of the anxiety deficit. However, some methodological considerations preclude such a conclusion, and the present study examined one such problem.

The present methodology tried to optimize the discovery of both bases for organization in the rhyme-associate list construction. It may be difficult to detect some intralist relationships when items are presented for study one at a time, and if semantic relationships are more salient than rhymes, then a fair test of deep vs. shallow encoding would not be obtained. In particular, if subjects must thoroughly study the words to detect the presence of rhymes, then that basis for organization can no longer be considered shallow. Therefore, to make the presence of rhymes and associates more obvious, all of the words were shown in a simultaneous display, providing a more definitive test of the depth analysis of the anxiety deficit.

A second reason for examining whole-list displays is that this procedure should minimize any intrinsic problem that high-anxiety subjects have in detecting intralist relationships per se, regardless of depth. Previous findings of an anxiety deficit for "shallow" bases of organization might derive from a failure to detect interitem relationships as much as a failure to utilize them, but whole-list display should enable subjects to discover the presence of all intralist relationships, reducing the contribution of discovery per se to the anxiety deficit. The ease of discovery should still be manipulable, for example, by arranging related items together in a spatial grouping vs. some noncontiguous manner. In fact, blocked simultaneous displays should be the optimal arrangement for detecting interitem relationships, and any remaining anxiety deficit would be due to differences in utilizing cues.

## METHOD

### Subjects and Design

Eighty subjects from introductory psychology courses participated: High-anxiety subjects had test-anxiety scores (Sarason,

Requests for reprints should be sent to John H. Mueller. This investigation was supported in part by Biomedical Research Support Grant RR 07053 from the National Institutes of Health and by funds from the Research Council of the Graduate School of the University of Missouri.

1972) greater than 23 (mean = 29.8); low-anxiety subjects had scores below 9 (mean = 5.4). The basic design was a 2 by 2 factorial, with two levels of anxiety and two arrangements for the display of related items (spatially blocked or random). There were 10 men and 10 women in each of the four main groups.

### Materials

The word list was composed of 12 triads. Each triad had a base word (e.g., CHAIR), its primary associate in the Shapiro and Palermo (1968) atlas (e.g., TABLE), and a rhyme to the base word (e.g., SHARE). Two versions of the list construction were developed. All words in a list were uniquely defined by their first two letters, nonhomophonic, three to six letters in length, and all except two were A or AA in Thorndike-Lorge frequency. The unrelated items used in previous experiments (Mueller, 1977, in press) were omitted here.

The entire 36-word list was presented at once for study. The blocked array had the words in a triad in the same row, with the base word in the middle. Six triads were arranged in a columnar fashion on the left in the display, and the other six were spatially distinct on the right; all those on the left had the triads ordered rhyme-base-associate, and those on the right reversed this order. There were four different arrangements, counterbalancing specific spatial location of the rhyme and associate within a line and across sides of the array: A different arrangement was presented for study each trial. The random array was constructed so that neither the rhyme nor associate was within two rows of the line containing the base word, nor were the rhyme and associate in the same line. Each line maintained the rhyme-base-associate (or vice versa) ordering. There were four arrangements of this array structure also. Subjects were not informed of the presence of the relationships or the spatial grouping.

### Procedure

Each subject had four trials on the list, involving a 90-sec display and an immediate 90-sec written test each time. The arrays were shown by slide projector, and subjects were instructed to devote 2-3 sec to each word. Instructions emphasized the option to recall in any order. During the test, the experimenter said "mark" after 45 sec had elapsed, at which signal subjects drew a line under the last word written at that point and then continued. This allowed some assessment of the effect of anxiety on rate of recall (cf. Eysenck, 1976).

After the fourth test, there was a delay of about 6 min, filled by digit-span tests and a questionnaire. A 90-sec delayed written recall test followed without further study. Finally, the subjects were contacted by telephone a week (6-8 days) later for an unannounced 90-sec oral recall test.

At the beginning of the experiment, and again after the digit-span tests, subjects filled out anxiety questionnaires: the 20-item State Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970) and the short form (20 items) of the Thayer Activation-Deactivation Adjective Checklist (Thayer, 1967), counterbalanced as first and second tasks. High and low scorers on the test-anxiety scale scored differently on the State Anxiety Inventory (means = 43.2 and 35.2, respectively). High- and low-test-anxiety subjects also differed on the High Activation subscale of the Thayer checklist (means = 9.6 and 7.1, respectively) and the Sleep subscale (11.3 vs. 9.3), though they did not differ on the General Activation (11.1 vs. 11.3) nor the Deactivation subscales (12.3 vs. 13.3). Thus, subjects selected by test-anxiety scores also differed in situational anxiety. This would be expected since most subjects perceive an experiment as an evaluation, and the general instructions encouraged this by indicating that the experimenter was interested in the relationship between intelligence, recall, and the screening inventory.

## RESULTS<sup>1</sup>

### Immediate Tests

**Recall.** Table 1 summarizes recall and organization

for Trials 1-4. Considering the entire test period, high-anxiety subjects recalled fewer rhymes, associates, and base words than low-anxiety subjects ( $F_s = 4.37, 4.68, 4.62$ ;  $MSE_s = 14.04, 7.93, 7.32$ , respectively), with no Anxiety by Trials interactions ( $F_s < 1.65$ ). Subjects with blocked arrays recalled more rhymes and base words than subjects with random arrays ( $F_s = 14.82, 13.22$ ;  $MSE_s = 14.04, 7.32$ , respectively), but not more associates ( $F = 2.46$ ). The Array by Trials interactions were not significant for the recall of rhymes or base words ( $F_s < 2.05$ ), with blocking superior from Trial 1. However, this interaction was marginally significant for associates recall ( $F = 3.22, MSE = 1.11, p < .10$ ), as blocking eventually led to superior recall though it initially was little better than the random array. Neither the Anxiety by Array nor the Anxiety by Array by Trials interactions were significant ( $F_s < 2.18$ ).

Table 1 also shows mean recall during the first half of the test period, in parentheses. High-anxiety subjects recalled fewer associates in the first half ( $F = 5.79, MSE = 5.29$ ), but the difference was not significant for base words ( $F = 3.57, MSE = 4.41, p < .06$ ) or rhymes ( $F = 1.11$ ). Blocking led to more recall of rhymes ( $F = 11.52, MSE = 8.50$ ), as if it did improve detection of the rhymes, but blocking had no significant effect on base word ( $F = 3.77, MSE = 4.41, p < .06$ ) or associates recall ( $F < 1$ ). The Anxiety by Array interactions were not significant ( $F_s < 1$ ). The trials main effect was significant in all three cases ( $F_s > 35.00$ ), with increasing recall during the first half, but there were no interactions with trials ( $F_s < 2.04$ ).

The early recall data may also be examined by considering the proportion of total recall which occurred during the first half of the test. Pooling across word types, the mean proportions were .78 vs. .81 for the high-anxiety blocked and random groups and .76 vs. .81 for the low-anxiety blocked and random groups. Thus, when differences in total recall were taken into account, there was no difference in early recall as a function of anxiety level, and blocking led to relatively more recall in the second half of the test period.

**Table 1**  
Average Recall, Clusters, and Triads Recalled Per Trial on Trials 1-4, by Test Anxiety and Display Arrangement

	High Anxiety		Low Anxiety	
	Blocked	Random	Blocked	Random
Recall				
Associate	7.51 (5.70)	7.40 (5.93)	8.58 (6.56)	7.70 (6.30)
Rhyme	6.79 (5.14)	5.08 (3.88)	7.56 (5.33)	6.05 (4.38)
Base Word	7.88 (6.05)	6.98 (5.68)	8.73 (6.58)	7.43 (6.04)
Clustering				
Associate	5.73 (4.39)	3.81 (3.14)	6.70 (5.08)	3.85 (3.21)
Rhyme	5.21 (4.09)	.61 (.49)	5.91 (4.45)	.51 (.40)
Triads				
Consecutive	.68 (.71)	.03 (.03)	.70 (.71)	.03 (.03)
Anywhere	.74 (.76)	.31 (.30)	.80 (.81)	.40 (.38)

*Note—Numbers in parentheses are the corresponding value for performance in the first half of the test period.*

**Clustering.** For the entire test period, blocking led to more associate and rhyme clusters than the random display ( $F_s = 29.40, 165.25$ ;  $MSe_s = 15.43, 12.10$ ). An Array by Trials interaction revealed the effect to be more pronounced on early trials for rhyme clusters ( $F = 15.35, MSe = 1.57$ ). The anxiety main effect was not significant for either type of clustering ( $F_s < 1.33$ ), and anxiety did not interact with array nor trials ( $F_s < 2.60$ ).

Clusters during the first half of the test revealed the same results. Blocking led to more clusters of each type ( $F_s = 18.49, 162.02$ ;  $MSe_s = 10.48, 7.22$ ), with an Array by Trials interaction for rhyme clusters ( $F = 6.79, MSe = 1.07$ ). There were no anxiety main effects ( $F_s < 1.11$ ) nor any interactions involving anxiety ( $F_s < 1.28$ ).

**Triad recall.** Another view of organization is provided by analyzing the recall of entire triads (base word, associate, and rhyme), either consecutively in output or anywhere in the output sequence (cf. Mueller, 1977, in press). The number of triads was divided by the number of base words recalled, that is, opportunities to recall a whole triad. Over the entire test period, blocking led to more triad recall for both consecutive and anywhere scorings ( $F_s = 286.80, 111.90$ ;  $MSe_s = .12, .12$ ). There was no anxiety main effect for the strict criterion ( $F < 1$ ) and only a marginal tendency for high-anxiety subjects to recall fewer triads by the lenient criterion ( $F = 3.71, MSe = .12, p < .06$ ). In neither case was the Anxiety by Array interaction significant ( $F_s < 1$ ). Anxiety did not interact with trials for either scoring ( $F_s < 1.82$ ), nor were the Anxiety by Array by Trials interactions significant ( $F_s < 2.63$ ). The strict criterion revealed an Array by Trials interaction ( $F = 8.44, MSe = .02$ ), as triads became more numerous over trials with blocked but not with random arrays. This interaction was not present for the lenient criterion ( $F < 1$ ).

The pattern of results was similar for the first half of the test. The array main effect was significant by either criterion ( $F_s = 284.32, 125.33$ ;  $MSe_s = .13, .13$ ), but neither the anxiety main effects ( $F_s < 2.70$ ) nor the Anxiety by Array interactions were significant ( $F_s < 1$ ).

**Delayed Tests**

Table 2 summarizes the results for the two delayed tests, with the means for the second test in parentheses. The first-half analyses will not be reported since no differences were apparent. The results should be interpreted cautiously in view of differences in degree of original learning.

On the first delayed test, while high-anxiety subjects consistently recalled less than low-anxiety subjects, the difference was not significant for any item type ( $F_s < 3.74$ ), nor were the Anxiety by Array interactions significant ( $F_s < 1.05$ ). Blocking led to more recall than the random display for rhymes and base words ( $F_s = 7.85, 5.19$ ;  $MSe_s = 5.54, 3.86$ ) but not for associates ( $F = 2.82$ ).

**Table 2**  
Average Recall, Clusters, and Triads Recalled on the Delayed Tests, by Anxiety and Display Arrangement

	High Anxiety		Low Anxiety	
	Blocked	Random	Blocked	Random
Recall				
Associate	9.00 (4.80)	8.20 (5.25)	9.55 (5.85)	9.00 (6.20)
Rhyme	8.00 (3.75)	6.65 (3.00)	8.75 (4.75)	7.15 (3.20)
Base Word	9.15 (4.95)	7.70 (4.80)	9.55 (5.75)	9.00 (6.10)
Clusters				
Associate	6.30 (3.55)	5.05 (3.40)	7.45 (4.35)	6.10 (4.10)
Rhyme	6.30 (2.65)	1.10 ( .55)	6.55 (3.15)	1.05 ( .35)
Triads				
Consecutive	.73 ( .50)	.04 ( .04)	.81 ( .58)	.07 ( .03)
Anywhere	.80 ( .57)	.40 ( .16)	.86 ( .67)	.50 ( .24)

*Note*—The mean for the first delayed test is shown and, in parentheses, is shown the corresponding mean for the second delayed test.

High-anxiety subjects also clustered less than low-anxiety subjects, but not significantly so ( $F_s < 3.42$ ), nor were the Anxiety by Array interactions significant ( $F_s < 1$ ). Blocked presentation led to more associates and rhyme clusters ( $F_s = 4.78, 100.72$ ;  $MSe_s = 7.08, 5.68$ ).

Neither the anxiety main effects nor Anxiety by Array interactions were significant for either method of scoring triads ( $F_s < 2.41$ ). Blocking led to more triad recall by either criterion ( $F_s = 226.58, 53.88$ ;  $MSe_s = .04, .05$ ).

On the second delayed test, blocking led to better performance for rhyme recall, rhyme clusters, and both triad measures ( $F_s = 5.60, 44.93, 116.71, 55.61$ ;  $MSe_s = 4.72, 2.67, .04, .06$ ) but not for associate or base-word recall ( $F_s < 1$ ) or associate clusters ( $F < 1$ ). High-anxiety subjects performed worse on most measures, but significantly so only for base-word recall ( $F = 4.37, MSe = 5.04$ ). There were no Anxiety by Array interactions.

**DISCUSSION**

The present results showed a clear effect due to the manner of blocking related items in whole-list displays. Therefore, even though simultaneous display should improve the discovery of intralist relationships compared to single-item presentation, it was still possible to vary the level of performance. This demonstrates that this procedure is sufficiently sensitive that the absence of an effect for some variable such as anxiety is not readily interpreted in terms of attenuation by a ceiling effect. The blocked-random effect is also relevant in that it allows an assessment of the anxiety deficit as a function of the salience of the organizational dimension.

With regard to the general anxiety deficit, the effects in the present data were in the direction observed in earlier studies (low-anxiety subjects performed better). The difference was significant in terms of number recalled for all three item types, and it was not reduced by blocking related items (i.e., no Anxiety by Array interactions). Therefore, even with simultaneous display and blocking, high-anxiety subjects still recalled less than low-anxiety subjects.

However, organization analyses indicated no significant

anxiety deficit in terms of clustering or triad recall, though the direction of any differences still favored low-anxiety subjects. This suggests that whole-list presentation did help high-anxiety subjects to detect both deep and shallow intralist relationships, even without blocking related items (i.e., no Anxiety by Array interactions). It can be noted that with blocked arrays rhyme clusters were roughly as numerous as associate clusters, unlike studies where items were presented one at a time (Mueller, 1977, in press). The gain for rhyme clustering was not limited to high-anxiety subjects, however.

With regard to the depth analysis of anxiety differences, there was a significant anxiety deficit for the recall of rhymes and associates and there was no deficit for either type of clustering. Neither outcome is consistent with the hypothesis that the anxiety deficit would be greater for the semantic dimension. There may still be some question as to whether rhymes can be truly superficial organizational schemes in a recall task, but simultaneous display with blocking would seem to nearly eliminate the problem of discovering both semantic and physical relationships. Other tasks may provide clearer tests of the hypothesis that the anxiety deficit is greater for semantic properties (i.e., situations where the "shallow" dimension can indeed be superficially processed). However, within the limits of procedures considered to date using free recall, the extension of the depth hypothesis receives little support. Instead, it appears that the anxiety deficit encompasses numerous bases for organization in memory, with no clear separation in terms of deep or shallow processing. It should be noted that this conclusion applies to depth as a characteristic of the to-be-remembered items. It does not necessarily mean that other conceptualizations of "levels of processing" (e.g., "breadth" vs. "depth" or "elaborative" vs. "maintenance") might be unable to incorporate these results. Furthermore, other research (e.g., Mueller, in press) has found that study tasks which orient processing to particular features may produce results more compatible with the extension of the depth viewpoint to the anxiety deficit. Such orienting tasks were not examined here, of course, since they cannot be rigorously controlled given a simultaneous display; however, they may represent a more fruitful area for further investigation of the cue utilization deficit.

## REFERENCES

- CRAIK, F. I. M., & LOCKHART, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 671-684.
- EYSENCK, M. W. Arousal, learning, and memory. *Psychological Bulletin*, 1976, 83, 389-404.
- KAUSLER, D. H., & TRAPP, E. P. Motivation and cue utilization in intentional and incidental learning. *Psychological Review*, 1960, 67, 373-379.
- MUELLER, J. H. Anxiety and cue utilization in human learning and memory. In M. Zuckerman & C. D. Spielberger (Eds.), *Emotions and anxiety: New concepts, methods and applications*. Hillsdale, N.J: Erlbaum, 1976.
- MUELLER, J. H. Test anxiety, input modality, and levels of organization in free recall. *Bulletin of the Psychonomic Society*, 1977, 9, 67-69.
- MUELLER, J. H. The effects of individual differences in test anxiety and type of orienting task on levels of organization in free recall. *Journal of Research in Personality*, in press.
- SARASON, I. G. Experimental approaches to test anxiety: Attention and the uses of information. In C. D. Spielberger (Ed.), *Anxiety: Current trends in theory and research* (Vol. 2). New York: Academic Press, 1972.
- SHAPIRO, S. I., & PALERMO, D. S. An atlas of normative free association data. *Psychonomic Monograph Supplements*, 1968, 2(Whole No. 28).
- SPIELBERGER, C. D., GORSUCH, R. L., & LUSHENE, R. E. *Manual for the state-trait anxiety inventory*. Palo Alto, Calif: Consulting Psychologists Press, 1970.
- THAYER, R. E. Measurement of activation through self-report. *Psychological Reports*, 1967, 20, 663-678.

## NOTE

1. Effects described as significant involve  $p < .05$ , unless noted otherwise. Effects involving the repeated-measures factor (trials) were evaluated with conservative degrees of freedom (Greenhouse-Geisser). All Fs have 1 and 72 df.

(Received for publication August 1, 1977.)