DOCUMENT RESUME

ED 237 339 SE 043 546

AUTHOR Pinker, Steven
TITLE Pattern Perception and the Comprehension of Graphs.
INSTITUTION Massachusetts Inst. of Tech., Cambridge. Dept. of Psychology.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Mar 83
CONTRACT NIE-400-79-0066
NOTE 46p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Cognitive Processes; *Comprehension; Educational Research; *Graphs; Higher Education; Instructional Materials; Reaction Time; Science Education; *Visual Perception; *Visual Stimuli

ABSTRACT

Three experiments tested the hypothesis that graphs convey information effectively because they can display global trends as geometric patterns that visual systems encode easily. A novel type of graph was invented in which angles/lengths of line segments joined end-to-end represented variables of rainfall and temperature of a set of months. It was expected that questions about single values of a variable would be easier to answer when the variable was encoded as a segment length whereas questions about global trends of a variable would be easier to answer when the variable was encoded as a segment angle. Subjects' (N=14) response times when answering questions pertaining to such graphs demonstrated the interaction hypothesized. This was true both when subjects construed stimuli as meaningless visual patterns and when they construed the same stimuli as graphs. Similar results were obtained regardless of whether subjects were explicitly instructed about how trends of the angle variable translated into geometric shapes. It is concluded that graph formats, and types of information conveyed by graphs, are not uniformly easy or difficult, but that a given type of information is conveyed efficiently in a graph format to the extent that it corresponds to a naturally perceivable visual pattern. (Author/JN)

PATTERN PERCEPTION AND THE COMPREHENSION OF GRAPHS

NIE 400-79-0066

Year 03

March 14, 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:

Steven Pinker, Ph.D.
Department of Psychology and Center for Cognitive Science
Massachusetts Institute of Technology
Cambridge, MA 02139

Abstract

Three experiments tested the hypothesis that graphs convey information effectively because they can display global trends as g.ometric patterns that our visual systems encode easily. A novel type of graph was invented, in which the lengths and angles of line segments joined end-to-end represented the variables of rainfall and temperature of a set of months. It was expected that questions about single values of a variable in the graph would be easier to answer when the variable was encoded as segment length, since single lengths are easier to perceive than single angles; whereas questions about global trends of a variable (e.g., whether or not it is consistantly above a reference level) would be easier to answer when the variable is encoded as segment angle, since global patterns -- angles cause the graph as a whole to assume recognizable shapes. Subjects' response times when answering questions pertaining to graphs of this type showed just that interaction: subjects were faster at extracting single values of the variable conveyed by segment length than of the variable conveyed by segment angel, but more slower at extracting global trends of the variable conveyed by length than of the variable conveyed by angle. This was true both when subjects construed the stimuli as meaningless visual patterns and had to report the lengths and angles of segments, and when they construed the same stimuli as graphs and had to report values of rainfall and temperature. Furthermore, the same results were obtained regardless of whether subjects were explicitly instructed about how trends of the angle variable translated into geometric shapes. It is concluded that graph formats, and types of information conveyed by graphs, are not uniformly easy or difficult, but that a given type of information is conveyed efficiently in a graph format to the extent that it corresponds to a naturally perceivable visual pattern.

Summary

Three experiments tested the hypothesis that the effectiveness of graphs as a means of communicating quantitative information stems from their ability to display global trends as geometric patterns that our visual systems encode easily. A novel graph format was invented, consisting of a chain of line segments joined end-to-end corresponding to the months of the year. The length of a segment represented the rainfall for that month relative to a reference level, and its angle with repect to the previous segment represented its temperature relative to a reference level (or vice versa). It was expected that single values for temperature or rainfall would be easier to extract when encoded as segment length than when encoded as segment angle, since the perception of segment angle requires attention to a pair of segments and normalization of the orientation of the first. In contrast, the detection of whether temperature or rainfall was consistently above or below the reference level, versus sometimes being above it and sometimes below, and a similar discrimination involving the detection of alternation, were predicted to be easier when the variable was encoded by segment angle. This is because for the angle variable, consistent years yield uniformly convex curves and inconsistent years yield curves with a concave region, whereas the length variable does not yield curves with recognizable shape differences contingent on the consistency or alternation of the variable. In the first experiment, subjects were shown the stimuli described as visual patterns, not as graphs, and answered questions about the lengths and angles of particular segments or the consistency and alternation of the lengths or angles of the entire sequence. As predicted, single lengths were recognized more quickly and accurately than single angles, but consistent sequences of lengths were

recognized more slowly and less accurately than consistent sequences of angles. This provides independent motivation for predictions about graph reading difficulty in Experiments 2 and 3. In the second experiment, the stimuli served as graphs, and subjects were told how segment angle and length conveyed information about temperature and rainfall, and were also told that consistency of the variable conveyed by angle translated into convexity, and alternation into zigzags. When answering questions about rainfall and temperature, subjects showed the same pattern of reaction times as did their counterparts in Experiment 1 did when answering questions about the corresponding geometric properties of the stimuli. This indicates that the subjects, as predicted, were able to exploit the correspondences betwen trends and shapes and recognize trends directly without examining individual point values. In Experiment 3, subjects were only told how the graphs conveyed information about temperature and rainfall for individual months, and showed similar patterns of response time and accuracy. It is concluded that graph formats, and types of information conveyed by graphs, are not uniformly easy or difficult, but a given type of information is conveyed efficiently in a graph format to the extent that it corresponds to a naturally perceivable visual pattern.

Patter[illegible] and the Comprehension of Graphs

Pictorial [illegible] [illegible]ious forms of human communication of quantitative info[illegible] [illegible] displayed as graphs, sets as circles, sentence str[illegible] as [illegible] [illegible]ograms as flowcharts. Perhaps visual displays are si[illegible] [illegible]ther [illegible] pleasing, but it is far more likely that they are popular [illegible] [illegible] [illegible]vey information in a form that is easier to perceive or r[illegible] [illegible] [illegible]owever, currently we have little understanding of what it is a[illegible] [illegible] human mind that makes graphs and other pictorial displays more e[illegible] than other formats containing the same information, such as tables of numbers. An answer to this question would not only shed light on a striking cognitive phenomenon, but it would be a prerequisite to solving the practical problems of devising more effective graph formats or of choosing the best existing format when deciding how a set of data should be displayed.

Pinker (1981) proposed a theory of graph comprehension that tried to explain this supposed advantage with the help of three claims: 1) The human visual system has the ability to recognize a large number of two-dimensional shape predicates quickly and easily. For example, we can detect the length, height, orientation, curvature, shape, parallelism, smoothness, compactness, etc., of a line or set of lines quickly and with a minimum of effort. 2) Depending on the graph format, different aspects of a data set will be translated into different types of visual patterns. Consider, for example, a standard line graph representing a dependent variable on the ordinate and two independent variables on the abscissa and as the parameter, respectively. The absence of an effect of the independent variables on the dependent variables translates into flat, overlapping lines; an effect of one of the independent variables translates into two lines with a slope, and an effect of the other independent variable translates into non-overlap of the two lines; additivity of the effects of the two independent variables translates into parallel lines and non-additivity into nonparallel lines, and so on. 3) Efficient graph readers know the correspondences between quantitative trends and visual patterns for a particular type of graph (e.g., for line graphs, no effect=flat line, effect =sloping line), and when they need to extract one such trend from a graph, they can look for the corresponding higher-order visual pattern and mentally translate it into the relevant trend without having to examine individual points one-by-one and compare their values against one another (e.g., subtracting and checking for a non-zero difference).

Pinker (1981) summarized the implications of this theory for predicting the degree of difficulty a reader will have in attempting to extract a particular sort of information from a particular type of graph. These implications can be summarized in a single principle: the ease of reading a certain type of information from a certain graph format will depend on the extent to which that graph format translates that trend into a single visual pattern that the visual system can automatically extract, and on the extent to which the reader knows that the correspondence in that format between the quantitative trend and the visual pattern holds. In other words, graphs, and types of information contained in graphs, are not easy or difficult across the board; rather, one graph format may be well-suited to yielding the answer to one sort of question, while ill-suited to yielding the answer to another, depending on the geometric pattern that conveys the answer and the visual system's ability to encode that pattern.

In support of this principle, Pinker cited experimental evidence comparing the ease of reading bar and line graphs. According to a number of studies (e.g., Carter, 1947; Culbertson & Powers, 1959; Schutz, 1961 a, b; Washburne, 1927), people make fewer errors reading line graphs than they do when reading bar graphs if they have to answer questions about data trends. However, if they have to answer questions about the difference between a pair of observations, or the absolute value of a single observation, then they made fewer errors reading bar graphs or tables. Pinker interpreted this interaction in terms of the differences between the two formats in the relative perceptibility of the patterns into which each sort of information is translated. In line graphs, trends translate into the shape of a line or of a configuration formed by a set of lines, which is an easily avoidable property (see Kubovy, 1981). However, in bar graphs, especially those that encode more than two variables, trends translate into a particular pattern of lengths of different bars, which, not forming a unitary Gestalt, must be examined and compared one or two at a time. However, matters are different when it comes to single values. In line graphs, the value of a single value of Y associated with a value of X translates into the height of an isolated point or portion of a continuous line, and if that point is not clearly demarcated on the line or segregated in some way from other points, focusing attention on that point may be an effortful process. On the other hand, in a bar graph, the value of a given Y translates into the length of a bar, which is a "good Gestalt" separated from adjacent bars by lines or space, hence easier to isolate perceptually.

Unfortunately, there are two problems with this argument. First, there is no independent evidence for the putative perceptual effects alluded to (e.g., effortless perception of line shape and single bar length vs.

effortful perception of set of relative bar lengths and height of segment of a curve). Second, there has arisen a considerable body of lore in graphics and style manuals conerning the appropriateness of different graph formats for conveying different types of information. One of the injunctions frequently mentioned is to use line graphs to convey trends (see Kosslyn & Pinker, in preparation). If graph designers follow such injunctions, they may use line graphs more often when the context of a graph requires the reader to extract trend information. Hence, such graphs may be more common in such contexts, giving readers more practice at looking for trends in line graphs. Hence, people in general may become faster at executing the sequence of operations necessary to verify trend information from line graphs, even if in fact those operations were intrinsically equally easy to carry out for all graph formats and the injunctions in style manuals no more than dogmatic conventional wisdom.

To obtain experimental evidence of the proper sort, a completely novel graph format was invented which conveyed information about two variables in two different ways. In this format, one type of question (the value of a datum relative to a reference value) translates into an easily perceivable visual pattern when it is encoded by one geometric attribute of the graph, but not the other attribute. On the other hand, a second type of question (whether or not a series of data is consistently above or below a reference value, or whether or not a series of data alternates between being above and below that value) should show the reverse pattern being encoded as an easily perceivable pattern when pertaining to the second geometric attribute of the graph, but not the first. If the principle about graph difficulty described above is correct, response times should show on interaction between the type of question asked and the geometric attribute of the graph that encodes the answer. Since these graphs are completely unfamiliar, people's past experience could not be the cause of any pattern of facility they are found to have at reading the graphs. In addition, the set of experiments included a test of the perceptibility of the relevant properties of the graphs when treated solely as an uninterpreted visual pattern, to serve as independent support for the perceptual effects allegedly influencing graph reading ease. Furthermore, the knowledge of the correspondence between trend and visual pattern was manipulated across experiments in an attempt to assess how robust the interaction is across different degrees of explicit training in how to read the graph.

Experiment 1

The experiments reported in this paper all employed the following type of graph: a string of eight line segments, joined end-to-end, represents the mean temperature and rainfall for each of eight successive months in a given year, temperature by the length of a line segment, and rainfall by the angle formed by one segment with respect to the previous one. The end of the chain representing January is indicated by a dot at the end of the segment; thus the graph can lie in any orientation on the page without affecting the information it conveys. Examples are illustrated in Figure 1. In each case, the temperature and rainfall are conveyed not in absolute terms, but with respect to a reference vlaue, specifically, the average or typical temperature or rainfall for that month (i.e., averaged over a number of years). Thus, a month whose temperature was at its average level would be depicted by a segment one inch long, whereas months below or above their average temperatures would be depicted by segments whose lengths would be less than or greater than one inch, respectively. On the other hand, a month whose rainfall was average would be represented as a line segment that was in line with the previous line segment ("previous" in the sequence from January to December); that is, that formed a 180 degree angle with its predecessor. A month whose rainfall level was below average would form a counterclockwise angle of between 135 and 180 degrees with its predecessor; and a month whose rainfall level was above average would form a counterclockwise angle of between 180 and 225 degrees with its predecessor. That is, rainfall is represented as the continuous variable of angle ranging from 135 to 225 degrees, with 180 degrees as the reference level. Another way of putting it is that, traversing a graph from January to December, a line pointing to the left with respect to its predecessor represents average rainfall, whereas one pointing to the right represents below average rainfall. Examples of long and short segments, and of large and small angles, are shown in Figure 2.

INSERT FIGURES 1 AND 2 ABOUT HERE

Note that in this graph format, whether a temperature of an individual month is above or below average should be easier to verify than whether its rainfall is above or below average. That is because rainfall is encoded by whether a segment forms an angle greater than or less than 180 degrees with respect to its predecessor, requiring that the reader attend to two line segments, not one. Furthermore, the angle discrimination is in effect a handedness judgement (i.e., whether the line points leftward or rightward). People cannot make handedness judgements independently of a figure's orientation, but they must mentally rotate the figure into a standard orientation first (Cooper & Shepard, 1973, 1975). For temperature (=line length), neither attention to a pair of line segments nor a mental rotation into a canonical orientation is necessary.

However, consider the question of whether all 12 months shown in the graph are consistently above or below average in rainfall, as opposed to being above average for some months and below average for others. For rainfall, a consistently wetter (or dryer) than average year translates into a string of segments each forming an obtuse angle (or each forming a reflex angle) with respect to its predecessor, yielding a consistently convex curve. If one of the months is below average and the rest are above (or vice-versa), that segment will cause the overall curve to be concave at that point. Hence, a consistency/ inconsistency judgement for rainfall can be made on the basis of whether the curve is uniformly convex, or convex in some parts and concave in others. Such changes in sign of curvature should be readily detectable by the visual system, since they play a role in defining how a contour is perceptually parsed into parts (Attneave, 1954; Rock, 1974; Hoffman, Note 2). However, no such shortcut should be available for temperature. If all the segments of a line are longer than one inch, than the line as a whole will be much longer than average, to be sure, but that cannot in general serve as a reliable cue: if the line is not much longer than average, it may be because one or more of the lines is below average or because all are above average by small amounts. When one or more segments is at a different length than the others, it does not change any perceptible qualitative property of the string as a whole, unlike a segment at a different angle. This is illustrated in Figures 1(a-d).

A similar line of reasoning can be applied to the case of determining whether a series of months alternates between being above average and below average in rainfall or temperature, or whether there is a pair of consecutive months in the sequence whose members are both above average or both below average. In the case of rainfall (angle), an alternating sequence translates into an approximately straight zigzag line, whereas a sequence that alternates except for one pair of adjacent months translates into a line with a convex bulge somewhere along its length and possibly a gross bend in the overall contour (see Figure 1(e-h)). For temperature (length), no such readily identifiable pattern emerges as a consequence of alternation versus non-alternation, and again, one might conjecture the reader would have little choice but to examine the length of each segment in turn and keep in mind the sequence of below and above average months. Thus, we have the same prediction for the detection of alternation as we had for the detection of consistency: it should be easier for rainfall (angle) than for temperature (length). And the detection of both of these properties, consistency and alternation, should show a different pattern from the detection of the value of a given month, which should be easier for line than for angle.

As mentioned, to pursue this line of argument it is imperative to obtain independent evidence for the perceptual phenomena alluded to, otherwise the explanation for why a given question is difficult to answer using a given graph is in danger of becoming circular. Experiment 1 is an attempt to obtain such evidence: patterns identical to the graphs described are shown to subjects as pure patterns, to be classified in terms of the length and angle of given segments or the consistency or alternation of their length and angle, with no reference to temperature or rainfall or any quantitative referent at all. The intent is to use chronometric data to verify that segment length is easier to encode than segment angle for individual segments, but that consistency or alternation of length is harder to encode than consistency or alternation of angle. If these perceptual phenomenals emerge, we can then see if they predict the difficulty of extracting different sorts of quantitative information from graphs.

Method

Subjects

Fourteen Stanford University undergraduates participated in a one-hour session either for pay or to fulfill an introductory psychology course requirement.

Materials

Nineteen graphs of the form described above and shown in Figure 1 were drawn on 8 1/2 x 11 in. paper, and four slides of each were made. Sixteen were for use in the experiment, and three were for use in practice trials. There were approximately equal numbers of graphs with January toward the bottom, top, left, and right sides. The line segments for below average months ranged from 1.75 to 2.38 cm length on the page, and those for above average months ranged from 2.86 to 3.49 cm. Segments formed counterclockwise angles of between 135 and 165 degrees, and between 194 and 225 degrees, with their predecessors (as measured with respect to a direction corresponding to a traversal of the line from January to August). In other words, no two consecutive segments formed an angle of 180 degrees or angles within 14 degrees of 180 degrees. The 16 experimental stimuli were drawn so that there were four exemplars each of stimuli with consistent and inconsistent angles, consistent and inconsistent lengths, and non-alternating angles, and alternating and non-alternating lengths. Of course, with only 16 patterns, a given pattern had to serve in several cells in this design; in fact, the patterns were drawn so that each one could serve once with a question about a single segment's length, once with a question about a single segment's angle, once with a question about consistency or alternation of length, and once with a question about consistency or alternation of angle.

Design

The experiment consisted of 64 trials, each consisting of a graph and a question. Half the trials involved questions about angle, and half about length (this will be called the "Dimension" factor); half the questions involved the value of a single month relative to its mean and half invoked the global sequence of eight months (this will be called "Locality"); and half the questions had "yes" as their answer and half had "no" ("Response Type"). All these factors were crossed orthogonally. In addition, for the questions involving a single month, half pertained to one of the first three months of the year (January was never in a question), and half pertained to one of the second four months; for the questions involving the global sequence of months, half asked whether the sequence was consistently above or below average, and half asked whether the sequence alternated between above and below average months (the latter subfactor will be called "Global Question Type"). Each of these subfactors was crossed orthogonally with Dimension and Response Type. Finally, as mentioned, each of the 16 graphs used in the experiment appeared an equal number of times in conjunction with questions involving the four combinations of Dimension x Locality. This ensured that the comparison of interest (the interactions between line/angle and single segment/global sequence) would not be affected by the idiosyncratic properties of individual graphs (e.g., how interesting a given graph looks).

Procedure

Subjects, tested individually, were read a set of instructions which was simultaneously available to them in printed form. They were told that they were participating in an experiment on pattern perception, and that on each trial they were to see a shape consisting of a chain of line segments of different lengths and forming different angles at each junction. One of the test graphs was shown on paper as an example. In order for it to be possible to refer to individual segments, they were told, each segment was given an arbitrary name corresponding to one of the months of the year, so that the names of the consecutive segments from one end of the chain to the other should form a sequence corresponding to the first eight months of the year, with a dot signifying the "beginning" of the chain. They were told that they would have to discriminate betwen two ranges of lengths of line segments, which were illustrated in a rear-projected slide consisting of six isolated "long" segments and six isolated "short" segments. The slide, similar to the left half of Figure 2, appeared as an approximately 21.6 x 27.9 cm vertical rectangle (i.e., the same size as the paper) on the screen, which was approximately 60 cm away. Subjects were asked to practice associating long and short segments with the right and left keys, respectively, as they pressed each key a number of times. Then the test slide was projected, and the experimenter recited a series of month names, to which the subject was to press one or the other key, depending on whether the relevant segment in the test pattern was long or short. A similar procedure was then followed for "large" versus "small" angles, which were first illustrated with a Y-shaped figure as well as with a collection of isolated segment pairs exemplifying large and small angles (similar to the right half of Figure 2). It was also pointed out to them that "large" and "small" angles pointed to the left and right, respectively.

Then the subjects were told that some graphs were "consistent" in length, meaning that the eight segments were either all long or all short, or all with large angles or all with small angles; others, which were "inconsistent", had one or more segments at a different length or angle from the rest.

It was also pointed out to them that patterns that were consistent on the angle dimension would appear as uniformly convex lines, whereas patterns that were inconsistent would have one or more concave or "pushed-in" regions in an otherwise convex curve. They were also told that patterns that were consistent on the length dimension would form a long or a short line, with uniformly long or short segments, whereas patterns that were inconsistent would have one or more shortish or longish segments interspread among others of uniform length (although it was anticipated that only the angle dimension would yield recognizable configurations or shapes contingent on its consistency, we gave these instructions for length as well to ensure that subjects would not simply devote more attention to the task when consistency questions pertained to angle). Examples of patterns consistent and inconsistent in length and angle were shown on paper to the subjects who were asked to associate consistency and inconsistency with the right and left keys, respectively, as they practiced pressing each one. Then three test slides were shown, and subjects had to press the appropriate key depending on its consistency of angle or of length, as indicated by the experimenter. A parallel set of instructions then followed for the remaining question type concerning the alternation of values for a dimension. In such cases, subjects were to press one key if the pattern consisted of alternating lines (short-long-short-long... or long-short-long-short... for length, or left-right-left-right... or right-left-right-left... for angle), and another key if the pattern contained one or more segments that broke the alternating sequence. It was pointed out that for angle, alternating sequences corresponded to a uniformly zigzag line, whereas non-alternating sequences corresponded to a zigzag line that also had one or more largish bulges somewhere along its length. As before, attention to line versus angle was equalized by also pointing out that for length, alternating sequences corresponded to a pattern with no longish or shortish series of consecutive segments in them, whereas non-alternating sequences corresponded to a graph with a sub-sequence of consecutive long or short segments somewhere along its length (of course, only for angle should there be an easily perceptible configuration or shape corresponding to alternation). As before, subjects were shown illustrations of graphs which were alternating and non-alternating in angle and length and were given practice at pressing the buttons and at discriminating alternation of length and angle in three test slides.

Finally, subjects were told about what they would have to do in the series of trials to come. Each trial would begin with a taped voice speaking two words. The first word would be either "length" or "angle", and the second would be either "consistent", "alternating", or the name of a month. Together, these words would constitute a question, which subjects would have to answer as quickly and as accurately as possible by pressing the appropriate key when a slide containing a graph appeared on the screen a short time later. The rightward key was appropriate for answers "large", "consistent", or "alternating", depending on the question. Subjects were urged to respond as quickly as possible without making errors.

The second word in the trial question was recorded on a second channel on the tape to serve as a timing signal, and 3.5 seconds after its onset, the tape recorder paused, the slide projector advanced, and a shutter opened, allowing the pattern to project onto the screen. The pattern was visible until the subject pressed one of the keys, which closed the shutter and initiated a three second intertrial interval. The tape had 68 trials, the first four of which were later treated as practice trials and excluded from the analyses. Trials representing different conditions were distributed evenly throughout the trial sequence, and there were never more than three consecutive "yes" or "no" trials. A microcomputer controlled the timing of the apparatus and recorded which key the subject pressed on each trial and the latency of each response from the opening of the shutter, with an accuracy of approximately msecs.

Results and Discussion

Subjects' error rates ranged from 3 percent to 11 percent, with a mean of 5 percent, not including four subjects who erred on more than 15 percent of the trials and hence were eliminated from the analysis. In analyzing the response times from the remaining subjects, we excluded response times for which the response was incorrect, and response times greater than twice the subject's mean (only 2 percent of the responses were excluded by this criterion). Such responses were replaced by the average of that subjects response times for other trials in which the same type of question was asked. The remaining times were submitted to an Analysis of Variance whose repeated measures factors were Dimension (Line vs. Angle), Locality (Global vs. Local question), Global Question Type (Consistent vs. Alternating; this was meaningful only for Global questions), Response Type (Yes vs. No), and Replications (First, Second, Third, or Fourth).

Results, plotted in Figure 3, indicate that questions about global properites take longer to answer than questions about local properties, $F(1,9)=6.28$, $p<.05$. Questions about consistency were answered more quickly than questions about alternation, resulting in significant effects of Global Question Type, $F(1,9)=34.78$, $p<.001$, and of the Global Question Type x Locality interaction, $F(1,9)=79.10$, $p<.001$.

INSERT FIGURE 3 ABOUT HERE

The principal prediction of the experiment--that local questions would be easier to answer when they pertained to line length, whereas global questions would be easier to answer when pertaining to angle--were strongly confirmed. The Dimension x Locality interaction was highly significant, $F(1,9)=24.76$, $p<.005$, reflecting the fact that for questions about local segment length, the mean response times were longer for Local than for Global questions, whereas for questions about angle, the mean response times were higher for Local than for Global questions. A close examination of the data show that this interaction holds far more strongly for Global questions concerning Consistency than for Global questions concerning Alternation. Questions concerning the Alternation of angles in fact took more time to answer than questions concerning a single angle (4713 vs 4492 msecs, respectively); the advantage of Global questions concerning angle can be seen only when the question pertained to Consistency, whose mean was 2473 msecs. (For length, questions about both Consistency and Alternation are harder than questions about individual segments).

The Replications factor was significant, $F(3,27)=3.56$, $p<.05$, with subjects responding more quickly in later replications. In addition, the Dimension x Locality x Global Question Type interaction in turn interacted with Replications, $F(3,27)=4.63$, $p<.005$, but this four-way interaction does not suggest the need for any qualifications of the conclusions reached in the previous paragraph. In all four replications, questions concerning length were answered more quickly when they pertained to a local segment that when they pertained to the global sequence (both Consistency and Alternation); questions concerning angle were answered more slowly when pertaining to a local segment than when pertaining to the Consistency of the global sequence, and angle questions about local segments were answered sometimes more quickly and sometimes more slowly than angle questions about Alternation in the global sequence. In other words, the qualitative nature of the pattern of response times depicted in Figure 3 did not change from replication to replication; the interaction simply reflected changes in the magnitudes of the differences from replication to replication. Several other interactions involving Response Type and Replications were significant, none of them incorporating the Dimension x Locality interaction, but since these higher-order interactions are partially confounded with individual graphs and questions, they are not easily interpretable.

Mean error rates are depicted by bar graphs at the bottom of Figure 3, and were analyzed in an Analysis of Variance parallel to the one performed on response times. Error rates mirrored the pattern found for reaction times: single lengths were detected more accurately than global patterns of lengths, whereas single angles were detected less accurately than global patterns of angles, resulting in a significant Dimension x Locality interaction, $F(1,9)=19.56$, $p<.005$. This shows that the response time interaction plotted in Figure 4 cannot be attributed to a speed-accuracy tradeoff, with subjects leaping to quick inaccurate responses in some conditions and slowly but accurately reasoning out others. Unlike response times, however, there was better performance in detecting both Consistency of angle and Alternation of angle compared to single angles. The Dimension x Locality x Question Type x Replications interaction was significant, $F(3,27)=3.47$, $p<.05$, but this reflects differences in the magnitude, not direction, of the Dimension x Locality interaction: Single lengths were detected as accurately or more accurately than either Consistency or Alternation of length in all four replications. The Replications factor itself was significant, $F(3,27)=3.47$, $p<.05$, reflecting a lower error rate in the last replication, as was Response Type, $F(1,9)=5.5$, $p<.05$, reflecting a slight bias toward "Yes" responses, as was their interaction, $F(3,27)=5.20$, $p<.01$.

The results of this experiment now allow us to make clearcut, noncircular predictions about the results of the next two experiments, involving presentation of these same patterns described as graphs that convey quantitative nonvisual information. As expected, people were faster and more accurate when determining the length of a single segment than when determining whether the lengths of an entire sequence of segments conform to some pattern, whereas for angle, they showed the reverse pattern. Presumably, this is because our visual system does not automatically compute global properties arising out of certain patterns of lengths of individual components, requiring subjects to assess the length of each segment in turn and then verify whether the list of encoded lengths conforms to the global property in question. This may have been what caused subjects to take longer in answering global questions: they may have repeatedly applied a process that need only have been applied once for questions about single segments. However, our visual systems may indeed automatically compute global properties arising out of the angles among individual segments. If one were to smooth out the stimuli so as to form continuous curves rather than chains of straight line segments, then inconsistent but not consistent stimuli would have a point of minimum curvature somewhere along their lengths. Hoffman (Note 2) points out that junctions of parts of objects (e.g., where a limb of an animal meets the torso or a knob meets a door) tend to yield points of minimum curvature along the outline of the object, and has gathered evidence that our visual system seeks out such minimal in order to help find part boundaries. It could be that our subjects detected inconsistency of angle rapidly because they were able to apply mechanisms of this sort to the task. That subjects were able to use some global strategy for angle can be seen from the fact that they took less time for global than for local questions about angle: if they were always examining individual segment angles, they would have taken longer for global questions, since these required several applications of a process that was applied only once for global questions.

For reasons that we cannot be sure of, these visual encoding processes were not able to compute alternation versus non-alternation of angle more quickly than the angle of a specific segment (though alternations was still detected more quickly for angle than for line, unlike the detection of single segment properties). Perhaps it is because the discrimination of number of minimal of curvature, or of part boundaries in general, conforms to Weber's Law, and a Consistency discrimination with these stimuli involves one versus two parts, whereas an Alternation discrimination involves three versus four parts; or perhaps it is because the range of angles employed in our stimuli was not broad enough to allow alternation versus non-alternation to translate into roughly straight versus grossly bent lines as we had anticipated. In any case, the results of this experiment lead us to expect large differences betwen the speed of perceiving values of local segments versus the speed of perceiving the consistency of the sequences, when we turn to the next experiments on graph reading per se. That is because only when one knows that a global property is easily perceivable in the first place can one confidently predict that the corresponding global quantitative trend will be easy to extract by a reader of the relevant graph.

Experiment 2

Although graph types are usually defined and taught in terms of the way they depict individual n-tuples of associated values of n variables (see Pinker, Note 1), often users are taught explicitly that trends of one variable with respect to another can be ascertained by looking for some higher-order geometric property rather than examining a succession of the units each encoding a single n-tuple. For example, calculus students are taught that if the first derivative of a function is positive, the graph of that function will rise as one examines it from left to right; that if its second derivative is positive, the graph will be curved with its concave side facing up; that the zero-crossings of the first and second derivatives can be located by looking for peaks and troughs or inflection points, respectively, and so on. Statistics students are taught that statistical main effects can be detected by looking for nonflat lines or separated lines, whereas interactions can be detected by looking for nonparallel lines.

However, it cannot be taken for granted that nonspecialists can exploit these sorts of correspondences. It may be that people will most naturally fall back upon the more basic definition of the graph type and how it conveys information about specific values, knowledge that is sufficient to compute any other sort of information necessary. On this hypothesis, the popularity of graphs (assuming it is not attributable to esthetic factors) might stem from, say, the fact that they display information about point values in specific locations on a page, or that related point values are perceptually grouped by the Gestalt principles, leading to greater ease at finding point values as compared to looking them up in a table. On this alternative account, trend-shape correspondences need play no role in the efficacy of graphs to ordinary readers.

This experiment tests the hypothesis that untrained people can immediately exploit correspondences between trends and shapes simply by having them pointed out to them. Subjects are shown the string-of-segment patterns described in the previous section for which we know (from Experiment 1) that there exists perceptual shortcuts for detecting the consistency and alternation of whatever variable is depicted by angle. The subjects are told that segment length corresponds to one variable and segment angle corresponds to another, and it is also pointed out to them that Consistency corresponds to uniform convexity and inconsistency to a concave region, and that alternation corresponds to a uniform zigzag pattern, whereas non-alternation results in the graph containing a bulge. If people, when learning about a graph type, can record the information that it is possible to exploit correspondences between visual patterns and quantitative trends, then in this experiment they should take the hints to heart, and when possible (i.e., for global questions concerning the variable represented by angle), ignore the basic definition of how the graph represents individual data values and encode the consistency directly from the convexity of the curve. That should result in an identical pattern of response time data as was observed in the previous experiment.

Method

Subjects. Fourteen Stanford University undergraduates participated either for pay or to fulfill a course requirement in Introductory Psychology.

Materials. The stimuli were identical to those of Experiment 1.

Procedure. The procedure was identical to that of Experiment 1, except that the subjects were told that they were being shown graphs of the temperature and rainfall of various regions in the first eight months of particular years, relative to their mean values. Half the subjects were told that angle represented rainfall and length temperature, and the other half were told that angle represented temperature and length rainfall (this was to ensure that subjects' naive meteorology would not contaminate the comparisons of interest). Instructions from the previous experiment were simply altered so that "rainfall" or "temperature" was substituted for "length" or "angle", "greater than average" or "less than average" was substituted for "long" and "short" lines or "large" and "small" angles, and so on. Thus subjects were taught how rainfall and temperature relative to their averages were represented in the graph, and also how consistency and alternation of temperature (or rainfall) translated into uniforn convexity and uniform zigzagging respectively. As before, subjects' attention was drawn to consistency and alternation of the variable represented by length as well so as not to confound the intrinsic perceptual distinction between length and angle with greater attention to angle. The trials were introduced as tests of their ability to read the graphs and answer questions about them; the two types of trials mentioned "temperature" or "angle" on each trial followed by "consistent", "alternating", or the name of a month. The subjects were told to indicate their answer to the question as it applies to the graph shown in the upcoming slide by pressing one of two keys (right hand key for greater than average, consistent, or alternating depending on the question). The stimuli, order of trials, and timing of events were identical to those of Experiment 1.

Results and Discussion

Subjects' error rates ranged from 2% to 14%, with a mean of 5%, except for three subjects who erred on more than 15% of the trials and whose data were not analyzed further. Three percent of the response times were more than twice the subject's mean; those response times and response times for trials on which errors occurred were replaced by the subject's mean response time for other trials representing the same combination of the Dimension, Locality, and Global Question type factor (as before).

INSERT FIGURE 4 ABOUT HERE

The main results of the experiment are shown in Figure 4. Response times to questions about the quantitative content of the graphs varied in the same way as the response times to questions about the lengths, angles, and shapes of these same patterns in Experiment 1. Questions about single months were answered more quickly than questions about global sequences, $F(1,10)=10.01$, $p<.005$, and questions about the global consistency of sequences were answered more quickly than questions about the global alternation of sequences, $F(1,10)=32.43$, $p<.005$ for Global Question Type; $F(1,10)=7.00$, $p<.05$ for Global Question Type X Locality. Most interestingly, it was easier to determine single point values for the variable signified by length than the variable signified by angle; however, it was easier to verify the global trends of the variable signified by angle than the variable signified by length. This interaction between Dimension and Locality was statistically significant, $F(1,10)=17.86$, $p<.005$. As Figure 4 shows, this interaction is mainly attributable to the ease of detecting the consistency, not the alternation, of the variable represented by angle, which is exactly what the results of the first experiment would lead us to expect (recall that alternation of angle was in fact no easier to perceive than single angles).

As before, Replications was significant, $F(3,30)=11.94$, $p<.001$, reflecting faster overall response times with each successive replication, as was Response Type, $F(1,10)=7.06$, $p<.05$, and the interaction among Replication, Dimension, Locality, Global Question Type, and Response Type, $F(3,30)=4.51$,

$p<.05$. This interaction is completely confounded with individual trials and hence is difficult to interpret, but it is reassuring that the Dimension X Locality interaction does not change qualitatively over the levels of the other factors: single values of the variable represented by length were detected more quickly than global trends in all cases; and consistency of the variable represented by angle was detected more quickly than single values of that variable in 7 of the 8 trials (alternation of the angle variable was detected more slowly than single values in 6 out of 8 trials, though still more quickly than alternation of the length variable on 5 out of 8 trials). Several other interactions were significant, though not clearly interpretable, also owing to their correlation with particular graphs.

As before, errors showed the same pattern as reaction times, with more errors occurring for global trends of the length variable than for single values, but fewer errors occurring for global trends of the angle variable than for single values (the interaction, however, was not statistically significant, nor were any other interpretable effects). This pattern of errors shows that a speed-accuracy tradeoff is unlikely to account for the interaction of interest in the response times.

This experiment shows that when graph readers know how a graph format conveys information about point values, and also that higher-order trends can be displayed as recognizable shapes in that format, they can perceive the shape and mentally translate it into the trend directly, without examining point values one-by-one and calculating the trend from those values. Had they in fact been examining point values, they would have taken longer for global questions about the angle variable than about local questions, since the former require several examinations of point values rather than a single examination. This cannot be attributed to the intrinsic ease of processing angle nor of detecting consistency, because the angle of individual segments and the consistency of segment length were detected more slowly than the consistency of angle. Similarly, subjects' ease at detecting consistency of angle cannot be attributed to their construing the graph solely as a means of communicating trends, ignoring completely its means of communicating point values, because subjects had to be processing point values to answer questions about individual months, and showed no advantage when processing trends whose associated visual pattern was shown in Experiment 1 to be difficult to perceive (i.e., alternation). Finally, the pattern cannot be attributed to subjects' finding certain of our graphs globally harder than others, since each graph was shown once in each of the combinations of the Dimension and Locality factors. It seems reasonable to conclude that people can extract quantitative trends from visual displays by encoding the corresponding shape and mentally translating it directly into that trend, and that one of the reasons that graphs are effective may be that they make this possible.

Experiment 3

Experiment 2 showed that people are capable of exploiting correspondences between trends and shapes when these are pointed out to them explicitly. However, if these were the only circumstances in which people could exploit such correspondences, it would not be obvious why graphs are considered so effective in general. Most graphs (e.g., line graphs) are potentially quite versatile, possessing a great many trend-shape correspondences, and if, as I have argued, one of the reasons they are effective is that they can exploit the process of shape perception, people would have to be taught each shape-trend correspondence explicitly, one at a time, to enjoy the full benefits of a graph type.

Pinker (1981) conjectured that mathematically sophisticated readers can deduce novel trend-shape correspondences from the analytic-geometric properties of the graph format, and more generally, that people can induce such correspondences by noticing that certain graph shapes always correlate with certain quantitative trends. This experiment tests people's ability to discover shape-trend correspondences in a graph format when they are not explicitly taught them. Subjects are only taught how the graph conveys information about single values, and are never told about the way that consistency and alternation translate into particular shapes. If the explicit instructions about such correspondences given in Experiment 2 but withheld here are necessary, subjects might show no advantage for detecting global trends of the variable encoded as angle, since they would be encoding the angles of segments separately rather than detecting overall convexity or zigzagging directly. However, to the extent that subjects can induce such correspondences by noticing that when they successfully compute a given trend from a graph, it has a certain shape, they should show a relative advantage for detecting global trends of the variable encoded by angle, as did their counterparts in the previous experiment.

Method

Subjects. Fourteen Stanford University undergraduates participated either for pay or to fulfill an introductory psychology course requirement.

Materials. Stimuli were identical to those of Experiments 1 and 2.

Procedure. The procedure was identical to that of Experiment 2, except that the parts of the instructions that pointed out how consistency and alternation of the angle variable translated into a uniformly convex or uniformly zigzag graph (and that drew attention to consistency and alternation of line length as well) were omitted. Subjects were only told what was meant by the term "consistency" and "alternation", and tested on 3 practice slides; no specific exemplars of consistent, inconsistent, alternating, or non-alternating graphs were shown, and nothing was said about how in general those trends would appear visually.

Results and Discussion

Error rates ranged from 2% to 14% with a mean of 10%, except for four subjects who erred on more than 15% of the trials and whose data were discarded. Extreme response times (which amounted to 2% of the trials) and those from trials on which an error was made were replaced by the mean of the other replications for the appropriate combination of Dimension, Locality and Global Question Type.

As in the previous experiments, questions about single months were answered more quickly than questions about global trends, $F(1,9)=6.45$, $p<.05$, and questions about consistency were answered more quickly than questions about alternation, $F(1,9)=30.76$, $p<.001$ for Global Question Type, $F(1,9)=31.16$, $p<.001$ for Global Question Type X Locality. However, the most interesting finding is a significant interaction between Dimension and Locality, $F(1,9)=5.88$, $p<.05$. As Figure 5 shows, subjects were far faster at detecting single values of the variable conveyed by length than global trends of that variable, whereas they were only slightly faster at detecting single values of the variable conveyed by angle than global trends of that variable. An examination of the points on the right hand side of Figure 5 (which break down global questions into their two types) shows that subjects did answer questions about the consistency of the angle variable more quickly than questions about single values of that variable; it was questions about alternation of the angle variable that they answered more slowly than single values of that variable. In other words, the pattern of results was very similar to that of the subjects in the previous experiment; the only apparent difference is that the questions about alternation of angle were answered so slowly in this experiment relative to questions about consistency that they brought the mean for global questions about the angle variable in general up to a level slightly higher than that for single values (resulting in the line representing the angle variable having a slightly positive slope in Figure 5, rather than a negative slope as before).

INSERT FIGURE 5 ABOUT HERE

Replications was once again significant, $F(3,27)=3.23$, $p<.05$, representing faster responses in later repetitions. In addition, the full 5-way interaction among all the variables of the experiment was significant, $F(3,27)=5.84$, $p<.01$, but since this interaction is confounded in complex ways with the individual graphs presented, it cannot be easily interpreted. Dimension X Locality did not enter into any other significant higher order interaction.

Results from this experiment were compared with those from Experiment 2, in which subjects' attention was drawn to the way in which global trends translated into shapes, by combining the two into a single analysis in which the two experiments are treated as a between subjects factor in a single experiment (to obtain equal sample sizes, one subject was dropped at random from Experiment 2). The between subjects factor, representing the effects of the explicit instructions concerning the shape-trend correspondence, did not interact significantly with the interaction between Dimension and Locality, nor with any of the higher-order interactions encompassing Dimension X Locality.

Error rates are shown in the bar graphs at the bottom of Figure 5, and once again there were more errors for questions about global trends of the length variable than for questions about single values, in contrast to questions about the angle variable, which were answered with more errors when pertaining to single values than global trends. This interaction between Dimension and Locality is only marginally significant, $F(1,9)=3.75$, $p<.09$, but what is important is that the error data show that the pattern of response times observed cannot be attributed to a speed-accuracy tradeoff. The Dimension X Locality interaction in turn interacted with Replications, $F(3,27)=3.74$, $p<.05$, reflecting the fact that the interaction of interest in error rates did not occur on the second replication, and with Response type, $F(1,9)=11.25$, $p<.01$, reflecting the fact that the interaction of interest in error rates did not occur on "no" trials. However, given the small numbers of errors involved and the correspoding likelihood of floor effects, it seems best not to try to interpret these interactions, nor the several others not involving Dimension and Locality.

This experiment shows that even in the absence of explicit instructions about trend-shape correspondences, readers are able to induce or deduce these correspondences from exemplars of graphs displaying and not displaying them. In many ways this is the most important of the 3 experiments, since here the subjects were given no suggestions about the perceptual shortcuts available, and were told only about how the graph format communicated point information. Yet they were still able to exploit these shortcuts, rather than reading point values one-by-one and computing the trend from them. In fact they were best able to do so when the trend was conveyed by a perceptual variable (uniform convexity) that we knew (from Experiment 1) would be easy to encode, showing no advantage when the trend was conveyed by a perceptual variable (uniform zigzagging) whose detection we knew to be more difficult.

General Discussion

These 3 experiments have tested people's ability to read information from a novel type of graph format, concocted so that certain trends of one variable would translate into what were hoped to be recognizable shape properties, whereas those same trends of a second variable would not. At the same time, point values of the first variable translated into a property that was hoped to be difficult to perceive, and point values of the second variable did not, allowing one to argue that subjects' speed at detecting global trends of the variable conveyed by angle is attributable to their exploiting a correspondence between trends and shapes, not to the relative difficulty of processing different visual dimensions in general or of processing questions about trends vs. point values in general. The first experiment tested the conjectures about pattern perception that were necessary to motivate predictions about the relative ease of reading trends from the different dimensions of the graph, and it was found that one of these conjectures (concerning the ease of perceiving uniform convexity) was correct (at least for the stimuli employed) and one of them (concerning the ease of perceiving uniform zigzagging) was at best equivocal. In the second experiment, subjects were taught how graph format conveyed information about point values, and it was also pointed out to them that certain trends corresponded to certain shapes. Subjects seemed to be able to use the latter information effectively, verifying trends directly from the graph rather than reading individual point values one-by-one in cases where the corresponding global shape property was the easily perceivable one (viz., uniform convexity). This was suggested by their taking less time to extract information about global properties of the variable conveyed by angle then single values, even though global questions would otherwise require multiple applications of a perceptual process applied only once for questions about a single value. In the third experiment, subjects were told how the graph symbolized point values but not how trends emerged as shape properties, and their response time and accuracy data suggested that they were able to induce and exploit the trend-shape correspondence in the same way as the subjects explicitly told about it did. Now that we know that readers are capable of exploiting trend-shape correspondences, it is reasonable to conclude that one of the reasons graphs are effective is that they make this possible, apart from any other advantage they might have over tables such as their esthetic value, their ability to group and localize data points where they can be easily found, their economy, and so on.

The set of experiments has several implications. First, it lends credence to the process model of graph comprehension proposed by Pinker (Note 1). According to that theory readers comprehend a graph by translating elements in their internal representation of the visual appearance of the display into conceptual "messages", guided by a schema for that graph format which defines possible translations at several levels of generality (i.e., separately for absolute point values, extremeness of point values, differences and trends of a variety of sorts; see Pinker, Note 1, for detailed proposals about the nature of the internal visual representations, conceptual messages, graph schemas, and translation processes that are used in graph comprehension). That theory predicted that the difficulty with which a given person will extract a given piece of information from a given graph format will depend on three things: how the format translates the information into a visual pattern, on how easily the visual system can encode that pattern, and on whether the reader knows the correspondence between conceptual information and visual pattern. If a piece of information is displayed as an easily encodable pattern and the reader knows the correspondence, the reader can extract information by making a simple mental substitution between internal symbols for visual and conceptual entities; if any of these preconditions does not hold, the reader must encode several pieces of more local information and then deduce the desired information through a series of mental comparisons or even mental arithmetic, which are presumably more time-consuming and error-prone.

These conclusions have important implications for the design of novel graph formats (e.g., those discussed in Wainer & Thisson, 1981), and for the choice of which familiar format to use for a particular set of data. Graph formats will in general not be easy or difficult across the board, but will be more or less difficult depending on the type of information that the reader must extract. A graph designer must decide which types of information are the ones worth communicating most clearly in the graph, and check to see what sort of visual discrimination the reader must make to extract that information from a candidate graph format. In many cases, this method may lead to conclusions that contradict conventional wisdom about graph formats. For example, a common injunction to graph designers is to avoid line graphs when the abscissa corresponds to a nominal or ordinal scale. However, according to the reasoning sketched above, line graphs might in fact be far preferable to bar graphs or tables regardless of the scale type, since only with line graphs do complex patterns of data translate into perceivable contours (as in MMPI "profiles"), and interactions between variables translate into patterns of parallelism, convergence, or intersection of contours. Whether or not this advantage outweighs the naturalness of associating discrete visual objects with nominal scales is of course an empirical question, one of many suggested by the present approach to graph comprehension.

One hopes that the choice of which graph format to use will often be dictated by existing psychological research on visual pattern perception. Kosslyn and Pinker (under review) and Pinker (1981) argue that in fact quite a large body of research on visual perception and memory is relevant to making such choices, including research on visual discrimination, magnitude estimation, grouping, integrality, shape encoding, coordinate frame selection, and schema formation. If so, the general moral of this research is a happy one: the study of graph comprehension and effective graphic design need not be an autonomous area of practical research, but can to a large extent be parasitic on what we already know about visual perception.

Footnote

This research was supported by NIE grant 400-79-0066 to CSI/Datacrown, Inc. and NSF grants BNS 81-14916, BNS 82-16546, and BNS 82-09450 to the author. I am grateful to Diane Gong for designing the stimuli and question sets and running the subjects, to Cathryn Downing and Laura Vanderpool for programming and other assistance, to Karin Stromswold for performing the statistical analyses, and to Stephen Kosslyn, Don Hoffman, Karin Stromswold, and Bill Simcox for helpful discussions. Reprints may be obtained from the author at the Department of Psychology E10-018, Massachusetts Institute of Technology, Cambridge, MA, 02139.

Reference Notes

Note 1. Pinker, S. A theory of graph comprehension. (Center for Cognitive Science Occasional Paper #10). Cambridge, MA: Massachusetts Institute of Technology, 1981.

Note 2. Hoffman, D.D. Representing shapes for visual recognition. Unpublished doctoral dissertation, Massachusetts Institute of Technology, 1983.

References

Attneave, F. Some informational aspects of visual perception. Psychological Review, 1954, 61, 183-193.

Carter, L.F. The relative effectiveness of presenting numerical data by the use of tables and graphs. In P.F. Fitts (Ed.), Psychological research on equipment design. Washington: U.S. Government Printing Office, 1947.

Cooper, L.A. & Shepard, R.N. Chronometric studies of the rotation of mental images. In W.G. Chase (Ed.), Visual information processing. New York: Academic Press, 1973.

Cooper, L.A., & Shepard, R.N. Mental transformations in the identification of left and right hands. Journal of Experimental Psychology: Human Perception and Performance, 1975, 1, 48-56.

Culbertson, H.M. & Powers, R.D. A study of graph comprehension difficulties. Audio Visual Communication Review, 1959, 7, 97-100.

Kosslyn, S.M. & Pinker, S. Understanding charts and graphs: A case study in applied cognitive science. Book manuscript submitted for publication.

Rock, I. Orientation and form. New York: Academic Press, 1974.

Schutz, H.G. An evaluation of methods for presentation of graphic multiple trends: Experiment III. Human Factors, 1961, 3, 108-119(a).

Schutz, H.G. An evaluation of formats for graphic trend displays: Experiment II. Human Factors, 1961, 3, 237-246(b).

Wainer, H., & Thissen, G. Graphical data analysis. Annual Review of Psychology, 1981, 32, 191-241.

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

Figure Captions

Figure 1. Examples of stimuli used in the experiments. 1(a) and 1(b) are "consistent" and "inconsistent" for length, respectively; 1(c) and 1(d) are "consistent" and "inconsistent" for angle; respectively; 1(e) and 1(f) are "alternating" and "non-alternating" for length, respectively; 1(g) and 1(h) are "alternating" and "non-alternating" for angle, respectively.

Figure 2. One of the displays shown to subjects to illustrate short and long line segments and small and large angles.

Figure 3. Mean response times to questions about lengths and angles of individual stimulus segments and about global properties of the stimulus. Questions about global properties are subdivided into questions about consistency and alternation, and the data from the 2 types are plotted as separate points. Error rates are plotted as bar graphs at the bottom of the figure.

Figure 4. Mean response times to questions about single values and global trends of the variables conveyed by length and angle of graph segments. Subjects were told about how the graph conveys point values and also about how global trends correspond to specific shapes. The data are plotted in the same way as in Figure 3.

Figure 5. Mean response times to questions about single values and global trends of the variables conveyed by length and angle of graph segments. Subjects were told only about the way in which the graph conveyed point values. Data are plotted as in Figures 3 and 4.

FIGURE 1

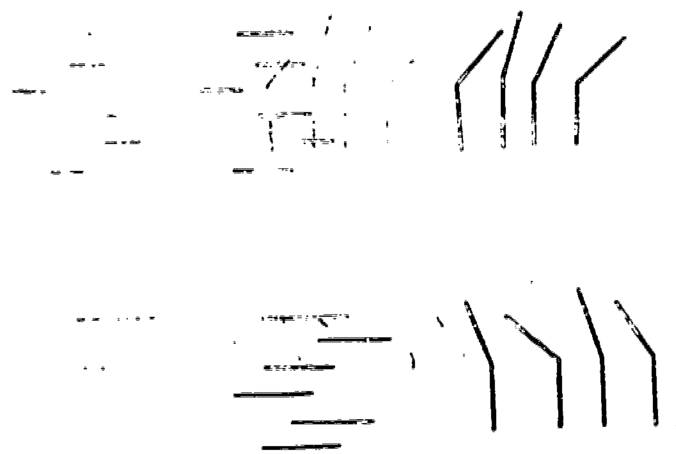


FIGURE 2

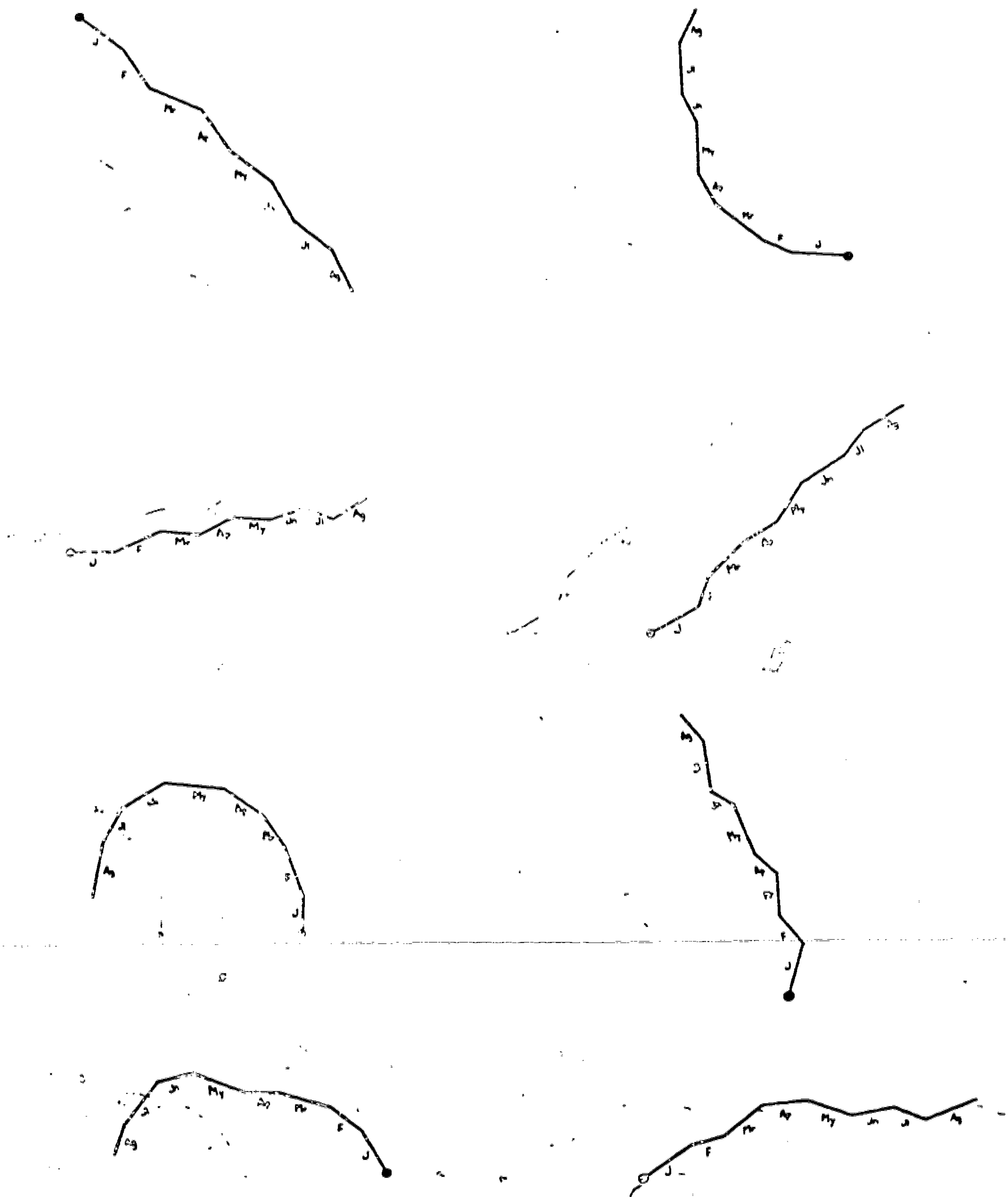


FIGURE 3

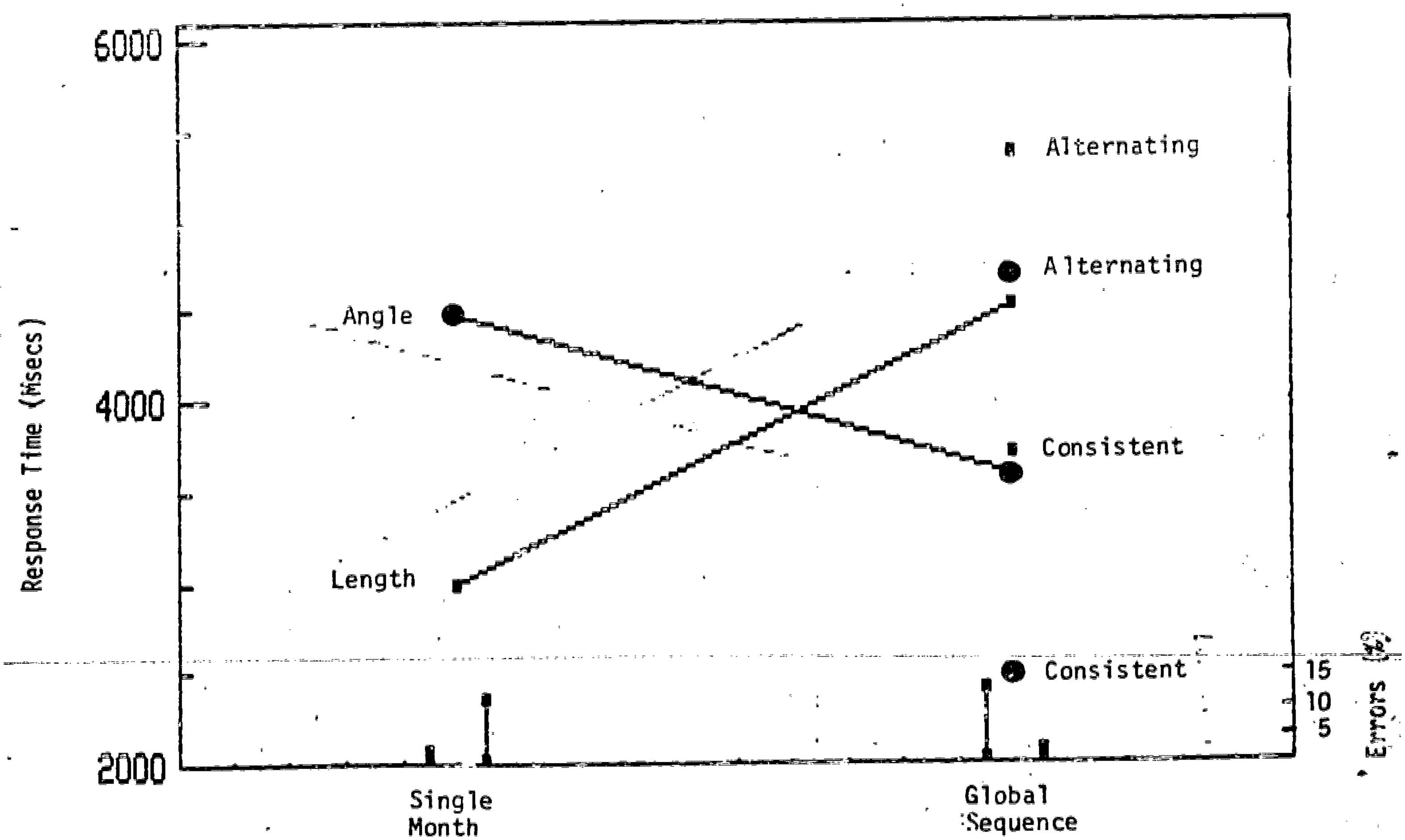
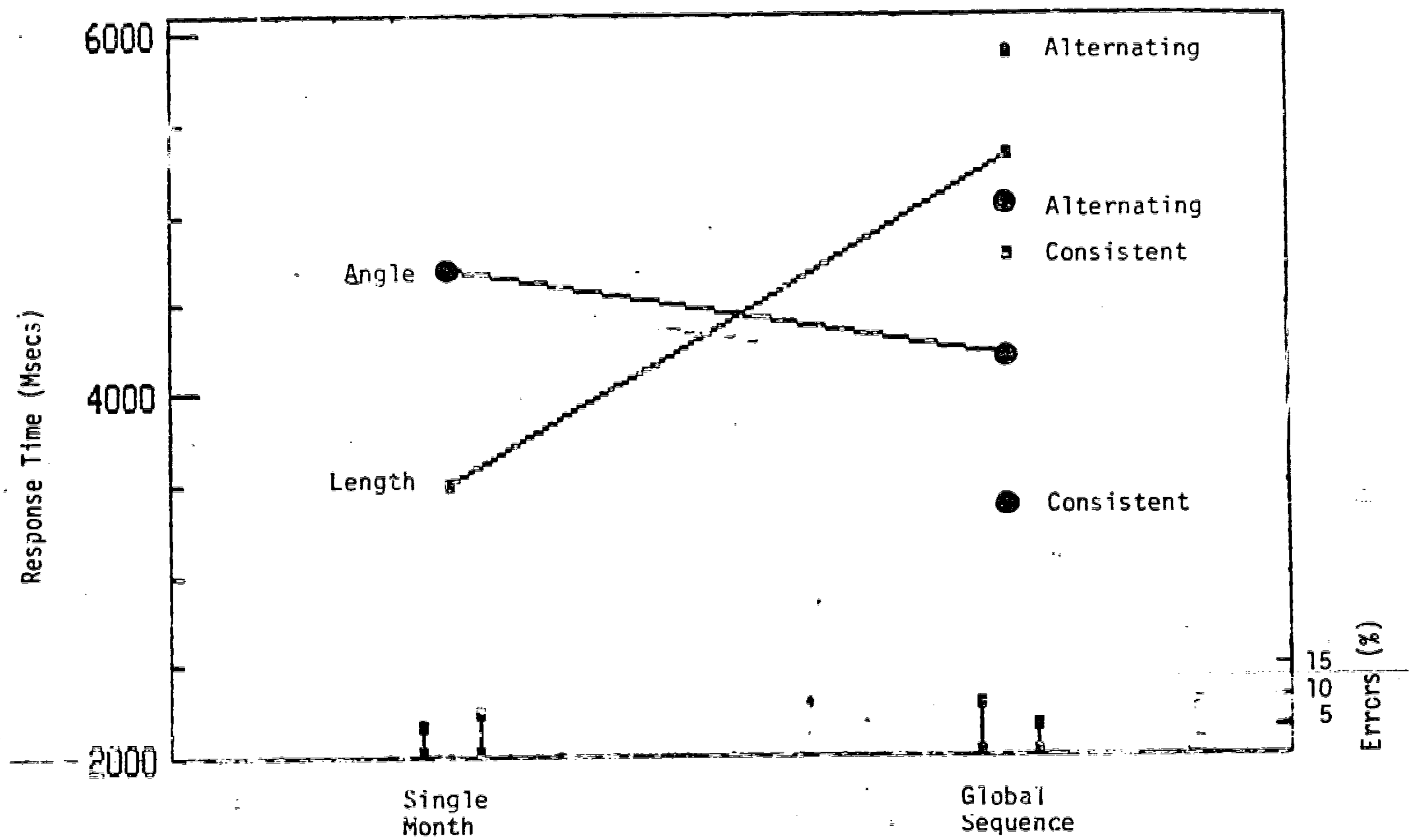
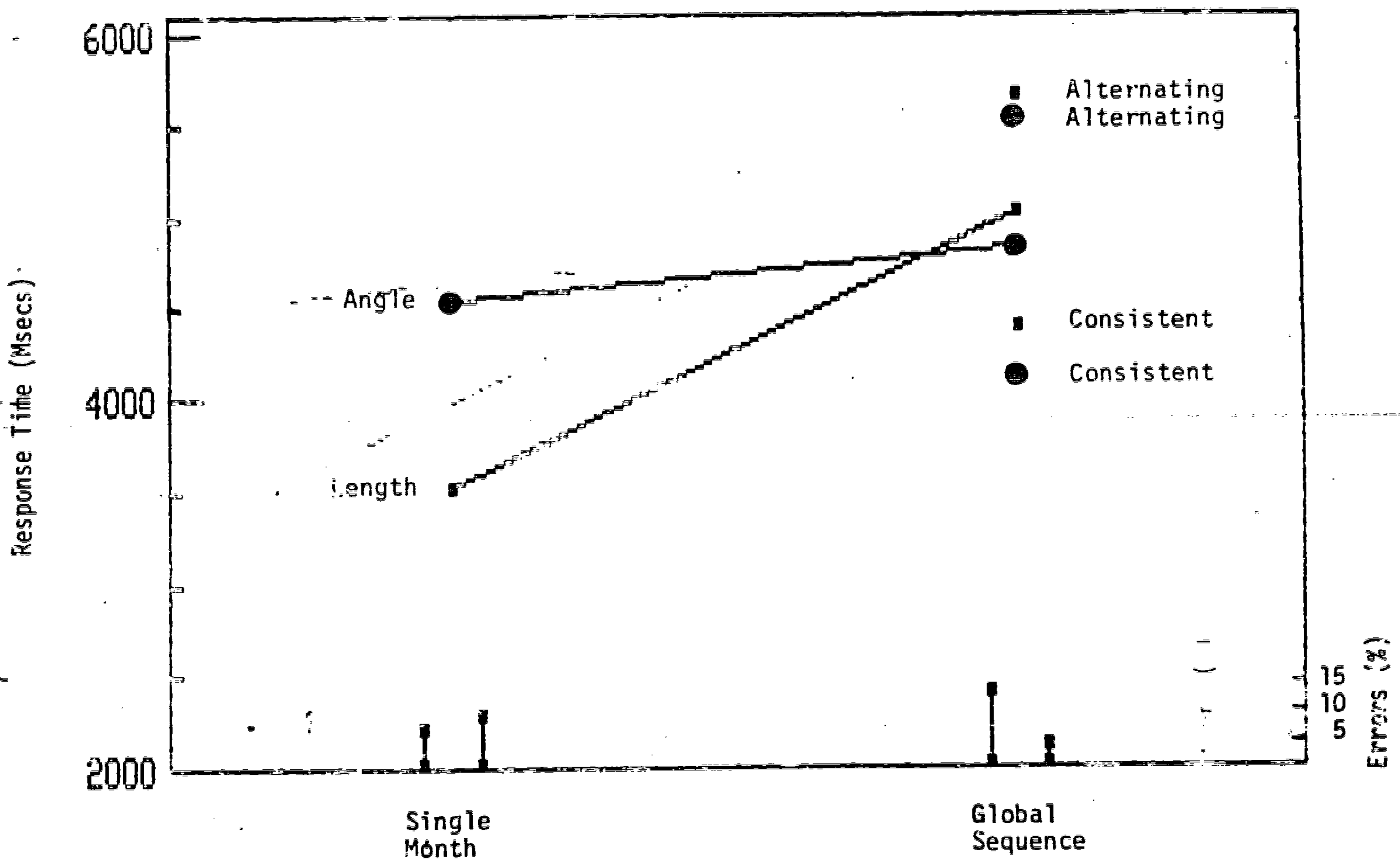

6000
4000
2000
Response Time (Msecs)
Alternating
Alternating
Angle
Consistent
Length
Consistent
15
10
5
Errors (%)
Single
Month
Global
Sequence

FIGURE 4


6000
4000
2000
Response Time (Msecs)
Alternating
Alternating
Consistent
Angle
Length
Consistent
15
10
5
Errors (%)
Single
Month
Global
Sequence


6000
4000
2000
Response Time (Msecs)
Alternating
Alternating
Consistent
Consistent
Angle
Length
15
10
5
Errors (%)
Single
Month
Global
Sequence