Models of anagram solution

JOHN T. E. RICHARDSON and PAUL B. JOHNSON
Brunel University, Uxbridge, Middlesex UB8 3PH, England

It is generally accepted that anagrams are solved by means of an interrogation of the lexicon on the basis of an orthographically regular subsequence of the letters given. This study was concerned with the effect upon the solution process of providing the semantic category that contained the solution. In this situation, the instance dominance of the solution in the specified category affected the solution time for anagrams drawn from small categories, but the number of letters in the anagrams did not. For anagrams drawn from large categories, the reverse was true. Moreover, cuing with the appropriate taxonomic category produced a significant reduction in solution time in the case of items from small categories, but not in the case of items from large categories. It is concluded that subjects employ a semantically directed interrogation of the lexicon when they are cued with a small category, but a structurally directed interrogation of the lexicon otherwise.

In an anagram task, the subjects are presented a sequence of letters and are required to produce a word that can be spelled using all and only the letters given. This is a well-defined task in which performance can be carefully measured, and it is of considerable interest to psychologists concerned with problem solving and with the morphological structure of language. In this paper, we discuss a series of models for the solution of anagrams.

The most trivial model that might be proposed for the anagram task suggests that subjects generate all possible permutations of the letters given and examine each in turn as a possible solution. This is well known to be an inadequate account of the solution process (Bourne, Ekstrand, & Dominowski, 1971, pp. 82-94). For example, it predicts that solution time is a positively accelerating function of the number of letters given and that all anagrams of a given length are equally difficult. Both of these predictions have been disproved: Solution time is actually a negatively accelerating function of the number of letters given (Kaplan & Carvellas, 1968), and a variety of structural properties other than anagram length predict solution time (Gilhooly & Johnson, 1978; Mayzner & Tresselt, 1958; Mendelsohn, 1976; Tresselt & Mayzner, 1968). It is now generally accepted that anagram solution is carried out in accordance with the subject's knowledge of the statistical structure of the language. Contemporary accounts suggest that the subject selects an orthographically regular sequence of letters from the set given and uses this as a probe to interrogate the lexicon and to retrieve possible solutions, which are then compared with the

full set of letters contained in the anagram (Bourne & Dominowski, 1972).

The principal question of interest in this paper is whether a similar solution process is undertaken when the subject is given information concerning the semantic category of the solution word (for example, "a flower: PITLU"). It might be supposed that in this situation the subject carries out the operations just described and then checks any possible solution for the relevant semantic feature. It is known that lexical decisions of this sort ("A tulip is a flower") take only a fraction of a second, and so they would not substantially affect the total solution time. Accordingly, this model would predict no effect of providing semantic information upon solution time and no effect of variables that selectively influence lexical decision, such as the number of exemplars of the category (cf. Wilkins, 1971); on the other hand, anagram solution time should continue to be influenced by structural properties of the anagram. This model is inadequate, since a substantial reduction in solution time may be obtained when the semantic category of the solution is given (Safren, 1962; Schuberth, Spoehr, & Haertel, 1979).

One might suppose that the provision of information concerning the solution's lexical category gives rise to a radically different solution process: that the subject interrogates the relevant category and compares exemplars of the category with the letters contained in the anagram. Most exemplars would produce an immediate mismatch, and so the most important component of the total solution time would be that devoted to lexical search. This model can handle the finding that semantic information may reduce the time to produce a solution. It would predict that variables that influence lexical search would affect anagram solution time. On the other hand, it would predict that structural properties of the anagram, such as the number of letters given, would be unimportant. Which stimulus attributes

should affect anagram solution time according to this model will depend upon the assumptions made about the lexical search process. If different subjects search through the instances of a lexical category in a different order, then the size of the category would be the most important factor influencing the average time taken to access a given item. On the other hand, if different subjects tend to search through the instances of a category in the same order, then the size of the category would be of less importance than the position of the critical item in the order of lexical search. In fact, experiments designed to produce category norms have shown that one can reliably order the instances of a lexical category in terms of the frequency with which they are produced and that this is highly correlated with the frequency with which different exemplars are given as the first instance of the category (e.g., Battig & Montague, 1969). One may therefore quantify the instance dominance of a given word with respect to a given lexical category, and one would predict that this factor would be of crucial importance in predicting the time to access a word within that category. This relationship was confirmed by Schuberth et al. (1979). On the other hand, words in different categories of the same instance dominance should be retrieved equally rapidly, irrespective of the size of the various categories.

The following experiment was designed to test these hypotheses by comparing the number of letters in an anagram, category size, and instance dominance as predictors of anagram solution time when subjects are told the semantic category of the solution word.

EXPERIMENT 1

Method

The stimuli for this experiment were selected from the category norms of Battig and Montague (1969). The size of each category was taken as the number of instances given by 10 or more of their sample of 442 subjects. Exemplars were selected from 10 "large" categories (defined as having more than 30 exemplars) and from 10 "small" categories (defined as having less than 26 exemplars). One exemplar of five letters and one exemplar of seven letters were selected from each category, such that both words were rated AA or A in frequency of usage according to Thorndike and Lorge (1944). Anagrams were constructed by randomly ordering the letters in each word, with the constraints that no letter appear in its original position and that obviously helpful sequences of letters be eliminated. The instance dominance of each solution word was defined as its position in the rank ordering of the exemplars of the relevant category in terms of decreasing frequency of emission.

Ten undergraduate students at Brunel University volunteered to participate in this experiment and were tested individually. Instructions were given as to the nature of the anagram task, and the 40 anagrams were presented in a different random order to each subject. They were presented in uppercase type on white cards, which also showed the category name in lowercase type. At the same time, the experimenter read the category name aloud. The subjects were allowed to use pencil and paper, but they did so only rarely. They responded orally and were told if their response was incorrect. The solution time was measured by a stopwatch to the nearest .5 sec. If the correct

response had not been produced within 90 sec, the trial was terminated and that time was recorded. The next trial followed immediately.

Results

The mean instance dominance, the mean percentage failure, and the harmonic mean solution time for each of the four conditions are shown in Table 1. Because of an obvious positive skewness in the raw data, an analysis of variance was carried out upon the reciprocal of each subject's solution time for each anagram, using the appropriate statistical techniques for generalizing across experimental subjects, across stimulus items, and across lexical categories (cf. Clark, 1973; Richardson, 1975). The effect of anagram length was significant [$F(1,18) = 4.86$, $p < .05$], but that of category size was not ($F < 1$), nor was there any interaction between the two effects ($F < 1$). A second analysis of variance carried out upon whether each subject had solved each anagram produced similar findings, but, of course, this was based upon dichotomous data.

Inspection of Table 1 shows that seven-letter anagrams were more difficult than five-letter anagrams, in terms of both the failure rate and the solution time, but the size of the category from which the solution was selected was quite unimportant. However, the analysis thus far has failed to consider the variable of instance dominance. It is reasonable to expect that any effect of instance dominance will be linear in solution time, rather than in the reciprocal of solution time. Therefore, an analysis of covariance was carried out on the harmonic mean solution time for each item, using the two fixed factors defined previously and the covariate of instance dominance. This found that the covariate did not explain a significant amount of the within-class variation in solution time ($r = +.05$; $F < 1$). However, tests of the assumption of homogeneity of within-class regression indicated that it was violated when the classes were distinguished in terms of category size [$F(1,32) = 3.41$, $p < .1$]. Accordingly, separate analyses of covariance were carried out on the stimulus items drawn from large and small categories. For small categories, the covariate of instance dominance explained a significant amount of the within-class variation [$r = +.49$; $F(1,17) = 5.55$, $p < .05$]. When the criterion data were

Table 1
Mean Instance Dominance, Mean Percentage of Failed Trials, and Harmonic Mean Solution Time (in Seconds), by Category Size and Anagram Length

Number of Letters	Dominance	Failure	Solution Time
Large Categories			
5	23.5	8	4.89
7	22.1	22	7.04
Small Categories			
5	9.6	4	3.91
7	18.0	20	7.13

adjusted for the linear trend on the covariate, the effect of anagram length was no longer significant $[F(1,17) = 2.32, p > .1]$. Thus, for these items, the effect of anagram length upon solution time was entirely explained in terms of the confounded variable of instance dominance. The within-class regression coefficient was +.20 sec/item. For large categories, the covariate of instance dominance did not explain a significant amount of the within-class variation $(r = -.12; F < 1)$, and no adjustment in the criterion data was undertaken. Thus, for these items, the effect of anagram length upon solution time could not be explained in terms of the variable of instance dominance.

Discussion

The predictions being evaluated in this experiment were derived from a model of anagram solution that emphasized the importance of semantically based search of the lexicon. This model predicted that the instance dominance of the solution within the specified category would be related to the time to arrive at the solution. On the other hand, it predicted that structural properties of an anagram, such as its length, would not be related to solution time. Moreover, it was expected that the size of the specified category would be unimportant when the instance dominance of the solution was controlled. The model was contrasted with an alternative based upon generally accepted accounts of the way in which subjects solve anagrams when they are not told the semantic category of the solution. This alternative theory predicted that structural properties of the anagram would affect the solution time, but that semantic properties of the solution would not.

In the case of the items taken from small lexical categories, there was a clear linear correlation between instance dominance and solution time (cf. Schuberth et al., 1979). However, there was no effect of anagram length upon performance when the effect of instance dominance was statistically controlled. These results are consistent with the model of anagram solution based upon semantically directed lexical search, but they are inconsistent with the idea that the subject constructs an orthographically regular sequence of letters from the set given and uses this to interrogate the entire lexicon and to retrieve possible solutions on a structural basis. In the case of the items taken from large lexical categories, there was no correlation between instance dominance and solution time, and so the effect of anagram length upon performance could not be explained in terms of the variable of instance dominance. These results are inconsistent with the model of anagram solution based upon semantically directed lexical search, but they are consistent with the theory of structurally directed interrogation of the lexicon.

It seems that two different solution strategies are employed in the anagram task investigated here, conditional upon the subject's estimation of the likelihood of arriving at a solution by means of semantically directed lexical search. If the lexical information indicates that the solution is drawn from a small semantic category, the subject interrogates the relevant category and compares exemplars of the category with the letters provided in the anagram. (It would appear that this process is carried out at the rate of 5 items/sec.) On the other hand, if the lexical information indicates that the solution is drawn from a large semantic category, the subject interrogates the whole lexicon, using as a probe a structurally regular sequence of letters taken from the set given in the anagram.

EXPERIMENT 2

The theoretical account that has just been specified has a direct corollary concerning the quantitative effects of cuing with the lexical category. The account suggests that this information affects the solution process only if the category is relatively small; accordingly, it would be predicted that cued and noncued solution times should be significantly different for exemplars of small categories, but not for exemplars of large categories. A further experiment was carried out, in which the material was selected in a manner similar to that employed in Experiment 1, and in which alternate anagrams presented to each subject were cued and not cued with the relevant lexical category. This found a reduced but still significant effect of cuing in the case of large categories. This might have been because alternating the presence and absence of a cue disrupted the subjects' spontaneous choice of solution strategy, but also because category size had not been adequately varied. The following experiment tested this prediction when the large and small categories were maximally separated, and when the effect of cuing was evaluated between independent groups of subjects.

Method

All of the anagrams contained six letters. Eight words were chosen as exemplars of small, definite categories: a day of the week, "Monday"; a month of the year, "August"; a season of the year, "spring"; a number between 11 and 20, "twelve"; an English king's Christian name, "Edward"; a make of English car, "Jaguar"; a racket game, "tennis"; and a British national newspaper, "Mirror." Eight words were chosen as exemplars of large, indefinite categories: an animate noun, "donkey"; an inanimate noun, "jacket"; a concrete noun, "carpet"; an abstract noun, "sorrow"; a famous person's surname, "Wilson"; a place, "Sweden"; a verb, "behave"; and an adjective, "pretty." Anagrams were constructed from these 16 words as in Experiment 1. Otherwise, no attempt was made to control the structural properties of each anagram, since the critical hypothesis concerned the effects of cuing within each anagram.

Twelve undergraduate students at Brunel University volunteered to participate in this experiment. They were tested individually, and the 16 anagrams were presented in a different random order to each subject. The stimuli were presented in uppercase type through a two-field tachistoscope. Once again, the subjects were allowed to use pencil and paper, but they did so only rarely. By means of a voice key, their oral responses triggered an electronic timer, which recorded the solution times in milliseconds. If the correct response had not been produced within 90 sec, the trial was terminated and that time was recorded. If the subject made an incorrect response, he was informed accordingly, and the stimulus was presented a second time at the end of the other trials. If he still produced an incorrect response, a time of 90 sec was recorded. For six of the subjects, chosen at random, the appropriate lexical category was read out by the experimenter before each anagram was presented for solution.

Results

The harmonic mean solution times for the subjects who were cued with the relevant lexical category were 1.29 sec for the items from small categories and 5.60 sec for the items from large categories. The corresponding times for the subjects who were not cued were 10.42 sec and 21.26 sec, respectively. An analysis of variance was once again carried out upon the reciprocal of each subject's solution time for each anagram, using the appropriate statistical techniques for generalizing across

experimental subjects and across stimulus items. This demonstrated a significant effect of cuing [quasi $F(1,20) = 20.46$, $p < .001$], a significant effect of category size [$F(1,14) = 10.42$, $p < .01$], and a significant interaction between these two factors [$F(1,14) = 19.31$, $p < .001$].

A posteriori tests showed that the effect of cuing was significant for items from small categories [quasi $F(1,33) = 29.39$, $p < .001$], but not for items from large categories [quasi $F(1,33) = 1.30$, $p > .25$]. This confirms the critical prediction stated earlier. The effect of category size was significant when subjects were cued with the relevant lexical category [$F(1,28) = 25.65$, $p < .001$], but not when the category was not provided ($F < 1$).

Discussion

The experiments reported here suggest a model of anagram solution that emphasizes the possibility of more than one solution strategy. In the absence of information concerning the semantic category of the solution word, the subject interrogates the lexicon on the basis of an orthographically regular sequence of letters selected from the set provided by the experimenter. When the relevant lexical category is supplied, however, the subject's choice of solution strategy appears to depend upon the number of exemplars in that category, and thus upon the likelihood of arriving at the solution by means of a search through that category. If the solution is drawn from a relatively small taxonomic category, the subject interrogates that category and compares exemplars of the category with the letters provided in the anagram. As a result, his solution time is reduced and varies with the instance dominance of the solution within the specified category, but apparently not with structural properties of the anagram or the solution. On the other hand, if the solution is drawn from a relatively large taxonomic category, the subject reverts to a structurally directed interrogation of the entire lexicon. His solution time is not significantly reduced, and it varies with structural properties of the anagram, but not with the instance dominance of the solution within the specified category.

This account of the strategies involved in solving anagrams is rather more complicated than the generally accepted model. Nevertheless, it is intuitively obvious that specifying the taxonomic category of an anagram's solution will be of little help to the subject if that category is indefinitely large (for example, "a concrete noun"), and it is perhaps reassuring that the subject's choice of solution strategy appears to take this into account. More generally, the results of the present investigation allow one to appreciate that heuristic methods in problem solving depend both upon the nature of the information provided and upon the relative likelihood of success of the solution strategies available.

REFERENCES

BATTIG, W. F., & MONTAGUE, W. E. Category norms for verbal items in 56 categories. *Journal of Experimental Psychology Monograph,* 1969, **80**(3, Pt. 2).

BOURNE, L. E., JR., & DOMINOWSKI, R. L. Thinking. *Annual Review of Psychology,* 1972, **23**, 105-130.

BOURNE, L. E., JR., EKSTRAND, B. R., & DOMINOWSKI, R. L. *The psychology of thinking.* Englewood Cliffs, N.J: Prentice-Hall, 1971.

CLARK, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12**, 335-359.

GILHOOLY, K. J., & JOHNSON, C. E. Effects of solution word attributes on anagram difficulty: A regression analysis. *Quarterly Journal of Experimental Psychology,* 1978, **30**, 57-70.

KAPLAN, I. T., & CARVELLAS, T. Effect of word length on anagram solution time. *Journal of Verbal Learning and Verbal Behavior,* 1968, **7**, 201-206.

MAYZNER, M. S., & TRESSELT, M. E. Anagram solution times: A function of letter order and word frequency. *Journal of Experimental Psychology,* 1958, **56**, 376-379.

MENDELSOHN, G. A. An hypothesis approach to the solution of anagrams. *Memory & Cognition,* 1976, **4**, 637-642.

RICHARDSON, J. T. E. Statistical analysis of experiments investigating stimulus attributes. *British Journal of Mathematical and Statistical Psychology,* 1975, **28**, 235-236.

SAFREN, M. A. Associations, sets, and the solution of word problems. *Journal of Experimental Psychology,* 1962, **64**, 40-45.

SCHUBERTH, R. E., SPOEHR, K. T., & HAERTEL, R. J. Solving anagrams: Category priming and the differential availability of category solutions. *Quarterly Journal of Experimental Psychology,* 1979, **31**, 599-607.

THORNDIKE, E. L., & LORGE, I. *The teacher's word book of 30,000 words.* New York: Teachers College, 1944.

TRESSELT, M. E., & MAYZNER, M. S. Anagram solution times: A function of single- and double-letter solution words. *Journal of Verbal Learning and Verbal Behavior,* 1968, **7**, 128-132.

WILKINS, A. J. Conjoint frequency, category size, and categorization time. *Journal of Verbal Learning and Verbal Behavior,* 1971, **10**, 382-385.