DOCUMENT RESUME

ED 238 683 SE 043 544

AUTHOR	Simcox, William A.
TITLE	Memorial Consequences of Display Coding.
INSTITUTION	Consulting Statisticians, Inc., Wellesley, MA.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	Jan 83
CONTRACT	400-79-0066
NOTE	30p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Adults; *Cognitive Processes; *Graphs; Mathematics; *Memory; Perception Tests; *Performance Factors; Psychological Studies; *Retention (Psychology); *Schemata (Cognition); Semiotics; Visual Learning; Visual Measures; Visual Perception

ABSTRACT

This investigation of cognitive effort begins with a review of cognitive principles relating to memory and graphic encoding of information. The cognitive framework described allows explanation of improved memory due to reparsing of graphic representations. It also provides the basis for optimizing graphical material design to insure best retention of encoded material. An experiment is described using a 4 (graph type) by 3 (question) within-subjects design, each subject (N=10) participating under all conditions. Thirty-six graphs were displayed completely at random. Subjects were informed that the purpose was to test perception of certain display properties and told to respond as quickly as possible without making errors. A graph was shown and subjects executed the response. The experimenter recorded latency and noted errors. Tests were then administered, one to recall a missing property, the other simple recognition of graphs rated on a confidence scale. The results support the hypothesis that incongruency between information to be communicated and its graphic representation leads to reparsing and better memory. A list of eight references, three figures, and five tables are attached. (JM)

ED238683

MEMORIAL CONSEQUENCES OF DISPLAY CODING

NIE 400-79-0066

Year 03

January 5, 1983

Submitted to:

The National Institute of Education
Contracts and Grants Management Division
1200 19th Street, N.W., Room 808
Washington, D.C. 20208

ATTN: Martin Engel, Ph.D.

Submitted by:

William A. Simcox
Senior Engineer
Consulting Statisticians, Inc.
20 William Street
Wellesley Hills, MA 02181

TE043544

MEMORIAL CONSEQUENCES OF DISPLAY CODING

Introduction

Situations frequently arise where demands are placed upon our memory for information previously encountered in graphic displays. These displays are commonly found in textbooks & technical documents, dial, meters, and computer displays. If the retention of the information from these sources is inadequate for the situation, processing and performance will suffer. For example, in a text or technical document, information from a graph seen earlier must be recalled at a later point and combined with newly acquired information if the main points of an argument are to be understood. It would be exceedingly difficult for the reader to form a coherent understanding of the argument if the relevent information could not be recalled since this would require going back and forth between pages. As another example, in process control situations, an operator may be required to integrate recently aquired system status information with current status information in order to respond decisively, especially if an off-normal condition has occured. If this recently acquired information has not been stored sufficiently, the appropriate decision will be delayed.

Situations may also occur where the retention of information is undesireable. For example, Hopkin (1980) points out that in the air traffic control (ATC) environment, much of the information, once used, is no longer needed by the decision process. This previous information would then constitute unwanted noise if persistently recalled. In a high stress environment such as the ATC, the presence of noise can have deleterious effects.

These examples point out the need in the display design process to identify user memory requirements imposed by the task early in the process and to take appropriate steps to ensure that the requirements are sastisfied. One such step would be to choose the display format to aid or hinder the retention process. Recent theories of memory (cf. Craik & Jacoby, 1979; Craik & Tulving, 1975; Winograd, 1978) have suggested that memory for stimuli is a function of the amount of perceptual analysis given the stimuli, the more extensive the analysis, the better the memory. With regard to graphic displays, Simcox (1982) has recently shown that the extent of perceptual analysis given the display depends on how the task relevent information is coded. Thus it appears that on the surface, a link exists between display coding and memory and the objective of the investigation is to verify this link empirically.

The connection between display coding and memory is provided by the amount of cognitive effort required to perform the perceptual analysis of the input. Cognitive effort is the amount of processing resources of the limited-capacity control processor allocated to an information-processing task in order to satisfy the demands of the task (Kahneman, 1973).

Memory and Cognitive Effort

The concept of cognitive effort was proposed by Tyler, Hertel, McCallom & Ellis (1979) to account for performance differences in a word-recall test. They had subjects perform both a semantic (sentence-completion) and nonsemantic (anagram) task in an incidental learning paradigm and later tested each subject in a free recall task. Within each task type, subjects were given a low-effort and high-effort condition. For example, the word doctor was transformed into the anagrams doctoc (low-effort) and croodt (high-effort). In the sentence completion task, the missing word in a low-effort sentence was judged to be almost redundant with other information in the sentence while for high-effort sentences, the word was less determined by other information. For example, given the word dream, a low-effort sentence would be: The girl was awakened by her frightening ________. A high-effort sentence would be the following: The man was alarmed by the frightening ________. Further, an independent measure of the level of effort required was provided by reaction time to an auditory probe (cf. Pew, 1979 a discussion of secondary task methodology). Tyler et. al. (1979) found that for each task, recall was reliably better in the high-effort condition than the low-effort condition. Additionally, recall performance was mirrored by probe reaction time data.

Cognitive effort can also be used to account for performance differences in recognition tasks. For example, Kolers (1975) had subjects recognize inverted and normal oriented sentences. Two groups of subjects were used. One group of subjects were relatively unskilled at reading the inverted typography. The other group of subjects were highly skilled in reading the inverted typography, so much so that their reading times were on par with times to read normal text. Both groups of subjects read decks of normal and inverted sentences and then were required to recognize these sentences amongst a set of distractors. He found that with respect to the unskilled subjects, sentences in inverted typography were more correctly recognized then sentences read in the normal typography. However, the outcome was quite different for the highly skilled subjects. There was no difference in recognition performance between inverted and normal typography. Thus, the extended practice at reading inverted text reduced its advantage to memory. Kolers (1975) argues that extended practice allowed automatic recognition procedures to develop which are relatively effortless (Shiffrin & Schneider, 1977), whereas such procedures were not possible with the unskilled subjects. These subjects instead had to resort to controlled processing strategy requiring greater processing resources which in turn increased memory performance.

Cognitive Effort and the Graphic Coding of Information

In a recent information processing account of how the choice of display formats can affect the ease or difficulty of processing information in graphs, Simcox (1982) describes the interaction in terms of the amount and nature of cognitive effort exerted in extracting the relevent information. When a graphic display is encountered, it must be perceptually represented by its primary encoding features. It is assumed that this representation is activated by the display properties themselves. The identification of the display will be accomplished by comparing this perceptual representation to an activated memory representation or graph schema set by the contextual situation. This identification then makes available the relevent information by a simple look up its value in the corresponding portion of the activated schema.

Suppose the context induces an expectancy or schema that is incongruent with the perceptual representation. The output of the comparison process then indicates a mismatch, implying that the relevent information is not in the perceptual representation. The schema then directs a reparsing of this representation in order to obtain the appropriate information. This means that we must look at the graph differently then we might be accustomed to. The reparsing can take on many forms for example, such as a search of the sensory buffer for the relevent information or a computation of the relevent information from the activated representation. But whatever form it takes, this reparsing requires additional processing resources or cognitive effort.

For example, suppose we are given the four basic graphic displays shown in Figure 1. The first display is defined by properties of pairs of lines on two L-shaped frameworks; the second display is defined by properties of pairs of lines on a single L-shaped framework; the third and fourth displays are defined by properties of pairs of unconnected and connected points respectively on L-shaped frameworks. Formally, we can define these properties as the dimensions of position and height of each point of the pair. However, the heights of both point are not the only properties that can perceptually define the four displays. For instance, as we move from left to right across the displays, from display 1A to display 1D, the defining properties intuitively seem to shift from these component dimensions to dimensions reflecting relationships between the pair of points, culminating in slope and overall height of the pair.

INSERT FIGURE 1 HERE

Using a speeded classification task in one experiment, Simcox (1982) had subjects classify stimuli of the form shown in Figure 1 on the basis of one or the other of the component dimensions, the individual heights, or one of the dimensions defined by the pair of objects at a perceptual unit, i.e. slope and overall height. He found that as the degree of incongruency between response basis (e.g. slope) and display format (e.g. lines on two frameworks) increased, classification time increased. This increased classification time was taken to reflect increased resources needed to reparse the data-driven representation in order to respond appropriately.

In another experiment, scale values were included in the displays of Figure 1, and subjects were asked conceptual questions about the information contained in these displays. The information referred to either trend, average level, or component values of a pair of variables. Latency of response was measured. The results showed a monotonic increase in latencies for a particular display type as the incongruency between response basis and display format increased. Since all displays were equated for legibility and type of response, the monotonicity in latencies observed was attributed to the increased effort needed to reparse the perceptual representation in order to extract the relevent information.

Memory and the Graphic Coding of Information

The studies of Tyler et al (1979) and Simcox (1982) suggest that cognitive effort is the mediating variable between graphic information coding and the retention of this information. However, there has been no systematic resea to date linking the two through a manipulation of effort. There has been a small amount of research on the retention of displayed information (Vernon, 1946, 1952; Washburn, 1927), but these studies contrast graphs with other media (primarily tables and prose paragraphs), are methodologically flawed, and confound the different cognitive components to such an extent that they are of little use.

Therefore, the purpose of this investigation was to test the hypothesis that an increase in cognitive effort necessitated by the type of display coding used will lead to better memory performance for the coding dimension of the display. Sets of displays of the types shown in Figure 1 were used. Using an incidental learning situation, subjects were initially asked questions about some conceptual property of the display. This property will be information concerning the trend, average level and component value of a variable Y. Subjects were then shown graphs of one of the four types from which to extract the information. Response latencies are measured. If the conceptual question induces an expectancy or representation other than that defined by the primary encoding features (i.e. the default visual representation), increased latency to the question answering task can be expected since a reparsing would be called for. Further, the more reparsing necessary, the more effort allocated, the greater the response time. Thus, for the displays shown in Figure 1, a monotonic increase in response latencies would be expected as we move from lines on two frameworks to connected points when responding in terms of a component value question for reasons outlined earlier. Similarly, the opposite expectation results when responding on the basis of either trend or average level. That is, a monotonic increase would result as we move from connected points to lines on two frameworks.

Upon completion of the task, subjects were given a retention test on the displays. If latency is an indicant of processing effort and if increased effort results in a greater memory advantage, then retention performance should exhibit the same monotonic relationship as the latency data for the graph types and questions.

METHOD

Subjects

Ten CSI/Datacrown coworkers served as subjects in this experiment. None of the ten subjects had performed in the first experiment.

Stimuli

The stimuli were sets of slides of displays representing each of the graph types, one display on each slide. For each of the four graph types (lines on two frameworks, lines on one framework, unconnected points, and connected points), sets of nine different displays were chosen. This results in a total of thirty-six different displays. Each display communicates two pairs of values between a dependent variable Y and an independent variable X. The variable Y ranged in value from 0 to 6, while the variable X took on the values 1 and 2. Within each graph type, twenty-five different displays were generated from the orthogonal combination of the set $\{2,3,4,5,6\}$ of Y values at each value of X in the set $\{1,2\}$. The twenty-five different combinations of values are shown in Table 1.

INSERT TABLE 1 HERE

Note from the table that ten displays have an increasing trend, ten displays have a decreasing trend, and five displays have an unchanging trend. Also, considering the displays in terms of the component properties, five displays each have a numeric value 2 thru 6 at X=1, and five displays each have values 2 thru 6 at X=2. Finally, note that with respect to the average value or level of the two component Y values, six displays have an average level equal to or less than 3, which is called low and six have an average level greater than or equal to 5, which is high. The remaining thirteen displays have an average level between 3 and 5, which is medium.

Thus, it is possible to use these twenty-five graphs to convey three different types of information, trend, absolute value, and average level. Each of these types of information can in turn be described by either three sets of linguistic values (increasing, decreasing, unchanging; low, medium, high), or by six numeric values (2, 3, 4, 5, 6). The nine displays representing each graph type were then chosen from each of these sets of twenty-five.

The slides were then projected on a screen located 250cm from the subject. At this viewing distance, the distance betwen the two points varied between 0.86 and 1.70 degrees of visual angle for the single framework graphs and between 2.2 and 3.2 degrees of visual angle for the two framework graphs.

Design

A 4(Graph Type) X 3(Question) within-subjects design was used, each subject participating under all conditions. Additionaly, the order of the thirty-six graphs was completely randomized for each of the ten subjects.

Three graphs each were assigned to each of the three question types for a total of nine graphs. Within each graph type, the nine graphs were randomly chosen subject to the following constraints. One graph of each trend value (increasing, decreasing, unchanging) was assigned to each question type. For trend and level questions, an additional constraint is that these trend values must occur with different level values (i.e. low, medium, high). For example, with respect to the trend question, one graph had a negative trend and a medium level, another graph had a positive trend and low level, and the third graph had an unchanging trend and a high level. Also within each graph type, the component values generating the graphs were unique. The same was true with respect to question type. For the value question, the only constraint was that each graph had to have a unique value.

Procedure

Subjects were seated at the appropriate distance from the screen and handed a booklet containing thirty-six questions. They were informed that the purpose of the experiment was to test their perception of certain display properties and were told to respond as quickly as possible but without making errors. The questions were of the form "The trend in Y is?", "The average level of Y is?", and "The value of Y at X=(1;2) is?". Subjects studied each question in turn and when they felt comfortable with the question and response categories, they signalled the experimenter. A graph of one of the four display types was then shown and the subject executed the appropriate response.

Time from onset of graph to response was recorded to the nearest millisecond using a Model 54035 (Lafayette Instrument Co.) clock/counter. A voice activated key was used to measure the latency of this response. Its activation stopped the response clock/counter and closed the shutter to terminate the stimulus presentation. The experimenter recorded the latency and noted any errors. Subjects were not informed of any times or errors, but all subjects seemed to know when they made an error. At the halfway mark of each session a brief rest period was allowed. Prior to each session, a practice trial comprising 18 questions and graphs was performed.

Upon completion of this part of the experiment, subjects were given a test booklet containing thirty-six frameworks, nine of which were multiple (i.e. two). Each framework contained the property of the display that was used to extract the answer. The subject was then a d to recall the other property. For example, a display representing pairs of lines on a single framework would be shown with a line of height 6 at X=2 corresponding to the prior question "The value of Y at X=2?" of which the answer was 6. The subject's task was then to recall the value at X=1. Similarly, a subject could be given the positive, negative, or unchanging slope and asked to recall the level or vice versa. A correct hit was scored for either the trend or level variable if the correct linguistic value was given, regardless of whether or not it was exact. That is, for a particular graph, if a high level occured with an increasing trend, then the recall of any increasing trend whether or not it was the actual trend seen was scored as a hit.

The displays were drawn such that at a reading distance of about 28 cm, they subtended the same visual angle as the displays seen on the slides. This was to ensure that all the relevent perceptual cues were operative.

After the recall tests, subjects are handed another booklet containing seventy-two graphs and told that the thirty-six graphs they had just seen were part of this set. They were instructed to indicate which of the seventy-two they had seen and to give a confidence rating defined on a 5-point scale of their answer. The scale was as follows: 1) certain; 2) highly confident; 3) confident; 4) somewhat confident; 5) not confident at all.[1] In this task, hits were scored only for a correct match.

---

[1]For purposes of further analysis, these ratings were collapsed on a three point scale with a 1, 2, and 3 rating mapped into 3, a 4 rating mapped into 2, and the 5 rating mapped into 1.

Results and Discussion

Response Times

The mean response tim averaged over subjects for question type and graph type are shown in Table 2 and can be summarized as follows: a) Level questions resulted in longer response times than did Trend or Value questions, and b) The pattern of results for a specific question seems to depend on the graph types chosen to represent the information. The first, but not the second statement is supported by an analysis of variance (ANOVA) on the response times (log transformed) in which Question Type was a reliable source of variation, $F(2,18) = 46.86$, $p < .001$, but the interaction between Question and Graph was not, $F(6,54) = 1.44$, $p < .21$.[2]

INSERT TABLE 2 HERE

The finding that level questions took longer than either trend or value questions was not surprising since a response to level questions involved a mental interpolation followed by a mapping to the Y axis and then applying the decision rule.[3] In all likelihood, the increased time reflects increased resources necessary to invoke the decision rule and not increased resources

---

[2]The absence of a Question X Graph interaction was not unexpected because of power considerations. The primary emphasis was in detecting differences in retention data, not in the time data. Thus, sample size was chosen on the basis of critical retention differences. If we want sufficient power to detect the differences shown in Table 2, sample size would have to be about twice the size used here.

[3]The decision rule was the following:

$$\text{Level} = \begin{cases} \text{Low} & Y \leqslant 3 \\ \text{Medium} & \underline{3}<Y<5 \\ \text{High} & \overline{Y}<5 \end{cases}$$

necessary to reparse the representation. Thus, we would not expect to see any overall advantage of this question type on memory performance. The advantage will only be seen within a question type to the extent that resources are allocated to reparse the perceptual representation, and that will depend on the incongruency between graph and question.

Response times as a function of graph type, for each of three types of questions are plotted in Figure 2A. While these are not reliably different in this situation, their values are indicative of the monotonicity we would expect as the degree of incongruency between question and graph increased. The dashed lines, drawn by visual inspection, emphasize the monotonic relationships suggested by the interaction. Thus, we would expect the retention data to follow these trends.

Recognition Response

The probabilities of correct recognition for the two experimental conditions are shown in Table 3 and also plotted in Figure 2B for illustrative purposes. These results are summarized as follows: a) It appears that no particular question type results in a greater memory advantage; and b) It does appear that recognition advantage is dependent on the specific question and graph type. Again, both observations are supported by analysis of variance (ANOVA) on the transformed (square root arc sine) scores. Main effects of question type were not reliable, $F(2,54) = 1.54$, $p < .24$, but a reliable interaction between question and graph type was indicated, $F(6,54) = 3.14$, $p < .01$.

One deviant point in need of explanation is apparent in Table 3. For the trend question, the graph type defined by the connected points shows a probability of recognition of 0.53, the highest value within the question type. The expectation was that this recognition score would be at least lower than either the score for the lines on two frameworks on the score for the lines on a single framework. A possible explanation of this apparent discrepancy is artifactual. Note that the component heights ranged from 2 to 6 and the slopes could be increasing, decreasing, or unchanging. One display in this group had as values for its encod' features (i.e. slope, overall height) a maximum level and an unchanging trend. Thus, when the display was encoded, it is quite possible that an 'extremeness' tag was also appended to the representation. Such a tag should, therefore, increase the memoriability for this specific display. Removing this potential confounding results in a probability of recognition of only 0.27, a fifty percent change.

INSERT TABLE 3 HERE

Figure 2B is shown with this confounding removed. Note from the figure that recognition probability shows the same monotonic relationship as response times. The trend lines, drawn by visual inspection, emphasize the direction of th relationship. One possible explanation that must be considered is again artifactual. Inspection of the false alarm probabilities, shown in Table 3, indicate that they are of the same magnitude as the hit probabilities. Thus, the curves of Figure 2B might be representative of a guessing strategy.

To test this possibility, the hit probabilities associated with the 1, 2, and 3 ratings were analyzed. It was thought that this subset would provide a more sensitive measure of recognition performance, since these ratings correspond to a relatively high confidence in the recognition. Table 4 shows the probability of recognition for this category of answers. Similar patterns of results emerge. That is, recognition advantage for a specific question type seems to depend on the graph type chosen to represent the information. This observation is again supported by an analysis of variance on the transformed scores. An interaction between question and graph type was the only reliable source of variance, $F(6,54) = 2.33$, $p < .045$.

INSERT TABLE 4 HERE

The preferred explanation for this Question and Graph Type interaction is that the more congnitive effort directed at the reparsing operation when an incongruency exists provides a more durable trace. This increased durability thus leads to a subsequent recognition advantage. Such an advantage does not result when the encoding features are congruent with the conceptual message.

Recall Response

Table 5 shows the recall probabilities for the different graph types and questions. Unlike recognition performance, recall performance is insensitive to the amount and nature of the increased resource allocation. An analysis of variance provides no reliable source of variation for this measure.

INSERT TABLE 5 HERE

Storage or retrieval. Why are the two measures of memory performance so different? In a model of the memory performance, Chechile and Meyer (1976) uses both recall and recognition to distinguish between storage failures and retrieval failures. Briefly, when a stimulus is encounted it can either be stored or not stored. Thus, at the time of test if the stimulus is not stored, then retention will not be evidenced.[4] However, even if the stimulus is stored, retention may not occur for yet another reason. At time of test,

---

[4]I have over simplified the model for discussion purposes. It is of course possible that the subject could invoke a guessing strategy and demonstrate correct retention. However, the model can account for such a strategy by an estimate of a guessing parameter generated from test data and rating data.

a stored representation is either retrieved or not retrieved. If it is not retrieved then retention will not be evidenced. Thus, we have two different processes that converge on the same result, low retention performance. The two processes are partitioned through the following assumption: In a recognition task, a stored stimulus will always be retrieved. Thus, by contrasting recognition and recall data, it is possible to assess retention performance as a consequence of either storage or retrieval failures.

The data in this experiment indicate that the reparsing operation with its increased use of resources increases the advantage of storing the stimuli, but not retrieving it. Broadly speaking, recall differs from recognition in that recall requires a cue or probe to be generated internally in order to activate the stored representation, while in recognition the cue is provided by the stimulus itself (Shiffrin, 1970). Thus, the reparsing operation has relatively little affect on generating the correct probe.

GENERAL DISCUSSION

The results of this experiment support the hypothesis that an incongruency between information to be graphically communicated and the graphic representation of this information leads to a reparsing of the corresponding perceptual representation. This reparsing is a resource demanding operation and shows a memory advantage having its greatest utility in terms of the storage process. The reparsing necessitated by the incongruency may lead to a more elaborated trace of the representation (Craik and Tulving, 1975). For example, when asked to attend to a component value of the graph depicted by the pair of connected points, there would be no parameter in the perceptual representation of that display that would allow one value to be processed (e.g. looked up) while the other value is left alone. Rather the slope and level parameters activated would implicitly contain information about the component values, and so both values would be processed to arrive at the correct representation. Thus, elaboration would lead to a representation of this display that is more 'distinctive' or dissimilar then others in memory since more of its attributes are involved in the processing. This distinctiveness would produce and advantage during recognition.

On the other hand, a congruency between encoding features and conceptual information would not produce an elaborated representation and so no advantage in recognition would be seen (see Winograd, 1981 for a similar argument with respect to memory for faces). This is because the correct parameter would be activated and could thus be processed while the others remained inactive.

Relevance to Displays

At first glance, the results seem conterintuitive to what should be the guiding tenet in display design, that is, design displays such that the least amount of cognitive effort is needed to extract of the relevent information. The results reported in this experiment suggest that if the retention of information is critical, then one should violate this tenet. However, before we run off and begin designing poor displays to represent our information, the results should be taken to mean that if retention or processing of other attributes of a display is desired, then some means of inducing a more extensive elaboration is in order. This could be nothing more then an accompanying context that directs attention to these other attributes of the display. All that is necessary is to have observers invoke a reparsing of the display. As stated, this can be done in a controlled manner.

REFERENCES

Chechile, R. & Meyer, D.L. A Bayesian Procedure for Separately Estimating Storage and Retrieval Components of Forgetting. Journal of Mathematical Psychology, 1976, 13(3), 269-295.

Craik, F.I.M. & Tulving, E.: Depth of Processing and the Retention of Words in Episodic Memory. Journal of Experimental Psychology: General, 1975, 104, 268-294.

Kolers, P.A.: Memorial Consequences of Automatized Encoding. Journal of Experimental Psychology: Human Learning and Memory, 1975, 1, 689-701.

Schneider, W. & Shiffrin, R.M.: Controlled and Automatic Human Information Processing: I. Detection, Search, and Attention. Psychological Review, 1977, 84(1), 1-66.

Shiffrin, R.M. & Schneider, W.: Controlled and Automatic Human Information Processing: II. Perceptual Learning, Automatic Attenting and a General Theory. Psychological Review, 1977, 84(2), 127-190.

Shiffrin, R.M.: Memory Search. In D.A. Norman (Ed.), Models of Human Memory. New York: Academic Press, 1970.

Simcox, W.: Cognitive Considerations in Display Design. NIE Final Report Contract No. 400-79-006, 1981.

Winograd, E.: Elaboration and Distinctiveness in Memory for Faces. Journal of Experimental Psychology: Human Learning and Memory, 1981, 7(3), 181-190.

List of Figures & Tables

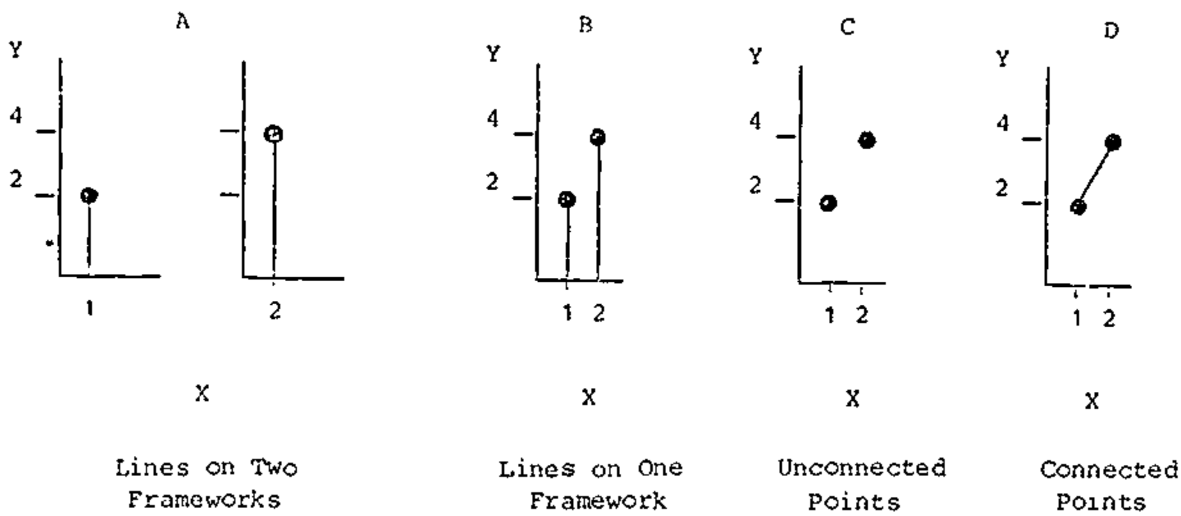

Figure 1   Four Symbolic Displays

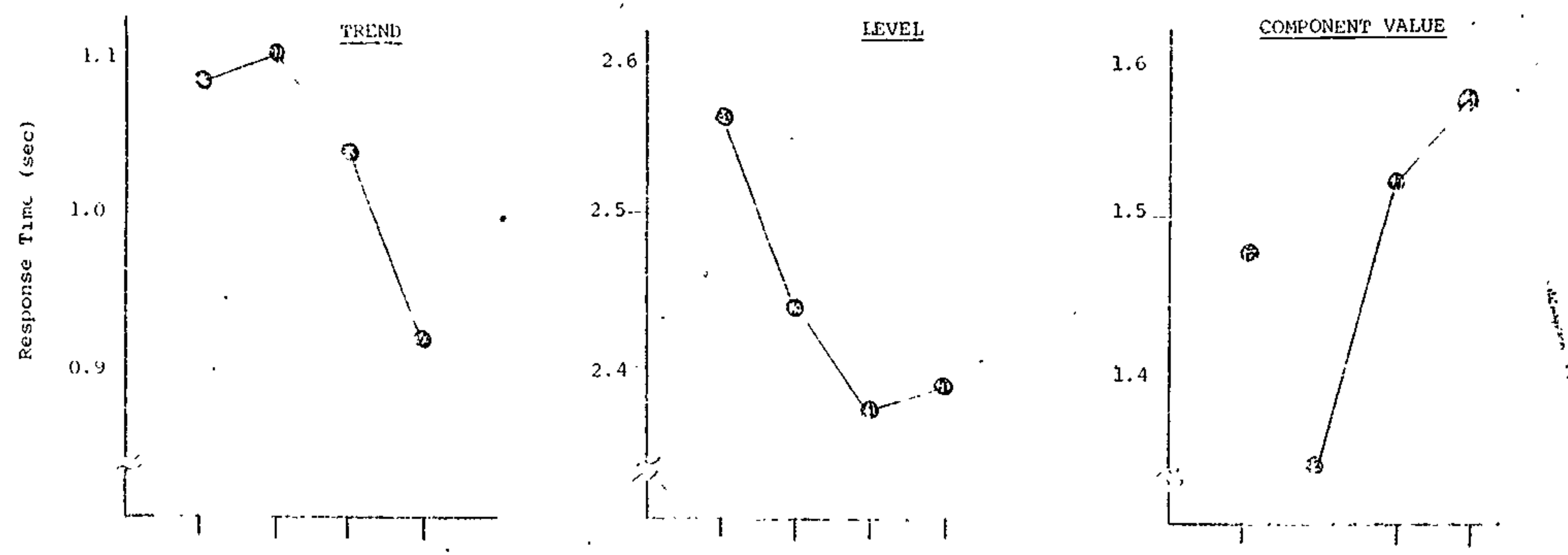

Figure 2A

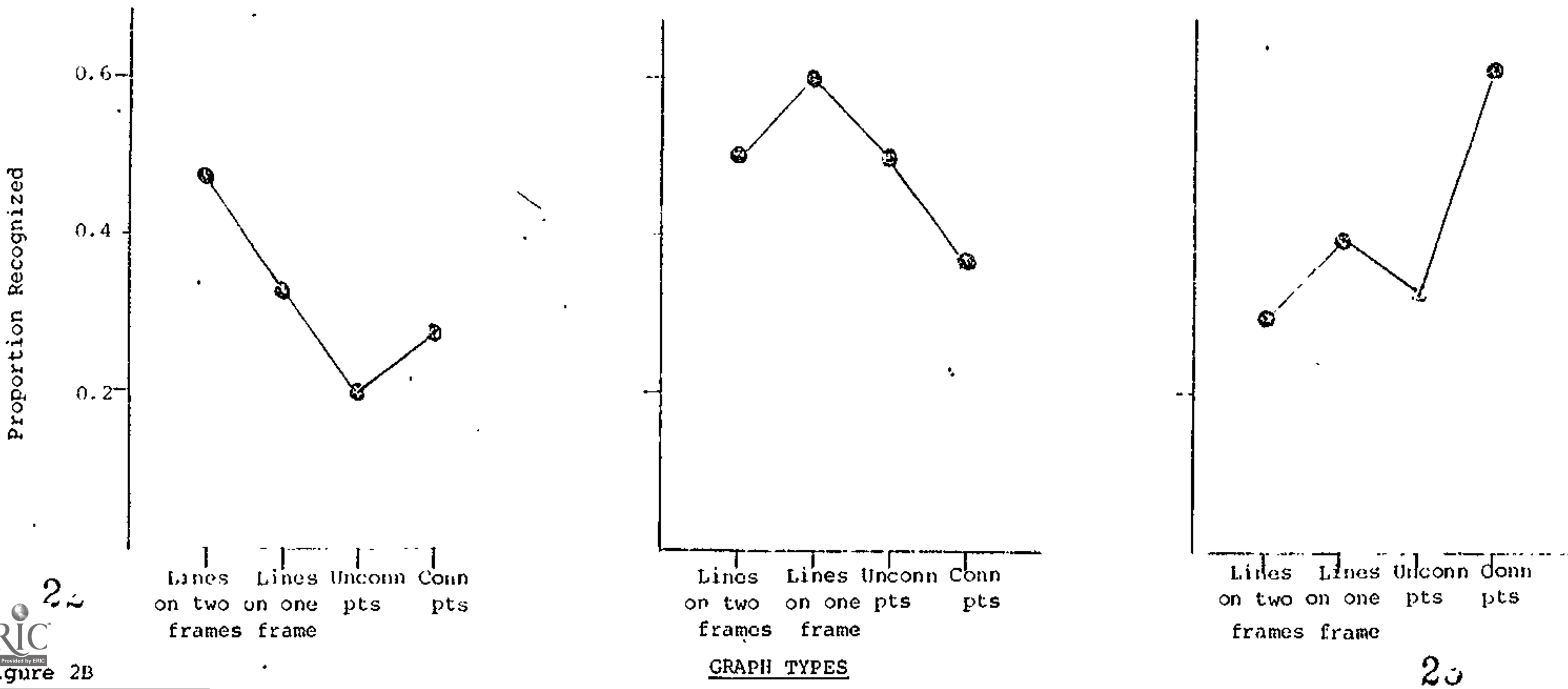

Figure 2B

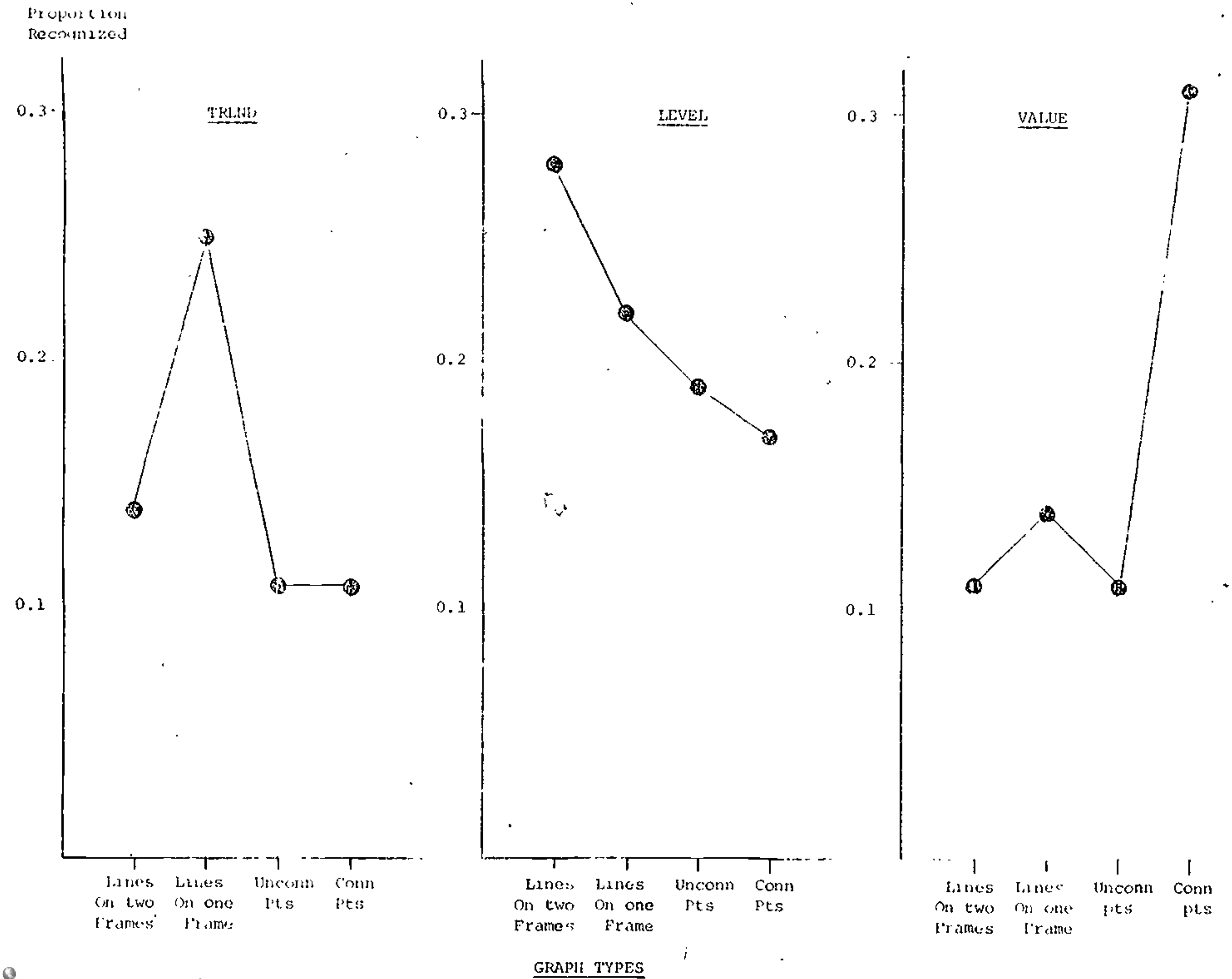

Figure 3 Probability of Recognition as a Function of Questions and Graph Types for Yes(3) Responses.

TABLE 1

Orthogonal Combinations of Defining Values

		Value of Right				
		2	3	4	5	6
	2	2,2	2,3	2,4	2,5	2,6
	3	3,2	3,3	3,4	3,5	3,6
Value of Left	4	4,2	4,3	4,4	4,5	4,6
	5	5,2	5,3	5,4	5,5	5,6
	6	6,3	6,3	6,4	6,5	6,6

TABLE 2

Mean Response Time as a Function of Graph and Question Type

MEAN RESPONSE TIME (sec)

	Graph Types			
Question	Lines on two frameworks	Lines on one framework	Unconnected points	Connected points
Trend	1.085	1.101	1.039	0.914
Level	2.565	2.415	2.360	2.398
Value	1.481	1.322	1.524	1.571

TABLE 3

Probability of Recognition as a Function of Graph and Question Type

PROPORTION RECOGNIZED

Graph Types

Question	Lines on two frameworks	Lines on one framework	Unconnected points	Connected points
Trend	0.47	0.33	0.20	0.53 (0.27)*
Level	0.50	0.60	0.50	0.37
Value	0.30	0.40	0.33	0.63
False Alarms	0.40	0.29	0.33	0.38

TABLE 4

Probability of Recognition as a Function of Question and Graph Type for Yes (3) Responses

	Graph Types			
Question	Lines on two frameworks	Lines on one framework	Unconnected points	Connected points
Trend	0.14	0.25	0.11	0.19
Level	0.28	0.22	0.19	0.17
Value	0.11	0.14	0.11	0.31

TABLE 5

Probability of Recall as a Function of Graph and Question Type

PROPORTION RECALLED

Graph Types

Question	Lines on two frameworks	Lines on one framework	Unconnected points	Connected points
Trend	0.10	0.13	0.10	0.03
Level	0.13	0.13	0.07	0.17
Value	0.00	0.07	0.07	0.03