

Anagram solving as influenced by solution word frequency, anagram transition probability, and subject's vocabulary level

ROY B. WEINSTOCK

Mary Washington College, Fredericksburg, Virginia 22401

A total of 128 female undergraduates served as volunteer subjects in an anagram problem solving experiment. The variables manipulated were Thorndike-Lorge frequency of the solution words (common, uncommon), anagram transition probability (low, high), and subject's vocabulary level (low, high). All combinations of these variables yielded a total of eight independent groups (each $n=16$). The main findings were: (1) anagram transition probability is effective when anagrams are based on Thorndike-Lorge common, but not uncommon, words; and (2) the vocabulary size effect (i.e., the finding that high-vocabulary subjects tend to solve more anagrams than low-vocabulary subjects) appears to be due strictly to the number of words stored in memory and not to any retrieval efficiency differences among these subjects.

By and large, psychologists have used the anagram task to examine three aspects of behavior. These have been the study of: (1) cognitive processes in general, (2) the effect language structure variables particularly exert on problem solving performance, and (3) the role(s) played by various individual difference factors in problem solving. In the research to be reported, aspects of the latter two categories were explored.

It has been proposed that "fundamentally the solution of anagrams is a retrieval problem, i.e., the subject's task is to retrieve a word which is already in memory" (Mendelsohn, 1976, p. 641). Such a statement seems to imply that retrieval-from-storage efficiency is uniform across people, and it would also lead to the prediction that differences in anagram solution performance would be positively related to manifest vocabulary size. The latter has, in fact, been found to be the case (Mendelsohn & Covington, 1972; Mendelsohn & Griswold, 1966; Mendelsohn, Griswold, & Anderson, 1966). However, it seems that another possibility should be considered: Could it not be the case that people having manifestly different vocabulary levels differ not in number of words stored but, rather, in retrieval efficiency? That is, it seems possible to imagine two people who have "stored" the same number of words in memory but obtain different vocabulary test scores due to differences in their retrieval skills. Such a situation would not only lead to differences in vocabulary scores but would affect anagram solution performance, as well, and in the direction mentioned just above. The main purpose of this research was to try to determine

whether people differing in anagram solution performance do so because of word storage differences or because of differences in retrieval efficiency. Accordingly, subjects differing in vocabulary test scores were given anagrams whose solutions were words of either low or high frequency, as reported by Thorndike and Lorge (T-L) (1944). If differences between these two sets of subjects are mediated by certain differences in the retrieval process, then high-vocabulary subjects should exceed low-vocabulary subjects at each level of the solution word frequency variable. That is, no interaction between vocabulary size and solution word frequency should be found. However, given that the critical differences between low- and high-vocabulary size subjects is related simply to a disparity in the number of words stored in memory, an interaction should be found between these variables. Such an interaction would be of the form that solution of high-frequency (i.e., common) word anagrams would be equivalent at each vocabulary level, whereas in the case of low-frequency (i.e., uncommon) words, high-vocabulary subjects would prevail. The latter would follow from the assumption that very common words have been stored by both low- and high-vocabulary subjects but that uncommon words are available only to high-vocabulary subjects. It was the intention of the present experiment to see which of these two alternatives would occur, thereby providing an answer to the question posed above.

In addition to the two variables of vocabulary size and solution word frequency, the structural variable of anagram transition probability (ATP) was investigated. Transition probability is a measure of the relative frequency of appearance in the language shown by various combinations of letter sequences. The efficacy of the ATP variable has recently been questioned (Mendelsohn & O'Brien, 1974) and, as such, it was felt that another examination of it was warranted.

The author wishes to thank Janet Barton and Suzanne Duperock for their assistance in data collection and analysis. Requests for reprints should be sent to: Roy B. Weinstock, Department of Psychology, Mary Washington College, Fredericksburg, Virginia 22401.

METHOD

Preexperimental Procedure

The student pool from which the subjects used in this study were drawn comprised approximately 400 female undergraduates enrolled in a course in general psychology. These students were prescreened with two paper-and-pencil test forms: a "student-opinion" questionnaire and a modified version of the Basic Word Vocabulary Test (BWVT; Dupuy, 1974). The ostensible purpose of the seven-item opinion survey (which was administered first) was to obtain information on student attitudes on a variety of issues. Each item could be answered by circling a number, on a scale from 1 to 9, which indicated the extent to which the student agreed or disagreed with a particular statement. Of the seven statements, the two of interest for this study were the following: "I enjoy doing crossword puzzles," and "I enjoy playing word games, such as anagrams (making words from scrambled letters)." Data relating to these two items will be presented and discussed below.

After the opinion survey sheet was collected, a modified version of the BWVT was administered. The modification consisted of a 60-item test made from Questions 50-109 of the original test (cf. Dupuy, 1974, Appendix 4) and was administered according to the procedures prescribed in the BWVT booklet. Data from both the opinion survey and the vocabulary test were collected in a group (classroom) setting.

Subjects

A total of 128 students from those given the opinion survey and modified BWVT took part in the experiment. This group comprised two sets of 64 students each who had scored at the upper and lower ends, respectively, of the vocabulary test distribution of scores. The mean scores for the two sets of students were 33.11 and 48.41 (maximum score = 60) for those designated low- and high-vocabulary size subjects, respectively. The difference between these means is highly reliable [$t(126) = 34.77$, $p < .001$]. All subjects were volunteers, were 18-22 years of age, and were experimentally naive at the time of testing.

Materials

A total of 30 five-letter words was selected from the T-L frequency norms and served as solution words. Half of these words were classified as common (A or AA rating), and the remaining 15 were classified as uncommon (frequencies of 1 or 2/million). None of the solution words contained repeated letters. Two sets of anagrams were formed from each list, one set being low in ATP level (cf. Mayzner & Tresselt, 1965), and the other, high. In this way, four different lists were constructed involving all combinations of low or high solution word T-L frequency and low or high ATP level. Every anagram used had a single solution outcome, as determined by Olson and Schwartz (1967).

Beyond the above manipulations, care was taken to equate the four lists of anagrams in several ways. For both frequency levels, each of the 15 solution words began with a different letter of the alphabet; this group of initial letters was identical for both sets of solution words. In addition, two other variables known to have a very strong effect on anagram solution performance were equated. Specifically, among the four anagram lists, number of letter moves to solution were matched; among the two sets of solution words (low and high frequency), total bigram rank scores (Mendelsohn & O'Brien, 1974) were equivalent (all $ps > .05$).

Design

The design for this study can be described as a 2 by 2 by 2 factorial arrangement. The variables were T-L solution word frequency (C = common; U = uncommon), ATP (L = low; H = high), and the subject's vocabulary level (L = low; H = high). All possible combinations of these variables yielded eight inde-

pendent groups (each $n = 16$); CLL, CHL, ULL, UHL, CLH, CHH, ULH, UHH, where the first letter represents T-L frequency, the second letter, ATP, and the third letter, vocabulary level.

Experimental Procedure

Each subject was tested individually by a thoroughly trained undergraduate who was unaware of a particular subject's vocabulary level at the time of testing. After being read standard anagram solution instructions (during which two sample anagrams were presented as practice tasks), the subject was presented with the set of anagrams appropriate to her preassigned condition. The 15 anagrams were presented one at a time in a fixed order, each anagram having been typed in capital letters on a 7.50 x 12.50 cm unlined index card. A 240-sec time limit was imposed on the solution of any given anagram. If a subject was unable to solve an anagram within this period of time, this was so recorded, and the solution was provided. Time to solution was recorded with the aid of a stopwatch. All solutions were given verbally.

RESULTS

The two dependent variables chosen to index anagram solution performance were number of anagrams solved and median time to solution. Means and standard deviations of these measures for the various conditions are presented in Table 1. Separate 2 by 2 by 2 factorial analyses of variance were performed on the data for each of these measures. The variables evaluated were solution word T-L frequency, ATP, and vocabulary level.

For the analysis of variance based on mean number of anagrams solved, each of the variables separately was statistically significant ($ps < .025$), but consideration of them must be qualified due to the fact that they were involved in two lower order interactions. In one case, T-L frequency interacted with ATP [$F(1,120) = 5.43$, $p < .025$]. A Newman-Keuls test of this interaction revealed that for uncommon words, ATP was ineffective, whereas for common words, solutions of low-ATP words exceeded those for high-ATP conditions ($p < .05$). The other interaction involved T-L frequency and vocabulary level [$F(1,120) = 4.63$, $p < .05$]. A Newman-Keuls breakdown of this interaction revealed that, while vocabulary level did not exert a differential

Table 1
Means and Standard Deviations for All
Groups for Each Dependent Variable

Group	Correct Solutions		Median Seconds to Solution*	
	Mean	SD	Mean	SD
CLL	12.94	1.06	23.78	20.28
CHL	10.88	2.22	40.72	25.95
ULL	8.38	2.92	45.28	24.96
UHL	8.88	1.96	57.19	18.31
CLH	13.50	1.37	16.50	12.92
CHH	12.25	1.81	36.41	23.53
ULH	11.38	1.75	45.31	23.10
UHH	10.81	2.14	44.31	15.92

Note—See text (Method section) for specification of group designations. *Solved anagrams only.

effect on the solution of common word anagrams, high-vocabulary subjects solved more uncommon word anagrams ($p < .05$) than did their low-vocabulary counterparts.

The second analysis of variance was performed on data showing mean median time to solution of solved anagrams only. The reason for working only with solved anagrams is that consideration of all anagrams (solved and unsolved) would have led to a spurious result, in that low-vocabulary subjects, having solved fewer anagrams than their high-vocabulary counterparts, would necessarily have shown very inflated time-to-solution scores. Such an outcome would have defeated the purpose of this analysis, which was to compare solution retrieval times for the various low- and high-vocabulary conditions. The results of this analysis are quite straightforward. It was found that T-L frequency was a significant source of variance [$F(1,120) = 25.14, p < .001$], indicating that uncommon words required considerably more working time to solution than did common word anagrams. Similarly, ATP was highly reliable [$F(1,120) = 10.27, p < .005$], showing that anagrams having a high transitional probability arrangement of letters took longer to solve than anagrams having a low transitional probability arrangement. Interestingly, vocabulary level was not a reliable source of variance [$F(1,120) = 2.69, p > .10$], showing that for solved anagrams only, low- and high-vocabulary subjects required equivalent amounts of time to achieve solution. Finally, although the T-L by ATP interaction only approached a conventional level of reliability [$F(1,120) = 3.03, .05 < p < .10$], the asymmetrical pattern of performance indicated by this interaction was virtually identical to that found for the same source of variance in the analysis of data reflecting number of anagrams correctly solved. It is felt that achieving the same pattern of performance with different dependent variable measures is behaviorally, if not statistically, significant.

Subsidiary Findings

A number of Pearson product-moment correlations were calculated that involved the "crossword puzzle" and "word games (such as anagrams)" items found on the opinion survey that had been conducted during the preexperimental phase of the study. It was found that crossword puzzle preference and vocabulary level were positively related ($r = +.19, p < .05$), as were crossword puzzle preference and number of anagrams solved ($r = +.18, p < .05$). No relationship was found between number of anagrams solved and the "word games" question ($r = +.08, p > .05$). Additionally, a correlation was computed between the subjects' vocabulary test scores and their college grade-point averages. Not unexpectedly, a strong positive relationship between these two variables was found ($r = +.41, p < .001$).

DISCUSSION

Before consideration is given to the principal findings of this

research, two of the less central results deserve comment. First, the finding that a positive correlation obtained between the subjects' reported enjoyment of doing crossword puzzles and anagram solution performance supports a finding originally reported by Hunter (1961). Hunter's data showed "a uniformly positive relation between amount of crossword experience and speed of solving the test anagrams" (1961, p 163). The fact that a reliable positive correlation was not obtained between the preexperimental opinion survey question involving "word games" and subsequent anagram solution performance is somewhat puzzling, but it could be due to the fact that people, in general, are probably less familiar with anagrams than with crossword puzzles, the result being the lack of relationship obtained. In any event, based on Hunter's and these data, knowledge of an individual's level of interest in crossword puzzles should be useful in predicting how well a subject will do in an anagram solution task.

Second, it was found that ATP was an effective variable only with common word anagrams, anagrams having low transitional probability totals being solved more frequently than high transitional probability anagrams. This finding would imply that the difficulty inherent in solving an uncommon word anagram overrides the potential effectiveness of ATP. There has been a history of conflicting findings with regard to the ATP variable; some researchers have reported that high transitional probability levels inhibit anagram solution performance (e.g., Warren & Thomson, 1969), whereas other reports have not supported such a result (e.g., Mendelsohn & O'Brien, 1974). The present data tend to strike a middle ground on this issue of ATP effectiveness and raise the possibility that one reason why it has been difficult to achieve consistent results with this variable is because the effect of ATP on solution performance is not unitary; rather, it depends importantly on the frequency characteristics of the solution words used.

Finally, the principal reason for conducting this study was to try to determine the mediating factor(s) linking the positive relationship found between vocabulary size and anagram solution performance. It seems clear from these data that the reason individuals with relatively large vocabularies solve more anagrams than those with smaller vocabularies is related strictly to the number of words stored in memory, there being no evidence indicating differences of a retrieval efficiency nature between these two types of individuals.

Support for this conclusion comes from two different, but complementary, results found in the analyses of variance cited above: (1) While vocabulary size was unrelated to the solution of common word anagrams, high-vocabulary subjects solved more uncommon word anagrams than did their low-vocabulary counterparts (this finding would tend to support the differential word storage hypothesis). (2) Based on time-to-solution data for solved anagrams only, vocabulary level was not a reliable source of variance, indicating that low- and high-vocabulary subjects solved anagrams in equivalent amounts of time (this finding would tend to refute the differential retrieval efficiency hypothesis). The joint presence of these two findings offers presumptive evidence that storage differences alone account for the vocabulary size effect in anagram solution performance.

REFERENCES

- DUPUY, H. J. *The rationale, development, and standardization of a basic word vocabulary test* (DHEW Publication No. HRA 74-1334). Washington, D.C.: U.S. Government Printing Office, 1974.
- HUNTER, I. M. L. Further studies on anagram solving. *British Journal of Psychology*, 1961, 2, 161-165.
- MAYZNER, M. S., & TRESSELT, M. E. Tables of single letter and digram frequency combinations for various word length and letter position combinations. *Psychonomic Monograph Supplements*, 1965, 1, 13-31.

- MENDELSON, G. A. An hypothesis approach to the solution of anagrams. *Memory & Cognition*, 1976, 4, 637-642.
- MENDELSON, G. A., & COVINGTON, M. V. Interval processes and perceptual factors in verbal problem solving: A study of sex and individual differences in cognition. *Journal of Personality*, 1972, 40, 451-471.
- MENDELSON, G. A., & GRISWOLD, B. B. Assessed creative potential, vocabulary level, and sex as predictors of the use of incidental cues in verbal problem solving. *Journal of Personality and Social Psychology*, 1966, 4, 423-431.
- MENDELSON, G. A., GRISWOLD, B. B., & ANDERSON, M. L. Individual differences in anagram-solving ability. *Psychological Reports*, 1966, 19, 799-809.
- MENDELSON, G. A., & O'BRIEN, A. T. The solution of anagrams: A reexamination of the effects of transition letter probabilities, letter moves, and word frequency on anagram difficulty. *Memory & Cognition*, 1974, 2, 566-574.
- OLSON, R., & SCHWARTZ, R. Single and multiple solution five-letter words. *Psychonomic Monograph Supplements*, 1967, 2, 105-152.
- THORNDIKE, E. L., & LORGE, I. *A teachers word book of 30,000 words*. New York: Teachers College, Columbia University, 1944.
- WARREN, M. W., & THOMSON, W. J. Anagram solution as a function of transition probabilities and solution word frequency. *Psychonomic Science*, 1969, 17, 333-334.

(Received for publication August 28, 1979.)