# Visual Graph Display Guidelines

Technical Report HFA-TR-0803
Atlanta, GA: Georgia Institute of Technology
School of Psychology – Human Factors and Aging Laboratory

November 2008

CARA BAILEY FAUSSET

WENDY A. ROGERS

ARTHUR D. FISK

**Requests for more information may be sent to Wendy A. Rogers, School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332-0170 (electronic mail to wendy@gatech.edu)**

## TABLE OF CONTENTS

## EXECUTIVE SUMMARY

The goal of this report is to summarize the current state of the literature on graph comprehension and the application of these data to visual graph design. Results from empirical studies have been integrated into a comprehensive guide for visual graph display designers. A brief introduction to visual graphs is given followed by a description of the factors that influence graph comprehension: task requirements, graph characteristics, data characteristics, and person characteristics. Each of these factors and their relevance to visual graph display design is discussed in detail. Visual graph display guidelines have been created based upon these empirical studies.

Our analysis of the literature led to the development of general graph design guidelines and graph-specific design guidelines. Once designers have decided to convey information in a graphical format, they should consider the following guidelines.

### General Graph Design Guidelines:

- **Graph characteristics can influence viewers' interpretations.**
    - Two-dimensional graphs lead to better performance overall than three-dimensional graphs.
    - Backgrounds, such as pictures, should be avoided.
    - Directly label the data instead of using a legend or key. However, beware of visual clutter.
    - Use color if it is meaningful to the data (e.g., red for negative cash flow vs. black for positive cash flow). However, be aware of multiple or domain-specific meanings; for example, green can mean "go" to most people but "infected" to healthcare workers.
- **The characteristics of the person also influences graph reading performance.**
    - As the reader's experience increases with the graph and task, so does accuracy.

- As the reader's experience increases with the graph and task, response time to complete the task decreases.
- Familiarity with the graph leads to better performance as measured by response time.
- Knowledge of the data content may lead to incorrect interpretations of the data when the data are not consistent with the reader's expectations.
- Having expert knowledge of graph types does not always predict accurate performance.
- Availability of cognitive resources (such as working memory) influences performance. Individuals with higher working memory abilities can perform more difficult tasks successfully. Design features (e.g., use of color and gridlines) can mitigate working memory difficulties.

- **General human factors principles must be followed to ensure optimal visual graph display design.**
  - Understand the physical, perceptual, and cognitive abilities of the target user group.
  - Test the users' ability to comprehend the graphs throughout the design process, not just with the final product. Be sure to conduct user testing with representative users, tasks, and contexts.
  - Consider training and instructional needs for the target population throughout the design process.

## Graph-Specific Design Guidelines:

### *LINE GRAPHS*

- **The task requirements will inform which graph type is best to use.**
  - Line graphs are best for trend reading, trend comparison, and judgments of change over time.
- **Graph characteristics can influence viewers' interpretations.**
  - Graph independent variables (causes) on the x-axis and dependent variables (effects) on the y-axis.
  - X- and y-axes should be used.

- Tick marks (or scales) in y-axis for line graphs should be avoided unless exact values are needed.

- **Data characteristics must be considered in graph design.**
  - Interpreting more than a one-way interaction is difficult. Presenting more than two variables should be avoided.
  - Y-z variable relationships presented in a line graph with x-y-z variables are difficult to interpret and error-prone.
  - As the number of data series presented increases, expect longer response times to complete tasks.
  - As the number of data points per series increases, expect longer response times to complete tasks.
  - Readers tend to focus on individual points when only a few points are plotted on a line graph.
  - Readers tend to focus on overall trends when many points are plotted on a line graph.

***BAR GRAPHS***

- **The task requirements will inform which graph type is best to use.**
  - Bar graphs are best for discrete data comparisons, exact point value extraction, comparing two points between data series, and identifying maximum or minimum point values.
- **Graph characteristics can influence viewers' interpretations.**
  - X- and y-axes should be used.
  - Tick marks (or scales) in y-axis should be avoided unless exact values are needed.
  - Use vertical instead of horizontal bars for bar graphs.
- **Data characteristics must be considered in graph design.**
  - Interpreting more than a one-way interaction is difficult. Presenting more than two variables should be avoided.
  - As the number of data series presented increases, expect longer response times to complete tasks.

***PIE GRAPHS (PIE CHARTS)***

- **The task requirements will inform which graph type is best to use.**
  - Pie graphs (or pie charts) are best for estimating proportions of the whole or for comparing more than one component.
- **Graph characteristics can influence viewers' interpretations.**
  - Tick marks for pie graphs should not be used.

## INTRODUCTION TO VISUAL GRAPHS

Visual graphs are a common format for conveying information and are intended to present data in formats that convey relationships between variables to a reader clearly and quickly (Gillan, Wickens, Hollands, & Carswell, 1998). Graphs are used in various domains, and there are myriad graph types that can display data such as line graphs, bar graphs, and pie graphs (pie charts). These numerous types of graphs offer display designers an array of options, but it is important to determine how designers can create a visual graphic display that best supports a successful interaction between the user and the display.

## FACTORS THAT INFLUENCE GRAPH COMPREHENSION

Current models of graph comprehension suggest multiple factors that must be considered in successful graph comprehension: the specific requirements of the task, the characteristics of the graph, the characteristics of the data, and the characteristics of the individual (e.g., Peebles & Cheng, 2003; Shah, Freedman, & Vekiri, 2005). Moreover, these factors have been shown to interact with each other and influence graph comprehension (e.g., Vessey, 1991). For example, if the user's task is to describe one variable as a proportion of the total, a line graph will be very difficult to use to successfully complete the task. Because of the potential demands that these factors and their interactions can impose on graph readers, these factors must be considered in designing effective visual graph displays.

### Task Requirements

Tasks such as reading points directly from the graph or comparing specific values shown directly in the graph are considered local tasks (Wickens & Hollands, 2000). In contrast are global tasks such as trend reading and trend comparison.  Other tasks include comparing quantities that must be derived from other quantities shown in the graph (implicitly available information) or using the relationships shown on the graph to predict how one variable may change as another variable is changed.  Tasks that require readers to derive information not explicitly displayed on a graph are likely to place a higher demand on a reader's cognitive resources (e.g., working memory) and to negatively affect performance.  Working memory is part of the memory system that is used for temporarily storing and manipulating information (Baddeley, 1986).  An ideal graph should facilitate task performance by reducing demands on readers' working memory (e.g., reducing the need for comparisons between legend and graph content by using direct labels).

Display designers should specify at the beginning of the design process the task relevant criteria (Meyer, Shinar, & Leiser, 1997).  That is, what is (are) the goal(s) of the display?  The designer must decide if the data should be described verbally, graphed, or detailed in a table.  If extraction of absolute values is required, a table is the best format to use (Kosslyn, 1994; Meyer, Shinar, & Leiser).  However, if perceiving relative values and the relationships between variables is the primary display goal, then a graph is the preferred format.  Other questions that should be asked by designers include:  Does the user need to extract information quickly?  Is performance accuracy a primary concern? Must the user interpret complex relations among variables?  By answering these questions early in the design process, designers can focus on the graph types that will support graph comprehension.  Figures 1 through 3 illustrate examples of matching and mismatching graphs to task requirements.

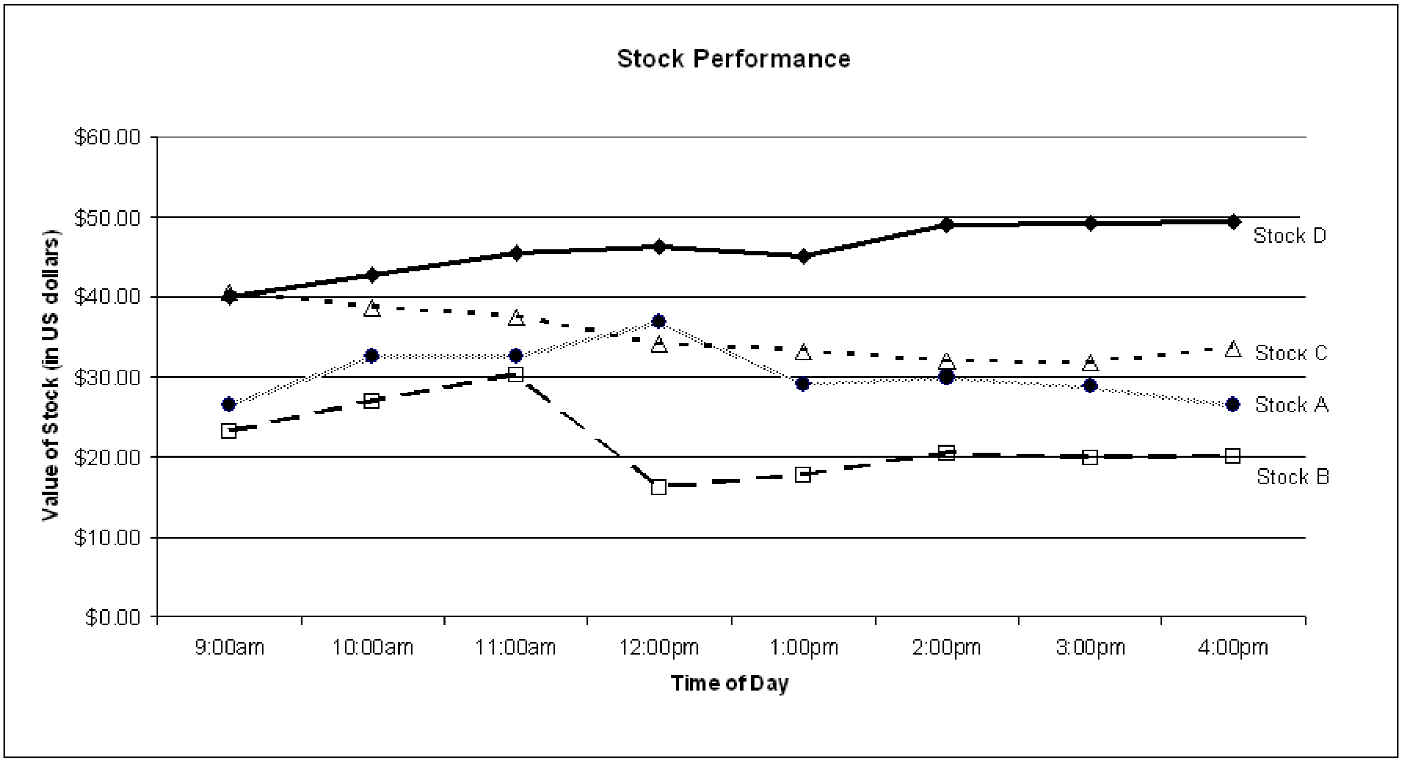


*Figure 1a.* Graph to task **match**. **Graph type**: Line graph. **Task**: Which stock has increased the most over the day? **Answer**: Stock D.

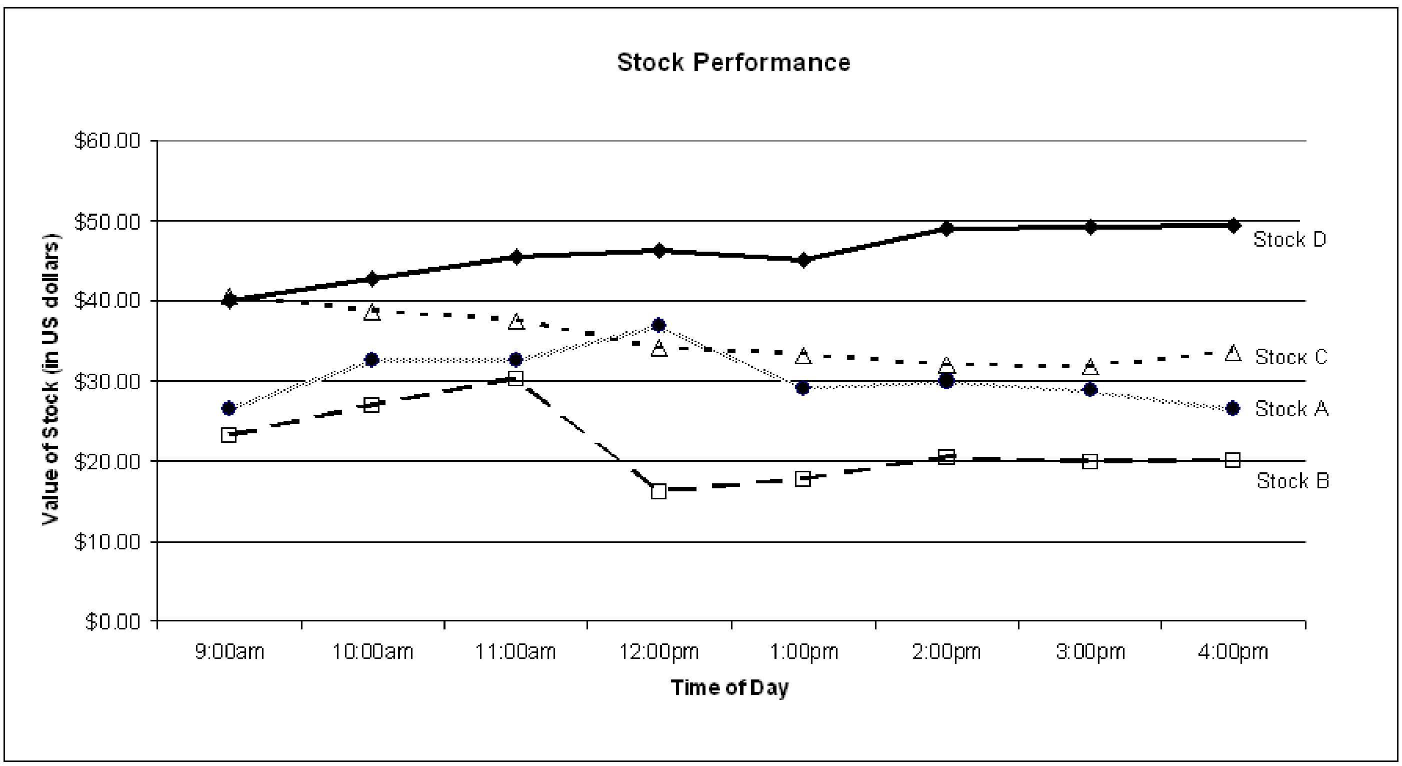


*Figure 1b.* Graph to task **mismatch**. **Graph type**: Line graph. **Task**: What is the value of Stock A at 1:00 p.m.? **Answer**: $29.10.

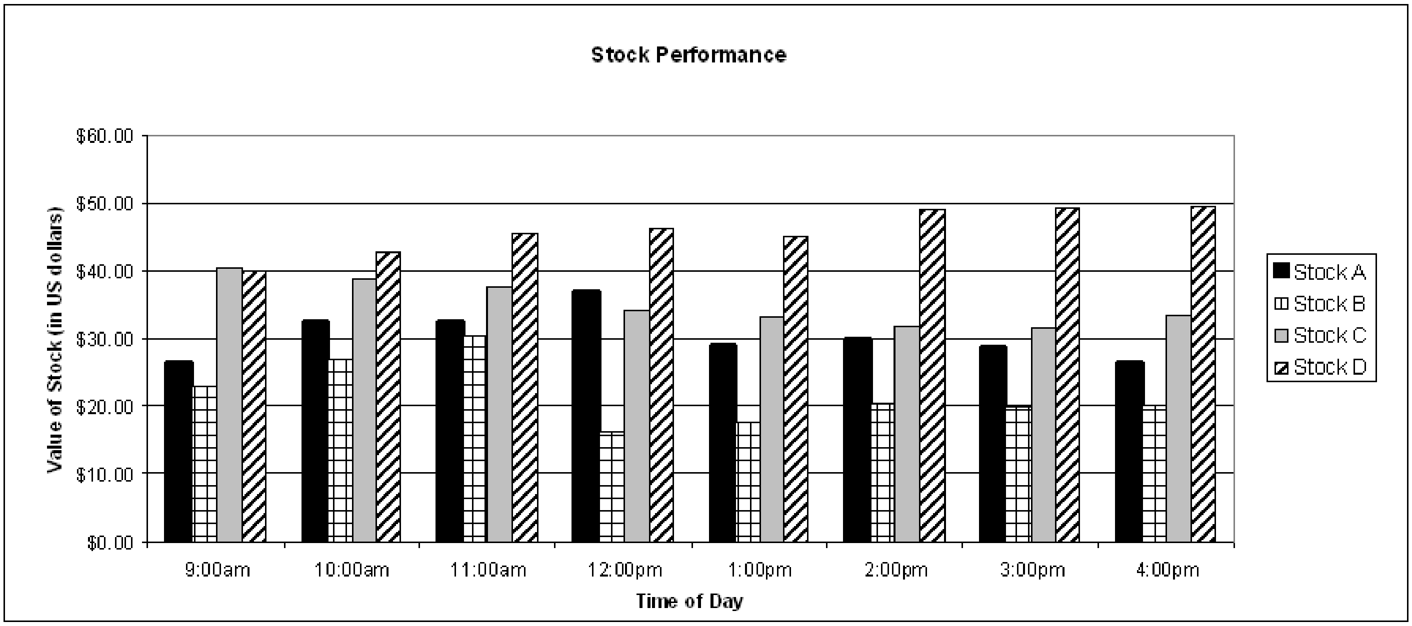


*Figure 2a.* Graph to task **match**. **Graph type**: Bar graph. **Task**: What is the value of Stock A at 1:00 p.m.? **Answer**: $29.10.

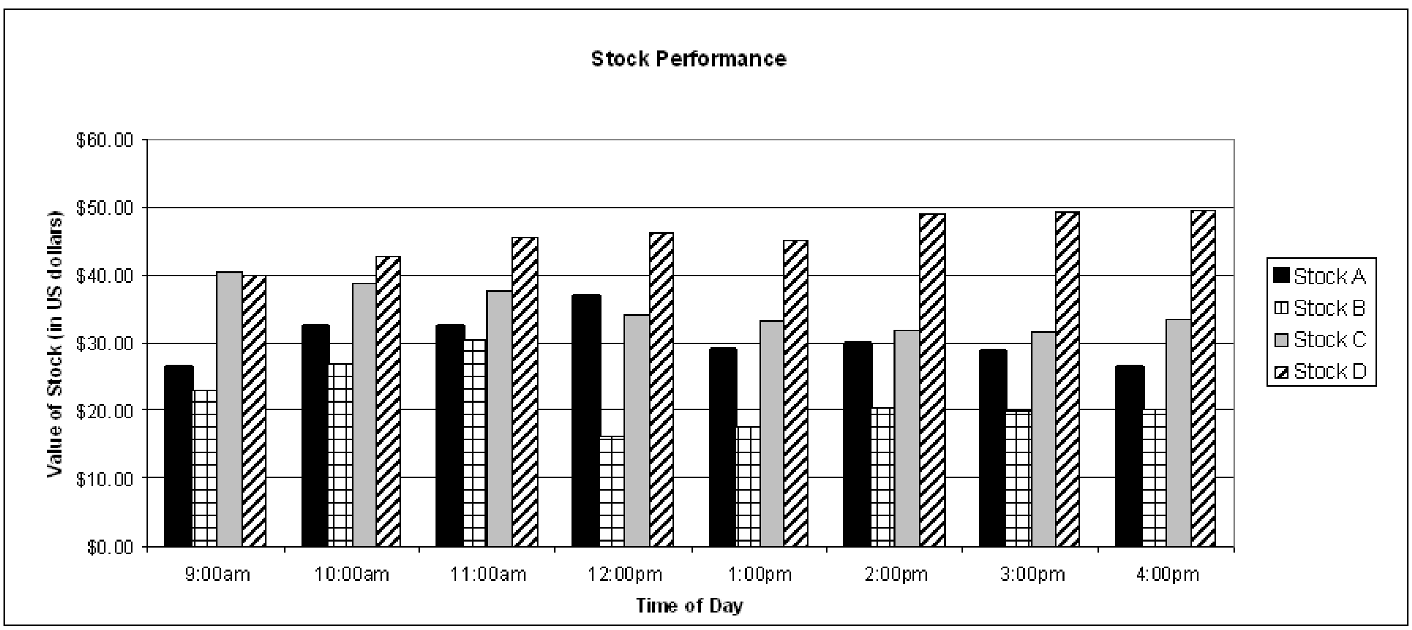


*Figure 2b.* Graph to task **mismatch**. **Graph type**: Bar graph. **Task**: Which stock has increased the most over the day? **Answer**: Stock D.

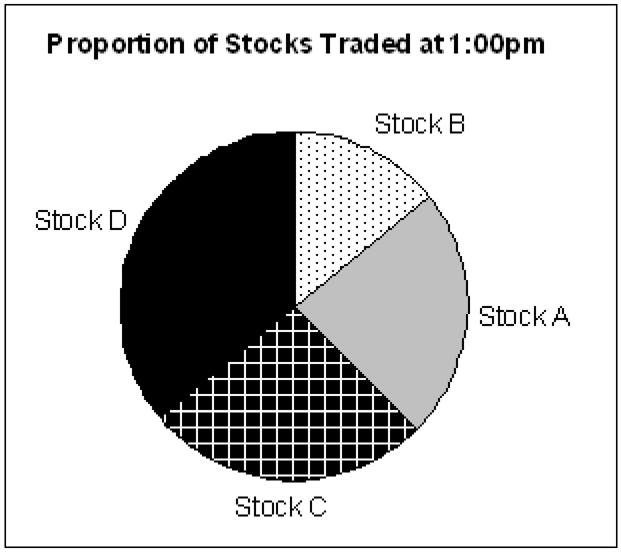


*Figure 3a.* Graph to task **match**. **Graph type**: Pie graph. **Task**: Which stock comprised the highest percentage of trading at 1:00pm? **Answer**: Stock D.

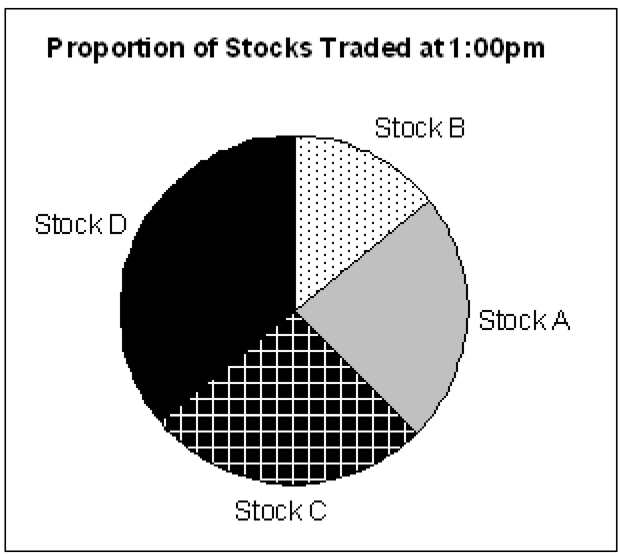


*Figure 3b.* Graph to task **mismatch**. **Graph type**: Pie graph. **Task**: What is the value of Stock A at 1:00 p.m.? **Answer**: Unable to answer with this graph.

Refer to Table 1 for the recommended graph types for specific tasks. Appendix A details the empirical findings of the visual graph types that support specific tasks.

Table 1

*Task Requirements by Graph Type*

| Task Requirements | Preferred Graph Type |
|---|---|
| Judgment of change over time | Line graph |
| Comparing trends | Line graph |
| Identification of trends | Line graph |
| Discrete data comparison | Bar graph |
| Comparing two points between data series | Bar graph |
| Exact point value extraction | Bar graph |
| Identifying maximum or minimum point values | Bar graph |
| Estimating proportion of the whole | Pie graph |
| Comparing more than one component | Pie graph |

**Graph Characteristics**

As discussed above, graph comprehension success depends upon the type of graph used to display information (see Table 1 and Appendix A). Additionally, specific characteristics of the graph itself, such as the use of dimensionality (two- versus three-dimensional), scales, background, axis orientation, color, gridlines, and the presence of a legend, have been shown to influence performance (e.g., Carpenter & Shah, 1998; Fischer, 2000; Gillan & Richman, 1994; Lohse, 1997; Rangecroft, 2003; Siegrist, 1996; Spence, 1990; Zacks, Levy, Tversky, & Schiano, 1998).

For example, if a legend is used, readers are likely to have to “refresh” their memory several times while trying to comprehend the graphs due to the limitations of working memory capacity, whereas directly labeling the graph can reduce the memory load on a reader (Carpenter & Shah, 1998). See Figures 4a and 4b for examples of graphs with and without a legend.

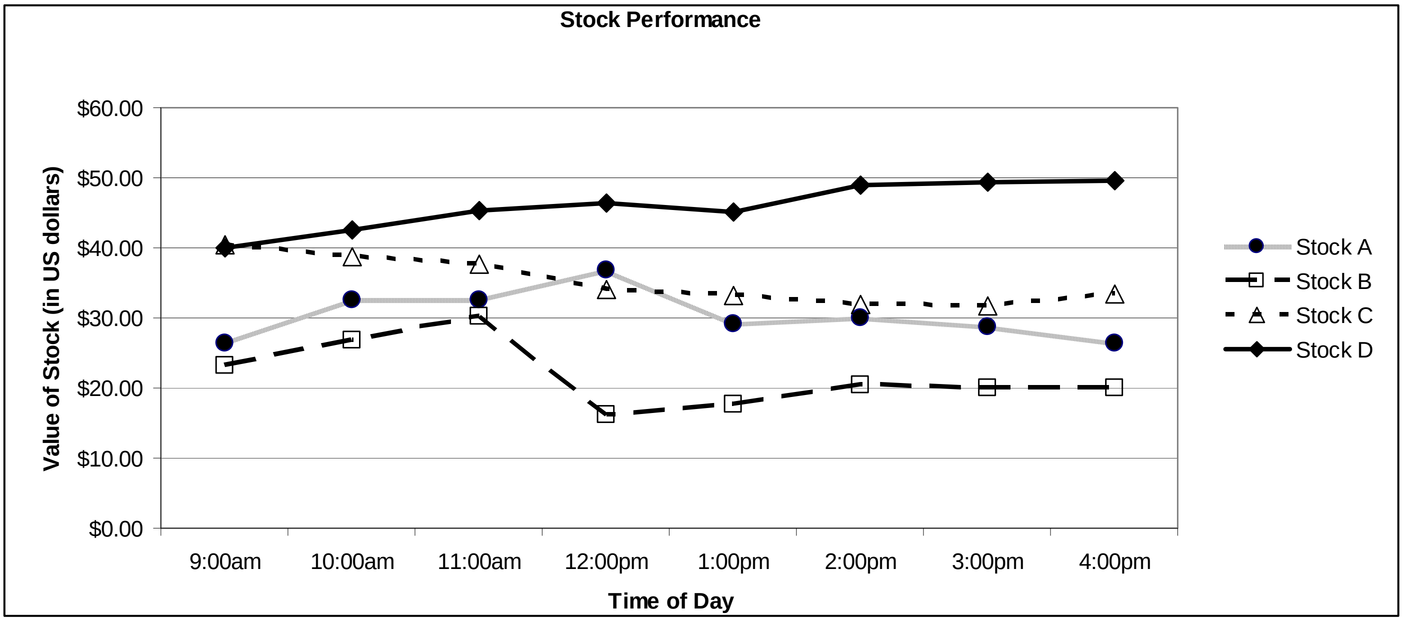


*Figure 4a.* Example of a line graph with a legend.

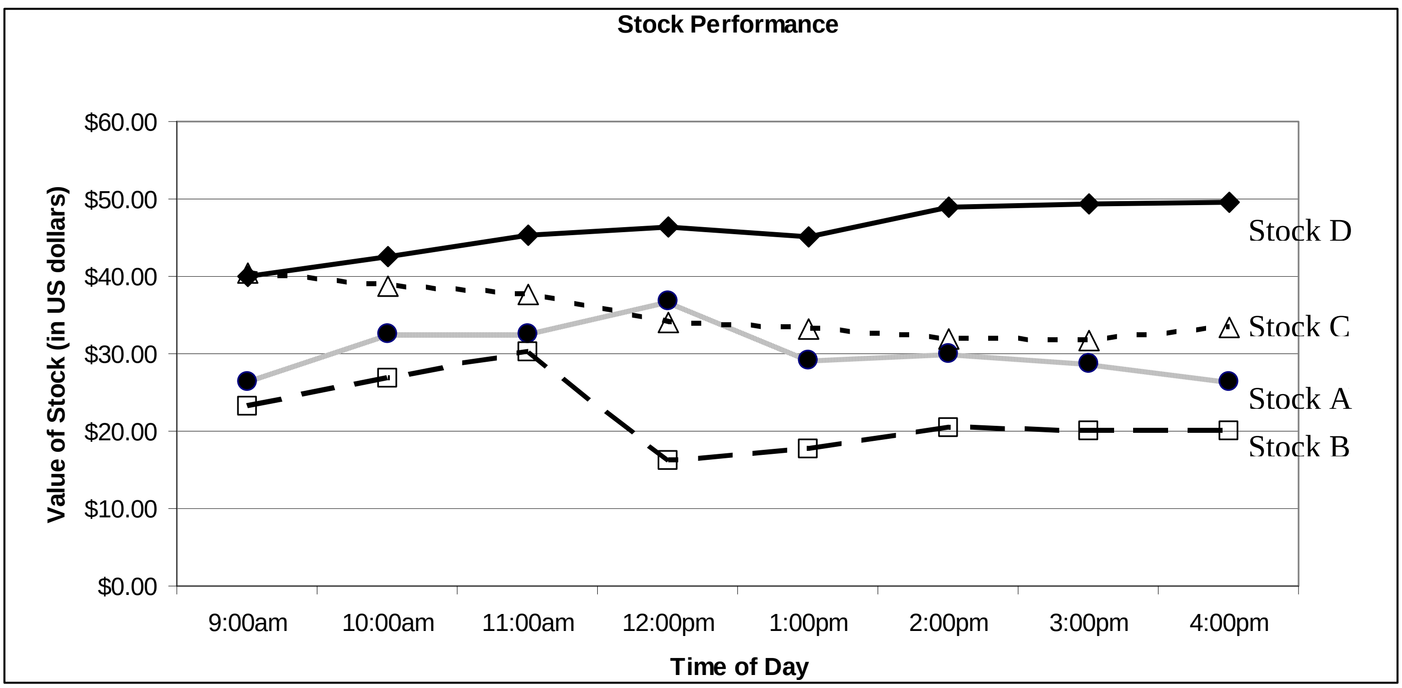


*Figure 4b.* Example of a line graph with each variable labeled directly.

Including a background image with the graph also impairs performance (Gillan & Richman, 1994). The presence of a background creates higher visual search demands on the reader; that is, there are more markings for the reader to sort through when looking at the graph to complete a specific task. A background image may also reduce the contrast between the target information and the background information making the graph more difficult to read. See Figure 5 for an example of a line graph that includes a background image and a legend. Appendix B summarizes studies that manipulated various graph characteristics and how those manipulations influenced graph reading performance.

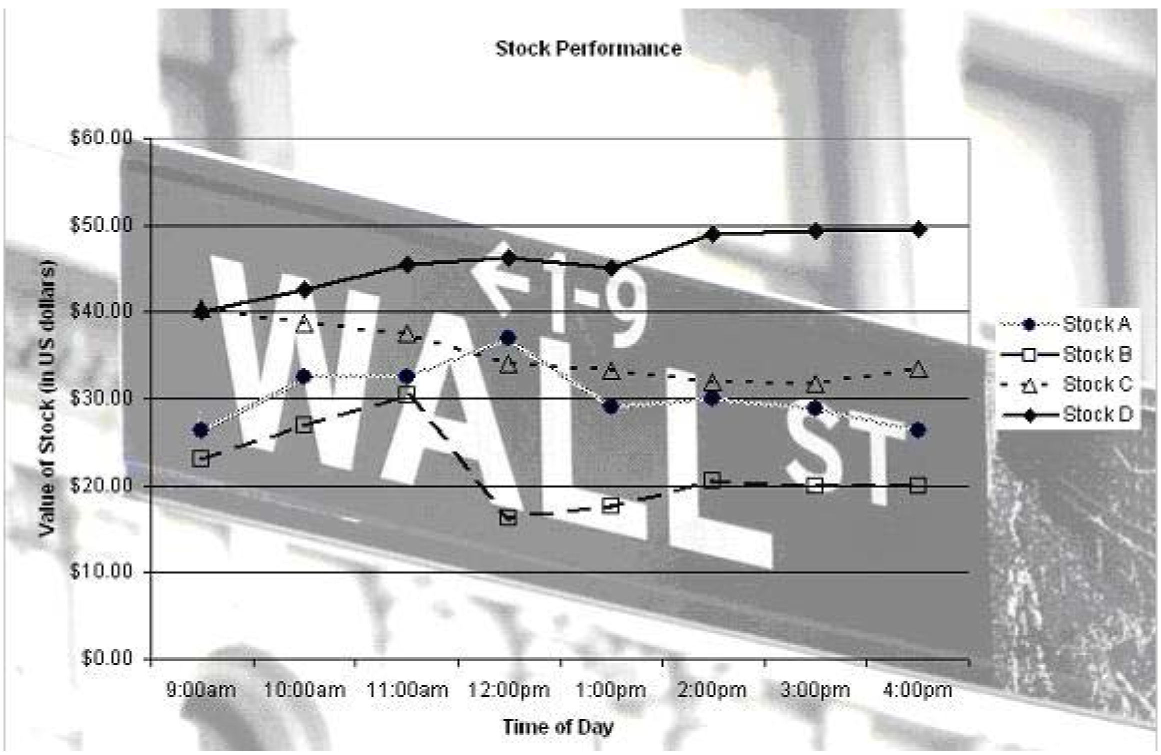


*Figure 5*. Example of a difficult-to-read line graph that includes a background image and a legend.

The use of dimensionality can enhance the attractiveness of a graph, which would meet the goal of drawing users' attention to the graph. However, rendering a graph in 3-D also adds extraneous marks to the graph, potentially distracting readers from performing their task accurately and quickly. In addition, 3-D graphs often create distortions in data that can lead to incorrect comprehension (Fischer, 2000; Rangecroft, 2003; Siegrist, 1996; Zacks, Levy, Tversky, & Schiano, 1998; but see Spence, 1990; Zacks, Levy, Tversky, & Schiano, 1998). Refer to Figures 6a and 6b for examples of 2-D and 3-D graphs, respectively. Appendix C summarizes the results of studies that investigated the influence of dimensionality on graph reading performance.

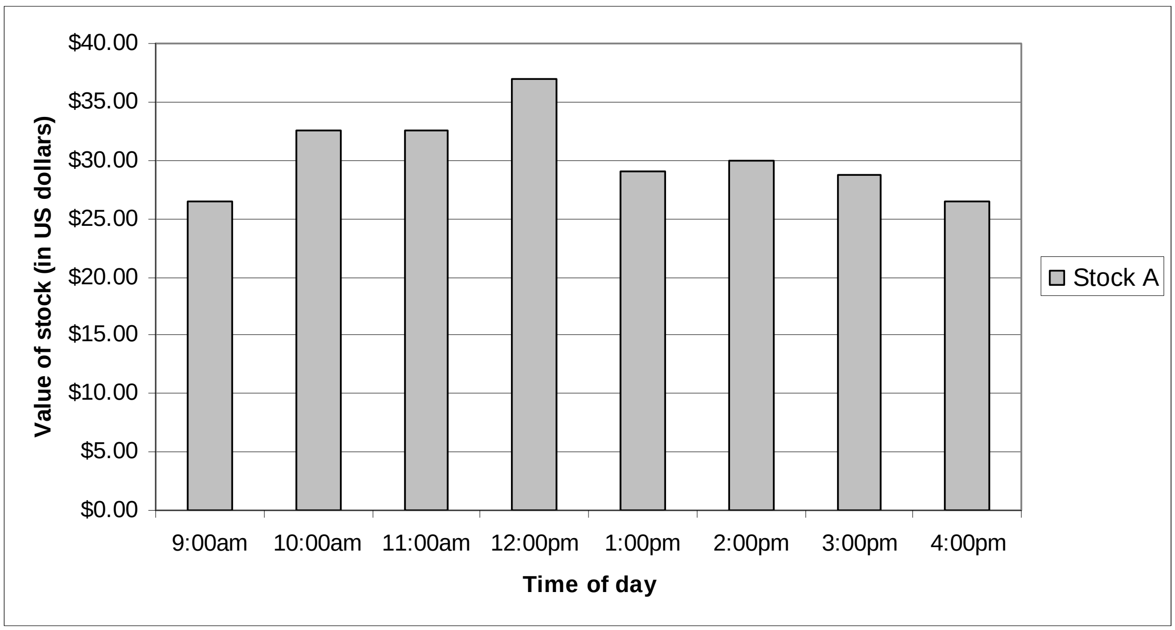


*Figure 6a.* Example of a bar graph with two-dimensions.

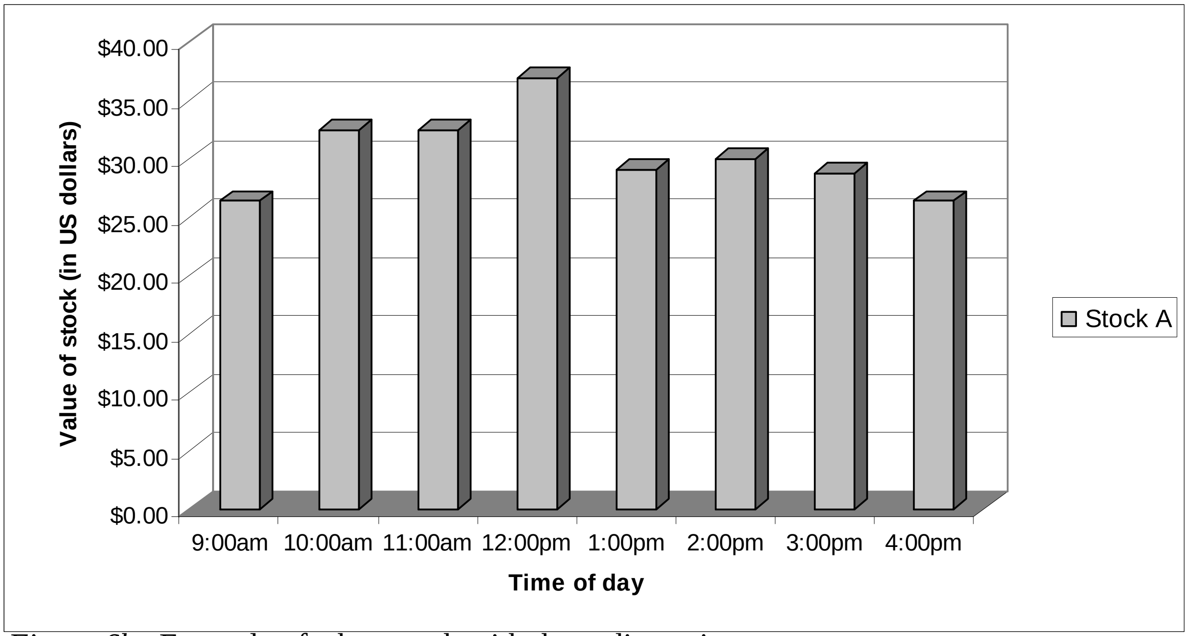

*Figure 6b.* Example of a bar graph with three-dimensions.

General graph display guidelines concerning graph characteristics may be summarized as follows:

1. Two-dimensional (2-D) graphs lead to better performance overall than three-dimensional (3-D) graphs. However, if the designer's goal is to attract attention, 3-D graphs may be used.

2. Backgrounds, such as pictures, should be avoided.

3. Directly label the data instead of using a legend or key. However, beware of visual clutter.

4. Graph independent variables (causes) on the x-axis and dependent variables (effects) on the y-axis.

5. X- and y-axes should be used for line and bar graphs.

6. Tick marks (or scales) in y-axis for bar and line graphs should be avoided unless exact values are needed.

7. Tick marks for pie graphs should not be used.

8. Use color if it is meaningful to the data (e.g., red for negative cash flow vs. black for positive cash flow). However, be aware of multiple or domain-specific meanings; for example, green can mean "go" to most people but "infected" to healthcare workers.

9. Use gridlines when specific values must be extracted.
10. Use vertical instead of horizontal bars for bar graphs.

**Data Characteristics**

The complexity of the actual data presented in a graph can also influence graph reading performance and comprehension. The number of variables, such as the number of lines displayed in a line graph, the number of trend reversals in a line (i.e., the up and down vacillations of one line), and the number of individual data points influence interpretations of graphs (Carpenter & Shah, 1998; Carswell, Emery, & Lonon, 1993; Halford, Baker, McCredden, & Bain, 2005).

Working memory limitations were implicated in a study that investigated the maximum number of variables that people can process simultaneously (Halford, Baker, McCredden, & Bain). Experienced graph readers were asked to interpret two-, three-, and four-way interactions represented in a bar chart. Task demands were manipulated by increasing the number of interactions that had to be interpreted. Accuracy for four variables decreased significantly from three variable accuracy, and performance with five variables was at chance. Thus, comprehending interactions between three variables was very difficult and comprehending interactions between four or more variables was next to impossible.

In a study where the number of data points on a single line was varied and viewers were asked to describe the graph, it was suggested that single data points on a single line functioned as separate entities until a certain data density was reached (Carswell, Emery, & Lonon, 1993). That is, as the number of data points increased, the cognitive resource demands on the reader increased until the working memory capacity of the viewer was overwhelmed.  Once this “critical” data density was reached, the viewer described the overall or global features of the line itself and not the individual points or local features comprising the line.

Having readers interpret line graphs that contained three separate variables researchers found that x-y relationships were salient and easily extracted (Carpenter & Shah, 1998).  However, the z-y relationship between variables was much less obvious to readers.  In fact, when presented with an alternative perspective of the same graph, participants failed to identify which graph was identical to the one they had been studying.  For example, Figures 7a and 7b provide two equivalent perspectives of a line graph depicting three variables but the fact that the data are the same across these figures is not immediately obvious.

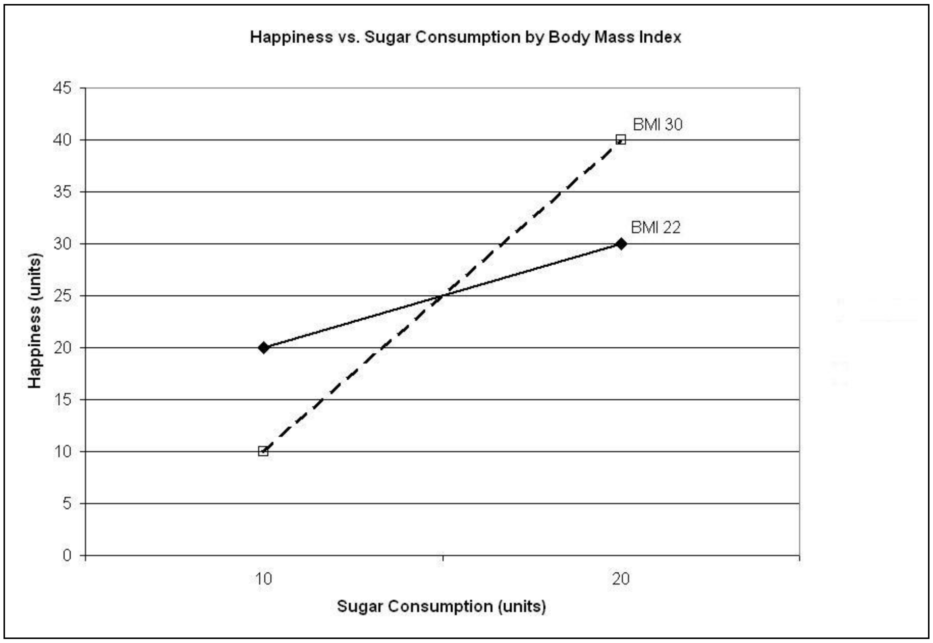


*Figure 7a.* Example of a line graph depicting three variables: happiness versus sugar consumption by body mass index (BMI).

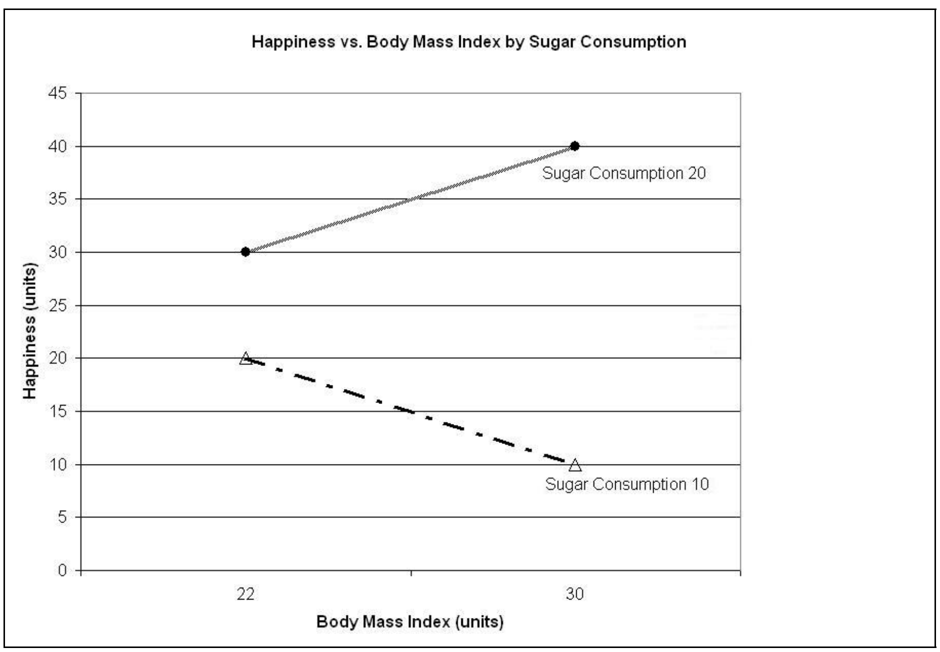


*Figure 7b.* Example of a line graph depicting the same data as in Figure 7a for three variables: happiness versus body mass index by sugar consumption.

Appendix D summarizes the empirical findings for various manipulations of data characteristics. The following recommendations are made regarding data characteristics:

1. Interpreting more than a one-way interaction is difficult. Presenting more than two variables should be avoided.

2. Y-z variable relationships presented in a line graph with x-y-z variables are difficult to interpret and error-prone.

3. As the number of data points per series increases, expect longer response times to complete tasks.

4. As the number of data series presented increases, expect longer response times to complete tasks.

5. Readers describe individual points when only a few points are plotted on a line graph.

6. Readers describe overall trends when many points are plotted on a line graph.

**Person Characteristics**

Characteristics of the individual user must also be considered in graph comprehension. An individual's knowledge of graphs and knowledge of the specific content presented in the graph greatly influences graph comprehension (Freedman & Shah, 2002; Meyer, Shinar, & Leiser, 1997; Roth & Bowen, 2003; Shah & Carpenter, 1995; Shah & Hoeffner, 2002). Experience with particular graphs and tasks was found to significantly improve both accuracy and response time (Meyer, Shinar, & Leiser). Practice effects were measured across three blocks of graph comprehension trials in which participants completed various tasks using line or bar graphs; accuracy increased and response time decreased over the course of the experiment.

Familiarity with the graphs themselves was also found to be significant in improving performance (Meyer, Shinar, & Leiser, 1997). To investigate the role of familiarity in graph comprehension, Meyer and colleagues presented half of the participants with the axes and the legend of line or bar graphs prior to the presentation of the graph complete with data, whereas the other group of participants saw the graph only during the experimental task. Response time was significantly faster for the group that was able to “preview” the graph suggesting that familiarity leads to faster comprehension of the data contained within the graph.

Although familiarity with graphs can improve performance, familiarity with the data contained within the graph does not always lead to better performance as the reader’s expectations can play a significant role (Shah, 2001). When participants were presented with data about the number of drinks consumed and the relationship to the number of car accidents, participants interpreted the graphs as they *expected*: More drinks leads to higher accident rates. However, the graphs did not indicate this relationship suggesting a powerful influence of readers’ expectations. A similar pattern emerged in a study showing that experienced professors’ descriptions of graphs common in their field were not accurate; the graphs did not support their descriptions suggesting a top-down influence on graph comprehension (Roth & Bowen, 2003).

Moreover, experience with graphs is not always predictive of accurate performance (Roth & Bowen, 2003; Shah & Carpenter, 1995). When graduate students with an average of five semesters of statistics were asked to interpret line graphs displaying variables on three axes (across the x, y, and z dimensions), they did not perform more accurately than undergraduate students (Shah & Carpenter, 1995).

Lastly, the availability of cognitive resources, such as working memory, also influences graph comprehension. Working memory was found to be a critical factor of success in a graph comprehension study that divided its participants into high and low working memory groups (Lohse, 1997). A complex budget task was to be completed using either monochrome line graphs without gridlines or color-coded line graphs with gridlines. The complexity of the graphs was further manipulated into high and low determined by the number of sales regions presented: nine or three, respectively. It was predicted that the color and the gridlines would reduce cognitive load and improve performance for the low working memory group in the high complexity condition. Performance on the high complexity task by the high working memory group using the monochrome graph without gridlines was equivalent to the performance of the low working memory group using color graphs with grid lines. This suggests that the addition of color and the gridlines did reduce the working memory demands for participants with lower working memory capacity and that the high working memory group had enough resources available to complete the complex task successfully with or without color and gridlines.

Appendix E summarizes the empirical findings of studies that investigated graph performance as a function of person characteristics. General graph display guidelines concerning person characteristics may be summarized as follows:

1. As the reader's experience increases with the graph and task, so does accuracy.

2. As the reader's experience increases with the graph and task, response time to complete the task decreases.

3. Familiarity with the graph leads to better performance as measured by response time.

4. Knowledge of the data content may lead to incorrect interpretations of the data when the data are not consistent with the reader's expectations.

5. Having expert knowledge of graph types does not always predict accurate performance.

6. Availability of cognitive resources (such as working memory) influences performance. Individuals with higher working memory abilities can perform more difficult tasks successfully. Design features (e.g., use of color and gridlines) can mitigate working memory difficulties.

**Human Factors Principles**

Task requirements, graph characteristics, data characteristics, and person characteristics must all be considered when designing a visual graph display. These factors that influence graph comprehension cannot be easily untangled. Graph reading is a complex process that involves multiple interactions. Each aforementioned factor can be studied and manipulated, but the whole graph-user interaction is more than the sum of the parts.

Consequently, general human factors principles should guide the design process. A primary tenet of human factors is to “know thy user.” It is imperative that designers understand who their target audience is from a physical, perceptual, and cognitive standpoint. Moreover, it is crucial that the target users are involved in testing the visual graph display throughout the design process. Such user testing must be conducted with representative users, tasks, and contexts. Design is an iterative process that can be informed by following the guidelines set forth in this report and by involving target users early in the design process.

**VISUAL GRAPH DISPLAY GUIDELINES**

Once designers have determined that the information they want to convey should be graphed, the guidelines for designing a visual graph display are as follows:

## General Graph Design Guidelines

- **Graph characteristics can influence viewers' interpretations.**
  - Two-dimensional graphs lead to better performance overall than three-dimensional graphs.
  - Backgrounds, such as pictures, should be avoided.
  - Directly label the data instead of using a legend or key. However, beware of visual clutter.
  - Use color if it is meaningful to the data (e.g., red for negative cash flow vs. black for positive cash flow). However, be aware of multiple or domain-specific meanings; for example, green can mean "go" to most people but "infected" to healthcare workers.
- **The characteristics of the person also influences graph reading performance.**
  - As the reader's experience increases with the graph and task, so does accuracy.
  - As the reader's experience increases with the graph and task, response time to complete the task decreases.
  - Familiarity with the graph leads to better performance as measured by response time.
  - Knowledge of the data content may lead to incorrect interpretations of the data when the data are not consistent with the reader's expectations.
  - Having expert knowledge of graph types does not always predict accurate performance.
  - Availability of cognitive resources (such as working memory) influences performance. Individuals with higher working memory abilities can perform more difficult tasks successfully. Design features (e.g., use of color and gridlines) can mitigate working memory difficulties.

- **General human factors principles must be followed to ensure optimal visual graph display design.**
    - Understand the physical, perceptual, and cognitive abilities of the target user group.
    - Test the users' ability to comprehend the graphs throughout the design process, not just with the final product.  Be sure to conduct user testing with representative users, tasks, and contexts.
    - Consider training and instructional needs for the target population throughout the design process.

## Graph-Specific Design Guidelines

***LINE GRAPHS***

- **The task requirements will inform which graph type is best to use.**
    - Line graphs are best for trend reading, trend comparison, and judgments of change over time.
- **Graph characteristics can influence viewers' interpretations.**
    - Graph independent variables (causes) on the x-axis and dependent variables (effects) on the y-axis.
    - X- and y-axes should be used.
    - Tick marks (or scales) in y-axis for line graphs should be avoided unless exact values are needed.
- **Data characteristics must be considered in graph design.**
    - Interpreting more than a one-way interaction is difficult.  Presenting more than two variables should be avoided.
    - Y-z variable relationships presented in a line graph with x-y-z variables are difficult to interpret and error-prone.
    - As the number of data series presented increases, expect longer response times to complete tasks.
    - As the number of data points per series increases, expect longer response times to complete tasks.

- Readers tend to focus on individual points when only a few points are plotted on a line graph.
- Readers tend to focus on overall trends when many points are plotted on a line graph.

***BAR GRAPHS***

- **The task requirements will inform which graph type is best to use.**
  - Bar graphs are best for discrete data comparisons, exact point value extraction, comparing two points between data series, and identifying maximum or minimum point values.
- **Graph characteristics can influence viewers' interpretations.**
  - X- and y-axes should be used.
  - Tick marks (or scales) in y-axis should be avoided unless exact values are needed.
  - Use vertical instead of horizontal bars for bar graphs.
- **Data characteristics must be considered in graph design.**
  - Interpreting more than a one-way interaction is difficult. Presenting more than two variables should be avoided.
  - As the number of data series presented increases, expect longer response times to complete tasks.

***PIE GRAPHS (PIE CHARTS)***

- **The task requirements will inform which graph type is best to use.**
  - Pie graphs (or pie charts) are best for estimating proportions of the whole or for comparing more than one component.
- **Graph characteristics can influence viewers' interpretations.**
  - Tick marks for pie graphs should not be used.

## ADDITIONAL RESOURCES


Anderson, M., Meyer, B., & Olivier, P. (Eds.). (2002). *Diagrammatic representation and reasoning.* London: Springer-Verlag.

Gillan, D. J., Wickens, C. D., Hollands, J. G., & Carswell, C. M. (1998). Guidelines for presenting quantitative data in HFES publications. *Human Factors, 40*, 28-41.

Kosslyn, S. M. (1989). Understanding charts and graphs. *Applied Cognitive Psychology, 3*(3), 185-225.

Kosslyn, S. M. (1994). *Elements of graph design*. New York: W.H. Freeman and Company.

Shah, P. & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review, 14,* 47-69.

Shah, P. & Miyake, A. (2005). *The Cambridge handbook of visuospatial thinking.* New York: Cambridge University Press.

Weitz, J., Rogers, W. A., & Fisk, A. D. (2006). *Display design guidelines* (HFA-TR-0601). Atlanta, GA: Georgia Institute of Technology, School of Psychology, Human Factors and Aging Laboratory.

## REFERENCES


Baddeley, A. (1986). *Working memory*. New York: Oxford University Press.

Carpenter, P. A. & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied, 4*, 75-100.

Carswell, C. M., Emery, C., & Lonon, A. M. (1993). Stimulus complexity and information integration in the spontaneous interpretations of line graphs. *Applied Cognitive Psychology, 7*, 341-357.

Fischer, M. H. (2000). Do irrelevant depth cues affect the comprehension of bar graphs? *Applied Cognitive Psychology, 14*, 151-162.

Fischer, M. H., Dewulf, N., & Hill, R. L. (2005). Designing bar graphs: Orientation matters. *Applied Cognitive Psychology, 19*, 953-962.

Freedman, E. G. & Shah, P. (2002). Toward a model of knowledge-based graph comprehension. In M. Hegarty, B. Meyer, & N. Hari Narayanan (Eds.), *Diagrammatic representation and inference* (pp. 8-31). Berlin: Springer-Verlag Berlin.

Gattis, M. & Holyoak, K. J. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology, 22*, 231-239.

Gillan, D. J. & Richman, E. H. (1994). Minimalism and the syntax of graphs. *Human Factors, 36*, 619-644.

Gillan, D. J., Wickens, C. D., Hollands, J. G., & Carswell, C. M. (1998). Guidelines for presenting quantitative data in HFES publications. *Human Factors, 40*, 28-41.

Halford, G.S., Baker, R., McCredden, J.E., & Bain, J.D. (2005). How may variables can humans process? *Psychological Science, 16*, 70-76.

Hollands, J. G. & Spence, I. (1992). Judgments of change and proportion in graphical perception. *Human Factors*, *34*, 313-334.

Kosslyn, S. M. (1994). *Elements of graph design*. New York: W.H. Freeman and Company.

Lohse, G. L. (1997). The role of working memory on graphical information processing. *Behaviour & Information Technology, 16*, 297-308.

Meyer, J., Shinar, D., & Leiser, D. (1997). Multiple factors that determine performance with tables and graphs. *Human Factors, 39,* 268-286.

Peebles, D. & Cheng, P. (2003). Modeling the effect of task and graphical representation on response latency in a graph reading task. *Human Factors, 45*, 28-46.

Rangecroft, M. (2003). As easy as pie. *Behaviour & Information Technology, 22*, 421-426.

Roth, W. M. & Bowen, G. M. (2003). When are graphs worth ten thousand words? An expert-expert study. *Cognition & Instruction, 21*, 429-473.

Schiano, D. J. & Tversky, B. (1992). Structure and strategy in encoding simplified graphs. *Memory & Cognition, 20*, 12-20.

Schutz, H. G. (1961). An evaluation of formats for graphic trend displays: Experiment ii. *Human Factors, 3*, 99-107.

Shah, P. (2001). Graph comprehension: The role of format, content, and individual differences. In M. Anderson, B. Meyer, & P. Olivier (Eds.), *Diagrammatic representation and reasoning* (pp. 173-185). Berlin: Springer-Verlag.

Shah, P. & Carpenter, P. A. (1995). Conceptual limitations in comprehending line graphs. *Journal of Experimental Psychology: General, 124*, 43-61.

Shah, P., Freedman, E. G., & Vekiri, I. (2005). The comprehension of quantitative information in graphical displays. In P. Shah. & A. Miyake (Eds.), *The Cambridge handbook of visuospatial thinking* (pp. 426-476). New York, NY: Cambridge University Press.

Shah, P., Hegarty, M., & Mayer, R. E. (1999). Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology, 91*, 690-702.

Siegrist, M. (1996). The use or misuse of three-dimensional graphs to represent lower-dimensional data. *Behaviour & Information Technology, 15*, 96-100.

Simkin, D. & Hastie, R. (1987). An information-processing analysis of graph perception. *Journal of the American Statistical Association, 82*, 454-465.

Spence, I. (1990). Visual psychophysics of simple graphical elements. *Journal of Experimental Psychology, 16,* 683-692.

Spence, I. & Lewandowsky, S. (1991). Displaying proportions and percentages. *Applied Cognitive Psychology, 5*, 61-77.

Vessey, I. (1991). Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decision Sciences, 22*, 219-240.

Wickens, C. D. & Hollands, J. G. (2000). *Engineering psychology and human performance*. Upper Saddle River, NJ: Prentice Hall.

Zacks, J., Levy, E., Tversky, B., & Schiano, D. J. (1998). Reading bar graphs: Effects of extraneous depth cues and graphical context. *Journal of Experimental Psychology: Applied, 4*, 119-138.

Zacks, J., & Tversky, B. (1999). Bars and lines: A study of graphic communication. *Memory & Cognition, 27*, 1073-1079.

## APPENDIX A: Empirical Studies of Graph Types

| Task | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|
| Judgment of change | Line graph<br>Horizontal bar graph<br>Vertical bar graph | Accuracy | Line graph > Vertical bar graph > Horizontal bar graph | Schutz (1961) |
| | | Response Time | Line graph < Vertical bar graph < Horizontal bar graph | |
| | | | | |
| Judgment of change | Line graph<br>Bar graph<br>Pie graph | Accuracy | Line graph = Bar graph > Pie graph | Experiment 1: Hollands & Spence (1992) |
| | | Response Time | Line graph < Bar graph < Pie graph | |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

## APPENDIX A: Empirical Studies of Graph Types (cont.)

| Task | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|
| Judgment of change | Line graph<br>Bar graph<br>Pie graph<br>Divided bar graph<br>Tiered bar graph | Accuracy | Line graph = Bar graph = Divided bar graph > Tiered bar graph = Pie graph | Experiment 2: Hollands & Spence (1992) |
| | | Response Time | Line graph < Divided bar graph < Bar graph < Tiered bar graph = Pie graph | |
| | | | | |
| Exact point extraction | Line graph<br>Bar graph | Response Time | Bar graphs < Line graphs | Meyer, Shinar, & Leiser (1997) |
| | | | | |
| Trend reading of a data series | Line graph<br>Bar graph | Accuracy | Line graphs > Bar graphs | Meyer, Shinar, & Leiser (1997) |
| | | Response Time | Line graphs < Bar graphs | |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

## APPENDIX A: Empirical Studies of Graph Types (cont.)

| Task | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|
| Point comparison in same data series but have different values on the x-axis | Line graph | Accuracy | Bar graphs = Line graphs | Meyer, Shinar, & Leiser (1997) |
| | Bar graph | Response Time | Bar graphs = Line graphs | |
| | | | | |
| Comparing two points that have the same value on the x-axis but belong to different data series | Line graph | Accuracy | Bar graphs > Line graphs | Meyer, Shinar, & Leiser (1997) |
| | Bar graph | Response Time | Bar graphs = Line graphs | |
| | | | | |
| Identifying the highest value for a specific data series | Line graph | Accuracy | Bar graphs = Line graphs | Meyer, Shinar, & Leiser (1997) |
| | Bar graph | Response Time | Bar graphs < Line graphs | |
| | | | | |
| Spontaneous interpretation of graph | Line graphs | Description of graph | Participants described line graphs as trends. | Experiment 1: Zacks & Tversky (1999) |
| | Bar graph | Description of graph | Participants described bar graphs using discrete contrasts. | |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

## APPENDIX A: Empirical Studies of Graph Types (cont.)

| Task | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|
| Spontaneous interpretation of graph | Line graphs | Description of graph | Participants described line graphs as trends. | Shah, Mayer, & Hegarty (1999) |
| | Bar graph | Description of graph | Participants described bar graphs using discrete contrasts. | |
| Comparison judgment | Bar graph<br>Line graph<br>Pie graph | Accuracy | Bar Graph > Line Graph > Pie graph | Simkin & Hastie (1987) |
| Estimating proportion of the whole | Bar graph<br>Line graph<br>Pie graph | Accuracy | Pie graph = Bar Graph > Line Graph | Simkin & Hastie (1987) |
| Judgment of proportion-Complex (comparing more than one component) | Pie Graph<br>Bar Graph | Accuracy | Pie graph > Bar Graph | Spence & Lewandowsky (1991) |
| Judgment of proportion-Simple | Pie Graph<br>Bar Graph | Accuacy | Pie graph = Bar Graph | Spence & Lewandowsky (1991) |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

**APPENDIX A: Empirical Evidence Related to Visual Graph Types (cont.)**

| Task | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|
| Judgment of proportion | Line graph<br>Bar Graph<br>Pie Graph | Accuracy | Line graph = Bar graph > Pie graph | Experiment 1: Hollands & Spence (1992) |
| | | Response Time | Line graph < Bar graph < Pie graph | |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

## APPENDIX B: Empirical Studies of Graph Characteristics

| Task | Graph Characteristic | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|---|
| Judgment of change | Scale (tick marks) | Line graph<br>Bar graph<br>Divided bar graph<br>Tiered bar graph<br>Pie graph | Accuracy | **Scale:** Line graph = Bar graph = Divided bar graph > Tiered bar graph = Pie graph | Experiment 2: Hollands & Spence (1992) |
| | | | | **No scale:** Line graph = Bar graph = Divided bar graph > Tiered bar graph = Pie graph. HOWEVER, lower accuracy with Tiered bar graphs and Pie graphs without scale. | |
| | | | | | |
| Judgment of proportion | Scale (tick marks) | Line graph<br>Bar graph<br>Divided bar graph<br>Pie graph | Accuracy | **Scale:** Line graph = Bar graph = Divided bar graph = Pie graph | Experiment 2: Hollands & Spence (1992) |
| | | | | **No scale**: Divided bar graph = Pie graph > Line graph = Bar graph | |
| | | | Response Time | **Scale:** Line graph = Bar graph = Divided bar graph = Pie graph | |
| | | | | **No scale**: Divided bar graph = Pie graph < Line graph = Bar graph | |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

## APPENDIX B: Empirical Studies of Graph Characteristics (cont.)

| Task | Graph Characteristic | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|---|
| Comparison, Difference, & Mean Questions | Scale (y-axis tick marks) | Line graph | Accuracy | Tick marks in y-axis decreased accuracy | Gillan & Richman (1994) |
| | | Bar graph | Response Time | Tick marks in y-axis increased response time | |
| | | | | | |
| Comparison, Difference, & Mean Questions | Background | Line graph | Accuracy | Background decreased accuracy | Gillan & Richman (1994) |
| | | Bar graph | Response Time | Background increased response time | |
| | | | | | |
| Comparison, Difference, & Mean Questions | Axis lines | Line graph | Accuracy | x- and y-axes increased accuracy | Gillan & Richman (1994) |
| | | Bar graph | Response Time | x- and y-axes decreased response time | |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

**APPENDIX B: Empirical Studies of Graph Characteristics (cont.)**

| Task | Graph Characteristic | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|---|
| Complex budget task | Color & Gridlines | Line graphs | Accuracy | Color & Gridlines > Monocolor & No gridlines | Lohse (1997) |
| | | | Response Time | Monocolor & No gridlines < Color & Gridlines | |
| Inequality decision (Did graph match inequality statement? E.g., A>B) | Bar orientation | Vertical bar graph<br><br>Horizontal bar graph | Response Time | Vertical bar graph < Horizontal bar graph | Fischer, Dewulf, & Hill (2005) |
| Describe graph | Placement of independent and dependent variables on the axes | Line graph | Accuracy | Participants described the variable on the y-axis as a function of the variable on the x-axis, regardless if the independent (cause) variable was on the x-axis. | Gattis & Holyoak (1996) |
| Interpret graph | Legend | Line graph | Eye gazes | Participants read and re-read axes and legends. Suggested that it was difficult to keep track of all graph details. | Carpenter & Shah (1998) |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

## APPENDIX C: Empirical Studies of Dimensionality

| Task | Graph Characteristic | Graph Type Investigated | Performance measurement | Results | Reference |
|---|---|---|---|---|---|
| Relative magnitude estimation | 1D; 2D; 3D | 1D: Vertical line<br>1D: Horizontal line<br>2D: Bar graph<br>2D: Pie graph<br>2D: Disk graph.<br>3D: Cylinder<br>3D: Box | Accuracy | 1D = 2D = 3D | Spence (1990) |
| | | | Response time | 2D = 3D < 1D | |
| | | | | | |
| Relative magnitude estimation | 2D; 3D | Pie graph | Accuracy | 2D > 3D | Siegrist (1996) |
| | | Bar graph | Accuracy | 2D = 3D | |
| | | | | | |
| Relative magnitude estimation | 2D; 3D | Bar graph | Response time | 2D < 3D | Siegrist (1996) |
| | | | | | |
| Height judgments | 2D; 3D | Bar graph | Accuracy | 2D > 3D | Zacks, Levy, Tversky, & Schiano (1998) |

NOTE: 1D = 1-dimensional; 2D = 2-dimensional; 3D = 3-dimensional.
> indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

**APPENDIX C: Empirical Studies of Dimensionality (cont.)**

| Task | Graph Characteristic | Graph Type Investigated | Performance measurement | Results | Reference |
|---|---|---|---|---|---|
| Identify largest and smallest segments | 2D; 3D | Pie graph | Accuracy | 2D > 3D | Rangecroft (2003) |
| | | | | | |
| Inequality decision (Did graph match inequality statement? E.g., A>B) | 2D; 3D | Bar graph | Response time | 2D < 3D | Experiment 1: Fischer (2000) |
| | | | | | |
| Comparison, Difference, & Mean Questions | 2D; 3D | Line graph<br>Bar graph | Accuracy | 3D = 2D | Gillan & Richman (1994) |
| | | | Response Time | 3D = 2D | |

NOTE: 1D = 1-dimensional; 2D = 2-dimensional; 3D = 3-dimensional.

> indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

## APPENDIX D: Empirical Studies of Data Characteristics

| Task | Data Characteristic | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|---|
| Trend reading | Data complexity: Low=3 data series with same trend across days; High=1 of 3 series had an opposite trend | Line graph | Accuracy | Low > High (Accuracy slightly better for simple than complex data displays) | Experiment 1: Meyer, Shinar, & Leiser (1997) |
| Comparisons between levels of x values | | | | Low > High | |
| Identify maximum values | | Bar Graph | | Low > High | |
| Comparisons between data series | | | | Low = High | |
| | | | | | |
| Trend reading | Data complexity: Low=3 data series with same trend across days; High=1 of 3 series had an opposite trend | Line graph | Response Time | Line graphs: Low = High (Data complexity had no effect on response time) | Experiment 1: Meyer, Shinar, & Leiser (1997) |
| Comparisons between levels of x values | | | | | |
| Identify maximum values | | Bar graph | | Bar graphs: Low < High (Simple displays had faster response times than complex displays) | |
| Comparisons between data series | | | | | |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

**APPENDIX D: Empirical Studies of Data Characteristics (cont.)**

| Task | Data Characteristic | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|---|
| Comparisons between levels of x values | Data complexity: Number of data points per series (3, 5, 7) | Line graph | Response Time | High data complexity (7 data points per series) led to longer RTs | Experiment 2: Meyer, Shinar, & Leiser (1997) |
| Identify maximum value | | Bar graph | | | |
| | | | | | |
| Exact point extraction | Data complexity: Number of data series presented (2, 4). | Line graph | Response Time | 2 data series < 4 data series presented. RT faster for 2 data series compared to 4 data series presented for all graph tasks and graph types. | Experiment 2: Meyer, Shinar, & Leiser (1997) |
| Trend reading of a data series | | | | | |
| Point comparison in same data series but have different values on the x-axis | | | | | |
| Comparing two points that have the same value on the x-axis but belong to different data series | | Bar graph | | Bar graphs < Line graphs. Difference between RTs was larger for line graphs than for bar graphs. | |
| Identify maximum value | | | | | |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

**APPENDIX D: Empirical Studies of Data Characteristics (cont.)**

| Task | Data Characteristic | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|---|
| Spontaneous interpretation of graph | Discrete data; Continuous data | Line graphs<br>Bar graph | Description of graph | Discrete data (gender) led to discrete comparisons. Continuous data (age) led to trend judgments. | Experiment 2: Zacks & Tversky (1999) |
| | | | | | |
| Interpret graphically displayed statistical interactions | Number of variables (2, 3, 4, 5) | Bar graph | Accuracy | Increasing number of variables leads to decreasing accuracy. | Halford, Baker, McCredden, & Bain (2005) |
| | | | Response Time | Increasing number of variables leads to increasing response time. | |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

**APPENDIX D: Empirical Studies of Data Characteristics (cont.)**

| Task | Data Characteristic | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|---|
| Spontaneous interpretation of graph | Number of trend reversals | Line graph | Description of graph | Local descriptions increased along with increasing number of trend reversals. | Carswell, Emery, & Lonon (1993) |
| | | | Study Time | Study time increased along with increasing number of trend reversals. | |
| | Number of data points comprising a data series | | Description of graph | Global descriptions increased along with increasing number of data points. | |
| | | | Study Time | Study time increased along with increasing number of data points. | |
| | | | | | |
| Interpret graph | Variables across three dimensions (x, y, and z) | Line graph | Accuracy (matching one graph to another) | Participants failed to recognize the same data when plotted in a different but equivalent graph. | Shah & Carpenter (1995) |
| | | | Verbal description | Participants failed to make z-y relationship inferences. | |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

## APPENDIX E: Empirical Studies of Person Characteristics

| Task | Person Characteristic | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|---|
| Comparisons between levels of x-value | Experience (practice effects across the three blocks of experimental trials) | Bar graph | Accuracy | Block 3 > Block 1. Accuracy increased over course of experiment. | Meyer, Shinar, & Leiser (1997) |
| Comparisons between series | | Line graph | | | |
| Trend reading | | | | | |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

**APPENDIX E: Empirical Studies of Person Characteristics (cont.)**

<table>
<tr><th>Task</th><th>Person Characteristic</th><th>Graph Types Investigated</th><th>Performance Measurement</th><th>Results</th><th>Reference</th></tr>
<tr><td>Exact point extraction</td><td rowspan="5">Experience (practice effects across the three blocks of experimental trials)</td><td rowspan="2">Bar graph</td><td rowspan="5">Response Time</td><td rowspan="2">Block 3 < Block 1. RTs decreased over course of experiment.</td><td rowspan="5">Meyer, Shinar, & Leiser (1997)</td></tr>
<tr><td>Trend reading of a data series</td></tr>
<tr><td>Point comparison in same data series but have different values on the x-axis</td><td rowspan="3">Line graph</td><td rowspan="3">Bar graph < Line graph (Bar graph RTs decreased more than line graph RT)</td></tr>
<tr><td>Comparing two points that have the same value on the x-axis but belong to different data series</td></tr>
<tr><td>Identifying the highest valuefor a specific data series</td></tr>
</table>

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

## APPENDIX E: Empirical Studies of Person Characteristics (cont.)

| Task | Person Characteristic | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|---|
| Comparison between data series | Familiarity (Half of ppts saw axes and legend for graphs prior to presentation) | Bar graph | Accuracy | Prior presentation = No prior presentation | Meyer, Shinar, & Leiser (1997) |
| | | Line graph | Response Time | Prior presentation < No prior presentation | |
| | | | | | |
| Reading exact value | Familiarity (Half of ppts saw axes and legend for graphs prior to presentation) | Bar graph | Accuracy | Prior presentation = No prior presentation | Meyer, Shinar, & Leiser (1997) |
| | | Line graph | Response Time | Prior presentation < No prior presentation | |
| | | | | | |
| Complex budget task | Working memory | Line graphs | Accuracy | High WM using monocolor graph without gridlines = Participants using color graphs with grid lines | Lohse (1997) |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.

**APPENDIX E: Empirical Studies of Person Characteristics (cont.)**

| Task | Person Characteristic | Graph Types Investigated | Performance Measurement | Results | Reference |
|---|---|---|---|---|---|
| Describe graph | Knowledge of graph type: Experts | Distribution graph<br>Population model graph<br>Isocline graph | Accuracy | "The experts were far from perfect in providing more than a literal reading and arriving at standard inferences from the graph…"(p. 466) | Roth & Bowen (2003) |
| | | | | | |
| Describe graph | Knowledge of data: Familiar vs. Unfamiliar | Line graph | Accuracy | When the graph contained data with which participants were familiar in an unexpected relationship, participants described the *expected* relationship and not the relationship shown. | Shah (2001) |
| | | | | | |
| Describe graph | Knowledge of graphs: Experienced vs. Unexperienced | Line graph | Accuracy | Graduate students who had much experience with graphs did not describe line graphs with x-y-z relationships more accurately than undergraduate students with less graph experience. | Shah & Carpenter (1995) |
| | | | | | |
| Draw graph from memory | Memory | Line graph | Accuracy | Memory for line graphs biased to 45 degrees. Suggests readers have expectations of data in line graphs. | Schiano & Tversky (1992) |

NOTE: > indicates better performance for accuracy; < indicates faster response time; = indicates equivalent performance.