DOCUMENT RESUME

ED 098 260 TM 004 044

AUTHOR Martuza, Victor R.; Bassett, Denise
TITLE The Effects of Organization and Adjunct Pre-Questions on Performance on Three Types of Test Items Based on the Information in Bar Graph Stimuli.
PUB DATE [Apr 74]
NOTE 17p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, Illinois, April 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS College Students; *Graphs; *Item Analysis; *Organization; *Questioning Techniques; Tables (Data); *Testing; Test Wiseness

ABSTRACT

Fifty-eight undergraduate and graduate student volunteers participated in a factorial experiment designed to test hypotheses about the effects of two levels of information organization and four types of adjunct prequestion treatments on performance on specific amount items and two types of items organized in the same manner as the graphical stimuli used. Significant organization x item type and prequestion x item type interactions suggest that these variables may differentially facilitate performance on the kinds of items used in this experiment. Potential implications for the educational use of graphical displays and suggestions for future research are presented and discussed. (Author)

The effects of organization and adjunct pre-questions on performance on three types of test items based on the information in bar graph stimuli

Victor R. Martuza Denise Bassett

University of Delaware

Fifty-eight undergraduate and graduate student volunteers participated in a factorial experiment designed to test hypotheses about the effects of two levels of information organization and four types of adjunct pre-question treatments on performance on specific amount items and two types of items organized in the same manner as the graphical stimuli used. Significant organization x item type and pre-question x item type interactions suggest that these variables may differentially facilitate performance on the kinds of items used in this experiment. Potential implications for the educational use of graphical displays and suggestions for future research are presented and discussed.

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRO
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRE
SENT OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

D R A F T

TM 004 044

Recent research has shown that the organization of information in a prose passage (e.g., Frase, 1969; Friedman & Grietzer, 1972; Meyers et al, 1972) and the use of adjunct aids, particularly the use of pre- and post-questions embedded in the passage (e.g., Rothkopf, 1966; Frase, 1968) are two variables which significantly affect the acquisition of information presented in prose form. While prose is the primary printed medium for presenting many kinds of information, graphs and tables are frequently used where quantitative information is involved. Yet, very few studies can be found in the literature which deal with the acquisition of information from these kinds of displays. The purpose of this study was to see whether these two variables, which have such a marked influence on prose learning, play a similar role in learning from graphical displays. The only published study dealing directly with this question was reported by Washburne (1927).

Washburne examined the effects of two types of graph organization on performance on three types of test items: (a) specific amount items, requiring recall of the price of a particular product at a particular point in time, (b) static comparison items, requiring comparison of the prices of two or more products at the same point in time, and (c) dynamic comparison items, requiring the comparison of price trends or the relative fluctuation in the prices of two or more products over a specific time interval. For convenience, we will refer to Washburne's two types of graph organization as static and dynamic organization. In bar graphs having a static organization, the bars depicting the prices of the different products are clustered together for each time point included in the display. In bar graphs having a dynamic organization, all bars depicting the prices of the same product at different times are clustered together so that the price trend per product across successive years is perceptually salient. Washburne's

data suggest that these two types of organization do not differentially affect performance on specific amount and dynamic comparison items, but static organization seems to facilitate performance on static comparison items.

The failure of dynamic organization to facilitate performance on dynamic comparison items is somewhat puzzling. If subjects encode bar graph information as it appears in the display, the spatial proximity of the bars as well as the gestalt of each cluster in a graph having dynamic organization should facilitate performance on items dealing with trend information as compared to a graph having static organization. Perhaps the explanation for Washburne's result with dynamic comparison items lies in the relative difficulty of the items used. Price, Martuza, and Crouse (in press) have shown that item difficulty is, to a large extent, dependent on the number of data points which must be recalled to answer an item correctly. To control for this possible confound in the present study, two-point static comparison and two-point trend items were used instead of the static and dynamic comparison items employed by Washburne. Assuming subjects encode the information as it is organized in the display, the static organization graph was expected to facilitate performance on the two-point static comparison items while the dynamic organization graph was expected to facilitate performance on two-point trend items. Since each specific amount item depends on the recall of just one data point and because the graphs used in the present study, regardless of organization, did not differ appreciably with respect to either the number of clusters of data points (bars) to be encoded or the number of data points per cluster, the two levels of organization were not expected to differentially affect performance on specific amount items.

In the same study, Washburne included several treatment conditions in which graphical and tabular displays were accompanied by a set of static comparison study questions. When used with a tabular display, these static comparison adjunct questions seemed to facilitate performance on static comparison test items, interfere with performance on dynamic comparison items, and exercise no effect on specific amount items. In conditions where static comparison adjunct questions were used with a bar graph or pictograph having a static organization, no additional increment in static comparison criterion test performance was obtained, suggesting that adjunct questions and organization may influence the encoding process in a similar way. The effects of static comparison adjunct questions on the acquisition of information from graphs with dynamic organization and the effects of other types of adjunct questions (e.g., dynamic comparison, specific amount) on information acquisition from other types of visual displays have not been studied to date. Thus, the second purpose of this study was to extend Washburne's findings by examining the effects of four adjunct prequestion conditions -- (1) two-point static comparison items only, (2) two-point trend questions only, (3) a combination of two-point static comparison and trend items, and (4) irrelevant items -- on acquisition of information from bar graphs having static and dynamic organization. Since the adjunct question literature suggests that these adjunct question conditions should affect performance somewhat differently on the three item types employed, a treatment by item type interaction was expected.

Subjects. Sixty-eight undergraduate education students who had volunteered for the experiment were randomly assigned in nearly equal numbers to 16 treatment conditions. Ten Ss were excluded prior to data analysis for failure to follow directions or to correctly answer the adjunct questions based on the sample graphs in their booklets. As a result, the analysis

is based on a sample of 58 Ss, with cell ns ranging from one to five.

Materials. All materials presented to Ss were organized into a booklet. Directions, a sample graph, and a set of adjunct questions appeared on the first page, the test graph appeared on the next page, and the retention test followed.

Four test graphs, depicting the prices of three fibers in four different years were constructed in the following manner: A data matrix of 12 cells (3 fibers by 4 years) was constructed. Cell entries were obtained from a table of random numbers. With these values, one graph was constructed in which the bars were grouped according to fiber (dynamic organization) and one graph was constructed in which the bars were grouped according to year (static organization). The values in the data matrix were then randomly reordered and two parallel form graphs were constructed using these values; one having static organization, the other having dynamic organization. The purpose for using these parallel forms was to increase the generablizability of results and to minimize idiosyncratic task-specific effects.

The organization of the sample graph in each booklet was the same as the organization of the test graph for that booklet. The sample graphs depicted the prices of two grains in each of three years and were constructed using the procedure described above for the test graphs.

Each sample graph was accompanied by a four-item set of adjunct pre-questions which defined the four levels of the adjunct pre-question independent variable. The four levels were: (1) static comparison questions (requiring comparison of the prices associated with two products in one year) e.g.,:

In 1910 the difference between the price of corn and the price of wheat was:
a) 1 cent per pound or less
b) 2 or 3 cents per pound
c) 4 or 5 cents per pound
d) 6 or more cents per pound

(2) trend questions (requiring comparison of the prices associated with one product in two years) e.g.:

From 1910 to 1920 the difference in the price of corn was:
a) 1 cent per pound or less
b) 2 or 3 cents per pound
c) 4 or 5 cents per pound
d) 6 or more cents per pound

(3) combination questions (two static comparison questions and two trend questions) and (4) irrelevant questions (questions unrelated to graph content) e.g.:

When I see graphs in texts, I _____ study them carefully.
a) almost always
b) usually
c) seldom

For both static comparison and trend questions two types of items were constructed: One type measured interval-level processing of the stimulus data and one type measured ordinal-level processing. Interval questions were of the type presented above, while ordinal questions were of the following type:

From 1900 to 1904 the price of corn:
a) increased
b) stayed the same
c) decreased

Directions accompanying the first three sets of adjunct questions stressed that the adjunct questions were like the test items and encouraged Ss to study the test graph so as to maximize performance of these types of questions. Ss who received irrelevant questions thus had no knowledge of type of test item to expect.

The dependent variable was a retention test consisting of one subtest of 18 interval items and one subtest of 18 ordinal items. Within each subtest equal numbers of three kinds of questions were included: (1) static comparison questions, (2) trend questions, and (3) specific amount questions.

The latter had the following form:

In 1100 the price of linen was

a) 1 cent or less per cunce
b) 2 Or 3 cents per ounce
c) 4 or 5 cents per ounce
d) 6 or 7 cents per ounce
e) 8 or 9 cents per ounce
f) 10 or more cents per ounce

In one-half of the booklets, the interval-item subtest appeared first on the retention test; in the other half, the ordinal-item subtest appeared first.

Procedure. Each subject was given a booklet and told to read the directions and answer the adjunct questions provided. After answering these questions Ss were given three minutes to study the test graph, after which they were given whatever time they needed to complete the test without referring to the expefimental graph. Upon completing the test, Ss were given verbal directions to write a description of the strategies they used to study the information in the experimental graph.

Design. Two types of graph organization (static, dynamic); four types of adjunct questions (static comparison, trend, combination, irrelevant); two parallel graph forms (I, II) and three types of test questions (static comparison, trend, and specific amount) resulted in a 2 x 4 x 2 x 3 factorial design with repeated measures on the last factor. Separate mixed model analyses of variance were performed on the data obtained from each subtest (i.e., the interval-level and ordinal-level subtests).

Results and Discussion

At the outset, it is important to note that 62 % of the subjects, regardless of treatment group membership, reported using a point learning strategy while studying the graphical stimuli used in this experiment. That is, most subjects tended to memorize lists of data point values and the associated data point labels for use in answering the subsequent criterion test questions. Evidently, most subjects perceived this point information memorization strategy as the most efficient and effective procedure for encoding this type of

information. With this in mind, consider the results of the interval item subtest analysis.

The only significant result observed in this analysis was the main effect of item type. The Newman-Keuls procedure was used to make all possible pair-wise comparisons. The results (Table I) showed that mean performance on the specific amount items significantly exceeded mean performance on each of the remaining two item types. Since both the static comparison and trend items require the recall of two data points, while the specific amount items are based

Table I about here

on just one data point, and because most Ss reported a point memorization learning strategy, this finding seems quite reasonable and is consistent with findings reported by Martuza and Wolfe (1973) and Martuza, Price, and Crouse (in press).

All other interval level results were non-significant, probably because of the difficulty of the task. Although the cell means (Table I) indicate above chance performance in all treatment conditions, the magnitudes of these means suggest that the Ss experienced a great deal of difficulty with the task, perhaps because of the three minute limit on study time employed here. If this explanation is correct, the expected effects should materialize on the ordinal level subtest because ordinal level items measure a lower level of acquisition. Let us now consider these results.

None of the ordinal level main effects were significant; however, the organization by item type, pre-question by item type, and graph form by item type interactions were significant.

Turning first to the organization X item type interaction, the cell means in Table II indicate that performance on the specific amount and

Table II about here

trend items is not differentially affected by organization, while static organization seems to facilitate performance and dynamic organization seems to interfere with performance on static comparison items ($t = 2.30$; within cells $df = 126$; $p < .05$). These data are quite consistent with Washburne's and, taken together, suggest the possible existence of a trend information learning set. Because of past experience with graphical displays, Ss may have formed a learning set which influences their encoding of point information in a way that facilitates subsequent retrieval for answering trend questions. Thus, when presented with dynamic organization, Ss are oriented both by learning set and by organization towards trend information; whereas, when presented with static organization, Ss are oriented by a learning set towards trend information but are oriented by organization to static comparison information as well. The net result is equivalent trend item performance under both conditions, but increased static comparison performance in the static organization condition. One must keep in mind that this explanation must still be regarded as speculative at this time; further study is required before firm conclusions about this phenomenon can be reached. It remains to be seen whether similar findings will result with: (a) increased study time, (b) bar graphs based on greater amounts of information, (c) bar graphs in which different organizations result in markedly different numbers or sizes of the clusters of bars, or (d) other types of graphical displays. However, under conditions approximately like those employed in Washburne's and our study, static organization of bar graph information is preferred because it results in performance equal to or better than that obtained using dynamic organization on all item types used so far.

A second important effect was the significant pre-question X item type interaction. First, the correlation between the specific amount

item means and the means of the two-point items (i.e., the means of the static comparison and trend item performances), regarding the pre-question treatment conditions as the sampling units, is significant ($p = 0.04$, permutation test) and provides additional evidence of the dependence of two-point item performance on specific amount learning. Second, mean specific amount performance in the static comparison pre-question group was significantly lower than the unweighted mean of the specific amount item means obtained in the other three pre-question treatment conditions (Bonferroni $t = 2.64$, within cells, $df = 126$, $p < .05$). An examination of the post-experiment inquiry indicated that the proportion of Ss employing a point-learning strategy did not differ significantly among the four groups in the experiment; thus, it appears that the Ss in the static comparison pre-question group simply did a poor job of learning the point information (See Table III). Whether this phenomenon can be attributed to intelligence or some other pertinent variable is not clear since data bearing on this question were not available.

Regardless of the reason for this anomaly, it does affect, in an important way, the effect of adjunct pre-question group performance on the static comparison and trend items contained in the criterion test. In the trend pre-question treatment group, where specific amount learning was relatively high, the expected result occurred -- i.e., the trend item mean was significantly higher than the static comparison item mean ($t = 2.78$; within cells, $df = 84$; $p < .05$). However, in the static pre-question condition where specific amount learning was very low, mean performance on static comparison criterion questions did not significantly exceed mean performance on the trend questions. The most plausible explanation for this nonsignificant result is the inadequate specific amount learning in this group.

The irrelevant pre-question group exhibited the same mean performance on both static comparison and trend items, as expected. An important implication of this result is that the static comparison and trend items are equally difficult, thus ruling out differential item difficulty (i.e., static comparison vs. trend item types) as a rival hypothesis for the organization results discussed earlier.

The combination pre-question treatment group produced the most surprising result, i.e., mean performance on static comparison items significantly exceeded mean performance on the trend items (Bonferroni $t = 2.81$, $df = 84$, $p < .05$). It was expected that the orienting effects of a combination of static comparison and trend prequestions would facilitate performance equally on both static comparison and trend test items; thus, the low trend item mean was unexpected. There are two possible explanations for this result: (a) the orienting effects of static comparison questions may have been greater because these items appeared last in all booklets for this group and were thus the last items encountered before the test graph was presented, and (b) when told to maximize performance on both trend and static comparison questions, Ss tended to perceive the static comparison questions as being dominant. Given the trend information learning set suggested above, static comparison pre-questions would have been somewhat incompatible with the learning set and hence may have "stood out;" thus exerting more influence on subject inspection of the test graph than the trend questions. These explanations are speculative at this point but can easily be tested in subsequent studies.

The third significant effect, item type X graph form interactions ($F = 6.87$; $df = 2,84$; $p < .01$) was not of primary concern in the present study. The two graph forms were employed only to insure greater generalizability of the results. Further analyses of the data will be carried out in the

future to determine the relationship between graphical configuration and performance on these three item types, but since this is not of concern in the present study, this interaction is not discussed further here.

In sum, the data suggest that subject to the demand characteristics of the present study, i.e., three-minute study time, three X four data matrices, etc., Ss tend to study specific point information regardless of the organization of the graph they are instructed to study or the type of adjunct pre-question used. However, performance on the two-point items employed here seemed to be influenced both by graph organization and by type of pre-question employed. The data suggest that organization is most important with respect to performance on static comparison items. Pre-questions have an effect; but the exact nature of this effect is not made clear by these data. Performance on trend questions seems to be facilitated by pre-questions; however, these data do not allow the same claim to be made for the static comparison pre-questions, principally because of the relatively poor overall performance of the group receiving these questions. Additionally, the effect of using several types of pre-questions simultaneously is in need of further study.

Conclusions

The results strongly suggest that further investigation should be carried out with respect to the organization and adjunct question variables since both may play an important role in shaping the acquisition of quantitative information from bar graphs. The end results of studies of this sort should provide clearer guidelines for curriculum builders and teachers interested in facilitating specific kinds of acquisition of information from graphical displays included in instructional materials. T limitations of this study do not permit the formulation of hard and

fast rules governing choice of organization and the optimal use of adjunct questions. However, the results are encouraging since they suggest that a formulation of such guidelines is within reach and could be developed within a reasonably short period of time.

Table I

Results of the Neuman-Keuls test on Differences among Means of Three Types of INTERVAL LEVEL ITEMS

		$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$
$\bar{X}_1$ (trend)	= 2.50		.14	.90*
$\bar{X}_2$ (static comparison)	= 2.64			.76*
$\bar{X}_3$ (specific amount)	= 3.40			

Table II

Mean Performance on Three Types of Ordinal Level Items under Two Types of Graph Organization

Organization / Item type	Dynamic (n=28)	Static (n=30)	Overall Mean
Trend	4.11	4.00	4.05
Static comparison	3.64	4.53	4.10
Specific amount	4.36	4.30	4.33
Overall mean	4.04	4.28	4.16

Table III

Mean Performance on Three Types of Ordinal Level Items under Four Adjunct Pre-question Conditions

Pre-question type / Item type	Trend (n=15)	Static Comparison (n=11)	Combination (n=15)	Irrelevant (n=17)
Trend	4.93	3.18	3.60	4.24
Static comparison	3.87	3.45	4.67	4.24
Specific amount	4.80	3.09	4.33	4.71
	4.53	3.24	4.20	4.40

References

Frase, L. T. Effects of question location, pacing, and mode upon retention of prose material. Journal of Educational Psychology, 1968, 59, 244-249.

Frase, L. T. Paragraph organization of written materials: The influence of conceptual clustering upon the level and organization of recall. Journal of Educational Psychology, 1969, 60, 396-401.

Friedman, M. P. & Grietzer, F. L. Organization and study time in learning from reading. Journal of Educational Psychology, 1972, 63, 609-616.

Martuza, V. R. & Wolfe, M. L. Effects of the amount of information, organization of information, and study time on the processing of quantitative information presented in graphical form. Presented at the American Educational Research Association Annual Meeting, 1973.

Myers, J. L., Pezdek, K. & Coulson, D. Effects of prose organization upon free recall. Cognitive Processes Laboratory Report 72-1, University of Massachusetts, 1972.

Price, J. R., Martuza, V. R. & Crouse, J. H. Construct validity of test items measuring acquisition of information from line graphs. Journal of Educational Psychology, in press.

Rothkopf, E. Z. Learning from written instructive material: An exploration of the control of inspection behavior by test-like events. American Educational Research Journal, 1966, 3, 241-249.

Washburne, J. N. An experimental study of various graphic, tabular, and textual methods of presenting quantitative material. Journal of Educational Psychology, 1927, 18, 361-376, 465-476.