Educational Psychology Review, Vol. 14, No. 1, March 2002 (© 2002)

Review of Graph Comprehension Research: Implications for Instruction

Priti Shah[1,2] **and James Hoeffner**[1]

Graphs are commonly used in textbooks and educational software, and can help students understand science and social science data. However, students sometimes have difficulty comprehending information depicted in graphs. What makes a graph better or worse at communicating relevant quantitative information? How can students learn to interpret graphs more effectively? This article reviews the cognitive literature on how viewers comprehend graphs and the factors that influence viewers' interpretations. Three major factors are considered: the visual characteristics of a graph (e.g., format, animation, color, use of legend, size, etc.), a viewer's knowledge about graphs, and a viewer's knowledge and expectations about the content of the data in a graph. This article provides a set of guidelines for the presentation of graphs to students and considers the implications of graph comprehension research for the teaching of graphical literacy skills. Finally, this article discusses unresolved questions and directions for future research relevant to data presentation and the teaching of graphical literacy skills.

KEY WORDS: graphs; graphical displays; graph comprehension; science education; graphical literacy.

Graphs are commonly used to depict mathematical functions, display data from social and natural sciences, and specify scientific theories in textbooks and other print media in and out of the classroom (Kaput, 1987; Lewandowsky and Behrens, 1999; Mayer, 1993). More recently, graphs have played an important role in teaching quantitative and scientific concepts in intelligent tutoring systems and other educational software (e.g., Nachmias

[1]Department of Psychology, University of Michigan, Ann Arbor, Michigan.
[2]Correspondence should be addressed to Priti Shah, Department of Psychology, University of Michigan, Ann Arbor, Michigan 48109-1109; e-mail: priti@umich.edu.

1040-726X/02/0300-0047/0

and Linn, 1987; Quintana *et al.*, 1999; Reiser *et al.*, in press; Scardamalia *et al.*, 1994).

One of the reasons graphs are so pervasive is that they seem to make quantitative information easy to understand (MacDonald-Ross, 1977; Tversky, in press; Winn, 1987). Graphs and other visual displays can be helpful in depicting a quantitative or scientific concept, particularly when the concept is expressed explicitly in the display (Larkin and Simon, 1987; Pinker, 1990). In some cases, however, the comprehension of graphs can be effortful and error prone (e.g., Bell and Janvier, 1981; Carpenter and Shah, 1998; Culbertson and Powers, 1959; Maichle, 1994; Shah and Carpenter, 1995). School-aged children, and even adults, commonly make systematic errors interpreting graphs, especially when graphs do not explicitly depict the relevant quantitative information (Gattis and Holyoak, 1995; Guthrie *et al.*, 1993; Leinhardt *et al.*, 1996; Shah *et al.*, 1999; Shah and Carpenter, 1995; Vernon, 1946, 1950).

Consider, for example, the bar graph shown in Fig. 1. Shah *et al.* (1999) studied participants' interpretations of graphs similar to this one, which is adapted from an American history textbook chapter on the Civil War (Armento *et al.*, 1991). The accompanying text indicates that the graph is intended to show that in the decades preceding the war, the North was urbanizing but the South remained largely rural. Unfortunately, when shown this graph, viewers rarely describe differences in the rate of urbanization in

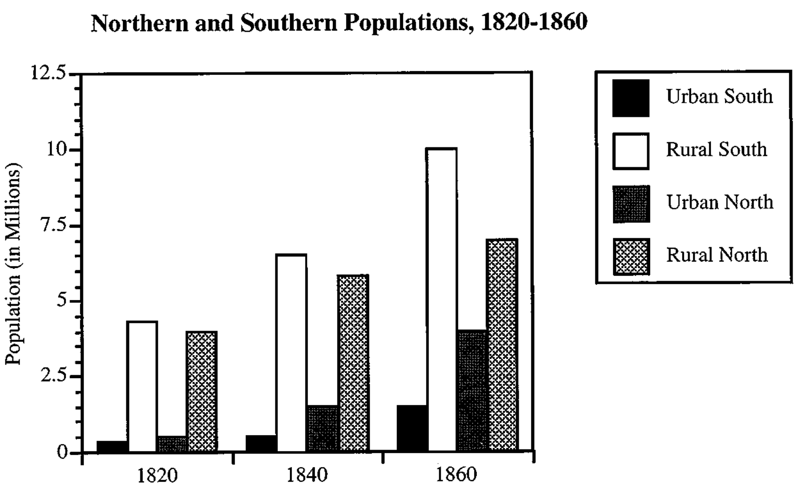

Fig. 1. A graph adapted from an eighth grade history textbook (Armento, Nash, Salter, and Wixson, 1991) that presents the population in the North and South leading up to the Civil War.

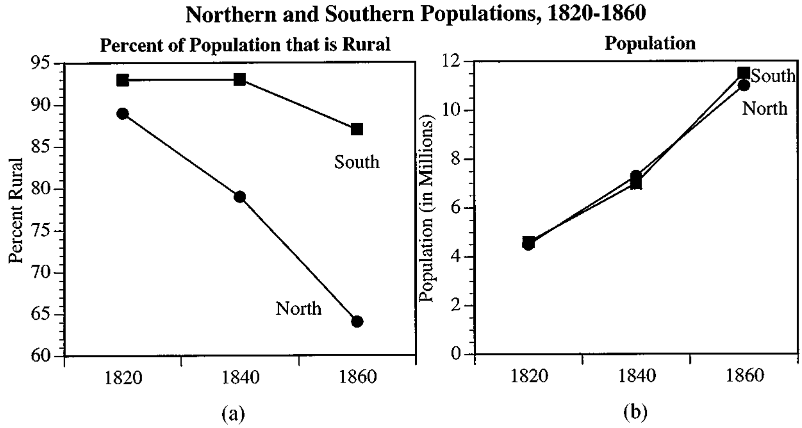

Fig. 2. A set of graphs that depict the identical quantitative information as in Fig. 1.

the North and South. Instead, viewers compare the relative populations in each of the four categories (Urban South, Rural South, etc.) in each time period. When the data were replotted to make the relevant trends more visually salient in the graphs in Fig. 2, viewers easily comprehended the relevant quantitative information.

Why are some graphs relatively easy for viewers to comprehend for a particular task, and other graphs more difficult? How do individual differences in graph reading skill and domain knowledge influence the kinds of interpretations that viewers give to graphs presented in texts? The first goal of this paper is to present a review of the empirical literature on how viewers interpret graphs and the factors that make graphs easy or difficult to understand. The difficulty of comprehension is not merely a function of characteristics of the graph itself but is also influenced by how those features interact with the viewer's knowledge and objective. The second goal of this article is to discuss the implications of graph comprehension research for designing of graphs for students and for teaching graphical literacy skills.

The scope of this review is limited, in two ways. First, it focuses on the *interpretation* of graphs depicting meaningful data (the task when graphs are "text adjuncts," the focus of this special issue), generally in the context of science and social science. Thus, this article does not review the large body of research examining the understanding of mathematical functions and the role of graphs in mathematics education.[3] Second, research on graph interpretation has taken place in a diverse variety of disciplines (Lewandowksy

[3]The reader is referred to an excellent review of students' understanding of functions and graphs by Leinhardt *et al.* (1990) for a more general review of graphs in educational contexts, including thinking about mathematical functions.

and Behrens, 1999), so it is impossible to cover all the research available. The research presented here focuses on a representative sample of those studies for which there are clear educational implications.

GRAPH COMPREHENSION

Task analyses of graph comprehension have identified three major component processes (Bertin, 1983; Carpenter and Shah, 1998; Cleveland and McGill, 1984, 1985; Kosslyn, 1989; Lohse, 1993; Pinker, 1990; Shah, in press). First, viewers must encode the visual array and identify the important visual features (such as a curved line). Encoding the visual information is influenced in two ways by the visual characteristics of the display. Specifically, displays may interact with the inherent biases and limitations of our perceptual apparatus to affect both the accuracy of information encoding (e.g., Cleveland and McGill, 1984, 1985; Legge *et al.*, 1989; Spence, 1990) and the way in which the information is grouped (e.g., Carpenter and Shah, 1998; Shah *et al.*, 1999). For example, viewers might encode absolute values more accurately for bar graphs than for pie charts (Cleveland and McGill, 1985). Or, a set of data points may be more likely to be "grouped" (into "Gestalt wholes"; Kosslyn, 1989) when they are connected by a line than when they are unconnected in a bar graph.

Second, viewers must relate the visual features to the conceptual relations that are represented by those features (Bertin, 1983; Kosslyn, 1989; Pinker, 1990). Several factors influence what conceptual relations are inferred from a visual display. One factor is the outcome of the first stage of graph comprehension: the encoding of visual features. For example, if three data points are grouped by being connected by a line, viewers are more likely to characterize the relationship among those three data points than if they are not grouped.

Even if viewers encode the relevant information accurately, their ability to map between different visual features and the meaning of those features may differ as a function of experience. In some cases, viewers may be able to retrieve, via a simple pattern-matching process, what a particular visual feature implies (e.g., a viewer knows that a curved line implies an accelerating relationship). In Pinker's terms, knowledge about the mapping of visual features to an interpretation is part of their *graph schema*, or general knowledge about graphs (Pinker, 1990). When a visual feature does not automatically evoke a particular fact or relationship, then that information is more difficult to comprehend (Cleveland, 1993; Kosslyn, 1989, 1994; Larkin and Simon, 1987; Pinker, 1990; Shah *et al.*, 1999; Stenning and Oberlander, 1995; Tversky, in press) and viewers may make an error in interpretation.

In one study, for example, viewers accurately drew various features in line graphs from memory, such as the distance between different lines on the graph. However, the viewers did not know what the distances implied about the quantitative relations they represented, and therefore were unable to recognize the same data plotted in another graph (Shah and Carpenter, 1995).

Another factor that may influence the second component process is the relative ease of some mappings between visual features and referents compared with other mappings. The easiest mappings are probably rooted in observations of, and experiences with, the physical world (Tversky, in press). For example, vertically oriented bars may be better than horizontally oriented bars in a graph when depicting quantities of physical material because a greater quantity of items usually corresponds to a taller pile. Similarly, horizontally oriented bars may be better for other dimensions for which the horizontal dimension is more meaningful, such as in depicting data about the distance traveled (Kosslyn, 1994).

The third component process of graph comprehension is that viewers must determine the referent of the concepts being quantified (e.g., Population, Rural South, etc.; in Fig. 1) and associate those referents to the encoded functions (Bertin, 1983). As discussed in more detail later, students' and/or novices' interpretations are often colored by their expectations of the content (Shah, 1995; Shah and Shellhammer, 1999). In addition, young students sometimes view graphs as representing literal pictures of situations rather than abstract quantitative information (e.g., Leinhardt *et al.*, 1990; McDermott, Rosenquist, and vanZee, 1987; Preece, 1990).

These three processes imply that three factors play an important role in determining a viewer's interpretation of data: the characteristics of the visual display (bar or line graph, color or black and white, etc.), knowledge about graphs (graph schemas), and content (e.g., age vs. height, time vs. distance). Below, we discuss research on these three factors and its implications for instruction.

VISUAL CHARACTERISTICS OF THE DISPLAY

Two-Dimensional Formats

A graph designer has the choice of many different formats that, as the research outlined below suggests, can have differential impacts on viewers' interpretations of data. One should keep in mind that any one experiment can only compare particular instantiations of a graphical format and include a relatively small number of possible interpretation tasks. Therefore, the best support for claims of format effect differences should come from agreement among multiple studies.

Viewers are more likely to describe x–y trends (e.g., as x increases, y decreases) when viewing line graphs than when viewing bar graphs (Carswell *et al.*, 1993; Shah *et al.*, 1999; Zacks and Tversky, 1999). Viewers are also more accurate in retrieving x–y trend information from line graphs than from bar graphs (Carswell and Wickens, 1987). Strikingly, even when two discrete data points are plotted in a line graph, viewers (college students at Stanford) sometimes describe the data as continuous. For example, a graph reader may interpret a line that connects two data points representing male and female height as saying, "The more male a person is, the taller he/she is" (Zacks and Tversky, 1999). This emphasis on the x–y trends can lead to incomplete interpretations of data when the data are complex (for example, multiple lines on a display representing a third variable). In one study in which viewers described such graphs, they focused entirely on the x–y trends and therefore were frequently unable to recognize the same data plotted when another variable was on the x axis (Shah and Carpenter, 1995). For example, participants typically described just the effects of noise level on achievement test scores when shown Fig. 3(a), but described the effects of room temperature on achievement test scores when shown Fig. 3(b).

Whereas line graphs emphasize x–y trends, bar graphs emphasize discrete comparisons (Carswell and Wickens, 1987; Shah *et al.*, 1999; Zacks and Tversky, 1999). Furthermore, bar graphs of multivariate data appear to be less biasing than line graphs. Consider, for example, the bar graph in Fig. 4 that shows the same data as the line graphs in Figs. 3(a) and (b). Viewers are

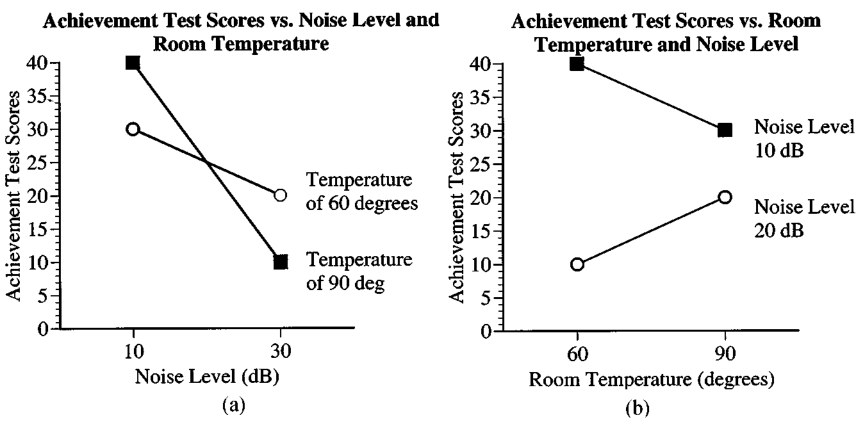

Fig. 3. Two views of the same data set in which different variables are plotted as a function of the x–y lines.

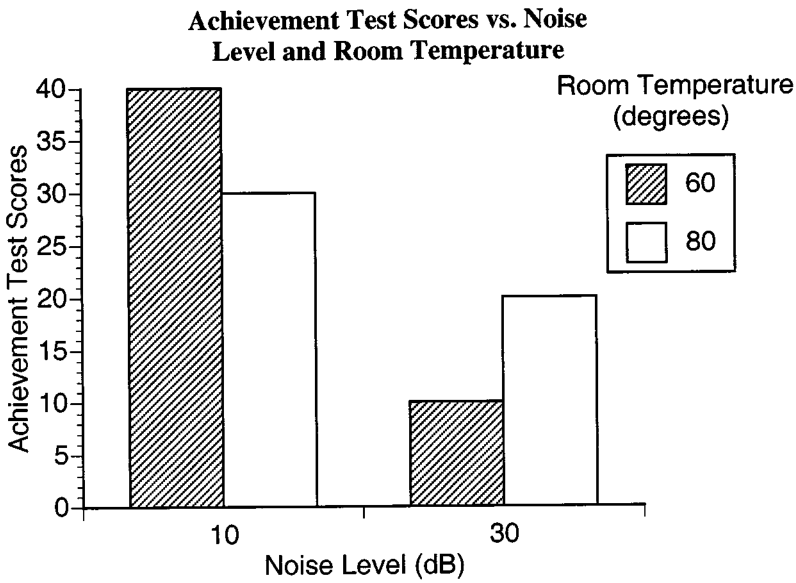

Fig. 4. A bar graph of the same data shown in Figs. 3(a) and (b).

much less biased in their descriptions of the relationships in this bar graph and describe the effects of room temperature and noise level on achievement test scores equally often (Shah and Shellhammer, 1999).

Although bar graphs and line graphs are typically used for presenting metric information about absolute scales, divided bar charts and pie charts are frequently used for the presentation of proportion data. In general, pie charts are more accurate for making part/whole judgments than divided bar graphs are because divided bar graphs often require adding up information from different parts of the bar (Simkin and Hastie, 1986). Thus, pie charts are generally preferred over divided bar charts when the goal is for viewers to comprehend relative proportions (Kosslyn, 1994; Wilkinson, 1999). However, divided bar charts may be better when absolute values as well as proportions are important to communicate (Kosslyn, 1994).

In considering format, one important additional finding is that there is a trade-off between the ability to accurately perceive specific quantitative facts and the ability to get a more qualitative gist of relationships depicted in the data. A table, for example, allows people to get single point values most accurately but provides the least integrative information (Guthrie *et al.*, 1993).

In summary, line graphs are good for depicting *x*–*y* trends, bar graphs for discrete comparisons, and pie charts for relative proportions. Thus, no

Table I. Rank Ordering of Visual Dimensions Used to Represent Quantitative Information Most Accurate to Least Accurate[a]

Rank	Visual dimension
1	Position along a common scale
2	Position along nonaligned scales
3	Length, direction, angle (tie)
4	Area
5	Volume, curvature (tie)
6	Shading, color saturation (tie)

[a] Adapted from Cleveland and McGill (1984).

graph format is necessarily better overall than any other format. Instead, there is an interaction of task and graphic format, called the *proximity compatibility principle* by Carswell and Wickens (1987; see also Simkin and Hastie, 1986). Integrated, object-like displays (e.g., a line graph) are better for integrative tasks, whereas more separable formats (e.g., bar graphs) are better for less integrative or synthetic tasks such as point reading (Carswell and Wickens, 1987).

It should be noted that the graphic formats considered here generally follow one important guideline regarding the visual dimensions that should be used to represent quantitative information. Bar graphs and line graphs represent quantities in terms of spatial extent, whereas pie charts represent quantities in terms of area or angle size. In principle, it is possible to represent quantitative information with a number of other possible visual features, such as color saturation. Both psychophysical studies (Cleveland and McGill, 1984, 1985) and theoretical analyses (Kubovy, 1981; Pinker, 1990) suggest a ranking of the best possible visual dimensions, and both agree that spatial extent or position is the best way to represent metric information accurately. A summary of these rankings is presented in Table I.

Three-Dimensional Displays

Three-dimensional displays (depicting multivariate data) and animations depicting quantitative information are becoming more common in the current computer graphics era. However, there has been relatively little research on these types of displays. In one study that compared the use of three-dimensional and two-dimensional scatterplot-like displays, three-dimensional displays proved better than two-dimensional displays when the questions required integrating information across all three dimensions (Wickens *et al.*, 1994). Another study (Shah, in press) examined the interpretation of three-dimensional "wireframe" graphs, an example of which is shown in Fig. 5(a)—next to a line graph depicting the same data in Fig. 5(b).

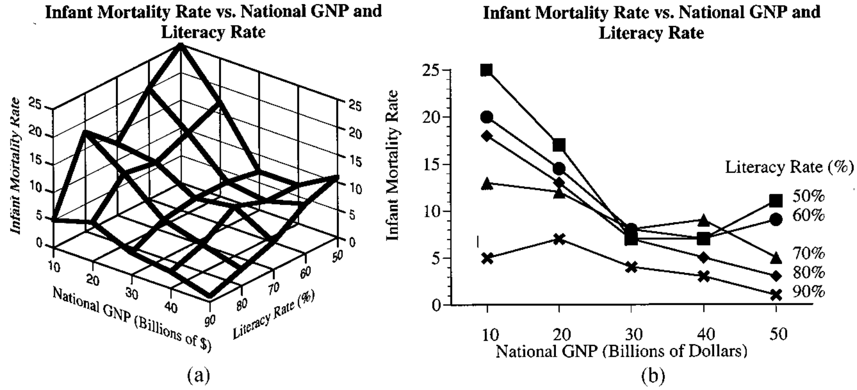

Fig. 5. A three-dimensional wireframe graph and a line graph depicting the same data.

In this study, viewers were more likely to describe the relationships among all three variables when viewing wireframe graphs than when viewing line graphs (in which they focused on the x–y trends). Despite the potential benefits of three-dimensional displays, the use of three-dimensional linear perspective drawings can degrade or occlude information (Merwin *et al.*, 1994). In our study, for example, viewers were inaccurate on tasks requiring reading individual data points in wireframe graphs, and were less likely to be familiar with how to interpret such displays (Shah, in press).

Animation

The implicit assumption behind the use of animation is that it is an ideal medium for communicating complex information, especially temporal information (e.g., quantitative trends over time). However, there is little empirical research on animation in the context of graphs, and the little that has been conducted has focused on the interpretation of fairly complex graphs depicting multivariate data for professionals such as statisticians. Some possible benefits of animation have been found for such specialized contexts. For example, motion cues can help people perceive the three-dimensional structure of graphs (Becker *et al.*, 1988) and animation may help viewers identify clusters of data in very limited circumstances in three-dimensional "pointclouds" (or three-dimensional scatterplots; Marchark and Marchak, 1991). Despite the possible benefits for the specialized tasks for which animation has been empirically examined, other researchers suggest that animation may also have harmful effects. Users of animation software for information visualization, describing their own experiences, claimed they had difficulty comprehending animated relationships and trends (Huber, 1987; Stuezle, 1987).

Furthermore, evidence from studies examining the interpretation of animations (not in the context of graphs) suggests that the perception of animation is often difficult and error prone (Hegarty *et al.*, 1999; Morrison *et al.*, 2000). In general, more research on the role of animation for graphs of quantitative information is needed before strong conclusions are drawn.

Other Perceptual Features

In addition to global decisions about the general format, a graph designer can choose a number of additional visual features, the most common being color, size, and aspect ratio (ratio of length to width).

Color

Color can be used to represent metric information (e.g., the closer to red the greater the debt, or the deeper the color saturation the higher the population) or to mark variables (e.g., the red bars represent girls and the blue bars represent boys; Brockmann, 1991; Cleveland and McGill, 1985; Hoffman *et al.*, 1993). However, color does not accurately represent precise quantitative information (Cleveland and McGill, 1985) and may even be misleading when used for continuous data. For example, certain colors are more likely to be interpreted as "higher" or "lower" in displays such as contour plots, influencing how well viewers are able to imagine these plots three-dimensionally. Furthermore, colors may be interpreted as representing categorical data even when they are intended to convey continuous information (Phillips, 1982).

However, color has certain benefits. One major use of color can be to group elements in a display. For example, color can help viewers group data in scatterplots (Lewandowsky and Spence, 1989); for example, different colors can be used for "boys" and "girls" in a scatterplot of height versus age. Similarly, if two pieces of data in separate graphs (such as two pie charts) are compared, they are more likely grouped if they are the same color (Kosslyn, 1994). Another potential benefit of color is to reduce the difficulty viewers face in keeping track of graphic referents because of the demands imposed on working memory (e.g., remembering that the line with open circles represents a temperature of 70° in Figs. 6(a) and (b); Fisher, 1982; Kosslyn, 1994; Schmid, 1983). Indeed, in a recent study examining viewers' eye fixations as they interpreted graphs, we showed that viewers must continuously reexamine the labels to refresh their memory (Carpenter and Shah, 1998). If those dimensions are represented with a meaningful color choice, such as red for warmer temperatures and blue for cooler temperatures, it might help viewers keep track of variable names (Brockman, 1991). Of course,

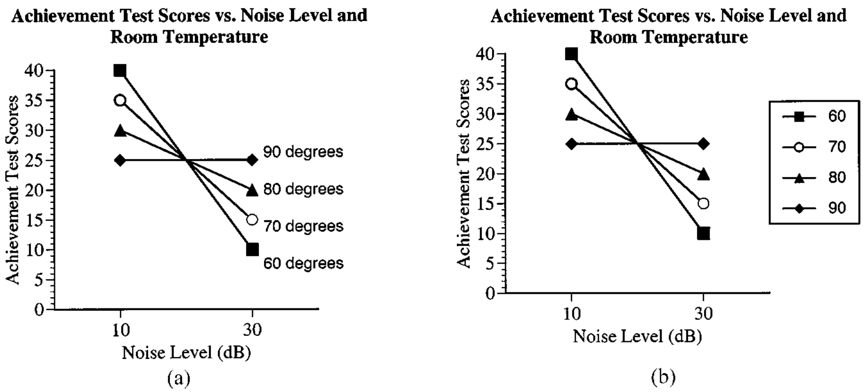

Fig. 6. A line graph in which the referents are labeled (a), and a line graph that uses a legend (or key) rather than labels (b).

the use of semantically related features is highly dependent on assumptions shared by the graphic designer and graph reader. For example, green means *profitable* for financial managers but *infected* for health care workers (Brockmann, 1991). Nonetheless, the use of semantically related features may be especially beneficial when presenting graphs to children, who, in general, have fewer working memory resources than adults (Halford *et al.*, 1998). In conclusion, color can provide helpful cues especially with respect to helping viewers keep track of quantitative referents, but is inadequate as the only source of precise quantitative information.

Legend (or Key) vs. Labels

Another choice a graph designer makes is whether to use a legend (or key) or to directly label graph features (such as lines and bars) according to their referents. Because legends or keys require that graph readers keep referents in memory, legends pose special demands on working memory; see, for example Fig. 6(b). Thus, the conventional wisdom is that graph designers should avoid using legends (except when labels would lead to too much visual clutter) and instead label graph features directly with their referents (Kosslyn, 1994) as in Fig. 6(a). Again, this advice may be particularly important when presenting graphs to children.

The "Third" Dimension

One commonly used perceptual characteristic is a third "spatial" dimension (that does not convey additional information) when used in a

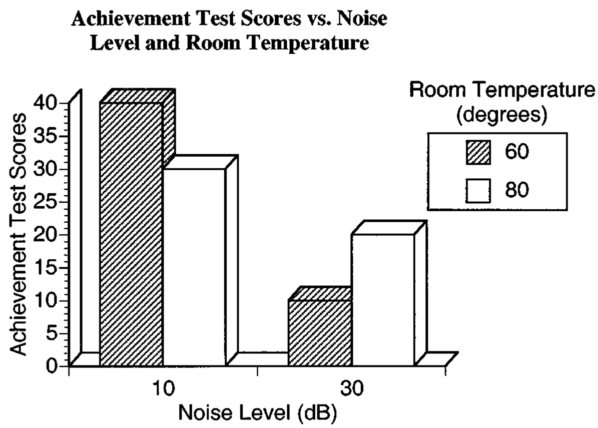

Fig. 7. A bar graph depicting the same data as in Fig. 4, but using extraneous depth cues.

two-variable graph. An example is the use of three-dimensional bars in a bar graph (see Fig. 7). Such graphs are commonly used in the media (Spence, 1990; Tufte, 1983) and are also fairly common in textbooks. Generally, the a priori advice is that additional, noninformative features are unhelpful and often distracting. Tufte (1983) even labeled an added third dimension as an example of *chart junk* and recommended keeping it to a minimum. Indeed, some studies have found that the accuracy of reading point values or making comparisons between data points from "three-dimensional" displays is slightly lower than for two-dimensional equivalents (Fischer, 2000; Zacks *et al.*, 1998). However, other studies show little or no difference in the accuracy or speed of making comparison judgments (Carswell *et al.*, 1991; Spence, 1990), at least from memory (Zacks *et al.*, 1998). In addition, viewers sometimes express preference for these three-dimensional displays (Levy *et al.*, 1996).

Density of Data (Graph Size) and Aspect Ratio

The density of a data set can also affect the perception of graphs. Viewers often mentally exaggerate the magnitude of correlations in scatterplots when the data appear dense (Cleveland *et al.*, 1982; Lauer and Post, 1989). This exaggeration occurs when the graph's perceived density is increased either adding data points, or reducing the graph's size. Similarly, the aspect ratio of a graph can have an influence on the patterns that viewers identify. Cleveland (1993), for example, argues that viewers can most easily detect cyclical patterns when an aspect ratio that makes the curve closest to 45° is used.

KNOWLEDGE ABOUT GRAPHS

Viewers' knowledge and expectations affect how they encode and remember pictures and diagrams. For example, verbal labels distort viewers' memory for ambiguous pictures (Carmichael *et al.*, 1932). More recent evidence suggests that viewers have schemas for graphs and maps that can distort their representations of them. For example, college students in one study were asked to draw line graphs from memory. They distorted the lines and drew them as being closer to 45° than the lines originally were (Schiano and Tversky, 1992; Tversky and Schiano, 1989). When they were told that the same display depicted a map, however, they distorted the lines so that they were closer to 0° or 90°. This suggests that a viewer's memory for data is distorted by general expectations about visually presented data. In the case of graphs, viewers have a canonical graph schema that tends to favor 45° lines rather than steeper or flatter lines.

Viewers have other expectations about graphical information that can also lead to systematic interpretation errors. For example, viewers typically expect dependent variables to be plotted as a function of the y axis in a line graph, and causal or independent variables to be plotted as a function of the x axis. When the dependent variable is plotted on the y axis, as is typical of most graphs, steeper slopes imply faster change rates. If a graph that violates this constraint is presented to graph viewers, college students often misinterpret the meaning of the slopes, incorrectly assuming that the steeper line represents a faster change rate (Gattis and Holyoak, 1995). Thus, viewers' knowledge about the mapping between slope and rate of change can lead to interpretation errors.

Elementary-school–aged graph viewers also make a similar error called the *slope–height confusion*. In this case, the viewer erroneously assumes that greater slope implies higher value, rather than different rate of change (Bell and Janvier, 1981). In this error, viewers whose task is to report on which line represents a greater rate of change, represented in terms of the slope of two lines, focus instead on the relative heights of the two lines.

Finally, I identified a related error made by novice college student graph viewers (Shah, 1995). In this study, I plotted three-variable data in line graphs such as the ones in Figs. 8(a) and (b). In Fig. 8(a) there is a systematic relationship between drunk drivers and car accidents, but in Fig. 8(b) there is no systematic relationship between drunk drivers and car accidents (because for each value of drunk driving in the graph there are multiple values of car accidents and vice versa). However, novice graph viewers often do not have the knowledge in their graph schemas to recognize that the two variables in Fig. 8(b) are independent. Therefore, they frequently included the statement, "As drunk drivers increase, car accidents increase" in their descriptions of

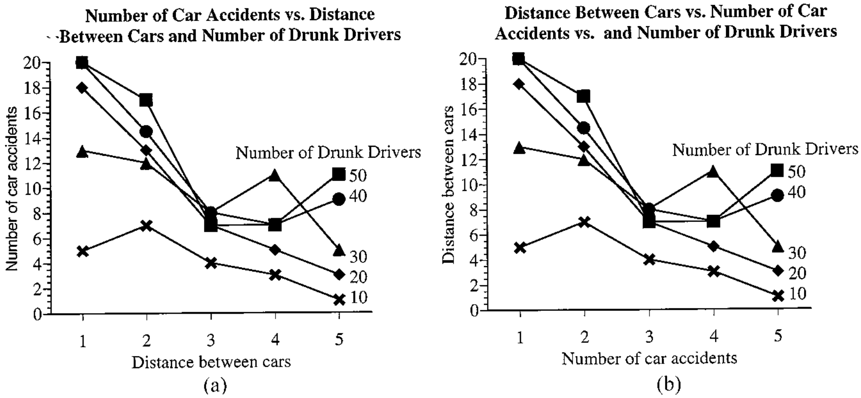

Fig. 8. The figure at left (a) shows a conventionally plotted graph that is consistent with most viewers' expectations of the causal relations between these variables. At right (b) the x and y axes are reversed.

the data in *both* graphs. Thus, novice viewers relied on their prior knowledge about the content rather than the information depicted in the graph. Experts rarely made this error.

KNOWLEDGE ABOUT CONTENT

The *drunk driver* study described above (Shah, 1995) suggests an additional factor that influences viewers' interpretations of graphs. In that study, viewers made systematic errors in interpretation only when their prior knowledge was inconsistent with information in the graph. Thus, viewers' knowledge about content also influences their data interpretations.

Viewers' expectations about data have long been known to affect their interpretations. In a classic study, Lord *et al.* (1979) gave people descriptions of studies that either supported or refuted their own prior beliefs about controversial topics like the death penalty. Overall, participants were more likely to notice problems in studies that were inconsistent with their beliefs than in studies that were consistent with their beliefs. Similarly, viewers' interpretations of graphically presented data are also influenced by their beliefs and expectations. Several studies, for example, have shown that viewers' estimation of correlations or covariation between variables is influenced by prior beliefs. That is, when viewers believe that two variables are related to one another (such as height and weight), they overestimate the correlation compared with when they have no prior expectations about those variables (e.g., Freedman and Smith, 1996; Jennings *et al.*, 1982).

Not only does the type of data influence viewers' quantitative estimates, but also their qualitative interpretations. In a recent study (Shah, in press), college undergraduates described graphs that depicted either familiar relationships (e.g., number of car accidents, number of drunk drivers, and traffic density) or graphs that depicted unfamiliar relationships (e.g., ice cream sales, fat content, and sugar content).[4] Overall, viewers were more likely to make inferences about general trends (main effects) when viewing graphs with familiar relationships than when viewing graphs with unfamiliar relationships. In another study, Shah and Shellhammer (1999) found that less-skilled graph viewers had more difficulty than highly skilled graph viewers in identifying general trends.[5]

Elementary-aged students are perhaps the most influenced by a graph's content. One common error is that viewers interpret abstract representations of data as an iconic representation of a real event (Bell and Janvier, 1981; Janvier, 1981; Leinhardt *et al.*, 1990; Preece, 1990). For example, students might misinterpret a graph representing the speed of a racecar to mean the position of the racecar on a track (Janvier, 1981). This error is particularly common in contexts for which there is an obvious iconic interpretation, usually when the graph is meant to represent change (such as growth, speed) and the concrete interpretation is the value on some dimension (such as height instead of growth, location instead of speed). Although young graph readers (until around fifth grade) make this error frequently, minimal graphing instruction helps viewers overcome this error (Leinhardt *et al.*, 1990).

The studies above suggest that graph viewers are better at understanding graphs with some types of content compared with other types of content, such as those representing change. In related studies, students have the most difficulty with graphs depicting acceleration, followed by graphs depicting velocity. Viewers have the easiest time comprehending graphs depicting distance or position. In addition, viewers often have difficulty translating between these types of graphs (Clement, 1985, as cited in Leinhardt *et al.*, 1990). Another factor is that students are much better at dealing with data in which time is one of the dimensions. One possibility is that time is easier to understand than other variables because often the variables used in instructional contexts are time-dependent (Leinhardt *et al.*, 1990).

Together, these studies suggest that peoples' knowledge of the content in graphs has an influence on their interpretations of, and memory for, data. Furthermore, several studies suggest that this is especially true for novice graph viewers who often do not have the graph schemas necessary to overcome the strong influence of their own content knowledge. This might be

[4]Viewers' expectations were determined by pretesting of our materials on a different, but similar population.
[5]As identified by their scores on the Test of Graphing in Science (McKenzie and Padilla, 1986).

particularly problematic in the context of science and social science, where a critical evaluation of the information in graphs, rather than mere fact retrieval, is often crucial (Hunter *et al.*, 1987; Lehrer and Romberg, 1996).

IMPLICATIONS FOR INSTRUCTION

The research reviewed in this article addresses two major educational goals. First, how can quantitative data be displayed effectively so that students can understand the relevant quantitative information? Second, what learning activities help students gain graphical literacy skills necessary to draw accurate quantitative and theoretical conclusions from the data in different graphical displays? Implications related to these two goals are outlined below.

Implications for Data Display

Below are the major principles of graph design implied by the research reviewed in this article. For additional advice, the reader is referred to Kosslyn (1994) and Tufte (1983).

1. *Choose the format depending on the communication goal.* Specifically, use line graphs for trends, bar graphs for discrete data, pie charts for percent data, divided bar charts when both absolute and percent data are relevant, and three-dimensional displays when integrating information is important but exact metric information is less important.
2. *Use multiple formats to communicate the same data.* Because different displays make different information salient, plotting the same data in multiple formats may be beneficial if there are multiple quantitative facts to be communicated, especially when the data are complex (e.g., a table to make exact quantities readable and a line graph to communicate a trend). Using multiple displays may have an instructional value as well; students may be able to build on prior knowledge and graph skills to form graph schemas relevant to novel or unfamiliar graph formats.
3. *Use the "best" visual dimensions to convey metric information whenever possible.* Follow the rankings of Cleveland and McGill (1985) shown in Table I to choose the best visual dimensions for representing quantity, but keep in mind that the ideal format is also a function of task (Simkin and Hastie, 1986).
4. *Use animation with caution.* Although some research suggests that animation has benefits for complex scientific visualization

tasks, other studies suggest that the comprehension of animation is difficult.

5. *Reduce working memory demands*. Keeping track of referents is a cognitively demanding task. Use meaningful colors, symbols, and labeling lines or bars, rather than legends, to reduce cognitive load. In addition, minimize the amount of information conveyed in any individual graph.
6. *Choose colors carefully*. Use different colors that are well separated in the spectrum when communicating categorical information. Use saturation or darkness of the same colors for more continuous data if necessary (but, if possible, use spatial extent for continuous data).
7. *The "third" dimension is okay unless precise metric information is needed*. Differences in accuracy in perceiving "3D" bars are minimally different from "2D" bars, and they are sometimes considered visually appealing.
8. *Choose aspect ratio and data density (graph size) carefully*. Viewers' interpretations of data are frequently influenced by the aspect ratio (relative height and width) or data density (of scatterplots, typically). Consider what size/aspect ratio communicates the main point without misleading the audience.
9. *Make graphs and text consistent*. Novice graph viewers often do not have the graph knowledge necessary to map between visual features and meaning. Describing the main point of the graph in the text in a format consistent with the graph might guide their interpretations. In addition, novice graph viewers are frequently influenced by their own prior knowledge. If data are inconsistent with prior knowledge, highlight the data in the text or caption so that the viewer is alerted.

Implications for Teaching Graphical Literacy

The research reviewed in this article also has implications for teaching graphical literacy skills. Although some conclusions can be drawn directly from the research, others listed here are more indirect implications of research.

1. *Graphical literacy skills should be taught in the context of science and social science*. Graph viewers are highly influenced by context in their interpretations of data, and students taught about graph reading in an abstract context may not be able to apply graph reading knowledge to real contexts in which their beliefs or expectations might influence their interpretations. An added benefit to teaching graphs in the

context of science and social science is that students may also learn that graphs are a tool for critically evaluating data, not just a tool for information delivery.

2. *Translating between representations may be beneficial*. The effect of format on viewers' interpretations of data suggests that activities that allow students to translate from one representation to another may enhance their ability to link visual feature information to the relevant quantitative information, and therefore be able to interpret graphs more easily.
3. *Explicitly focus on the links between visual features and meaning*. Just as in the mathematics classroom, where there is an emphasis on understanding the mapping between equations and graphs (e.g., Leinhardt *et al.*, 1990), it may be beneficial to have students explicitly focus on the mapping between visual features and meaning in other contexts. One such context is fostered by *Model-it is*, a software for scientific modeling and simulation. In this software, students can manipulate either text (e.g., "As air pollution increases, water pollution increases vs. water pollution increases a lot") or a graph (e.g., with a greater or lesser slope) to explicitly learn the mapping between the graph and the statements (Jackson, Stratford, Krajcik, and Solloway, in press).
4. *Make graph reading metacognitive*. Many of the errors in graph reading, such as interpreting data iconically or being misled by prior expectations, may be reduced by training viewers to think of graph reading as an interpretation and evaluation task as opposed to a mere fact retrieval task.

REMAINING QUESTIONS

Despite the relatively large body of research on graph comprehension, a number of questions remain. First, most cognitive research on graphs has focused on graph comprehension, rather than graph construction (especially outside of the mathematics context; Leinhardt *et al.*, 1990). There are a number of questions about how students with different levels of expertise construct graphs. For example, what kinds of errors are made in constructing graphs? Do students choose different kinds of graph formats depending on the goal (to communicate specific facts, to test a hypothesis, etc.), or are they biased to use familiar formats? Students' abilities to construct graphs, and how this relates to their ability to comprehend graphs, is particularly relevant for project- or design-based science activities in which students create graphs of data that they collect.

A related question concerns metaknowledge about graphs in science and social science. For example, do students recognize the potential of misleading graphs? How might one foster a student's ability to think critically about the data presented in graphs, rather than to merely retrieve information presented in them? What do students think about the uses or benefits of graphs? In a recent pilot study of middle-school students' conceptions of graphs, I found that most students recognized the communicative function of graphs. However, graphs were less commonly viewed as a tool for thinking about data, for example, to see if data are consistent with hypotheses or to test a model. What kinds of activities might foster students' thinking about graphs in terms of scientific reasoning? One possibility is that the use of graphs embedded in larger science or social science activities may lead students not only to a better ability to interpret graphs but also a better understanding of the function of graphs in science and social science (see Lehrer and Romberg, 1996, for an example).

Finally, given that we have highlighted the importance of graph comprehension as a *science or social science reasoning process*, rather than as a fact retrieval process, the role of graphs in scientific reasoning needs to be examined in greater detail. How might different ways of presenting data promote or inhibit different kinds of scientific reasoning? One component of expertise in scientific reasoning is the ability to think about data by relating the data directly to theories, conclusions, hypotheses, or explanations (e.g., Schunn and Anderson, 1999). When interpreting graphs, are expert graph readers more likely to provide explanations, rather than mere descriptions of graphs? How might we foster students' ability to explain, rather than merely describe information in a graph? One possibility that we are currently examining is whether tasks in which students are asked to draw graphs predicting results based on certain theories may promote their later ability to explain data and relate data to theories and hypotheses.

Another component of scientific reasoning is the ability to ask the right questions, and to design experiments that test the right questions (Schunn and Anderson, 1999). Do certain kinds of data displays support reasoning about the right questions? Many of the current approaches to teaching scientific inquiry skills incorporate various types of data displays (either constructed by the students or presented to them), and there are several possible benefits to some of these activities. In BGuILE, for example (Reiser *et al.*, in press), the overarching goal is to teach students about scientific reasoning in the context of evolution. Students form and test hypotheses about why certain characteristics of animals survived via evolution and others did not. One aspect of BGuILE is that the data are plotted in scatterplots, and students can select data points or sets of data points in which an animal has certain features or sets of features. It is possible that the interactive features of this display highlight the possible questions that students might ask, and thus promote their scientific reasoning skills. An important remaining question is how might different features, especially interactive features in which the data are not merely presented but *manipulated*, promote thinking about data.

In summary, three classes of factors have been shown to affect the interpretation of graphs: (1) the characteristics of the visual display, (2) the viewer's knowledge of graphical schemas and conventions, and (3) the content of the graph and the viewer's prior knowledge and expectations about that content. Research on graph interpretation can be used to help develop guidelines for the design of graphs as text adjuncts or in educational software, as well as for the teaching of graphical literacy skills.

ACKNOWLEDGMENTS

We would like to thank Eric Freedman, Steve Kosslyn, Akira Miyake, Dan Robinson, and Howard Wainer for their helpful comments on an earlier version of this paper.

REFERENCES

Armento, B. J., Nash, G. B., Salter, C. L., and Wixson, K. K. (1991). *A More Perfect Union*, Houghton Mifflin, Boston.

Becker, R. A., Cleveland, W. S., and Wilks, A. R. (1988). Dynamic graphics for data analysis. In Cleveland, W. S., and McGill, M. E. (eds.), *Dynamic Graphics for Statistics*, Wadsworth, Belmont, CA, pp. 1–.

Bell, A., and Janvier, C. (1981). The interpretation of graphs representing situations. *Learn. Math.* 2: 34–42.

Bertin, J. (1983). In Berg, W. (trans.), *Semiology of Graphics: Diagrams Networks Maps*, The University of Wisconsin Press, Madison, WI.

Brockmann, R. J. (1991). The unbearable distraction of color. *IEEE Trans. Profession. Commun.* 34: 153–159.

Bryant, P. E., and Somerville, S. C. (1986). The spatial demands of graphs. *Br. J. Psychol.* 77: 187–197.

Carmichael, L. C., Hogan, H. P., and Walters, A. A. (1932). An experimental study of the effect of language on the reproduction of visually perceived form. *J. Exp. Psychol.* 15: 73–86.

Carpenter, P. A., and Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *J. Exp. Psychol. Appl.* 4: 75–100.

Carswell, C. M., Emery, C., and Lonon, A. M. (1993). Stimulus complexity and information integration in the spontaneous interpretation of line graphs. *Appl. Cogn. Psychol.* 7: 341–357.

Carwell, C. M., Frankenberger, S., and Bernhard, D. (1991). Graphing in depth: Perspectives on the use of three-dimensional graphs to represent lower-dimensional data. *Behav. Inform. Technol.* 10: 459–474.

Carswell, C. M., and Wickens, C. D. (1987). Information integration and the object display: An interaction of task demands and display superiority. *Ergonomics* 30: 511–527.

Chernoff, H. (1973). Using faces to represent points in *k*-dimensional space graphically. *J. Am. Stat. Assoc.* 68: 361–368.
Clement, J. (1985). Misconceptions in graphing. In Streetfland, L. (ed.), *Proceedings of the Ninth International Conference of the International Group for the Psychology of Mathematics Education*, IGPME, Utrecht, The Netherlands, Vol. 1, pp. 369–375.
Cleveland, W. (1993). *Visualizing Data*, AT&T Bell Laboratories, Murray Hill, NJ.
Cleveland, W. S., Diaconis, P., and McGill, R. (1982). Variables on scatterplots look more highly correlated when the scales are increased. *Science* 216: 1138–1141.
Cleveland, W. S., and McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *J. Am. Stat. Assoc.* 77: 541–547.
Cleveland, W. S., and McGill, R. (1985). Graphical perception and graphical methods for analyzing scientific data. *Science* 229: 828–833.
Fisher, H. T. (1982). *Mapping Information*, Abt Books, Cambridge, MA.
Fischer, M. H. (2000). Do irrelevant depth cues affect the comprehension of bar graphs? *Appl. Cogn. Psychol.* 14: 151–162.
Freedman, E. G., and Smith, L. D. (1996). The role of data and theory in covariation assesesment: Implications for the theory-ladenness of observation. *J. Mind Behav.* 17: 321–343.
Gattis, M., and Holyoak, K. (1995). Mapping conceptual to spatial relations in visual reasoning. *J. Exp. Psychol. Learn. Mem. Cogn.* 22: 1–9.
Guthrie, J. T., Weber, S., and Kimmerly, N. (1993). Searching documents: Cognitive process and deficits in understanding graphs, tables, and illustrations. *Contemp. Educ. Psychol.* 18: 186–221.
Halford, G. S., Wilson, W. H., and Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behav. Brain Sci.* 21: 803–864.
Hegarty, M., Quillici, J., Narayanan, N. H., Holmquist, S., and Moreno, R. (1999). Multimedia instruction: Lessons from evaluation of a theory-based design. *J. Educ. Multimed. Hypermed.* 8: 119–150.
Hoffman, R. R., Deitweiler, M., and Lipton, K. S. (1993). General guidance for establishing color standards for meteorological displays. *Weather Forecast.*
Howell, W. C. (1993). Engineering psychology in a changing world. *Ann. Rev. Psychol.* 231–263.
Huber, P. J. (1987). Experiences with three-dimensional scatterplots. *J. Am. Stat. Assoc.* 82: 448–453.
Hunter, B., Crismore, A., and Pearson, P. D. (1987). Visual displays in basal readers and social science textbooks. In Willows, D., and Houghton, H. A. (eds.), *The Psychology of Illustration, Vol. 2: Instructional Issues*, Springer, New York, pp. 116–135.
Janvier, C. (1981). Use of situations in mathematics education. *Educ. Stud. Math.* 12: 113–122.
Jennings, D. L., Amabile, T., and Ross, L. (1982). Informal covariation assessment: Data-based versus theory-based judgements. In Kahneman, D., Slovic, P., and Tversky, A. (eds.), *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge, England, pp. 211–230.
Kaput, J. J. (1987). Representation and mathematics. In Janvier, C. (ed.), *Problems of Representation in Mathematics Learning and Problem Solving*, Erlbaum, Hillsdale, NJ, pp. 19–26.
Kosslyn, S. (1985). Graphics and human information processing: A review of five books. *J. Am. Stat. Assoc.* 80: 499–512.
Kosslyn, S. (1989). Understanding charts and graphs. *Appl. Cogn. Psychol.* 3: 185–225.
Kosslyn, S. M. (1994). *Elements of Graph Design*, Freeman, New York.
Kubovy, M. (1981). Concurrent pitch segregation and the theory of indispensable attributes. In Kubovy, M., and Pomerantz, J. (eds.), *Perceptual Organization*, Erlbaum, Hillsdale, NJ.
Larkin, J. H., and Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cogn. Sci.* 11: 65–99.
Lauer, T. W., and Post, G. V. (1989). Density in scatterplots and the estimation of correlation. *Behav. Inform. Technol.* 8: 235–244.
Legge, G. E., Gu, Y., and Luebker, A. (1989). Efficiency of graphical perception. *Percept. Psychophys.* 46: 365–374.
Lehrer, R., and Romberg, T. (1996). Exploring children's data modeling. *Cogn. Instr.* 14: 69–108.

Leinhardt, G., Zaslavsky, O., and Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Rev. Educ. Res.* 60: 1–64.

Levy, E., Zacks, J., Tversky, B., and Schiano, D. (1996). Gratuitous graphics: Putting preferences in perspective. *Human Factors in Computing Systems: Conference Proceedings*, ACM, New York, pp. 42–49.

Lewandowsky, S., and Behrens, J. T. (1999). Statistical graphs and maps. In Durso, F., Dumais, S., Nickerson, R., Schvaneveldt, R., Chi, M., and Lindsay, S. (eds.), *The Handbook of Applied Cognitive Psychology*, Wiley, Chichester, England, pp. 513–549.

Lewandowsky, S., and Spence, I. (1989). Discriminating strata in scatterplots. *J. Am. Stat. Assoc.* 84: 682–688.

Lohse, G. L. (1993). A cognitive model of understanding graphical perception. *Hum. Comp. Interact.* 8: 353–388.

Lord, C. G., Ross, L., and Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequent evidence. *J. Pers. Soc. Psychol.* 37: 2098–2110.

MacDonald-Ross, M. (1977). Graphics in texts. *Rev. Res. Educ.* 5: 49–85.

Maichle, U. (1994). Cognitive processes in understanding line graphs. In Schnotz, W., and Kulhavy, R. W. (eds.), *Comprehension of Graphics*, Elsevier, Amsterdam.

Marchak, F. M., and Marchak, L. C. (1991). Interactive versus passive dynamics and the exploratory analysis of multivariate data. *Behav. Res. Methods Instrum. Comput.* 23: 296–300.

McDermott, L., Rosenquist, M., and vanZee, E. (1987). Students difficulties in connecting graphs and physics: Example from kinematics. *Am. J. Phys.* 55: 503–513.

McKenzie, D. L., and Padilla, M. J. (1986). The construction and validation of the Test of Graphing in Science (TOGS). *J. Res. Sci. Teach.* 23: 571–579.

Merwin, D. H., Vincow, M., and Wickens, C. D. (1994). Visual analysis of scientific data; Comparison of 3d-topographic, color and gray scale displays in a feature detection task. In *Proceedings of the 38th Annual Meeting of the Human Factors and Ergonomics Society*, Human Factors and Ergonomics Society, Santa Monica, CA.

Morrison, J. B., Tversky, B., and Betrancourt, M. (2000). Animation: Does it facilitate learning? Paper presented at the *AAAI Spring Symposium on Smart Graphics, March, 2000*, AAAI Press.

Nachmias, R., and Linn, M. C. (1987). Evaluations of science laboratory data: The role of computer-presented information. *J. Res. Sci. Teach.* 24: 491–505.

Phillips, R. J. (1982). An experimental investigation of layer tints for relief maps in school atlases. *Ergonomics* 25: 1143–1154.

Pinker, S. (1990). A theory of graph comprehension. In Freedle, R. (ed.), *Artificial Intelligence and the Future of Testing*, Erlbaum, Hillsdale, NJ, pp. 73–126.

Preece, J. (1990). Some HCI issues concerned with displaying quantitative information graphically. In Gorny, P., and Tauber, M. J. (eds.), *Visualization in Human–Computer Interaction*, Springer, New York.

Quintana, C., Eng, J., Carra, A., Wu, H., and Soloway, E. (1999). *Symphony: A Case Study in Extending Learner-Centered Design Through Process Space Analysis*, Paper presented at CHI 99: Conference on Human Factors in Computing Systems, May 19–21, 1999, Pittsburgh, Pennsylvania.

Reiser, B. J., Tabak, I., Sandoval, W. A., Smith, B., Steinmuller, F., and Leone, T. J. (in press). BGuILE: Stategic and conceptual scaffolds for scientific inquiry in biology classrooms. In Carver, S. M., and Klahr, D. (eds.), *Cognition and Instruction: Twenty Five Years of Progress*, Erlbaum, Mahvah, NJ.

Scardamalia, N., Bereiter, C., and Lamon, M. (1994). The CSILE Project: Trying to bring the classroom into the world. In McGilly, K. (ed.), *Classroom Lessons: Integrating Cognitive Theory and Classroom Practice*, MIT Press, Cambridge, MA.

Schiano, J. D., and Tversky, B. (1992). Structure and strategy in encoding simplified graphs. *Mem. Cogn.* 20: 12–20.

Schmid, C. (1983). *Statistical Graphics: Design Principles and Practices*, Wiley, New York.

Schunn, C. D., and Anderson, J. R. (1999). The generality/specificity of expertise in scientific reasoning. *Cogn. Sci.* 23: 337–370.

Shah, P., and Carpenter, P. A. (1995). Conceptual limitations in comprehending line graphs. *J. Exp. Psychol. Gen.* 124: 43–61.

Shah, P., Mayer, R. E., and Hegarty, M. (1999). Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. *J. Educ. Psychol.* 91: 690–702.

Shah, P. (in press). Graph comprehension: The role of format, content, and individual differences. In Anderson, M., Meyer, B., and Olivier, P. (eds.), *Diagrammatic Representation and Reasoning*, Springer, New York.

Shah, P. (1995). *Cognitive Processes in Graph Comprehension*, Unpublished doctoral dissertation.

Shah, P., and Shellhammer, D. (1999). *The Role of Domain Knowledge and Graph Reading Skills in Graph Comprehension*, Paper presented at the 1999 Meeting of the Society for Applied Research in Memory and Cognition, Boulder, CO.

Simkin, D. K., and Hastie, R. (1986). An information processing analysis of graph perception. *J. Am. Stat. Assoc.* 82: 454–465.

Somerville, S. C., and Bryant, P. E. (1985). Young children's use of spatial coordinates. *Child Dev.* 56: 604–613.

Spence, I. (1990). Visual psychophysics of simple graphical elements, *J. Exp. Psychol. Hum. Percept. Perform.* 16: 683–692.

Stenning, K., and Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cogn. Sci.* 19: 97–140.

Stuetzle, W. (1987). Plot windows. *J. Am. Stat. Assoc.* 82: 466–475.

Tufte, E. R. (1983). *The Visual Display of Quantitative Information*, Graphics, Cheshire, CT.

Tversky, B. (in press). Spatial schemas in depictions. In Gattis, M. (ed.), *Spatial Schemas and Abstract Thought*, MIT Press, Cambridge.

Tversky, B., Kugelmass, S., and Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cogn. Psychol.* 23: 515–557.

Tversky, B., and Schiano, D. J. (1989). Perceptual and conceptual factors in distortions in memory for graphs and maps. *J. Exp. Psychol. Gen.* 118: 387–398.

Vernon, M. D. (1950). The visual presentation of factual information. *Br. J. Educ. Psychol.* 20: 174–185.

Wickens, C. D., Merwin, D. H., and Lin, E. L. (1994). Implications of graphics enhancements for the visualization of scientific data: Dimensional integrality, stereopsis, motion, and mesh. *Hum. Fact.* 36: 44–61.

Wilkinson, L. (1999). Graphs for research in counseling psychology. *Counsel. Psychol.* 27: 384–407.

Winn, B. (1987). Charts, graphs, and diagrams in educational materials. In Willows, D., and Houghton, H. A. (eds.), *The Psychology of Illustration*, Springer, New York.

Zacks, J., and Tversky, B. (1999). Bars and lines: A study of graphic communication. *Mem. Cogn.* 27: 1073–1079.

Zacks, J., Levy, E., Tversky, B., and Schiano, D. J. (1998). Reading bar graphs: Effects of depth cues and graphical context. *J. Exp. Psychol. Appl.* 4: 119–138.