9. ANALYSIS AND VISUALIZATION

Chaired by Doris Aaronson, *New York University*

Visualization of multivariate data: Human-factors considerations

EDWARD COLET and DORIS AARONSON
New York University, New York

Recent software provides new tools for visualizing multivariate data that facilitate data analysis. We focus on (1) the learnability and use of visualization systems, and (2) the perceptual and cognitive processes involved in viewing visualizations. Effective visualization systems support a broad range of user tasks and abilities, are easy to learn, and provide powerful and flexible output formatting. Effective visualizations incorporate Gestalt and other perceptual and cognitive principles that encourage more rapid, automatic processing, and less slow, controlled processing.

Computers have always provided powerful analytical tools and efficient means of “number crunching.” Today, computers are also efficient at generating and presenting information graphically. This combination of computing and graphical power makes them almost ideal tools for combining data analysis and scientific visualization. Visualization can be thought of as the process of using computers to display information visually for purposes of analysis (Hummel, 1994), using a combination of points, lines, coordinate systems, numbers, symbols, words, shading, and color to represent measured quantities (Tufte, 1983). In essence, visualizations can serve as graphical descriptions and summaries of numerical information. Although it is often treated as a separate topic in and of itself, visualization has the same purpose as mathematical and statistical procedures in data analysis—namely, to make apparent any trends and patterns in the data that indicate underlying relationships among variables. Data analysis is therefore a continuum of activity ranging from exploratory work to the presentation of results.

In this paper, we first discuss the value of visualization using both simple (consisting of few variables) and complex (multivariate) data sets. Several examples generated from currently available software are presented. Despite the fact that visualization can add value to data analysis and that the software to generate visualizations is commercially available, there seems to be limited use of vi-

sualization techniques as a regular component of data analysis in the social and behavioral sciences. The latter part of this paper presents a human-factors perspective to account for the limited use of visualization, focusing first on the factors relevant to the use of computer-based systems to generate visualizations, and then on the cognitive and perceptual factors that pertain to the comprehension and interpretation of visual views of data.

THE VALUE OF VISUALIZATION

Visualization can aid in the theoretical interpretation of data by guiding the process of math modeling. For example, a Pearson correlation coefficient is often used to quantitatively describe a linear relationship between two variables. But it is important to view the data in order to assess the appropriateness of this measure. A scatterplot can easily expose a curvilinear trend in the data, or the presence of an outlier that highly influences the Pearson coefficient. Thus, for both simple and complex data, visualization is useful for guiding researchers toward appropriate math models, and away from inappropriate ones.

Visualization often enables and encourages extensive exploratory analysis of data to discover trends that may not be evident with traditional analyses. Rabenhorst (1993) makes this point as follows: Given a multidimensional data set of N variables, the influence of one or more of the variables may be limited to a particular subset of the observations, and therefore it is important to be able to identify that subset of cases or variables. Users should easily be able to generate a variety of “views” of the data in order to discover relationships of all kinds. These underlying relationships may be overlooked in an analysis that relies simply on p values and basic hypothesis testing (Cohen, 1994; Loftus, 1993).

This work was supported in part by USPHS Grant MH 16, 496 to D.A. We thank Tom Donnelly, Don Icken, Lloyd Kaufman, Neal Keller, Robert Proctor, Linda Colet, and Jennifer Thomas for helpful comments on earlier drafts, and Lucien Van Elsen for technical help. Correspondence should be addressed to E. Colet or D. Aaronson, New York University Psychology Department, 6 Washington Place, 8th Floor, New York, NY 10003 (e-mail: edc@xp.psych.nyu.edu or doris@xp.psych.nyu.edu).

CURRENT VISUALIZATION TECHNIQUES

Some powerful visualization techniques (partially reviewed in Marchak, 1994) are illustrated with simulated data in Figures 1–3. The *surface plot* in Figure 1A is one technique used to view three variables. Given a data set that contains three variables, x , y , and z , we can represent the magnitude of z by the height above the intersection of a particular x and y value in a two-dimensional projection of three-dimensional data. In this example, our x -, y -, and z -axes represent age (in years), hours of practice, and pronunciation accuracy in a second language. Unless the viewpoint can be rotated, a drawback of this representation is that a peak may obscure other interesting visual features in the data.

The *contour plot* in Figure 1B contains the same data as Figure 1A. It also projects three dimensions (variables) onto a two-dimensional space. It is possible to generate a contour plot for any data set that contains three variables, x , y , and z , which can be arranged in a grid so that each (x,y) pair has exactly one z value. In Figure 1B, the horizontal and vertical axes represent values for the x and y variables, respectively. The contour lines correspond to discrete data levels of the third variable, z (z can be thought of as a position on a surface given by the values of x and y). As in a topographical map, the distance between contour lines represents a rate of change, with close contours indicating rapidly changing values.

Both the surface and contour plots in Figure 1 show that the greatest pronunciation accuracy (i.e., where the peaks are highest) occurs after a certain age (beyond the halfway point of the x -axis), and after a certain amount of practice (beyond the halfway point of the y -axis).

Figure 2 illustrates a three-dimensional view of a scatterplot in a *transparent cubical box*. As in the surface plot, each of the variables has an independent axis, but unlike in a surface plot, a corner of the cube (the lower-left corner) is the origin for the three axes. The ranges of values for each of the variables are normalized to cover the same plotting area, so that all data points fit within the cube's space. (Might future software allow transparent rectangles?) The cube can be rotated to be oriented in any direction, and rapidly rotating the cube (dynamic spinning) provides an effect similar to that of animation. Since each variable has a unique axis, the addition of color can be used to represent a fourth variable, such as experimental condition. In this example, the axes represent subject number (subject), reaction time (rt), and a performance score (score). If performance scores were perfectly associated with an experimental condition, each plane in the figure would be uniquely colored.

The *scatterplot matrix* in Figure 3 consists of individual scatterplots of all pairs of variables arranged in rows and columns. The simultaneous view of all pairs of variables makes it easy to detect trends and patterns (linear or otherwise) in data. It can also encourage exploratory analysis. For example, if two variables appear closely related, one can look for other variables that may be related to either member of the original pair, and then these

combinations of variables can be viewed in a separate visualization. In this example, it appears that the correlation between starting salary and job satisfaction (first row, second column) is stronger than that between job satisfaction and education level (second row, third column).

HUMAN FACTORS ISSUES

Given the current capability of computers as tools for analysis and the value that can be added through the use of visualization, why is visualization of multivariate data not commonly used in data analysis and for reporting results? What factors contribute to the limited use, and limited effectiveness, of visualization? These questions can be answered from a human-factors (HF) perspective.

Figure 1. (A) A surface plot showing pronunciation accuracy (z -axis) as a function of age in years (x -axis) and hours of practice (y -axis). (The default in SAS/Graph plots the values from more to less.) (B) A contour plot uses contour lines to represent discrete levels of pronunciation accuracy as a function of practice and age.

The HF perspective used here (in the context of a discussion of visualization) is the view that the person represents an integral element of a system made up of system hardware, system software, and the cognitive and perceptual processes that are part of the human interface. Thus, considerations of user requirements and user capabilities in the design, use, and effectiveness of software are critically important (Gould & Lewis, 1983; Kosslyn, Gershon, Lefkowitz, & Pearlman, 1992; Rogowitz & Treinish, 1993). The importance of considering the user's needs and requirements has formed part of the principles and guidelines for the design of interactive systems (see Schneiderman, 1980, for a review).

From this perspective, the following HF dichotomy becomes useful: (1) HF issues that are relevant to a user interacting with the computer-based system to generate visualizations; and (2) HF issues for the user (and others) in comprehending the output of the system (i.e., in making use of the generated visualizations).

HUMAN FACTORS INVOLVED IN USING COMPUTER-BASED TOOLS

In terms of systems characteristics necessary for generating effective visualizations, the first part of our dichotomy is concerned with four aspects: (1) Systems should support the activities and requirements needed for data analyses; (2) systems should provide the user with sufficient control and power to customize the formatting of graphical output; (3) visualization systems and tools should be easy to learn and use, regardless of the user's computer experience; and (4) systems should be accessible for novices and remain useful for experts.

Supporting User Requirements in Data Analysis

Most current visualization tools (and familiar graphics packages) do not support the entire process involved in data analysis. Most tools are designed only to support the task of presenting results. However, data analysis is really a continuum of activity ranging from the initial

exploratory activity through to the final presentation of results, and as such, the various parts along this continuum are associated with characteristically distinct cognitive processes that depend on the task at hand.

Springmeyer, Blattner, and Max (1992) characterize the process of scientific analysis as involving the following: (1) queries for quantitative information, such as the generation of summary statistics; (2) comparisons among values, or even shapes and objects (e.g., as in comparing heights of bar graphs) to find trends and patterns in data; (3) mathematical processes, such as transforming variables and values and/or creating derived measures (interestingly enough, much of this activity may represent attempts to "see" one's data without the use of visualization tools); (4) data management, such as creating new data sets from summary statistics of particular subsets; and (5) record keeping and documentation of existing routines. While researchers occasionally perform all or most of the above activities using different modules within a single application (e.g., SAS [SAS Institute, Cary, NC], SAS/Graph, saving SAS logs), more often they use several packages that may not interact easily with each other, but that are easier to use for the specific task at hand (e.g., using Harvard Graphics [Software Publishing Corp., Santa Clara, CA] can be easier than using SAS/Graph for generating graphs).

User Control for Formatting Output

The objective of visualization software is to create images of one's data (e.g., as a graph). In general, a graph has three areas: the axes, the data area within the axes, and the annotations. Effective visualization tools should provide the user with ample formatting control over each area. Marsh (1994) lists a variety of examples of desired formatting controls (and recommended packages) relevant to each of the three areas in a graph.

For flexible control over the *axes*, Axum (TriMetrix, Seattle, WA) and EasyPlot (Spiral Software, Brookline, MA) are noteworthy. Their desirable formatting features include: (1) the ability to draw multiple *x*- and *y*-axes in order to plot the same variable on multiple measurement scales or a second variable added to one of the dimensions; (2) the ability to determine the number, spacing, and appearance of tick marks on both axes; and (3) the ability to scale the *x*- and *y*-axes with respect to each other.

To format the *data area* of a graph, the desired features include: (1) the ability to specify size and shape of plotting symbols (most packages have limited symbol libraries, but some of the recent graphics tools allow one to control the size of the plotting symbol on the basis of values of a third variable); (2) the ability to control the line style and line width; (3) the ability to draw error bars for data points with respect to both the *x*- and *y*-axes simultaneously; and (4) the ability to easily move/delete parts of a graph with a mouse and to have the data automatically updated for reexamination (these features would facilitate exploratory analyses of one's data).

Controlling the *annotation* (legends, titles, etc.) involves the ability to use Greek symbols, superscripts,

Figure 2. A transparent cube, or a three-dimensional scatterplot, showing the relationships between individual subjects (subject), reaction times (rt), and performance (score).

Figure 3. A scatterplot matrix showing individual scatterplots between starting salary, job satisfaction, and education level.

and subscripts, as well as the ability to size and place text anywhere (Axum, SigmaPlot [Jandel Scientific, San Rafael, CA], and SlideWrite [Advanced Graphics Software, Inc., Carlsbad, CA] are recommended; these are difficult to do in Quattro [WordPerfect Corp., Orem, UT] and with the graphics of SAS/BASE). An additional possibility for annotation may be to use hypertext links, which would enable one to switch from the graph to expanded notes and text that summarize the main findings, or even to switch to a file containing the raw-data values.

Ease of Learning the Software

The critical question here is whether a user can generate something useful in a short amount of time. Regardless of the method of interaction (i.e., whether choosing menu items, entering command-line arguments, using keystroke combinations, or dragging and dropping icons across the screen), actions should be intuitive and feel natural for the user. As a nonintuitive example, in the Diamond package (BMDP/Diamond, Los Angeles, CA), color is applied by a mouse click and a keystroke (Rabenhorst, 1993); in some cases, the keystroke is a certain lowercase letter key, while in others, it is the "!" key, and in yet others, it is the "home" key (all of which also require the mouse to be positioned in certain areas). Each of these combinations is valid in one situation or another, but knowing which one to apply may be difficult to remember when first learning to use the system.

Documentation can be an important learning aid. It should always be available, despite the finding that many people typically do not read it (Carroll 1990). An example of task-oriented documentation is Carroll's "Guided Exploration Cards," in which information is organized and presented in the form of answers to questions about common tasks (Carroll). Additional on-line help can be context sensitive, perhaps using hypertext links (Nielsen, 1990) to make it easier for a user to find related information.

Research on individual differences has HF implications for systems learnability. Aaronson (1982) found that prior experience and general skill level with computers was an important factor in learning to use a new system. This implies that the methods of teaching or learning to use the tools should be adaptive and individualized, as opposed to standardized approaches that might demand too much or too little from a user (on the basis of one's skill level). An educational approach using on-line, self-paced tutorials may be more effective.

Regardless of the educational approach, feedback and guidance are useful for the prevention of, and recovery from, errors. Further, feedback could communicate which features may be useful/applicable on the basis of the user's current context. Since some of the recent visualization tools allow for the creation of novel representations of data, this type of feedback may be especially useful. At present, the users often do not know which of several possibly novel representations may be applicable, given their data; current tools provide little guidance (Rogowitz & Treinish, 1993).

Improving Long-Term Usefulness

An effective visualization system should be useful both for new users and for experienced experts. An example of such a software package is Mathematica (Wolfram Research, Inc., Champaign, IL), because it can be used to perform both simple analyses (those that are useful for novices) as well as complex ones (those that are useful for experts; observed by Marchak & Zulager, 1992). Software with this capability may depend on a "flexible" interface—one that provides the user with varying degrees of power and control. For example, some packages allow users to select menu-driven options (for novices or for shortcuts), as well as to generate program code (for experts or for handling unusual situations; e.g., SAS/ASSIST and SAS's program editor). A trade-off typically exists between the power of a software tool (as measured by the number of features that are available to the user to customize/control the results) and the ease of learning to use the tool (as discussed above). This is partly a result of the fact that the amount a user may need to learn increases with the number of available features.

HUMAN FACTORS IN PERCEIVING AND COMPREHENDING SYSTEM OUTPUT

The second part of our dichotomy deals with human-factors issues relevant to perceiving and comprehending graphic output. Although such HF research is limited, the following are relevant questions to consider: (1) What characterizes a "good" (and therefore effective) graphical image? (2) Can these characteristics be generalized as a set of principles? (3) What psychological mechanisms operate when viewing graphs? (4) Do the same principles and processes apply for the visualizations of both simple and more complex data sets? According to Tufte (1983), "graphical excellence is that which gives the viewer the greatest number of ideas in the shortest time with the

least ink in the smallest space. Graphical excellence is nearly always multivariate. Graphical excellence requires telling the truth about the data" (p. 51). This definition implies that a "good" image clearly and simply communicates trends and patterns in the data with a minimal amount of encoding effort on the part of the user.

Characteristics of "Good" Visualizations: Gestalt Principles

The degree to which clarity and meaning become apparent depends in part on a graph's consistency with Gestalt principles (Foley & Moray, 1987). These principles include: (1) proximity (objects that are close are readily seen as a group, as in some graphs with multiple points or curves); (2) similarity (objects of the same or similar color or shape are readily seen as a group; this has implications for coding multivariate information to highlight relatedness or contrast); (3) common fate ("elements in the field undergoing simultaneous, correlated changes are grouped together" [Pomerantz & Kubovy, 1986, p. 36-14]; e.g., a slope or direction is perceived if two or more curves are approximately parallel); (4) continuation (this attribute includes smoothness—as opposed to jaggedness—of a curve [Pomerantz & Kubovy], closure of space within lines or curves, and symmetry or balance among objects in the space); and (5) conformity with user set (the user may be "set"—i.e., by innate perceptual mechanisms, past conventions, related graphs, or textual discussion) for certain featural aspects.

In the spirit of recent HF research (see, e.g., Gillan & Lewis, 1994), one might postulate two stages of graphical information processing—namely, *perceptual* processing of the *visual* image, followed by *cognitive/interpretive* processing of the *scientific* message. Processing will be most effective when visual characteristics that contribute to a "good" image also communicate information that matches the user's interpretive intent. These two stages may be differentially facilitated by the various visualizations in Figure 4 of exactly the same (hypothetical) data. The data are RTs from a recognition memory task which can be viewed as having any one of three 2×2 designs, each involving two of the following three factors defined as main effects: study mode (pictures, words), test mode (pictures, words), and mode agreement (study and test agree or disagree).

First, purely in terms of Gestalt visualization principles—and *ignoring* the scientific interpretations—the "pictures" in Figure 4 might be ranked as A, B, E, F, D, and C. These rankings would be determined by the combined effects of the number, strength, and presence of the Gestalt features that operate on the graph's lines and spaces (figure and ground, respectively). Panel A ranks best, as it reflects proximity, similarity, common fate, and symmetry (across a diagonal axis), with panels B and E less "good" on these attributes. In contrast, panels F, D, and C rank poorly because proximity, similarity, and common fate are less prominent, and the lines break the space into several jagged shaped pieces (although continuity of the crossed lines is prominent). Thus the arrangement or

organization of *parts* in panels A and B provides a simpler whole than is provided by the other panels, consistent with Gestalt views (Pomerantz & Kubovy, 1986).

When the scientific information is being cognitively processed, if there are no strong a priori hypotheses, the statistical main effects are generally evaluated before the interactions (e.g., Hays, 1994). Note that what is defined and graphed as a main effect in one experimental design can be represented as an interaction in another. For example, the main effect of mode agreement in panels E, D, B, and A is the study $\times$ test-mode interaction in panels F and C. Thus, a ranking of the six components of Figure 4 on the basis of the salience of main effects relative to interactions appears to be A, B, E, D, C, and F. Panels A and B both highlight the main effects of test mode and mode agreement, but the steeper slopes and closer spacing in panel A add salience. Panels E, D, and C each highlight one main effect and one interaction. Finally, panel F highlights the study $\times$ test-mode interaction, but it does not make salient either the main effects or the two interactions (between mode agreement and either study or test mode).

Interestingly, in this example, there are some similarities (panels A, B, and E) between visual (i.e., using perceptual processing) and scientific (i.e., using cognitive processing) rank orderings and also some differences (panels F, D, and C for visual ordering vs. panels D, C, and F for scientific ordering). Thus, researchers should explicitly evaluate both visual and scientific factors with an eye toward optimizing the ease and effectiveness of the total set of psychological processes. Software that facilitates exploratory data analysis by providing user-friendly ways in which to display natural "families" of graphs, perhaps automatically with a single instruction set, would enhance the research tools that are currently available.

Operative Psychological Mechanisms and Principles

Mechanisms of automatic and controlled processing. The HF perspective on Gestalt principles is that the psychological processes induced by "good" images are a function of underlying mechanisms that are "hard-wired" in neural circuitry (Foley & Moray, 1987), and thus imply the existence of universal principles of perceptual organization. Thus, displays that are easy to read and that are correctly perceived with little effort should be consistent with these principles. Conversely, displays that are hard to interpret, that preoccupy attention, and that result in perceptual errors are likely to involve forced and misleading correspondences. Thus, we suggest a continuum in graphical processing, with the endpoints being consistent with the notions of *automatic* and *controlled* processing (Fisk, Ackerman, & Schneider, 1987; Shiffrin & Schneider, 1977).

Principles to minimize encoding effort. For visualizations of multivariate and multifactor experiments, global attributes, such as the following, may gain importance: (1) figure/ground separation, for discrimination; (2) figure boundaries, for clarity; and (3) closure, to enhance the perceptual processes (McCormick, 1976).

Figure 4. Six graphs of the same hypothetical recognition-time data, which vary in their use of Gestalt visualization principles and the scientific messages conveyed.

Further, past research (Foley & Moray, 1987) suggests that 95% accuracy in distinguishing between different objects can be achieved using the following maximum perceptual limits: for qualitative dimensions, 15 shapes and 10 colors; for quantitative dimensions, 6 area sizes, 6 lengths, and 4 intensities. These attributes may interact with each other, as well as with other attributes, such as illumination. However, because these criteria neglect the effects of context and task considerations, they represent a simplified view of a user's perceptual limits.

The complex representation in Figure 5, a *foam plot*, is designed to be consistent with yet other perceptual principles. Specifically, these are that (1) important features and characteristics of the data should be made readily perceptible or salient; and (2) these perceptual features should be easy to associate with meaningful interpretations. Figure 5 is a graphical view of a correlation matrix with respect to a particular focus variable selected by the user. Variables in the foam plot normally appear as color-coded "bubbles," with the focus variable appearing in the center and bubbles of the same color representing the same variable (in Figure 5, color has been recoded as shading and pattern). Nonfocus variables are arranged (counterclockwise) in descending order on the basis of their area size, which is an index of their correlation with the focus variable. These nonfocus variables, in turn, serve as foci in remote correlation patterns. Univariate statistics for each variable are communicated by features of that variable's "bubble" representation. Specifically, kurtosis is indicated by the deviation from circularity, skewness is indicated by orientation

(i.e., the deviation of major axes of ellipses or variable labels of circles from the horizontal), and the standard deviation is indicated by degree of perimeter smoothness (greater jagged deviations from elliptical smoothness indexing greater standard deviations). Thus, in a foam plot, visually salient features such as color (or shading/pattern), size, shape, and orientation of bubbles communicate information regarding identity, correlation magnitude, kurtosis, and skewness. Appropriate heuristic use of such salient features can minimize the encoding effort and mental load for the viewer.

Layout principles for presenting information. Most of our HF discussion has concerned the data area of a graph. Additional layout principles can enhance the other two aspects of a graph—that is, the axes and graphical annotations (Marsh, 1994). Such principles include (1) balance (objects—including verbal labels—should frame and draw the eyes to the center); (2) gridding (borders should be used in consistent ways to separate areas); and (3) proportion (comparative sizes of segments should be consistent).

In addition to visual principles regarding objects, empirical guidelines, based on research on readability, are recommended for fonts and text that affect readability. For example, capital letters and numerals are read most accurately when the ratio of stroke width to height is between 1:6 and 1:8 for black on white (1:10 for white on black) and when character width is at least two-thirds of the height (McCormick, 1976). A mixture of upper- and lowercase letters is more effective than all uppercase letters, because lowercase letters provide more discrim-

Figure 5. A foam plot showing the three variables most highly correlated with hours of television watching. Color (recoded here as shading and texture), size, shape, and orientation of each bubble correspond to identity, correlation magnitude, kurtosis, and skewness of that variable.

inable features, as well as a distinctive perceptual "envelope" for the word as a whole.

CONCLUSION

The HF perspective put forth here maintains that the effectiveness of visualizations depends on developing systems that allow researchers to easily generate visualizations that viewers perceive and comprehend quickly. This is true both for the simple and familiar data views (line and bar graphs) and for the more complex visualizations (transparent cubes and foam plots). To view multiple variables simultaneously, novel ways of presenting and viewing data can be useful and effective.

Because of the novelty of some of the visualization procedures, there are many open research directions for psychologists in several subareas. Cognitive psychology has much to contribute to understanding visualizations. For example, what are the critical factors involved in deciding which data view to use, given a data set and a set of theoretical questions? And, what types of visualizations are more memorable for the viewer? Psychologists concerned with learning might focus on the question of which attributes of user interfaces facilitate learning to use visualization software, and on the qualitative ways in which such learning differs between novices and expert computer users. The perception researcher would be faced with many interesting questions, as visualization obviously involves visual processes, as well as psychophysical relationships. Even social psychologists would find some interesting research issues, as visualizations are used to communicate information to social groups with varying social communication styles. Can research suggest principles for optimizing these seemingly diverse styles of vi-

sual communication? As Woods, Johannessen, and Potter (1992) have rightly pointed out, applications of basic research must also involve considerations of both tasks and contexts.

REFERENCES

- AARONSON, D. (1982). An evaluation of Unix in the psychology computer laboratory. *Behavior Research Methods & Instrumentation*, **14**, 137-141.
- CARROLL, J. M. (1990). *The Nurnberg funnel: Designing minimalist instruction for practical computer skill*. Cambridge, MA: MIT Press.
- COHEN, J. (1994). The earth is round ($p < .05$). *American Psychologist*, **49**, 997-1003.
- FISS, A. D., ACKERMAN, P. L., & SCHNEIDER, W. (1987). Automatic and controlled processing. In P. A. Hancock (Ed.), *Human factors psychology* (pp. 159-197). Amsterdam: North-Holland.
- FOLEY, P., & MORAY, N. (1987). Sensation, perception and systems design. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 45-73). New York: Wiley.
- GILLAN, D. J., & LEWIS, R. L. (1994). A componential model of human interaction with graphs: Linear regression modeling. *Human Factors*, **36**, 419-440.
- GOULD, J. D., & LEWIS, C. (1983). Designing for usability: Key principles and what designers think. *Human Factors in Computing Systems*. New York: ACM.
- HAYS, W. L. (1994). *Statistics* (5th ed.). New York: Harcourt Brace.
- HUMMEL, R. (1994, Fall). *Data visualization* (Course description). New York University, Department of Computer Science, New York.
- KOSSLYN, S. M., GERSHON, N. D., LEFKOWITZ, H., & PEARLMAN, J. D. (1992). Improving visualization: Theoretical and empirical foundations. In *Conference Proceedings: Visualization 1992* (pp. 372-374).
- LOFTUS, G. R. (1993). A picture is worth a thousand p values: On the irrelevance of hypothesis testing in the microcomputer age. *Behavior Research Methods, Instruments, & Computers*, **25**, 250-256.
- MARCHAK, F. M. (1994). An overview of scientific visualization techniques applied to experimental psychology. *Behavior Research Methods, Instruments, & Computers*, **26**, 177-180.
- MARCHAK, F. M., & ZULAGER, D. D. (1992). A review of scientific visualization software for the Macintosh computer. *Behavior Research Methods, Instruments, & Computers*, **24**, 328-335.
- MARSH, P. L. (1994). A review and comparison of five graphics programs for PC-DOS and MS-DOS computers. *American Statistician*, **48**, 44-51.
- MCCORMICK, E. J. (1976). *Human factors in engineering and design*. New York: McGraw-Hill.
- NIELSEN, J. (1990). *Hypertext and hypermedia*. San Diego, CA: Academic Press.
- POMERANTZ, J. R., & KUBOVY, M. (1986). Theoretical approaches to perceptual organization. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 2, pp. 36-1-36-46). New York: Wiley.
- RABENHORST, D. A. (1993). *Interactive exploration of multidimensional data* (Research Rep. No. 19416). Yorktown Heights, NY: IBM T. J. Watson Research Center.
- ROGOWITZ, B. E., & TREINISH, L. A. (1993). *Data structures and perceptual structures* (Research Rep. No. 18878). Yorktown Heights, NY: IBM T. J. Watson Research Center.
- SCHNEIDERMAN, B. (1980). *Software psychology: Human factors in computer and information systems*. Cambridge, MA: Winthrop.
- SHIFFRIN, R. M., & SCHNEIDER, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, **84**, 127-190.
- SPRINGMEYER, R. R., BLATTNER, M. M., & MAX, N. L. (1992). *A characterization of the scientific data analysis process*. In *Conference Proceedings: Visualization 1992* (pp. 235-242). Washington, DC: IEEE Computer Society Press.
- TUFT, E. R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics Press.
- WOODS, D. D., JOHANNESSEN, L., & POTTER, S. S. (1992). The sophistry of guidelines: Revisiting recipes for color use in human-computer interface design. In *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 418-422).