An Information-Processing Analysis of Graph Perception

DAVID SIMKIN and REID HASTIE*

Recent work on graph perception has focused on the nature of the processes that operate when people decode the information represented in graphs. We began our investigations by gathering evidence that people have generic expectations about what types of information will be the major messages in various types of graphs. These graph schemata suggested how graph type and judgment type would interact to determine the speed and accuracy of quantitative information extraction. These predictions were confirmed by the finding that a comparison judgment was most accurate when the judgment required assessing position along a common scale (simple bar chart), had intermediate accuracy on length judgments (divided bar chart), and was least accurate when assessing angles (pie chart). In contrast, when the judgment was an estimate of the proportion of the whole, angle assessments (pie chart) were as accurate as position (simple bar chart) and more accurate than length (divided bar chart). Proposals for elementary information processes involving anchoring, scanning, projection, superimposition, and detection operators were made to explain this interaction.

KEY WORDS: Cognitive processing; Schemata; Statistical graphics.

1. INTRODUCTION

Research on the perception of graphs has been dominated by practical questions concerning the efficiency of various types of graphs as sources of information to perform several common judgment and inference tasks. Pioneering researchers (e.g., Eells 1926) tended to simplify issues and propose blanket empirical hypotheses; for example, bar charts are better graphs than pie charts. Recently, researchers (e.g., Cleveland and McGill 1984, 1985; Follettie 1986) have begun to focus on the nature of the processes that operate when people decode the information represented in a graph. The approach of Cleveland and McGill begins with the isolation of the particular aspects that encode the quantitative information. Identification of these elementary codes of graphs has given rise to more subtle hypotheses stated in terms of interactions between the aspect of the graph that must be decoded and the judgment task; the position of elements on a graph will support more efficient performance of judgment tasks requiring the estimation of absolute quantitative magnitudes than will the same information encoded as an angle, but the two codes will not differ in proportional estimation tasks. This research also emphasizes the need to consider not only the graphs but also the human beings for whom the graphs are intended. Thus we should draw on modern human information-processing theories (Kosslyn 1985; Pinker 1981). This article is an exercise in this recent tradition. We report three experimental studies of graph perception that demonstrate some interactions between elementary code and judgment task, on which speed and accuracy measures depend.

* David Simkin is a graduate student and Reid Hastie is Professor, both in the Department of Psychology, Northwestern University, Evanston, IL 60201. The authors wish to thank William Revelle and James Hall for their generous investment of time and equipment.

2. SURVEY STUDY

Our guiding precept was that the usefulness of a graph would depend on the judgment task that was being performed. Our first empirical study was a survey of intelligent but unsophisticated (undergraduate) respondents' reactions to several types of graphs. The methodological assumption was that spontaneous judgments would provide a clue to the tasks that could be performed most efficiently for the graph type. Two hundred undergraduates were shown bar charts, divided bar charts, pie charts, and line graphs and asked to provide written summaries of the information in each display.

When presented with a bar chart, most respondents spontaneously made comparisons between the absolute lengths of the bars (referred to as *comparison judgments*). In contrast, when presented with a pie chart, most people compared individual slices with the whole, making proportion-of-the-whole judgments (referred to as *proportion judgments*). Based on these findings, we predict that length and position, the elementary codes of bar charts, would yield superior performance in a comparison judgment task. Angle, however, the code used in a pie chart, would support performance in proportion judgments. Thus we are predicting an interaction between elementary code (position or length vs. angle) and judgment task (comparative vs. proportional judgment) on accuracy and speed measures of performance. An experiment was conducted to test the hypothesis.

3. EXPERIMENT 1

3.1 Design

There were two judgment conditions in the experiment. Subjects viewed graphs and either made a discrimination judgment followed by a comparison judgment or made a proportion judgment. All graphs were presented on a cathode-ray tube (CRT) screen controlled by a microcomputer. Subjects were undergraduate students enrolled in an introductory psychology class participating for course credit. There were 40 subjects, run in groups of 1–4, in each of the judgment conditions.

In the comparison judgment condition, each subject saw the three types of graphs depicted in Figure 1. All bars were 54 mm long, and the pie had a diameter of 25 mm. A division of each bar or pie was marked with a dot, and subjects were asked to judge what percentage the smaller division was of the larger. For the simple bar chart depicted in the top panel of Figure 1, the dot always appeared in the lowest division of the bar. To visually compare these

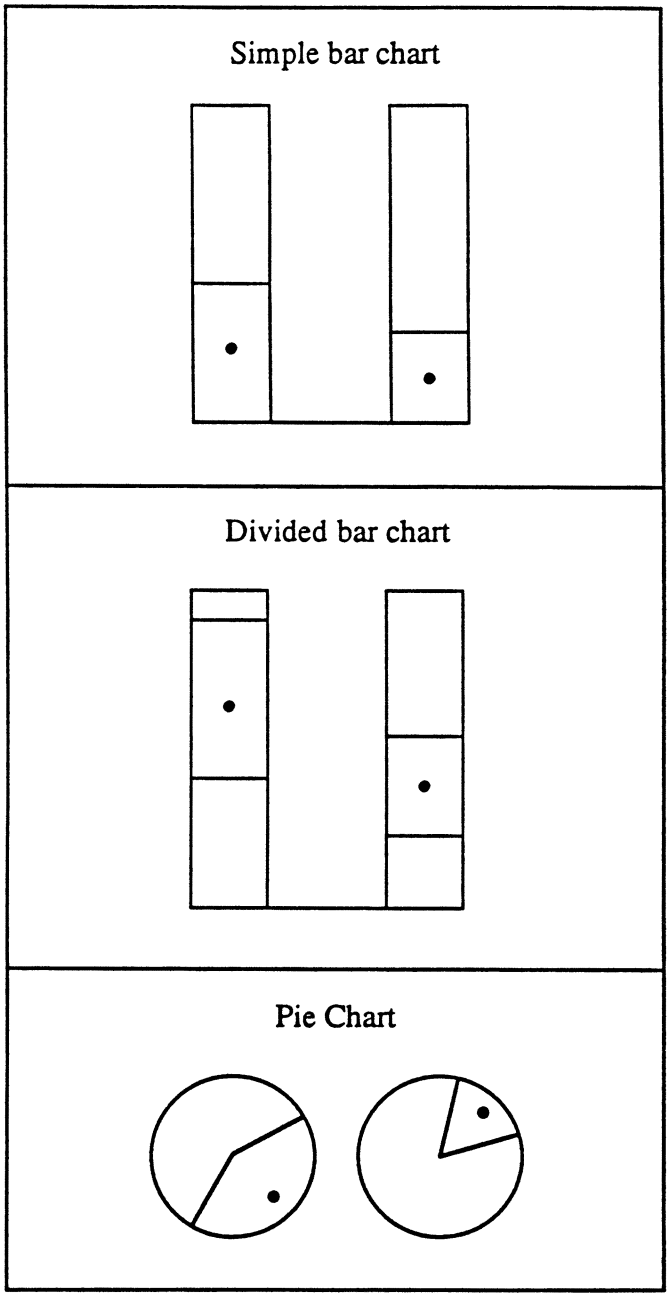

Figure 1. Graphs From Experiments.

divisions the subject must judge position along a common scale. The marked divisions of the divided bars in the middle panel never align with the bottom or top of the bar, and the bottom or tops of the two divisions never line up. Here the elementary code is length. Each pie chart contains one division that is marked with a dot, and the subject must judge angles.

Thirty pairs of values were randomly generated by a uniform random-number generator with the constraints that no division would be larger than 47% of the bar or pie or smaller than 3%, and the ratio of the values had to be between 5% and 95%. For the values that actually arose, the ratios ranged from 8.45% to 94.76%. Each set was depicted as a pie chart and both types of divided bar charts, yielding 90 graphs. The presentation order of the graphs was randomized for each subject.

Subjects were seated at a computer terminal and the task was explained to them. There were 10 practice trials during which all questions were answered. The subjects then worked uninterrupted through the 90 experimental trials. During all trials the subjects were asked to sit at the terminal with their left forefinger on the "Z" key and their right forefinger on the "/" key. They began each trial by pressing the space bar with their thumb, causing the graph to be displayed. As quickly as possible, they were to indicate if the division on the left or right was smaller by pressing the left or right key (the discrimination judgment). The graph was replaced with a masking grid after one second, whether they had responded by then or not. Presentation time was controlled in an attempt to limit noise in the data. Pilot work had indicated that the subjects could achieve good performance levels with this (1 second) stimulus duration.

After the discrimination response, the text signaled them to enter a number from 1 to 100 that was their judgment of the percentage the smaller division was of the larger (the comparison judgment). Subjects had been given the following instructions: "You are to make a quick visual judgment and not try to make precise measurements, either mentally or with a physical object such as a pencil or your finger. You should enter this estimate as soon as you make it, but don't force yourself to respond very quickly as you did in choosing which division was smaller." After making this response the subjects were instructed to prepare for the next trial by placing their fingers back on the response keys and pressing the space bar. Responses and reaction times for the discrimination response as well as the comparison response were recorded. The discrimination reaction time was the length of the interval from the onset of the graph until the "Z" key or "/" key was pressed. The comparison reaction time was the interval from when the discrimination response was made until the first digit of the estimate was entered. To measure comparison accuracy we used

$$|\text{judged percent} - \text{true percent}|.$$

All data from trials on which the discrimination response was in error were eliminated from the analysis.

The procedure for the proportion judgment was quite similar. The stimuli consisted of single bars or pies of the type shown in pairs in Figure 1. The divisions ranged from 4% to 47.33%. The subjects pressed the space bar, and the stimulus was shown and then replaced with a masking grid after .5 seconds. This stimulus duration was set at half the duration of the comparison judgment condition. The task was to judge the percentage that the division represented of the whole bar or pie. Responses in this proportion task were made by entering a number from 1 to 100. The proportion reaction time is the length of the interval from the presentation of the graph until the first digit of the estimate was entered. The accuracy measure was the same as defined for the comparison judgment. Again, the quick visual judgment was requested.

It was our expectation that the results of the comparison condition would replicate the experiments of Cleveland

and McGill (1984, 1985). Subjects would be most accurate decoding position, second most accurate judging length, and least accurate judging angle. In addition, the reaction time data provide a second criterion by which to order the codes. In contrast, on a proportion judgment we expect the superiority of position and length over angle to disappear and possibly even reverse. Finally, the data could provide a basis for ordering the elementary codes when the discrimination response is being made. Cleveland and McGill found few errors on the discrimination decision, and we expect good performance with a stimulus duration of one second. Thus the accuracy of the discrimination decision does not provide a basis for the ordering of these codes. The reaction time of the discrimination judgment is the primary basis for ordering.

In the full design, each of the three judgments (discrimination, comparison, and proportion) was made for three elementary codes—position along a common scale, length, and angle—which are the prominent codes for the simple bar chart, divided bar chart, and pie chart, respectively.

3.2 Data Analysis

Recall that there were 40 subjects in each judgment condition. A subject made 90 judgments, 30 for each elementary code. We estimated the location of the distribution of the 40 reaction time or absolute error values for each judgment by the midmean, a robust estimate of location (Mosteller and Tukey 1977). The mean of the 30 midmeans for each elementary code by judgment type combination is presented in Table 1. The data for the discrimination judgments, comparison judgments, and proportion judgments were each analyzed separately in a 3 (elementary code) × 10 (blocks) design. The 30 true percentages were blocked into 10 groups of size 3 to pull out variance attributable to the magnitude of the true percentages. The residual variance after removal of the variance attributable to elementary code, blocks, and the elementary code by blocks interaction provided the error term and corresponds to the within-groups mean squared. The elementary-code main effect is of primary interest in this discussion. Later analyses will bear on the blocks and elementary code by block effects. All paired comparisons are the results of Newman–Keuls tests.

Table 1. Mean Discrimination Reaction Times, Comparison Errors, Comparison Reaction Times, Proportion Errors, and Proportion Reaction Times for Experiment 1

	Elementary code		
	Position	Length	Angle
Discrimination reaction time (msec)	790_a (2.58)	900_b (4.58)	943_b (7.25)
Comparison error	5.93_a	6.88_b	7.74_c
Comparison reaction time (msec)	$2{,}520_a$	$2{,}649_b$	$2{,}667_b$
Proportion error	3.29_a	3.96_b	3.18_a
Proportion reaction time (msec)	$2{,}389_a$	$2{,}623_b$	$2{,}882_c$

NOTE: The numbers in parentheses after the discrimination reaction times are the percentage of errors out of 1,200 for each elementary code. Means on a row not sharing the same subscript are significantly different by the Newman–Keuls test ($p < .05$, two-tailed).

Elementary code was a significant factor in the analysis of the discrimination reaction time [$F(2, 60) = 16.37$, $p > .001$]. The reaction time for the discrimination judgment of the position code was significantly faster than that of the other two codes. In Table 1, the number in parentheses after the discrimination reaction time is the percentage of errors out of 1,200 for each elementary code (subjects made few errors).

Elementary code was a significant factor in the comparison absolute-error analysis [$F(2, 60) = 7.49$, $p < .01$] and in the comparison reaction-time analysis [$F(2, 60) = 3.49$, $p < .05$]. The comparison absolute errors ordered the three elementary codes as expected. Position yielded the most accurate comparison judgments, followed by length and then angle, which was least accurate. The comparison reaction-time data show the same ordering, but the difference between length and angle is not significant.

Not surprisingly, elementary code accounted for a significant amount of variance in both the proportion absolute-error analysis [$F(2, 60) = 11.47$, $p < .001$] and the proportion reaction-time analysis [$F(2, 60) = 18.79$, $p < .001$]. There was a reversal in the ordering of the codes by accuracy for the proportion judgment from the comparison judgment. Length led to significantly fewer accurate judgments than the other two codes, which did not differ from each other. Although the angle judgments were the most accurate, they also took the most time to make. It is possible then that the proportion judgment accuracy advantage of angle over length would be reduced if the subjects were not trading off accuracy for speed.

When making a comparison judgment, the position along a common-scale elementary code (simple bar chart) produced more accurate judgments than length (divided bar chart), and angle judgments (pie chart) were least accurate. The discrimination reaction time ordered the codes in the same way. In contrast, the ordering of codes by accuracy when a proportion judgment was made was angle (pie chart) = position along a common scale (simple bar chart) > length (divided bar chart). The reaction time for this judgment, however, shows the same ordering as the comparison judgment. Essentially, we obtained the predicted elementary code by judgment-task interaction. A second experiment was conducted with more natural presentation conditions: when the display did not terminate quickly.

4. EXPERIMENT 2

The experiment uses the same graphs, design, and procedure as the prior experiment with the exception that the graphs were left on the CRT screen until the subjects entered their estimates. Table 2 presents the means of the midmeans for each elementary code by judgment combination. The analysis again revealed elementary code to be a significant factor for all dependent measures [for discrimination reaction time, $F(2, 60) = 11.83$ and $p > .001$; for comparison absolute error, $F(2, 60) = 6.24$ and $p > .01$; for comparison reaction time, $F(2, 60) = 6.78$ and $p > .01$; for proportion absolute error, $F(2, 60) = 3.99$ and $p > .05$; and for proportion reaction time, $F(2, 60) =$

Table 2. Mean Discrimination Reaction Times, Comparison Errors, Comparison Reaction Times, Proportion Errors, and Proportion Reaction Times for Experiment 2

	Elementary code		
	Position	Length	Angle
Discrimination reaction time (msec)	684_a (3.50)	813_b (5.92)	883_b (5.83)
Comparison error	5.11_a	6.00_b	6.74_b
Comparison reaction time (msec)	$3{,}689_a$	$4{,}140_b$	$4{,}093_b$
Proportion error	2.48_a	2.78_a	2.47_a
Proportion reaction time (msec)	$4{,}867_a$	$5{,}872_b$	$5{,}957_b$

NOTE: The numbers in parentheses after the discrimination reaction times are the percentage of errors out of 1,200 for each elementary code. Means on a row not sharing the same subscript are significantly different by the Newman–Keuls test ($p < .05$, two-tailed).

32.19 and $p > .001$]. The direction of the differences replicated Experiment 1 completely, but several of the paired comparisons did not reach significance. Experiment 2 replicated the elementary code by judgment-task interaction from Experiment 1. Note also that the subjects took more time in Experiment 2 than in Experiment 1 to make their proportion and comparison judgments, resulting in more accurate responses.

Our secondary goal is to present an information-processing theoretical analysis of our graph-perception findings. We will return to a fine-grained analysis of the data after we introduce our theoretical approach.

5. THEORY: ELEMENTARY PROCESSES

We want to develop a vocabulary of elementary mental processes that can be combined to build information-processing models of performance in common graph-perception tasks. The cognitive analysis of graph perception started in the work of Bertin (1983), Cleveland and McGill (1984, 1985), Follettie (1986), and Pinker (1981, 1983). These investigators have addressed two key cognitive questions: How is the information from a graph represented mentally? What mental processes intervene between early vision and the establishment of the mental representation, operate on the representation to infer non-obvious properties, and operate on the representation and the inferences to generate a task-appropriate response? Our research concentrates on the nature of processes that operate on the representation of the graph (referred to as the image) to perform simple magnitude estimation and larger/smaller or proportional size-comparison tasks. Like other studies in the extant literature, our conclusions concerning elementary processes in graph perception should be taken as serious speculation with some empirical support that merits further experimental evaluation.

Table 3. Elementary Codes and an Ordering

Rank	Elementary code
1	Position along a common scale
2	Position along nonaligned scales
3	Length
4	Angle
4–10	Slope
6	Area
7	Volume
8	Density (amount of black)
9	Color saturation
10	Color hue

There are several lists of elementary perceptual processes that provide vocabularies with which to write algorithms to account for performance in simple graph-perception tasks. Cleveland and McGill (1984, 1985, in press) have isolated 10 elementary codes that correspond to geometric and textural aspects of graphs. We judge these aspects to extract visually information about the relative magnitudes of quantities shown on a graph. These elementary codes are listed in Table 3 with their proposed ranking based on the accuracy with which they are judged. They discuss the most common types of graphs, hypothesize which elementary codes are prominent in perception of each type, and predict relative error rates in perception across some example graphs. There has been no formal experimental verification of the codes of Ranks 8–10, so their ranking relative to one another is conjectural, but informal evidence places them below the other codes.

Cleveland and McGill (1984) assessed accuracy of per-

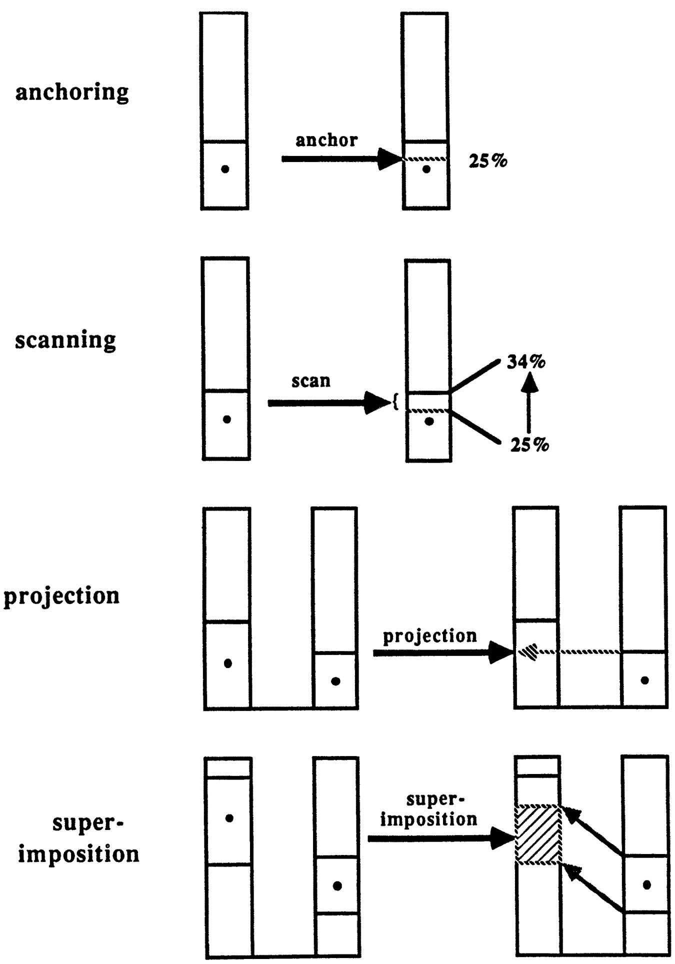

Figure 2. Schematic Summaries of Proposed Elementary Mental Processes That Can Be Combined to Explain Performance in the Experimental Tasks.

formance in one graph-perception judgment for several elementary codes. People performed a comparison judgment with bar charts, divided bar charts, and pie charts. The key elementary code for each graph type was position along a common scale, length, and angle, respectively. For the comparison judgment, they found that the hypothesized difficulty ordering was reflected in subjects' error rates, with position most accurate and angle least accurate.

Follettie (1986) distinguished between measurement (e.g., responding with an absolute length value), discrimination (e.g., indicating which bar is longer), and comparative estimation (e.g., responding with a percentage indicating the relative length of two bars) tasks. He noted that the taxonomy is preliminary, nonexhaustive, and general. Directions for future development are implied in his breakdown of the global-measurement task into four possible types, according to the nature of the magnitude extraction process (direct reading from the image vs. interpolation) and the nature of the scale used to calculate response values [organic (values integral to the image representation) vs. yardstick (a scale external to the image

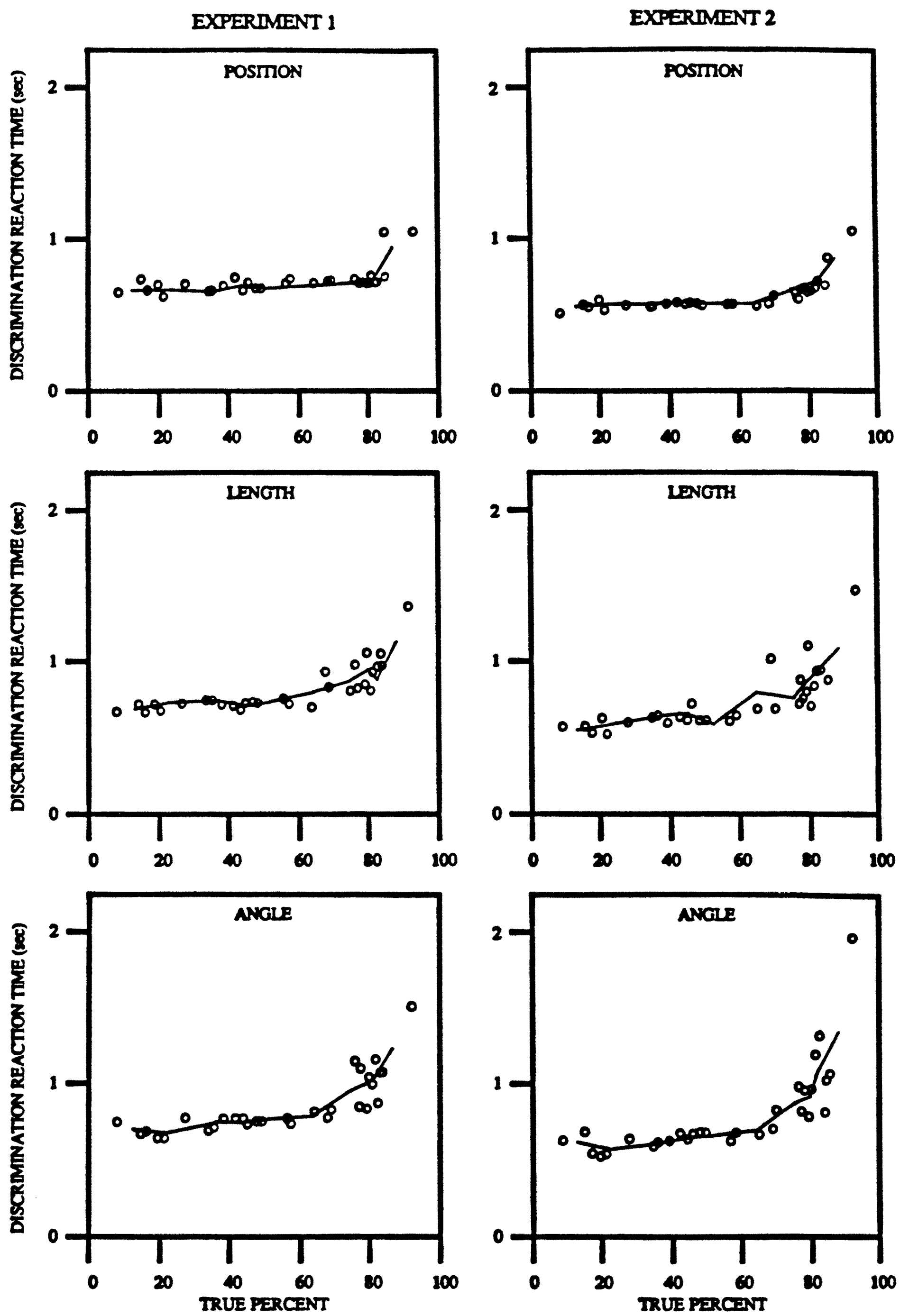

Figure 3. Midmeans of Discrimination Reaction Times Against True Percentages for the Three Elementary Codes in Experiments 1 and 2.

must be applied to the image to infer a quantitative response value)].

The contrast between Follettie's focus on the judgment task and Cleveland and McGill's focus on elementary codes suggests an integrative speculation: The ordering of codes by accuracy (from Cleveland and McGill) may depend on the analytic task (Follettie) performed by the graph reader. Of course, this is a restatement of our basic precept that graph type and judgment task will interact to determine performance.

Pinker (1981) provided a theoretical framework that guides the development of hypotheses about elementary code by analytic task effects on performance accuracy. The key is Pinker's concept of graph schema. A schema is a generic cognitive structure, learned from past experience in a domain and stored in long-term memory, that guides a perceiver in organizing incoming information into a complex knowledge representation (e.g., Hastie 1981; Rumelhart 1984; Schank and Abelson 1977). The concept is often described as a generic mental scaffold on which we hang new information: The schema prescribes relationships among ideas, tells the preceiver what is missing, and provides general "default" values for missing information that "fill gaps" and set up expectations for as yet unperceived objects and events. Pinker's graph schemata define general classes of graphs (e.g., bar charts, divided bar charts, pie charts, etc.), and they enable the perceiver to translate information from the retina into an orderly conceptual representation (image). These schematically organized images tend to be especially useful in generating answers to conceptual questions that are defining characteristics of graph perception tasks.

Our initial survey study was an attempt to identify some of the graph schemata that are shared by moderately sophisticated graph perceivers in our culture [see Bower, Black, and Turner (1979) for an analogous study of schemata for the comprehension of narrative discourse]. The results of this survey study implied that there are distinct, consistent schemata associated with bar charts and pie charts. The lack of clear interrespondent agreement in the responses to divided bar charts and line graphs leaves us uncertain about the nature of generic mental representations for these graph types.

Pinker (1981) also outlined a taxonomy of perception processes that is closest to the list of computational operators that would be sufficient to write a cognitive model for graph perception. Pinker's system is described at a general level, as classes of procedures, rather than at the level of specific information-processing operators. He defines four classes of processes: (a) Match processes that determine which category of graph (e.g., bar chart, pie chart, divided bar chart) is presented and that activate schemata for top-down processing of the graph; (b) message assembly processes that translate information from early visual processes into a conceptual image guided by the active graph schema; (c) integration processes that infer information using retinal inputs and the generic graph-schema information as a premise to derive relationships and values that are missing from the mental image (e.g., compute a larger-than/smaller-than relation of two pie-chart divisions); (d) inferential processes that use the mental image and the generic graph schema as a premise to derive new relationships and values (e.g., calculating the relative proportion that a smaller division is of a larger division after both have been encoded in the image).

Our list of processes is based on the systems proposed by Pinker (1981) for graph perception, by Kosslyn (1980) for processing mental images, and by Ullman (1984) for visual scene perception. In our experiments, simple charts are presented and responses occur while the display is still visible or while an almost veridical image is still available to consciousness. Our processes are simple anchoring, scanning, projection, superimposition, and detection (larger/smaller) operators that are intuitively plausible and directly analogous to the Kosslyn and Ullman processes. The elementary processes can be combined in lists that would be sufficient instructions to another person to perform the experimental tasks. A next step would be to embed the processes in a theoretical language to write computer-program models to perform the tasks. Kosslyn's image-processing model is an example of such a theoretical program (Kosslyn 1980, pp. 112–173).

Figure 2 presents schematic summaries of each of our hypothetical processes:

Anchoring. Segmenting a component of the image that is a standard for some estimate. These segments will act

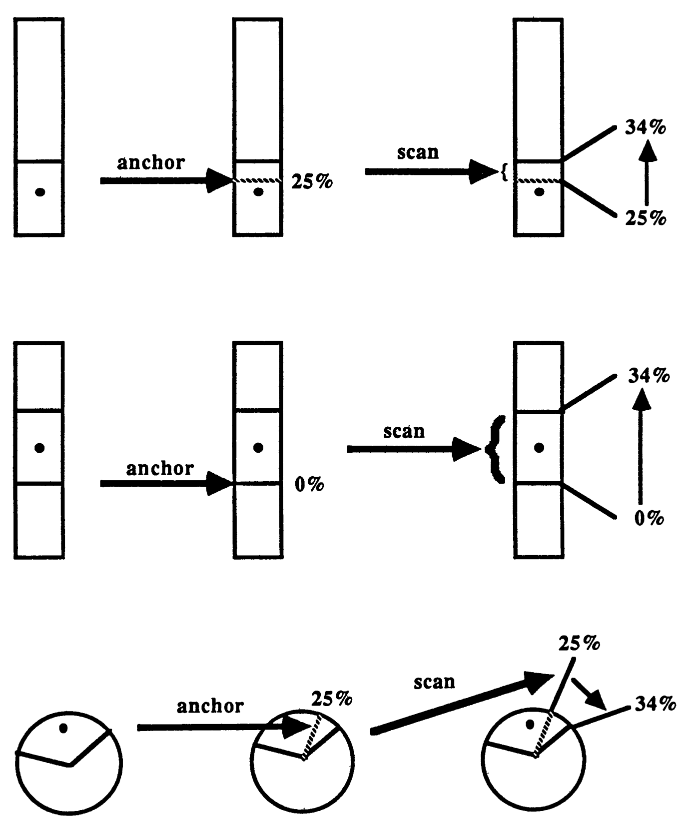

Figure 4. Proposed Sequence of Elementary Mental Processes to Explain Performance in the Proportion Judgment Task for Position (top panel), Length (middle panel), and Angle (bottom panel).

as anchors, providing an initial value that is adjusted to yield the estimate (see Kosslyn's "find" process and Ullman's indexing and marking operations).

Scanning. "Sweeping" across the distance in the image being estimated. Some characteristic of the scan, possibly the duration, acts as a "tape measure" for the estimate (see Kosslyn's "scan" process and Ullman's shifting-the-processing-focus operation). The accuracy decreases as the scan distance required to make the estimate increases. This may be an inherent characteristic of the process or may be due to a tendency to be conservative in the adjustment from the initial anchor.

Projection. Sending out a ray from one point in the image to another. This process is most accurate when the projection is oriented horizontally or vertically (see Kosslyn's "scan" and "compare" processes and Ullman's ray-intersection operation).

Superimposition. Moving elements of the image to a new location often so that the elements overlap another component of the image (see Kosslyn's "rotate" process

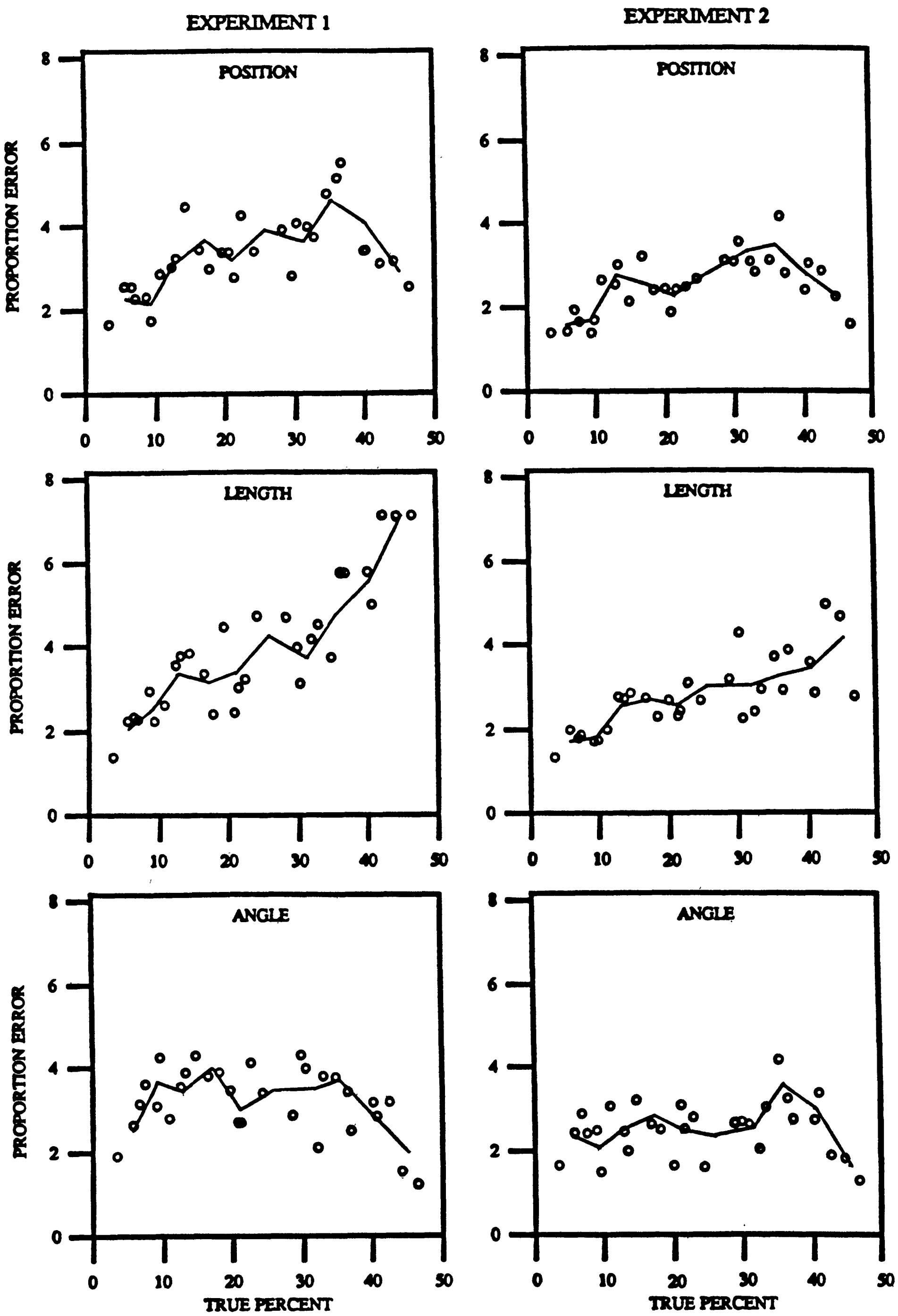

Figure 5. Midmeans of Absolute Proportion Errors Against True Percentages for the Three Elementary Codes in Experiments 1 and 2.

and Ullman's boundary-trace operation). It is used when the simpler and more accurate projection process is not adequate.

Detection Operators. These operators detect differences in the sizes of two components in the image (see the Kosslyn and Ullman comparison operations). The analysis is much simpler than that of the previous processors, and it returns simple dichotomous decisions such as larger/smaller. This simpler analysis allows for fast responses.

We now suggest how these processes may be combined to allow performance of the experimental tasks. Much of our speculation is based on qualitative analysis of the absolute-error data plotted against the true percentages from the experiments. We now turn to this analysis.

6. DETAILED ANALYSIS OF EXPERIMENTS

6.1 Discrimination Judgment Task

Figure 3 shows plots of the 30 midmeans of the discrimination reaction times for each elementary code from both experiments. Superimposed on each plot are curves indicating the cell means when the 30 true percentages were blocked into 10 groups of size 3. For each of the 10 points,

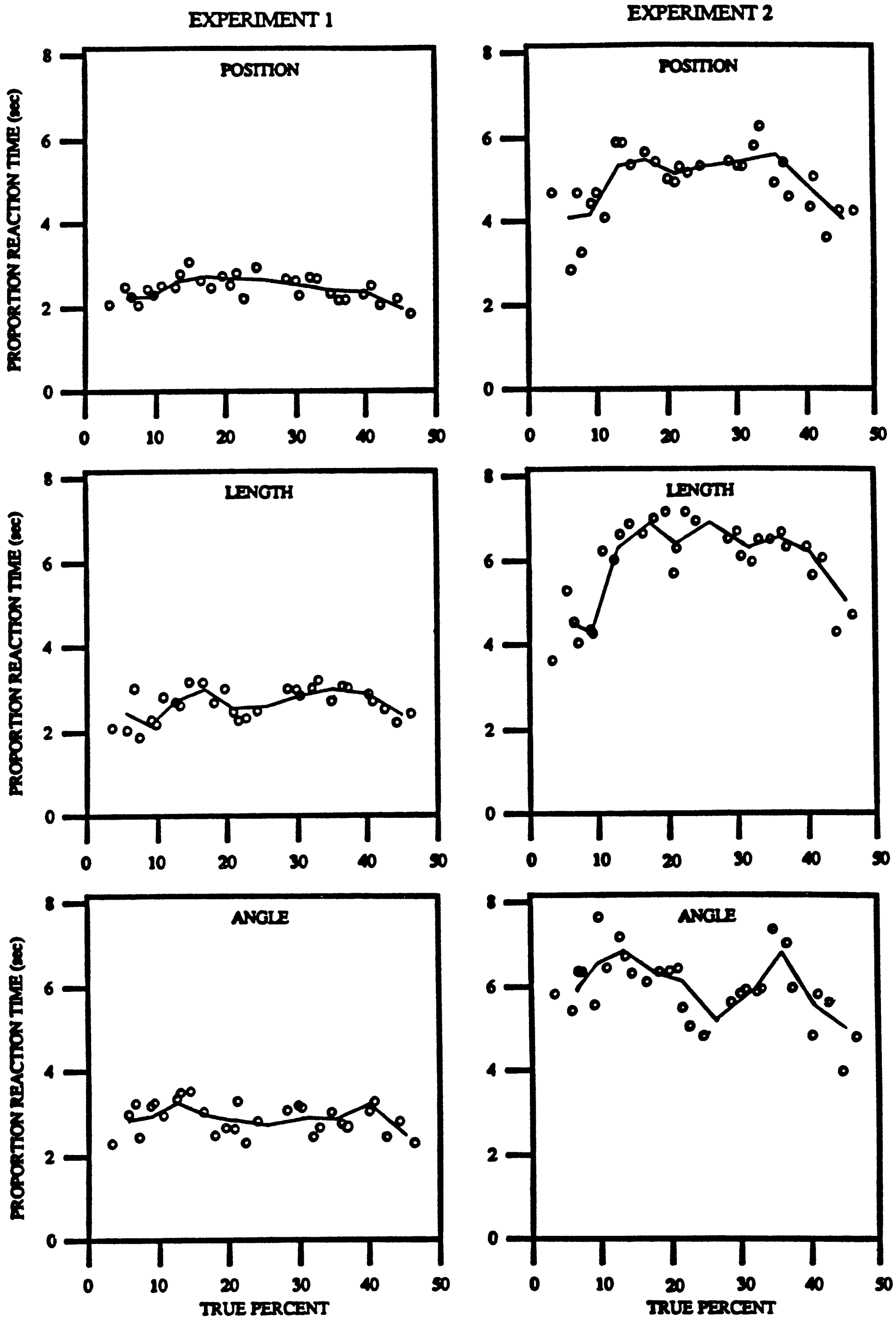

Figure 6. Midmeans of Proportion Reaction Time Against True Percentages for the Three Elementary Codes of Experiments 1 and 2.

the X coordinate is the mean of the 3 true percentages in that block and the Y coordinate is the mean reaction time for these 3 true percentages. All three elementary codes show an increase in reaction time as a function of true percentage. We suspect that this relationship reflects a decrease in absolute difference between the divisions in the image. Length and angle show a steeper increase in reaction time. This is not surprising, since position discrimination only requires that the perceiver detect a difference between the tops of the two bars. The response time does not rise significantly until the difference is very small. The earlier rises for angle and length indicate that the analysis is more complicated.

6.2 Proportion Judgment Task

The sequence of processes for the proportion judgment of all elementary codes is

$$\text{anchoring} \rightarrow \text{scanning}.$$

This sequence is depicted for all codes in Figure 4. The standard, which in this case is the whole bar or pie, is segmented to provide anchors. The available anchor nearest the division is chosen, and the distance between the anchor and the element being evaluated is scanned to adjust the estimate. The plots of the proportion absolute-error data for position from the two experiments, displayed in the top two panels of Figure 5, indicate that accuracy is highest at 0% and 50%. Accuracy then drops off as longer scans from the anchors are required or as other anchors that cannot be established as accurately are being used. The plots of the 30 proportion reaction-time midmeans for position, displayed at the top of Figure 6, lend support to this view. The curves indicate that the fastest reaction times occur at 0% and 50%. The slight dip at intermediate percentages in Experiment 2 suggests additional anchoring locations. Recall that in Experiment 2 the graph was displayed until a response was made. The greater accuracy observed with longer viewing and response times may be a result of better use of these intermediate anchors, which the additional time allows. This is not to say that anchoring at intermediate values such as 25% will necessarily manifest itself through local minima. The intermediate anchoring may simply pull everything down.

The plots of the proportion absolute errors for length displayed in the middle two panels of Figure 5 show an increase in error as the true percentage increases. Indications are that, except for the easily preceived anchor of 0%, the anchoring process is much less accurate for length than position. Even with the additional time in Experiment 2, anchoring at 50% is inferior to that at 0%. The plots of the proportion reaction times for length displayed in the middle two panels of Figure 6 again show faster reaction times at 0% and 50%. Surprisingly, it is Experiment 1 that shows a dip at intermediate percentages. The accuracy advantage of position is a reflection of the more accurate anchoring possible.

The proportion absolute-error midmeans and the proportion reaction-time midmeans for angle are displayed in the bottom panels of Figures 5 and 6. We argue, as Cleve-

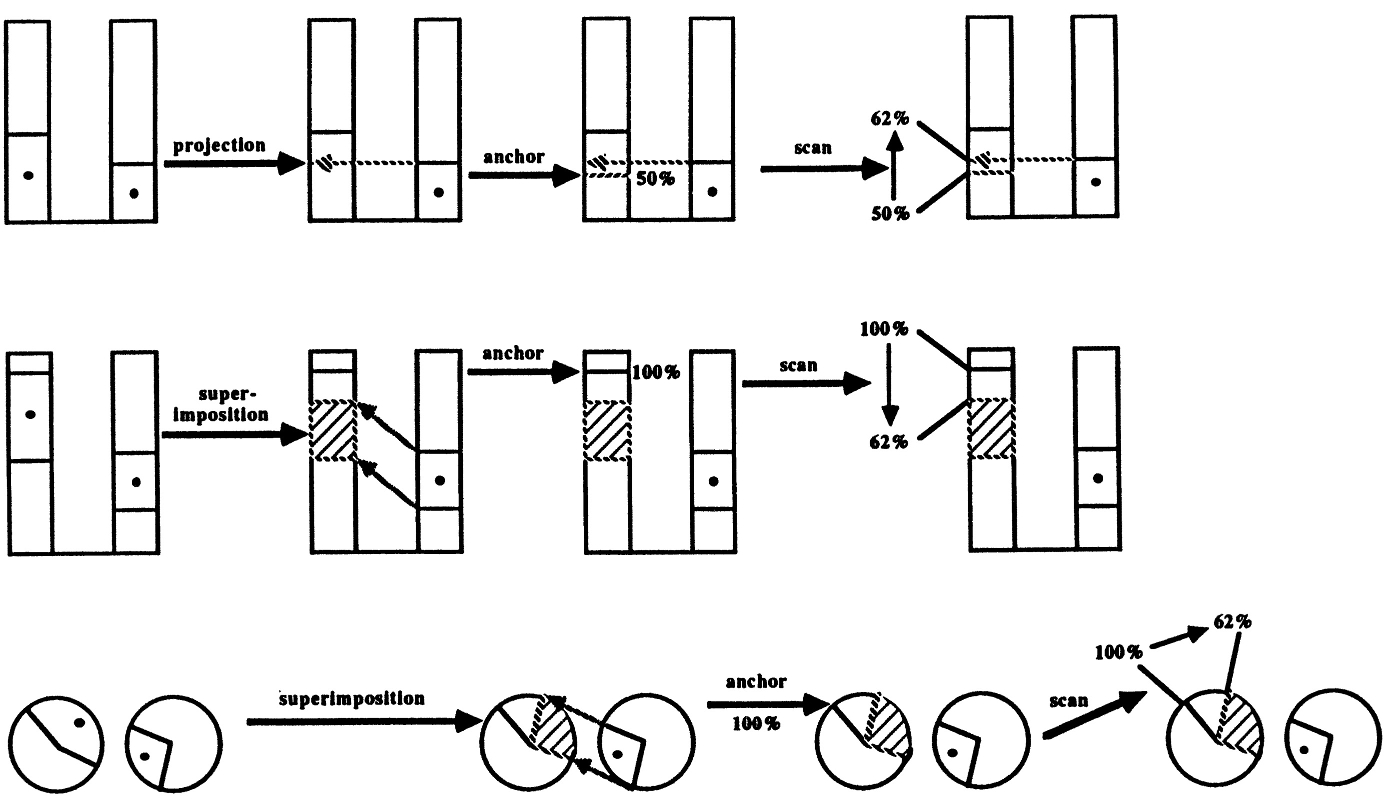

Figure 7. Proposed Sequence of Elementary Mental Processes to Explain Performance in the Comparison Judgment Task for Position (top panel), Length (middle panel), and Angle (bottom panel).

land and McGill have, that the processing of angles is more difficult than length or position. In our system we expect both scanning and anchoring to be inferior for angles. When the anchors are part of a whole, however, the anchoring accuracy is enhanced because the angles formed by the anchors of 0%, 25%, and 50% are 0°, 90°, and 180°, respectively. These angles have a particular perceptual salience and seem to "jump out" at the person performing the task. Because of this, angle accuracy is more similar to position accuracy than to length accuracy. The angle reaction-time data, particularly from Experiment 2, suggest that for angles the 25% anchor is as easily established as the 0% and 50% anchors. These assumptions have implications for the comparison task to be discussed shortly.

The trends apparent in the accuracy data for position and angle are not a result of the subjects responding more frequently with values of 0%, 25%, and 50%. Of 1,200 estimates for each graph, the frequency of these three values combined was 91, 101, and 78 for position, length, and angle, respectively, in Experiment 1, and it was 50, 94, and 61 in Experiment 2. These are too infrequent to have created the trends, and the greatest frequency of anchor-value estimates was observed for length that did not show the trend. In terms of proportion reaction-time predictions, position is faster than length because the an-

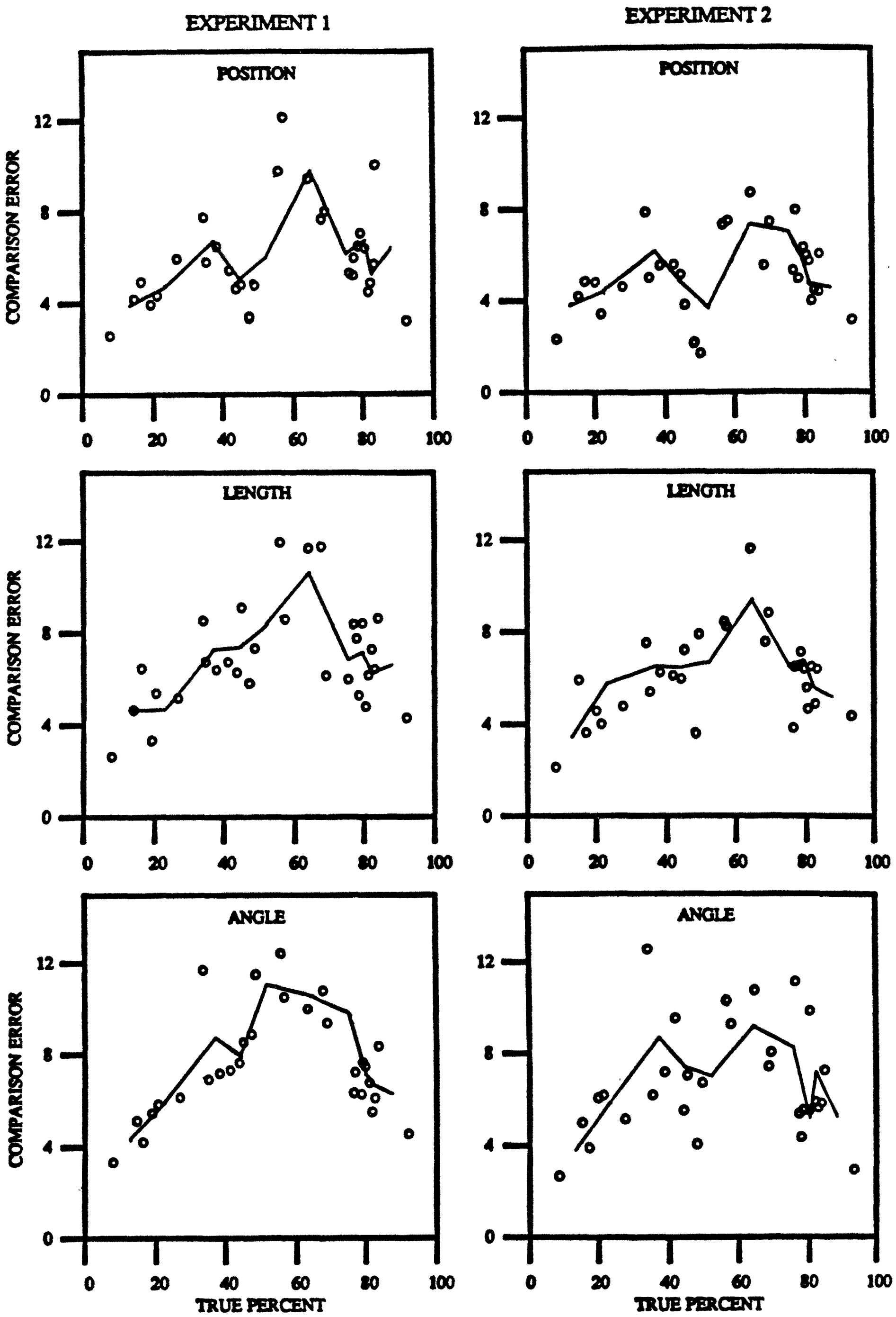

Figure 8. Midmeans of Comparison Absolute Errors Against True Percentages for the Three Elementary Codes in Experiments 1 and 2.

chors can be established more quickly or because it requires less scanning (more frequent anchors). The slowest is angle because of the more difficult scanning involved.

6.3 Comparison Judgment Task

The processing involved in performing the comparison task is a direct extension of that proposed for the proportion task. The sequence for position is

projection → anchoring → scanning.

This sequence is depicted in the top panel of Figure 7. Here the location of the top of the smaller division is projected to the larger division. As was the case in the proportion task, anchors are found in the standard; but, instead of the standard being the entire bar, the larger division is the standard. For length and angle the projection process is replaced with superimposition so that the sequence is

superimposition → anchoring → scanning.

This sequence is depicted for length in the middle panel of Figure 7 and for angle in the bottom panel of Figure 7.

The plots of the comparison absolute-error data shown in the top two panels of Figure 8 indicate that position is most accurate because anchoring can occur at 0%, 50%, and 100%. The length data in the middle panels of Figure

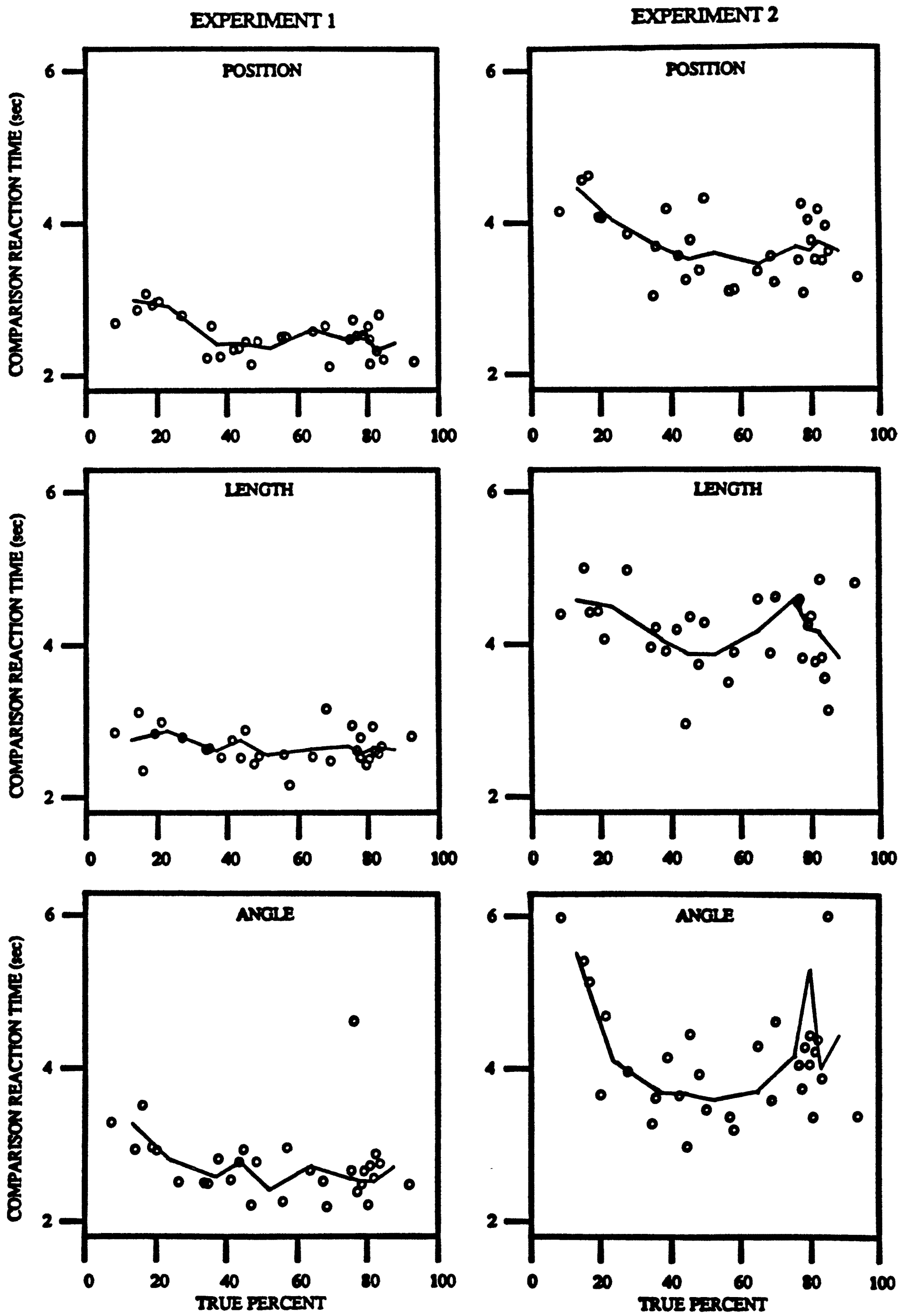

Figure 9. Midmeans of Comparison Reaction Times Against True Percentages for the Three Elementary Codes of Experiments 1 and 2.

8 and the angle data in the lower panels of Figure 8 indicate that, for these two codes, anchoring is far less accurate at 50%. As was the case with the proportion judgment, anchoring at the intermediate values was more apparent in Experiment 2. Again the additional time allowed for more fine-grained anchoring. The accuracy of anchoring the angles in the pie chart is markedly reduced in comparison with the proportion task, because the standard is no longer the whole pie and the anchors are not at 90° and 180°. The bars and pies are now on even footing, and the processing of position and length wins out.

The ordering of the codes on comparison reaction time is the same as that for proportion reaction time. These data are displayed in Figure 9, but interpretation of the plots of reaction time should be guarded, since our scanning process makes predictions depend partially on absolute differences in the size of divisions. The true percentages in the comparison task do not fully represent this characteristic. The fact that the data consistently show longer reaction times at very small and very large percentages, however, is at odds with our system.

6.4 Summary

This analysis is obviously tentative, but the evidence consistently points to anchoring as the key process for proportion and comparison judgments. When making a proportion-of-the-whole judgment, the more accurate anchoring possible with position and angle codes accounts for their superiority over the length code. Although processing angles is more difficult than processing linear aspects, this judgment for the pie chart is a special case in which the anchors are at the perceptually salient angles of 0°, 90°, and 180°. When making a comparison judgment, the position code is superior to the other two codes. Length again suffers from less accurate anchoring. Angles provide the least accurate estimates because of the inferior anchoring when these anchors are no longer at perceptually salient angles.

7. CONCLUSION

We have tried to make three major points in our analysis. First, people have schemata for graphs that include slots for the conceptual message of the graph. Second, we demonstrated that elementary code and judgment task interact to determine performance. Third, we proposed elementary processes of anchoring, scanning, projection, superimposition, and detection operators to explain these interactions.

The information-processing approach to human cognition provides a promising vocabulary to summarize these relationships between task, learned skills, and knowledge, and mental capacities to yield a useful theoretical treatment of graph perception.

[*Received July 1986. Revised October 1986.*]

REFERENCES

Bertin, J. (1983), *Semiology of Graphics,* Madison: University of Wisconsin Press.

Bower, G. H., Black, J. B., and Turner, T. J. (1979), "Scripts in Memory for Text," *Cognitive Psychology,* 11, 177–220.

Cleveland, W. S., and McGill, R. (1984), "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods," *Journal of the American Statistical Association,* 79, 531–554.

——— (1985), "Graphical Perception and Graphical Methods for Analyzing and Presenting Scientific Data," *Science,* 229, 828–833.

——— (in press), "Graphical Perception: The Visual Decoding of Quantitative Information on Graphical Displays of Data" (with discussion), *Journal of the Royal Statistical Society.*

Eells, W. C. (1926), "The Relative Merits of Circles and Bars for Representing Component Parts," *Journal of the American Statistical Association,* 21, 119–132.

Follettie, J. F. (1986), "Real-World Tasks of Statistical Graph-Using and Analytic Tasks of Graphics Research," unpublished paper presented at the annual meeting of the National Computer Graphics Association, Anaheim, CA.

Hastie, R. (1981), "Schematic Principles in Human Memory," in *Social Cognition: The Ontario Symposium on Personality and Social Psychology,* eds. E. T. Higgins, C. P. Herman, and M. P. Zanna, Hillsdale, NJ: Lawrence Erlbaum.

Kosslyn, S. M. (1980), *Image and Mind,* Cambridge, MA: Harvard University Press.

——— (1985), "Graphics and Human Information Processing," *Journal of the American Statistical Association,* 80, 449–512.

Mosteller, F., and Tukey, J. W. (1977), *Data Analysis and Regression,* Reading, MA: Addison-Wesley.

Pinker, S. (1981), "A Theory of Graph Comprehension," Occasional Paper 10, Massachusetts Institute of Technology, Center for Cognitive Science.

——— (1983), "Pattern Perception and Comprehension of Graphs," Report ED 237-339, Massachusetts Institute of Technology, Psychology Dept.

Rumelhart, D. E. (1984), "Schemata and the Cognitive System," in *Handbook of Social Cognition* (Vol. 1), eds. R. S. Wyer, Jr., and T. K. Srull, Hillsdale, NJ: Lawrence Erlbaum.

Schank, R., and Abelson, R. (1977), *Scripts, Plans, Goals, and Understanding: An Inquiry Into Human Knowledge Structures,* Hillsdale NJ: Lawrence Erlbaum.

Ullman, S. (1984), "Visual Routines," *Cognition,* 18, 97–159.