

Topics in Cognitive Science 3 (2011) 446–474

Copyright © 2011 Cognitive Science Society, Inc. All rights reserved.

ISSN: 1756-8757 print / 1756-8765 online

DOI: 10.1111/j.1756-8765.2011.01150.x

## The Cognitive Science of Visual-Spatial Displays: Implications for Design

Mary Hegarty

*Department of Psychological & Brain Sciences, University of California, Santa Barbara*

Received 15 July 2010; received in revised form 6 January 2011; accepted 27 February 2011

---

### Abstract

This paper reviews cognitive science perspectives on the design of visual-spatial displays and introduces the other papers in this topic. It begins by classifying different types of visual-spatial displays, followed by a discussion of ways in which visual-spatial displays augment cognition and an overview of the perceptual and cognitive processes involved in using displays. The paper then argues for the importance of cognitive science methods to the design of visual displays and reviews some of the main principles of display design that have emerged from these approaches to date. Cognitive scientists have had good success in characterizing the performance of well-defined tasks with relatively simple visual displays, but many challenges remain in understanding the use of complex displays for ill-defined tasks. Current research exemplified by the papers in this topic extends empirical approaches to new displays and domains, informs the development of general principles of graphic design, and addresses current challenges in display design raised by the recent explosion in availability of complex data sets and new technologies for visualizing and interacting with these data.

*Keywords:* Visual displays; Graphics; Design; Iconic displays; Relational displays; Visualization

---

### 1. Introduction

Visual-spatial displays are ubiquitous in human communication. We can trace their origin to caveman drawings (Tversky, 2001). They evolved throughout history with the development of perspective drawing in the Renaissance (Panofsky, 1960), mapping techniques in the age of exploration (Brown, 1979), drawing techniques during the industrial age (Ferguson, 2001), and graphing techniques by Playfair and others (Wainer, 2005). In recent years, with advances in computer graphics and human–computer interaction techniques, dynamic and interactive displays have become commonplace. We now watch animations of

---

Correspondence should be sent to Mary Hegarty, Department of Psychological & Brain Sciences, University of California, Santa Barbara, CA 93106. E-mail: [hegarty@psych.ucsb.edu](mailto:hegarty@psych.ucsb.edu)

developing weather patterns on the news every night, and we can pan and zoom to any location on the earth's surface using programs such as Google Earth.

Cognitive scientists have long argued that representations that are informationally equivalent (contain the same information) are not necessarily computationally equivalent (Larkin & Simon, 1987). In the case of visual-spatial displays, specifically, there is much evidence that task performance can be dramatically different with different visual displays of the same information (e.g., Breslow, Trafton, & Ratwani, 2009; Gattis & Holyoak, 1996; Hegarty, Canham, & Fabrikant, 2010; Novick & Catley, 2007; Peebles & Cheng, 2003; Sanfey & Hastie, 1998; Schwartz, 1995; Shah & Carpenter, 1995; Simkin & Hastie, 1986; Yeh & Wickens, 2001; Zhang & Norman, 1994). This evidence argues for the importance of visual display design.

Many different disciplines and research communities are concerned with questions of the design and effectiveness of visual displays. The design of maps is the domain of cartography (Dent, 1999). Statisticians are concerned with how to best design graphs to gain insights into data patterns (Cleveland, 1985; Tufte, 2001; Tukey, 1977; Wainer, 2005). Researchers in human factors and engineering psychology develop and test principles of design for domains such as process control, aviation, medicine, and strategic route planning (Smallman & St. John, 2005; Vicente, 2002; Wickens & Hollands, 2000). Educational researchers develop dynamic interactive visualizations to teach students about scientific processes such as chemical and mechanical interactions or to relate different representations of the same phenomenon (e.g., Ainsworth, 2006; White, 1993; Wu, Krajcik, & Soloway, 2001). Research communities in scientific visualization (McCormick, DeFanti, & Brown, 1987), information visualization (Card, Mackinlay, & Schneiderman, 1999), and geovisualization (MacEachren & Kraak, 2000) have developed around questions of how to best use new information technologies to reveal patterns in complex data sets. In recent years, a new multidisciplinary field known as *Visual Analytics* has emerged and is characterized by the use of dynamic visualizations to support analytic thinking with large data sets (Thomas & Cook, 2005; Fisher, this volume). However, with their own conference series, etc., these communities have remained somewhat isolated and although cognitive science is central to the design of visual displays, the development of new display technologies is often uninformed by cognitive science.

This paper begins by outlining the types of visual-spatial displays with which we are concerned. This is followed by a discussion of why visual-spatial displays are said to “augment cognition” and an overview of the perceptual and cognitive processes involved in using displays. I then argue for the importance of cognitive science methods to the design of visual displays and review some of the main principles of display design that have emerged from these methods. Finally I introduce the other papers in the topic and identify some current challenges of research on this topic.

## 2. Types of visual-spatial displays

This topic is concerned with external visual-spatial representations, that is, visual-spatial arrays that *represent*, or are symbols for, objects, events, or more abstract information. We

are not concerned with verbal representations. To understand the nature of a representation, we must specify the nature of the object or entity that is represented (i.e., the referent or *represented* world), the representational medium (i.e., the *representing* world), and the correspondence between these two (Palmer, 1978). In the case of visual displays, the representational medium is a display that is perceived by the visual system and distributed over space. It can be printed on paper or displayed on a computer screen. Visual displays can be categorized into different types based on the relation between the representation and its referent and the complexity of the information represented.

2.1. *Iconic displays: Visual-spatial displays that represent visual-spatial entities*

First, iconic displays are representations of objects that themselves are visual spatial entities. Examples of these types of displays are a diagram of a machine (Fig. 1A) or a road map (Fig. 1B). In iconic displays, space on the page represents space in the world and the properties displayed (shape, color, etc.) are also visible properties. However, most visual displays abstract from the reality that they represent, displaying some aspects of the information but not others, for example, when a mechanical diagram represents pulleys as circles as in Fig. 1A. Iconic displays can also distort reality. For example, many subway maps show

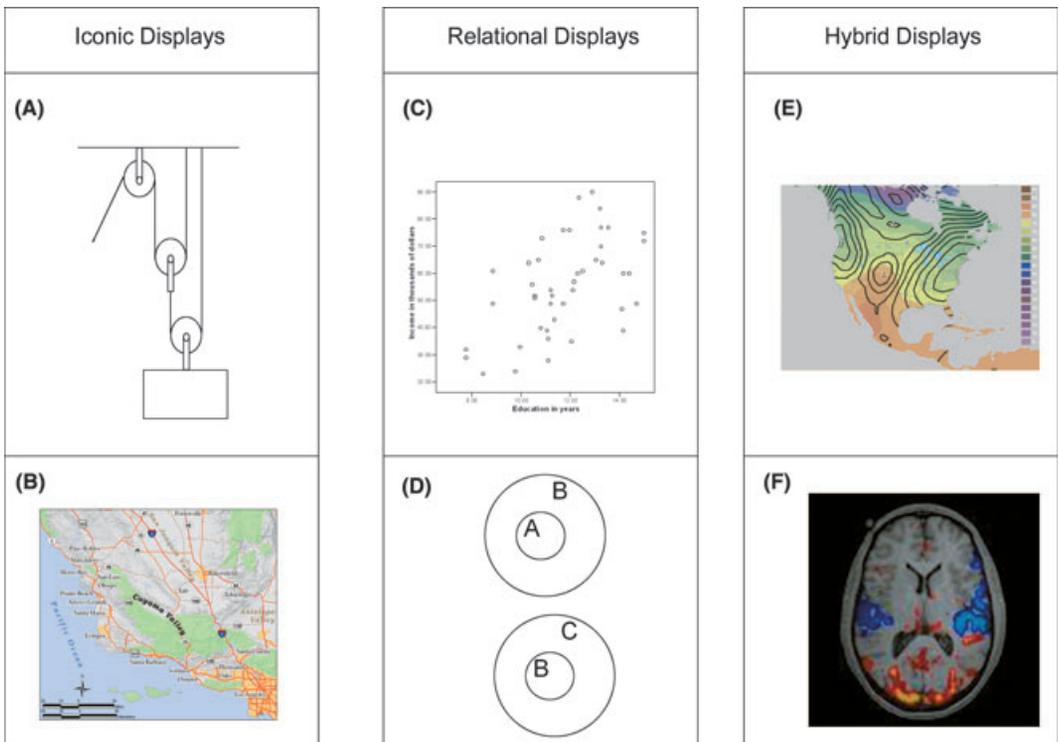


Fig. 1. Examples of iconic, relational, and hybrid displays.

the connectivity between subway stations while distorting the distance, and maps of the earth's surface involve distortions because they project a three-dimensional entity on a two-dimensional space. Finally, iconic displays can show views of their referents that are not visible in real-world viewing, such as a cross section of the liver or a diagram of a molecule.

## 2.2. Relational displays: Visual-spatial displays that represent abstract relationships

Other visual displays represent entities that do not have spatial extent and are not visible, such as when a scatterplot shows the relationship between two variables (Fig. 1C), Euler circles represent the premises in a reasoning problem (Fig. 1D), and a linguistic tree diagram shows the structure of a sentence. These are often referred to as *relational displays* (e.g., Zhang, 1996). In these displays, visual and spatial properties represent entities and properties that are not necessarily visible or distributed over space. Variables, such as color, shape, and location are the *representing* dimensions of the display (Bertin, 1983) but the *represented* dimensions can be any category or quantity. Color can represent heat, location can represent importance, etc. Space and visual properties are used as metaphors for other properties or relations.

Note that some displays can be a *hybrid of* iconic and relational displays in displaying nonvisible properties overlaid on a representation of a visual-spatial entity. For example, a meteorologist might use a map showing levels of pressure and temperature (which are not visible in the natural world, see Fig. 1E). Thematic maps are *geospatial* displays that represent entities at the geographic scale of space (Montello, 1993). Hybrid displays can also represent other scales of space—for example, when fMRI data are visualized to show a map of activity across the brain (Fig. 1F). In these displays, there is a direct mapping between space in the representation and space in the referent, but nonvisual properties are represented by visual variables, such as color, shading, etc.

## 2.3. Complex displays: Multiple representations, animation, and interactivity

Orthogonal to the distinction between iconic and relational displays, external visual-spatial displays vary in the complexity of the information that they represent, and displaying complex information often necessitates the use of visual-spatial representations other than a single static image. One type of complex display may show multiple related representations, for example, when the interface to a nuclear power plant shows representations of the current states of several components of the system, or a computer-aided design system allows the user to view different orthographic projections and sections of the structure of a building or machine. Complex displays also include animations, in which a sequence of images shows, for example, how a process unfolds over time (Tversky, Morrison, & Betrancourt, 2002). They also include interactive displays that allow the user to add or subtract variables displayed on a graph or map, rotate the display to view an object or three-dimensional data set from different perspectives, filter out task-irrelevant information, or zoom into a data set to view some aspect of the data in more detail (Card et al., 1999; Robertson, Czerwinski,

Fisher, & Lee, 2009). Developments in information technologies have led to new challenges of how to visualize large and complex data sets with researchers in scientific visualization focusing on displays of spatially distributed data (e.g., the development of a thunderstorm) and researchers in information visualization focusing more on visualization of abstract information spaces (e.g., semantic relations between documents).

### **3. How visual-spatial displays augment cognition**

Visual displays are often said to enhance or “augment” cognition (Card et al., 1999; Larkin & Simon, 1987; MacEachren, 1995; Norman, 1993; Scaife & Rogers, 1996). The following sections summarize some of the main advantages they afford for cognitive tasks.

#### *3.1. External storage of information*

First, all types of visual displays are external representations, and therefore store information externally, freeing up working memory resources for other aspects of thinking (Card et al., 1999; Scaife & Rogers, 1996). This does not mean that there are no internal representations or processes when people use graphical displays. However, the external display can serve as the information store, so that the internal representation at a given time can be quite sparse, perhaps containing only detailed information about a single location of the display being currently viewed and pointers to locations of other important information in the display (Pylyshyn, 2003). Thus, the representation is distributed over sparse internal representations and detailed external representations (Scaife & Rogers, 1996; Zhang & Norman, 1994).

#### *3.2. Organization of information*

A second advantage of visual-spatial over sentential (language-like) representations is that they organize information by indexing it spatially (Larkin & Simon, 1987). Grouping information that is related is a natural property of iconic displays. In these displays, space in the display represents space in the world, so that if the representation of two items is close in the display, it is likely that those items are also close in the represented world. Things that are close in the natural world tend to be more highly related (proposed as “the first law of Geography” by Tobler, 1970, p. 236). For example, objects exert forces on objects they touch, and diseases spread to closer regions before they spread to more distant regions. Therefore, information that needs to be related in interpreting and making inferences from iconic displays is likely to be represented by visual features that are close in the display.

In the more abstract world of digital information, related information is not necessarily physically closer. However, relational displays often organize information such that the representations of related entities are close, facilitating search and integration. Graphs organize entities by placing them in a space defined by the x and y axes. As a result, similar entities are visualized as close together. For example, in a scatter plot relating a sample of people’s

income to their education (e.g., Fig. 1C), the dots representing people with similar levels of income and education are located close together in the display. Similarly, in an organizational chart of a company, the representations of people with similar levels of authority are close in the vertical dimension. Moreover, closeness in spatial location is just one aspect of “display proximity” according to Wickens and Carswell (1995). Proximity in other dimensions such as color, or achieved by graphical devices (e.g., connecting related items by lines, or enclosing related items with contours) can also facilitate search and integration of disparate sources of information. The latter often reflect Gestalt principles of perceptual organization.

### 3.3. Offloading of cognition on perception

In addition to offloading storage, visual displays can allow the offloading of cognitive processes onto perceptual processes (Scaife & Rogers, 1996), referred to by Card et al. (1999) as “using vision to think.” When nonvisual data are mapped onto visual variables, patterns often emerge that were not explicitly built in, but which are easily picked up by the visual system. These are referred to as emergent features (Pomerantz & Pristach, 1989), that is, visual properties of a group of objects that are more salient than properties of the individual objects themselves. They can enable complex computations to be replaced by simple pattern recognition processes. For example, consider the data shown in a table and graph in Fig. 2 representing the population of San Francisco from 1860 to 2000. It is clear from

Year	Population
1860	56,802
1870	149,473
1880	233,959
1890	298,997
1900	342,782
1910	416,912
1920	506,676
1930	634,394
1940	634,536
1950	775,357
1960	740,316
1970	715,674
1980	678,974
1990	723,959
2000	776,733

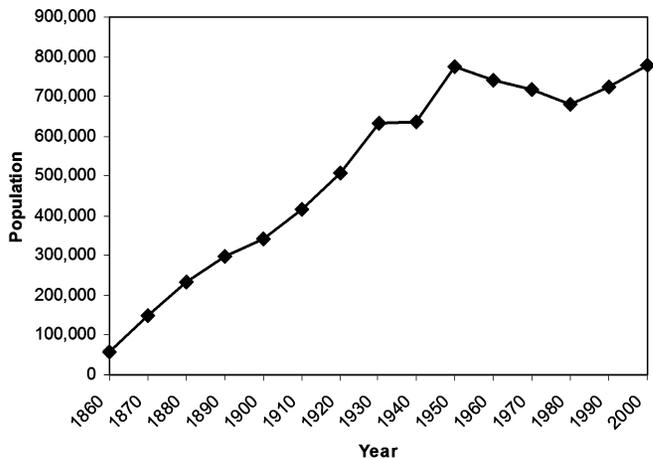


Fig. 2. Table and graph of the same data showing the population of the city of San Francisco from 1860 to 2000 (source: Wikipedia). The near linear increase in population from 1860 to 1930 is easily detected by the visual system when the data are graphed, but it requires a series of computations to be detected from the table.

looking at the graph that there was an almost linear increase in population from 1860 to 1930. The visual system easily detects that the line joining the data points for these years is almost straight. This linear trend is an emergent feature of the graph. To discover the same trend from the table, one would have to subtract the population for each year from each preceding year, remember the intermediate products and notice that they are approximately equal.

Designing displays that capitalize on emergent features is a central principle of the ecological approach to the design of human machine interfaces to complex systems (Vicente, 2002). For example, say it is important for safety that the levels of fluid in four reservoirs in a power plant remain equal. Individual sensors measure the fluid in individual reservoirs, but higher-level relations between the levels in the different reservoirs are more important to the goals of the individual monitoring the system. Departures from equality will be detected more easily with the configuration of rectangles in Fig. 3B than with that in Fig. 3A, because the alignment or misalignment of fluid levels is an emergent property of Fig. 3B.

Offloading cognition on perception also occurs when the right representation for a problem constrains possible inferences to those which are most valid (Scaife & Rogers, 1996). For example, solving syllogisms is easier when the premises are represented by Euler circles, because the representation places constraints on the types of inferences that can be made (Stenning & Oberlander, 1995). To take a simple example, it is obvious that if “All A are B” and “All B are C” are represented as in Fig. 1D, then “All A are C” is a valid inference, because combining these diagrams cannot lead to a situation in which the representation of set A is not inside set C. This works because the constraints of the diagrams correspond to the rules of logic, so that matching the representation to the domain is critical.

### 3.4. Offloading cognition on action

Finally, when a display is interactive, people can offload internal mental computations on external manipulations of the display itself (Card et al., 1999). For example, Kirsh and

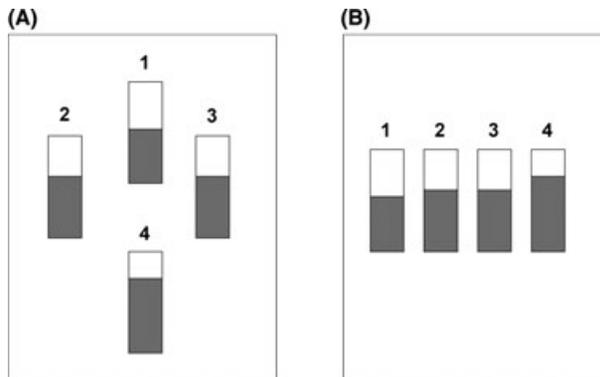


Fig. 3. Two possible configurations of displays showing the level of fluid in different reservoirs in a hypothetical power plant. The nonequality of the fluid levels is more easily detected in configuration B because it is an emergent feature of the display.

Maglio (1994) found that experienced Tetris players often rotated the Tetris shapes more than they needed to, and proposed that they learned to rotate shapes in the external display and observe the results rather than performing cognitively effortful mental rotations. Kirsh (1997) referred to these types of interactions with external displays as *complementary* actions, that is, actions performed in the world that relieve the individual of the need to perform an internal computation. Similarly, one might use an interactive interface to filter information in a complex database so that only task-relevant information is displayed. For example, Schneiderman (1994) developed a system that allows real-estate professionals and their clients to display the locations of homes, filtering by such features as price and number of bedrooms. In this example, an interaction with the display does some of the cognitive work for the viewer so that he or she does not have to engage effortful selective attentional processes to sift relevant from irrelevant information in a complex visual display.

#### 4. Cognitive processes in using visual-spatial displays

Although visual-spatial displays can enhance thinking in many ways, this does not mean that their use is necessarily easy or transparent. Display comprehension involves a complex interaction between bottom-up and top-down processes, diagrammed in Fig. 4 (adapted from Freedman & Shah, 2002; Pinker, 1990), which are not guaranteed to be successful. First, the visual system senses the features of the display, such as color and shape, and encodes these features to construct an internal representation of the display. Exactly which of these features are encoded depends on attention, which might be directed by the viewer’s goals and expectations or what is salient in the display. For example, one difficulty in display comprehension might arise if the viewer is distracted by highly salient but task-irrelevant information so that he or she fails to encode the critical information, although it is presented.

In addition to basic perceptual, attentional, and encoding processes, which construct a representation of the external display, the user of a display typically has to apply knowledge

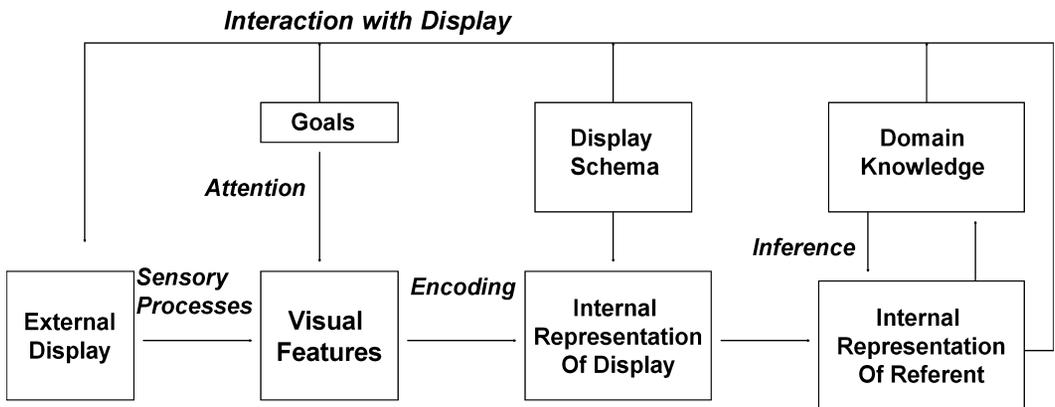


Fig. 4. Diagram of the different representations (indicated by boxes) and processes (indicated by arrows) involved in understanding a visual display.

to construct a representation of its referent. This can include knowledge of the display conventions, such as the meaning of the x and y axes in a graph, which types of information are typically included in this type of display and which are omitted, which aspects of the display are to be taken literally (such as the relative length and configuration of roads on a road map) and which are not (such as the color and width). This type of knowledge has been discussed primarily in the context of graphs and referred to as a graph schema (Pinker, 1990; Ratwani & Trafton, 2008), but in Fig. 4, I use the more general term “display schema.” Again comprehension can fail if the user’s display schema is incomplete.

Understanding a graphic can also include making further inferences based on domain knowledge or visual-spatial processes (comparison, mental rotation, etc.) so that the resulting internal representation comes to contain information that is not presented explicitly in the external display. If the display is interactive, the individual may also choose to change it, for example, by annotating it, zooming in, rotating it, etc. In this case, not just the internal representation, but the external representation changes constantly during the comprehension process. The decision to interact, and choice of how to interact with the graphic depends on meta-knowledge of the affordances of that type of display, such as whether it can be zoomed, rotated, or animated. It also depends on meta-knowledge of which interactions with the display are task relevant. This type of understanding has been referred to as *meta-representational competence* (DiSessa, 2004) and cannot always be assumed. Moreover, even if users understand the affordances of the interactive display, they might become disoriented as they use interactive features like rotation or zooming.

Not all difficulties with display use can be solved by design. For example, display design cannot compensate for lack of relevant knowledge or meta-representational competence. But good display design can help alleviate some of the problems outlined above. For example, a display can be designed to make task-relevant information salient or eliminate irrelevant information. It can capitalize on cultural conventions (e.g., higher is better; red signifies danger) so that the mapping between the display and its referent is more transparent. It can include landmarks to prevent users from getting disoriented when they rotate or zoom into displays. There are many decisions that a designer has to make in creating a new display, including how much detail to present, what type of graphic (table or graph, network or matrix, etc.), how to map visual variables to the conceptual variables that they represent, and the amount of interactivity to allow. We now turn to a discussion of how cognitive science can inform these decisions.

## 5. The importance of cognitive science to display design

The design of visual displays is often based on intuitions, approaches that analyze the basic dimensions or “visual variables” that make up visual-spatial displays (e.g., Bertin, 1983), or deriving principles from examples of good and poor graphics (e.g., Tufte, 2001). Display design is both an art and a science and these approaches are important. However, there are many cases in which intuitions and expert opinions about displays do not conform to their actual effectiveness when performance with displays is objectively measured.

Empirical methods used in cognitive science and related fields (e.g., human factors) are central to testing and revising design principles based on objective data. Moreover, knowledge of human information processing and empirical measures of performance with visual displays can inform the development of cognitive models that make a priori predictions about display effectiveness.

### 5.1. *The importance of objective measures*

Intuitions about the effectiveness of displays do not always conform to their actual effectiveness. Animations of physical processes, such as the workings of machines, and biological mechanisms, provide a prominent example. Intuitively it would seem that animation should be a good means of communicating how these processes work. For example, in animations of these processes, the shapes, locations, and movements of parts of the representation correspond directly to the shapes, locations, and movements of their referents. However, in a review of several papers comparing animated to static displays, Tversky et al. (2002) indicated that there was no advantage to animations over static displays, making the point that animations are often ineffective because they are too fast or too complex. Realistic animations of a mechanical or biological system often show several different components moving at once, and critical phases often happen very quickly, but visual attention is limited, so it is not possible to encode and relate the movements of the components in the time available. A possible solution is to give users interactive control over the animation (allowing them to control the speed, pause, rewind, etc.), but even with such interactive controls, a recent study found that students constructed the wrong mental model of how a mechanical system works, a model that was actually inconsistent with the information displayed (Kriz & Hegarty, 2007). At least in educational situations, a series of ‘‘small multiples’’ showing key frames in the process can be as effective or more effective (Mayer, Hegarty, Mayer, & Campbell, 2005).

Intuitions about animations are an example of a more general intuition that iconic displays should resemble their referents as much as possible, when in fact the power of visual spatial displays often comes from their ability to simplify and abstract from reality. Smallman and St. John (2005) have provided extensive evidence that people have a strong preference for displays that emphasize high-fidelity spatio-temporal realism, even when these displays result in poor performance. They term this misplaced faith in realistic displays ‘‘Naïve Realism’’ and theorize it is rooted in metacognitive errors (folk fallacies) about the nature of perception. Specifically they argue that folk psychology is that perception is simple, accurate, and complete, accounting for the intuitions that a realistic information display be internalized easily when in fact, perception is hard, flawed, and sparse accounting for the poor performance with realistic information displays.

In summary, preference for animation, and naïve realism more generally, provide a strong argument for empirically testing the effectiveness of displays rather than relying on users’ or designers’ intuitions. One might argue that expert intuitions are more likely to be accurate than those of novices. But even expert intuitions have been found to be erroneous. For example, during the 20th century, statisticians developed a strong bias against the pie chart

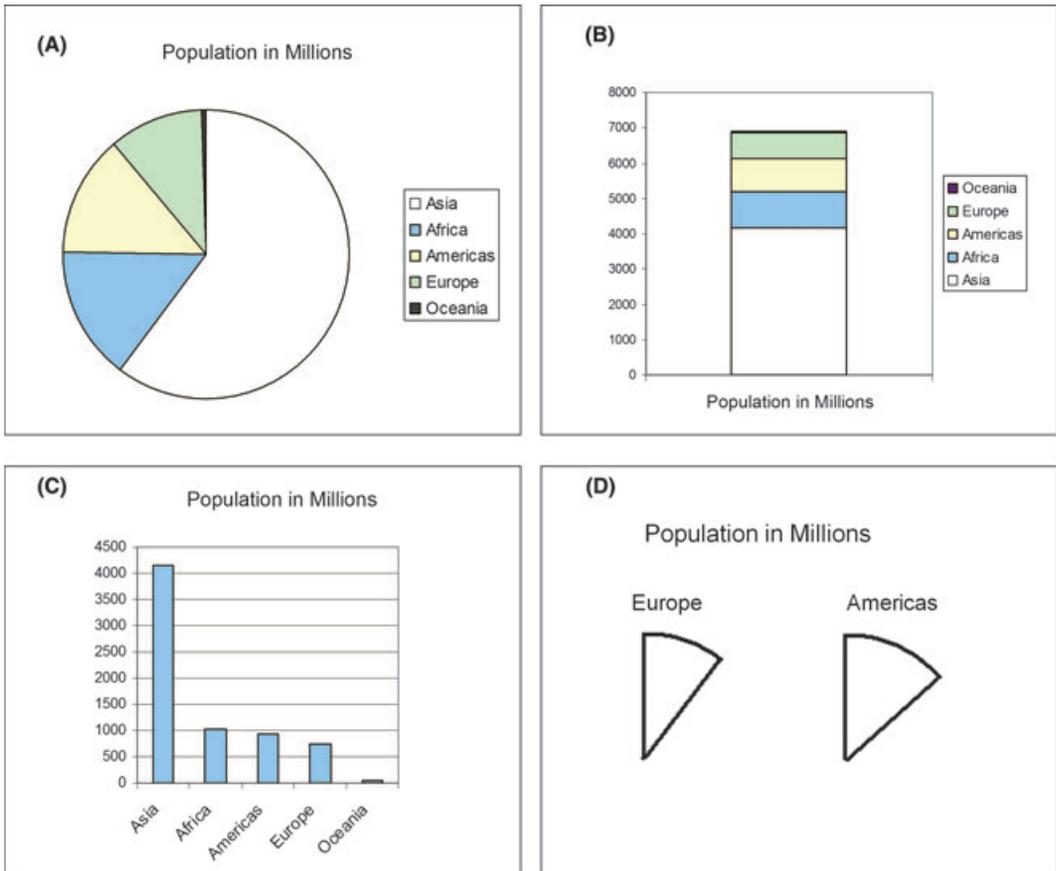


Fig. 5. Examples of (A) a pie chart, (B) a divided bar chart, (C) a bar chart showing world population by continent, and (D) an aligned pie graph showing a subset of the data.

(see Fig. 5A) preferring divided bars (see Fig. 5B) or bar graphs as a means of displaying proportions. However, careful experiments indicated that for some tasks, pie charts are as effective as divided bar charts, and for other tasks they are actually more effective (Hollands & Spence, 1998; Simkin & Hastie, 1986; Spence & Lewandowsky, 1991). Simple judgments (e.g., comparing the population of Europe and Africa in Fig. 5) were slightly more effective with bar graphs, but complex comparisons (e.g., comparing the combinations of components) were more efficient with pie charts (Spence & Lewandowsky, 1991).

## 5.2. How objective measures inform display design

Researchers collect several types of empirical data to compare the effectiveness and different types of visual displays. The most common method is to record accuracy and response times of individuals as they answer specific questions with different displays of the

same information, with the assumption that more effective displays are those that produce more accurate and efficient question answering. This approach has been used extensively in examining graph comprehension in particular (see reviews by Lewandowsky & Behrens, 1999; Shah, Freedman, & Vekiri, 2005). Objective performance measures of accuracy and response time are also increasingly being used to examine the effectiveness of geospatial displays (e.g., Fabrikant, Rebich-Hespanha, & Hegarty, 2010; Smallman & Cook, this volume; Yeh & Wickens, 2001), reflecting an increasing emphasis on the use of cognitive methods in cartography (e.g., Fabrikant & Lobben, 2009; MacEachren, 1995).

Another approach is to show people different displays and ask them to describe what the graph shows. Spontaneous descriptions can reveal what information is salient in a graphic, how much of the displayed information different individuals actually encode, and their schemas for what types of information a particular type of graphic communicates. For example, Shah and Carpenter (1995) asked people to describe line graphs showing the effects of two variables (e.g., stress and hours of study) on a measured variable (e.g., scores in an achievement test). They varied which of the independent variables was displayed on the x axis and which shown by lines with different colors and markers (referred to as the z-variable, see Fig. 6). Participants described the same data differently, depending on how it was displayed and their descriptions emphasized the x-y trends. When shown two graphs of the same data, as in Fig. 6, they were unable to tell that they were informationally equivalent.

A related approach is to examine what visual-spatial representations people spontaneously produce when asked to communicate different forms of information. For example, Tversky, Kugelmass, and Winter (1991) asked children to place stickers on a page to represent spatial, temporal, quantitative, and preference dimensions. For a temporal judgment they might have to place stickers for breakfast, lunch, or dinner and for a preference

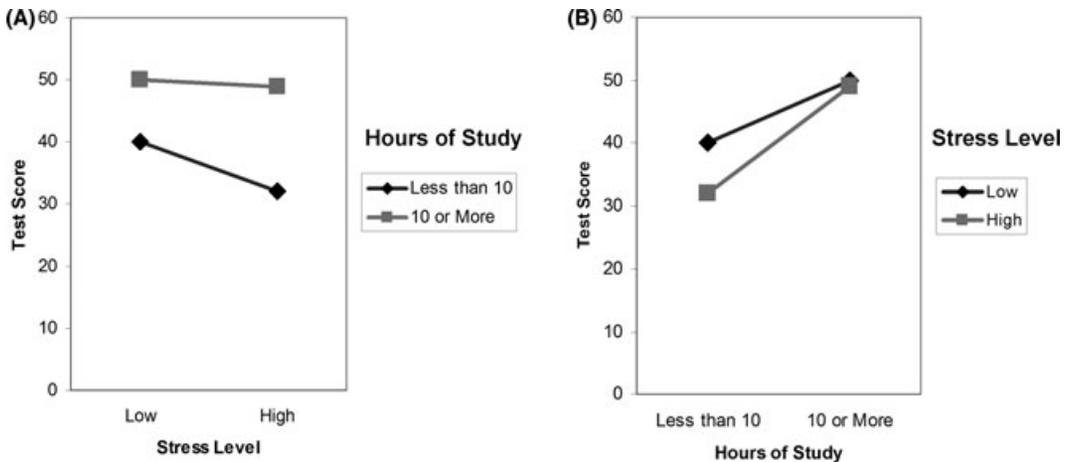


Fig. 6. Examples of different line graphs of the same fictitious data showing the relationship between stress, hours of study, and score on a test. The graphs differ in which variable is on the x axis and which is indicated by different lines. Shah and Carpenter (1995) found that when shown graphs like this, students were unable to tell that they showed the same data.

dimension they might have to place stickers for their least favorite food, a food they like, and their favorite food. Most children placed the stickers in a line that preserved the relationships, indicating that they naturally mapped more abstract relations to space. They mapped spatial and temporal dimensions to space at an earlier age than they mapped quantitative and preference dimensions and their mappings were affected by writing order in their cultures (see examples in Fig. 7). Spontaneous depictions reveal natural mappings between meaning and space that can be capitalized in the design of visual displays.

With the development of user-friendly eye trackers, eye fixations are increasingly being used to inform the design of visual displays by cognitive scientists (e.g., Carpenter & Shah, 1998; Peebles & Cheng, 2003; Ratwani, Trafton, & Boehm-Davis, 2008) as well as in related domains such as education (van Gog & Scheiter, 2010) and cartography (Fabrikant et al., 2010). Eye fixations can be interpreted as a measure of overt visual attention (cf. Henderson & Ferreira, 2004). While reaction times provide information about the general efficiency of task performance with a display, eye fixations can provide more diagnostic information, for example, identify areas of a display that attract attention although they are not task relevant. Observing the eye fixations of more expert or more successful users of a display may also lead to the design of displays that direct less successful users' attention to the task relevant information. For example, Grant and Spivey (2003) examined eye fixations on a diagram of Duncker's (1945) classic tumor problem, while people solved this problem and found that successful problem solvers made more eye-fixations on the outline of the

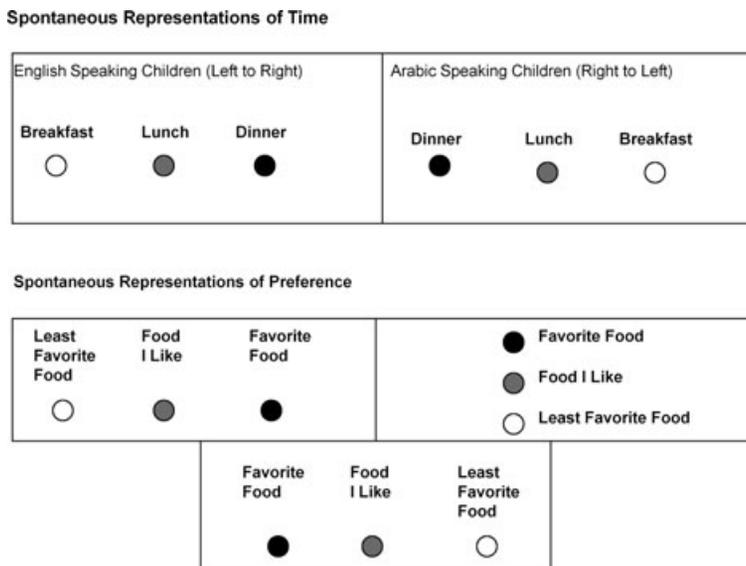


Fig. 7. Examples of configurations spontaneously produced by students when they were asked to place stickers representing temporal and preference dimension by Tversky et al. (1991). The top panel shows that children naturally mapped time to the horizontal dimension, with the dominant direction influenced by the order of writing in their cultures. The bottom panel shows that they naturally mapped preference dimensions to space, using both horizontal and vertical dimensions; when the vertical dimension was used, preference was from top to bottom.

body (the skin). They then redesigned the display to make the skin more visually salient by animating it, and thus improved problem solving with the display.

Finally, with the increased availability of interactive displays, methods that log users' interactions with these displays are increasingly important (Robertson et al., 2009). Interaction logs indicate the extent to which people use the different functions afforded by interactive displays and can be related to measures of performance to reveal which interactions are most effective. For example, Keehner, Hegarty, Cohen, Khooshabeh, and Montello (2008) examined use of an interactive visualization to perform a task that involved imagining the cross section of a three-dimensional object. Participants were provided with a computer model of the object that could be rotated in any direction using an intuitive 3-degrees-of-freedom interface. Interaction logs indicated that the most common user interaction was to rotate the model to view the object from a perspective perpendicular to the cross section to be imagined. Those who used the interactive models in this way had better task performance, but many participants did not use the models in this way. This study makes it clear that just providing people with an interactive visual display does not ensure that they will use it effectively.

### *5.3. The importance of cognitive models*

While empirical methods provide objective measures of the effectiveness of different visual displays, they are time consuming and expensive to carry out. Besides empirical investigations, an important contribution of cognitive science and related disciplines is the development of models that can be used to predict the effectiveness of visual displays a priori and inform the design of new displays. These include cognitive task analyses, computational models, and more recently, models of visual salience.

### *5.4. How cognitive models inform display design*

Some cognitive science and human factors researchers make prescriptions about the design of displays on the basis of task analyses and knowledge of perception and cognition. In a classic task analysis, Cleveland first analyzed the basic perceptual tasks that had to be carried out to encode the information in different common kinds of statistical graphs, such as pie charts, bar charts, scatter plots, etc. (Cleveland, 1985; Cleveland & McGill, 1984). For example, perception of angles is necessary to understand pie charts, perception of position along a common axis is necessary to understand bar charts, and position along non-aligned scales is necessary to compare corresponding elements in stacked bar charts (see Fig. 5). On the basis of psychophysics research and their own empirical studies, Cleveland and colleagues ordered the basic perceptual tasks in terms of accuracy. Perceiving position along a common scale was judged as the most accurate, followed by position along nonaligned scales, comparisons of line lengths, angles, areas, and volumes in that order. The ordering of the necessary perceptual tasks was used to predict the effectiveness of different types of graphs, for example, that bar charts would be more effective than pie charts for presenting relative magnitudes because position along a common scale is a more accurate

perceptual judgment than is angle. A meta-analysis by Carswell (1992), in addition to Cleveland's own research, provided good support for the model, when the graph comprehension tasks involved extracting of specific data points and local comparisons (e.g., comparing the proportions for Europe and Africa in Fig. 5), although Carswell suggested that the model was less effective in explaining performance for tasks that involved making global comparison and synthesis judgments (e.g., comparing combinations of data points or judging the general variability of the data points in the graph).

Other task analyses provide models of the elementary perceptual and cognitive processes necessary to carry out various data interpretation tasks with different types of displays (e.g., Gillan & Callahan, 2000; Gillan & Lewis, 1994; Hollands & Spence, 1998; Lohse, 1993; Simkin & Hastie, 1986; Spence & Lewandowsky, 1991). Elementary processes might include visual search to find an element in a display, scanning to estimate the distance between two components, and mental superimposition to compare the size of two components. The number of basic processes (and estimates of their duration) is then used to predict the efficiency of carrying out different graph comprehension tasks with various types of displays, with the assumption that displays that minimize the number of basic processes will be more efficient. These models have been quite successful in predicting the efficiency of specific graph comprehension tasks. For example, Gillan and Lewis (1994) found that a simple componential model accounted for up to 85% of individuals' response times to answer different questions (identifying single values, comparing values, and calculating values) from common graph types (line graphs, scatter plots, and stacked bar graphs). Models can also guide the design of new displays. For example, on the basis of a task analysis, Gillan and Callahan (2000) redesigned the pie graph to create a new format, the aligned pie graph (see Fig. 5D), which proved to be more efficient for the specific task of comparing proportions. They argued that this comparison involves both mental and rotation and superimposition of the two elements with the standard bar chart (Fig. 5A) but only superimposition for the aligned pie graph in Fig. 5D.

Task analytic models have also been used to develop computational models that predict the sequence of eye fixations that a person will make while answering a question from a visual display as well as response time (Lohse, 1993; Peebles & Cheng, 2003). For example, Peebles and Cheng (2003) developed production system models of optimal scan paths for reading values from different types of graphs and evaluated these models using both reaction time and eye fixation data. The model accounted for 87% and 66% of the variance in reaction times for two different graph formats and demonstrated that a less familiar graph (parametric graph) that is better tuned to the task requirements can be more effective than a more familiar type (function graph). Similarly, Trafton and colleagues (Breslow et al., 2009; Ratwani et al., 2008; Trafton et al., 2000) have used a combination of task analysis, cognitive modeling, eye-tracking, and verbal protocols to study how people extract information from geospatial displays, integrate information across different displays and variables, and to explain interactions between tasks and display format.

A relatively new modeling approach is to use general models of visual salience (e.g., Itti & Koch, 2000) or visual clutter (e.g., Lohrenz, Trafton, Beck, & Gendron, 2009; Rosenholtz, Li, & Nakano, 2007) to guide the design of displays. For example, the Itti and Koch

model uses information about how visual features (color, intensity, and orientation) are processed by the visual system to derive a “salience map” for any image, assuming that salient areas are those that are most different from their surrounding regions on the visual features. Fabrikant et al. (2010) used the Itti and Koch (2000) model in conjunction with an informal task analysis to redesign weather maps to make task-relevant information salient. An original weather map (downloaded from the Web) is shown in Fig. 8A and the task studied was to infer wind direction, which is based on pressure, so that pressure is task relevant and temperature is irrelevant. To redesign the map, Fabrikant et al. used cartographic principles (Bertin, 1983) to make the task-irrelevant temperature information less salient (by muting the colors showing temperature) and make the task-relevant pressure systems more salient. The resulting maps were tested by applying the salience model, and the redesign and test cycle was repeated until the arrow and pressure systems were identified as the most salient display regions by the model (Fig. 8B). Empirical testing indicated that people performed the inference task more efficiently (Fabrikant et al., 2010) and more accurately (Hegarty et al., 2010) with the redesigned maps.

In summary, empirical studies have made it clear that one should not rely on intuitions alone to judge the effectiveness of visual displays, as people’s intuitions about displays are not necessarily a good indication of their effectiveness. Cognitive scientists have had good success in characterizing the cognitive processes involved in performing tasks with visual displays and in developing cognitive models that can predict the relative effectiveness of different displays. However, to date most of this research has focused on relatively simple displays of quantitative data and on well-defined tasks, such as extracting specific values, comparing values, or detecting expected trends. These simple tasks contrast sharply with the types of tasks of interest to the visual analytics community where the goals are much broader, and ill defined, including data exploration, sense making, and reasoning with visualizations of complex data sets with thousands of data points (Thomas & Cook, 2005). Research to date therefore points to both the promise of cognitive science approaches and

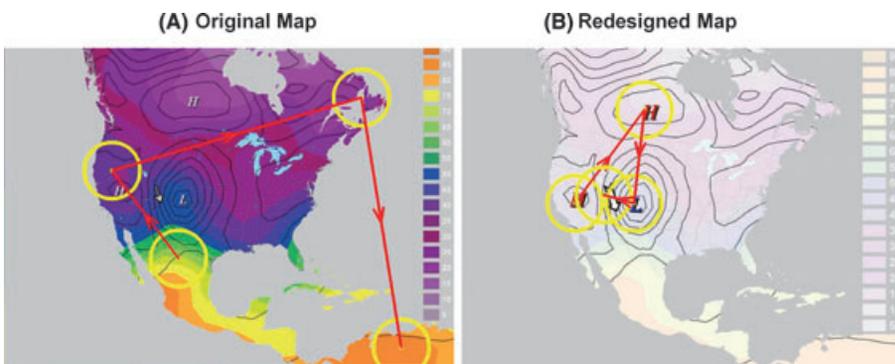


Fig. 8. Examples of original weather map and a redesigned weather map from the studies of Fabrikant et al. (2010) and Hegarty et al. (2010). In each case, the yellow circles show the four most salient regions of the map, according to the Itti and Koch (2000) salience model.

the challenges that lie ahead in scaling up cognitive approaches to the design of displays for more complex tasks.

## **6. Principles of effective graphics**

Although there has been relatively little research on performance of ill-defined tasks and more complex displays, research to date on both focused tasks with visual displays and on visual-spatial cognition more generally has produced some basic principles of effective visual displays that have received some empirical validation and can guide the design of both simple and complex visualizations. These are best considered as heuristics for the design of displays that need to be empirically tested, especially in more complex contexts. Many of them have been documented by Kosslyn (1989, 1993, 2006) in a series of articles and books about graph design (for excellent practical advice on the design of graphs see also Gillan, Wickens, Hollands, & Carswell, 1998). Here, I attempt to extend them to the design of visual-spatial displays more generally. They are summarized in Table 1.

### *6.1. Principles related to task specificity*

We start with the meta-principle that there is no such thing as a “best” visual-spatial display, independent of the task to be carried out with this display. As Liben (2001) put it, a map (or any other visual display) has a purpose as well as a referent: “it is not only of something, it is for something” (p. 50). Visual displays are used for many different purposes such as recording and storing information, serving as computational aids, data exploration, and conveying information to others (Liben, 2001; Tversky, 2001). Displays that are effective for one task may be ineffective for another. For example, while animation has not been shown to be effective for communicating about physical processes (Tversky et al., 2002), as reviewed earlier, animation does have advantages over static diagrams in showing transitions between different views of an information space and avoiding disorientation in complex visualizations (Robertson et al., 2009). Similarly, tables are better than graphs for extracting specific values, whereas graphs are better than tables for noticing trends (Gillan et al., 1998); and bar graphs are better than pie charts for simple judgments of proportion, whereas pie charts are more effective for more complex comparisons, as discussed earlier (Spence & Lewandowsky, 1991).

One version of this meta-principle is the proximity compatibility principle, proposed by Wickens and Carswell (1995). These researchers argued that tasks to be carried out with multivariate displays can be ordered from more integrative or high in proximity (e.g., when two or more variables have to be combined in the analysis) to more focused or low in proximity (e.g., when considering one variable requires others to be filtered out). They also argued that displays can be manipulated to make the visual variables that represent the different dimensions in the display high or low in proximity, including spatial proximity and other grouping variables (similarity in color, connecting lines, etc.). The proximity compatibility principle states that more integrative tasks are facilitated by displays that are high in

Table 1  
Principles of effective graphics

---

*Principles related to task specificity*

- There is no such thing as a “best” or “worst” display independently of the task to be performed.
- The Proximity Compatibility Principle: More integrated tasks are facilitated by displays that are high in display proximity; more focused tasks are facilitated by displays that are low in display proximity (Wickens & Carswell, 1995).

*Principles related to the expressiveness of displays*

- The Relevance Principle of Graphics should present no more or no less information than is needed by the user (Kosslyn, 2006).
- The Principle of Capacity Limitations: Displays should be designed to take account of limitations in working memory and attention (Kosslyn, 2006).
- Minimize the data-ink ratio within reason (Tufte, 2001).

*Principles related to the perception of displays*

- The Apprehension Principle: To be effective, a visual display has to be accurately perceived. Present animations at a speed that can be apprehended (Tversky et al., 2002). Use visual dimensions that are accurately judged (Cleveland & McGill, 1983; Wickens & Hollands, 2000).
- The Principle of Discriminability: Visual forms indicating a difference between two variables should differ by a large enough amount to be perceived as different (Kosslyn, 2006).
- The Principle of Perceptual Organization. Ensure that groupings based on Gestalt principles, and emergent features more generally, are compatible with the tasks to be carried out with a display (Kosslyn, 2006; Vicente, 2002).

*Principles related to the semantics of displays*

- The Principle of Compatibility: A visual display is easier to understand if its form is compatible with its meaning (Kosslyn, 2006).
- Choose natural mappings between graphic forms and their meanings and between spatial arrangements of these forms and their meanings (Tversky, this volume).
- Match the dimensions of the representing and represented variables in a display in terms of scales of measurement (Bertin, 1983; Mackinlay, 1986; Zhang, 1996).
- The Representational Epistemic approach: Encode the fundamental conceptual structure of the domain in the display (Cheng, 2002, this volume; Cheng & Barone, 2004, 2007).

*Principles related to the pragmatics of displays*

- The Principle of Salience: Design displays to make the most important thematic information salient (Bertin, 1983; Dent, 1999; Kosslyn, 2006).
- The Principle of Informative Changes: Avoid large changes across properties of a display that do not carry information (Kosslyn, 2006).

*Principles related to the usability of displays*

- The Principle of Appropriate Knowledge: Ensure that the viewer has the necessary knowledge to extract and interpret the information in the display (Kosslyn, 2006). For example, provide information regarding the conventions of the graphic form, in a legend, or caption.
  - The Principle of Visual Momentum: Code multiple displays consistently and provide visual aids to help users make referential connections between different displays and avoid disorientation in animated and interactive displays (Wickens & Hollands, 2000; Woods, 1984).
- 

display proximity, whereas more focused tasks are facilitated by displays that are low in display proximity.

Given the task-specific nature of display effectiveness, an important question for display designers is when they should create optimal displays for each specific task that a user will perform, and when they should create displays that, while not optimal for any specific task,

support a wider range of tasks. For example, in the case of the proximity compatibility principle, research has shown that there is more evidence for the positive effects of display proximity for integrated tasks than the negative effects of display proximity on more focused tasks. This suggests that more integrated displays may provide satisfactory (if not optimal) support for both integrated and focused tasks (Bennett & Flach, 1992; Wickens & Carswell, 1995).

A related question is whether to provide one optimal display of information or to give the viewer flexibility to select and modify different displays of the same data. The answer to this question depends on several factors such as the purpose of the display, the viewer's task, and his or her meta-representational competence. For example, a scientist designing a display of data for a scientific publication should choose the single display that best communicates the most important trends in the data (Gillan et al., 1998). In contrast, when designing displays for people who monitor complex systems, such as power plants, the designer may need to provide several displays and flexibility in the interface to represent the whole work domain and support a range of tasks, including more automatic skill-based behavior and more knowledge-based problem-solving behaviors (Smith, Bennett, & Stone, 2006; Vicente, 2002).

Finally, when it is necessary to present multiple displays of the same information, it is important to design the displays to aid users in making referential connections between the different displays and avoiding disorientation. Researchers have proposed a set of techniques known as visual momentum (Woods, 1984) to accomplish this goal. These include keeping display elements as consistent as possible across displays, using graceful transitions such as animation, zooming, or morphing to show the transitions, highlighting common landmarks or anchors in the different displays, and continuously displaying an overview of the natural or information space, with an indication of the location of the currently viewable information within that space (Wickens & Hollands, 2000; Woods, 1984).

## 6.2. Principles related to the expressiveness of displays

The next set of principles is related to the expressiveness of visual displays, that is, their information content. The *relevance* principle (Kosslyn, 2006) states that graphics should present no more or no less information than is needed by the user. Presenting all of the relevant information in the display relieves the user of the need to maintain a detailed representation of this information in working memory, whereas presenting too much information in the display leads to visual clutter or distraction by irrelevant information (Rosenholtz et al., 2007; Wickens & Carswell, 1995). Thus, the relevance principle is highly related to another of Kosslyn's principles, *the principle of capacity limitations*, which points out that graphics should be designed to take account of limitations in working memory and attention.

The relevance principle also echoes Tufte's (2001) principle of maximizing the data ink ratio in graphics (that is, the ratio of ink showing data to all of the ink used in printing the graphic. Tufte advocated deleting all non-data ink and all redundant data ink "within reason" (p. 96), including deleting background pictures that are often included in newspaper

graphics, deleting the lines and tick marks on the axes and deleting the bars and filler patterns in bar graphs (which he referred to as redundant coding). Gillan and Richman (1994) provided empirical evidence that increasing the data ink ratio in graphs improved accuracy and decreased response time, but their research also indicated that Tufte's principle was too simplistic and different types of non-data ink varied in their effects. For example, background pictures were generally disruptive, including the x and y axes was generally beneficial, and the effects of redundant coding in displaying the data (e.g., fill patterns for bars in bar graphs) were inconsistent and depended on the task and type of graph. This is an excellent example of how empirical testing of a designer's intuitions can refine a design principle.

### 6.3. Principles related to the perception of displays

The expressiveness of a representation is not a sufficient condition for effectiveness of that representation. To be effective, a visual display has to be accurately perceived. Tversky et al. (2002) refer to this as the apprehension principle of visual displays. They formulated this principle to explain why animations of physical processes are not more effective than static graphics that show key phases in the operation of these devices, as discussed earlier. The apprehension principle is also related to the general principle that one should use visual dimensions that are accurately perceived (Cleveland & McGill, 1983; see Table 1) and avoid using visual variables that lead to biased judgments (Wickens & Hollands, 2000). For example, Smallman and Cook (this volume, see also Smallman & St. John, 2005) highlight a perceptual error that is relevant to rendering geospatial displays in 3-D. They point out that 3-D displays are not effective for judging absolute distances because people misperceive distances in depth. In reality, widths decay into the scene in inverse proportion to distance, whereas depths decay faster, in proportion to the square of distance. However, people's judgments of distance suggest that they use a simplifying heuristic that the decay function is the same in both dimensions. This is true in viewing of the real 3-D world (e.g., Loomis & Knapp, 2003) and not just visual displays of the world, so that adding more depth cues (e.g., stereoscopic viewing) does not solve the problem.

Kosslyn's principle of discriminability (Kosslyn, 2006) is another example of the apprehension principle. This principle draws on basic research in psychophysics to make the point that visual forms indicating a difference between two variables need to differ by a large enough amount to be perceived as different. Finally another important perceptual principle is what Kosslyn refers to as the *Principle of perceptual organization*, that people automatically group elements of displays into units. This principle is based on the Gestalt principles of perceptual organization, which determine which elements of displays are grouped and can be compatible or incompatible with the tasks to be carried out with a display. For example, line graphs facilitate comparisons between the units plotted on the x axis (see Fig. 6) because the lines group data points as a function of this variable (reflecting the Gestalt principle of good continuation). As discussed earlier, grouping is just one example of an emergent property, and a more general principle, central to ecological interface design (Vicente, 2002) is to match the emergent properties of the display with the

higher-order relations between components in a complex system and the users' mental model.

#### 6.4. Principles related to the semantics of displays

Another class of principles relate to what Kosslyn (2006) refers to as the *principle of compatibility*, that a visual display is easier to understand if its form is compatible with its meaning. One aspect of this principle is choosing mappings between the representing and represented variables in a display that reflect common metaphors in our culture, for example, up is good and down is bad (Lakoff & Johnson, 1980) or common associations, for example, that lines indicate connections, circles indicate cyclic processes, and the horizontal dimension is naturally mapped to time (Tversky, this volume).

In the case of relational displays, an important principle emphasized by several theorists (Bertin, 1983; Mackinlay, 1986; Zhang, 1996) is matching the dimensions of the visual variables with the underlying variables that they represent in terms of scales of measurement. Both representing and represented dimensions can vary in scale from categorical to interval, ordinal, or ratio (Stevens, 1946). For example, shape is a categorical variable, shading is an ordinal dimension, orientation is an interval dimension, and length is a ratio dimension. Zhang (1996) proposed that representations are most *accurate* and *efficient* when the scale of the representing variable corresponds to the scale of the represented variable. *Efficiency* refers to the fact that the relevant information can be perceived because it is represented in the external display. For example, the use of line length to represent the size of a computer file is an efficient representation because both line length and size are ratio variables. If a less powerful representing variable (e.g., shape, a categorical variable) is used to represent size, not all of the information is represented externally, so information about the correspondence between specific shapes and the sizes they represent must be internally represented. The match, in terms of scales of measurement, between the representing and represented variables was a central principle coded by Mackinlay (1986) in a system that automated the design of relational graphics.

The representational epistemic approach (Cheng, 2002; Cheng & Barone, 2004, 2007) emphasizes the importance of encoding the fundamental conceptual structure of a domain in the visual display. Cheng (this volume) argues that when displays express the fundamental conceptual structure of a domain, semantic transparency automatically results. Other researchers have emphasized matching the structure of the external representations with the user's representation of a data structure or problem. For example, Durning, Becker, and Gould (1977) found that college students were generally good at matching a list of words with an inherent structure (list, hierarchy, network, or matrix) to the appropriate spatial representation and in research on problem solving, Novick and Hurley (2001) found that college students were able to match the structure of a problem to a type of diagram (a matrix, a network, or a hierarchy). Novick and Hurley argued that people have schemas that include applicability conditions for the different

types of diagrams. Matching the display of information to these natural schemas has obvious advantages.

### 6.5. Pragmatic principles

As in language, pragmatics of graphics refers to the broader context in which visual displays communicate and their rhetorical function. One general pragmatic principle, the principle of salience, can be summarized by the general prescription that displays should be designed to make the most important thematic information salient (Bertin, 1983; Dent, 1999; Kosslyn, 2006). The converse of this principle, the principle of informative changes (Kosslyn, 2006), is that people expect changes across properties of a display to carry information. The relevance principle, described in the section on expressiveness of displays, can also be categorized as a pragmatic principle.

More broadly, the ways in which information is visually displayed can subtly communicate information. People have more confidence in geospatial data that are presented in realistic displays (Fabrikant & Boughman, 2006), a factor which probably underlies naïve realism (Smallman & St. John, 2005). The proportion of space taken up by graphs in journal articles varies across the sciences, with more space devoted to graphs in disciplines rated as “hard sciences” (Cleveland, 1984). The same is true of subdisciplines within psychology (Smith, Best, Stubbs, Archibald, & Roberson-Nay, 2002) and the quality of scientific reasoning in neuroscience articles is rated as higher if the article includes brain images rather than bar graphs or no graphics (McCabe & Castel, 2008). Thus, how data are displayed might give an impression of their reliability or scientific nature that may or may not be warranted.

### 6.6. Usability principles

The final set of cognitive principles relate to the usability of visual displays, where usability refers to ensuring that the viewer has the necessary knowledge to extract and interpret the information in the display. Visual displays are highly conventional representations, so at a minimum, users need to know the conventions of a particular graphic form in order to comprehend it. Kosslyn (2006) refers to this as the *principle of appropriate knowledge*; communication requires prior knowledge of relevant concepts, jargon, and symbols. This knowledge is often thought of as being part of the display schema (Pinker, 1990; Ratwani & Trafton, 2008).

The conventions of a visual display are often provided in a legend and in cartography a legend is considered to be an obligatory component of every map (Dent, 1999). But providing a legend is not sufficient for understanding a visual display. Displays often depend on implicit conventions which become evident only when we find people that do not know them. For example, children often confuse referential and incidental features of maps, believing that red lines on a map represent roads that are really red (Liben, 2001). This mistake would never be made by adults, but it exemplifies the type of implicit convention that might be adopted by an expert in any domain and not evident to a novice.

Furthermore, knowledge of the conventions of a display is not sufficient information for fully interpreting a visual display, so that novices do not encode the same information or make the same inferences from a display compared to experts, even when they are very familiar with the conventions (e.g., Lowe, 1996; Myles-Worsley, Johnston, & Simons, 1988).

With the advent of interactive displays, usability becomes even more critical. Interactive displays now give users the opportunity to choose the representation they will use to perform a task, add and subtract variables, rotate, pan, and zoom displays. This type of interactivity is central to the emerging field of visual analytics (Thomas & Cook, 2005; Fisher, this volume). It is often assumed that this interactivity will enhance performance. However, it is important to realize that interactive displays put the burden of display choice and design on the user. Choosing displays depends on meta-knowledge of which specific displays are optimal for specific tasks. Interactivity can have costs. For example, Yeh and Wickens (2001) found that the time and attention costs needed to decide to subtract irrelevant variables from a display outweighed the performance decrement due to visual clutter. Furthermore, as discussed earlier, users do not always discover or use the interactive functions provided (Keehner et al., 2008).

Finally, disorientation is a common problem in displays of complex data, typical of information visualizations and scientific visualizations. In these displays, there is not enough space in a single screen to show all the data, so that when users zoom in to view some part of the data in detail, they may lose track of their location in the information space. Researchers in information visualization have invented a variety of different techniques to allow users to focus in on items of interest, while also maintaining their orientation in the whole information space (for reviews, see Card et al., 1999; Robertson et al., 2009). For example, fisheye views (Furnas, 1986) filter or distort views so that the most important information is rendered as larger and less important information is rendered as smaller or omitted. Visual momentum techniques such as using graceful transitions and continuously displaying an overview of the natural or information space can also be important in alleviating problems of disorientation in complex scientific and information visualizations (Wickens & Hollands, 2000).

## **7. Introduction to the other papers in this topic**

The authors of the papers in this topic were invited to contribute papers that addressed guidelines for the design of visual displays that were based on empirical research. They represent a variety of approaches. Cheng's article emphasizes the expressiveness of visual displays, focusing on epistemic rather than cognitive principles for the design of displays. He argues that if you model the fundamental conceptual structure of a knowledge domain directly in the representational system, semantic transparency (ease of comprehension and learning) and "plastic generativity" (ease of manipulation and inference) automatically result. In this paper, he enumerates his epistemic design principles, illustrates the use of these principles to create novel displays in the domain of probability theory, and validates

his approach in a new experiment that shows that the novel displays are superior to currently used representations in this domain.

Tversky's paper is centrally concerned with the semantics of displays, that is, the relationship between the form of a display and its meaning. She argues that visual displays are cultural artifacts that reflect natural ways of using visual forms and spatial arrangements to communicate meaning or "visualize thought" and which work because they have been "user tested" in natural use. She draws from a very broad literature in making her case, including research on how people naturally communicate with visual forms and spatial arrangements, in gesture as well as more permanent marks on a page, and also incorporates insights from the history of art and design. The take-home message of her paper is that visual displays that conform to the many natural mappings that she has identified will ultimately be more effective.

The article by Novick et al. provides new evidence for the influence of the type of natural mappings discussed by Tversky, and how a mismatch between the format of a diagram and the meaning it conveys can reinforce students' misconceptions rather than promoting accurate understanding. Several parallels can also be found between Novick et al.'s paper and that of Shah and Friedman, which examines interpretations of line versus bar graphs showing multivariate data. Both papers continue the tradition of empirical studies examining the relative effectiveness of different displays of the same data. Both are concerned with perceptual grouping effects and provide evidence for how Gestalt principles influence what is salient in a display and both papers make specific prescriptions of when to use different display formats. They also demonstrate the general principle that the design of visual displays must take the knowledge and expectations of the user into account, because interpretation of a visual display involves an interaction between bottom-up and top-down processes. Both papers challenge a commonly held idea that knowledge makes the viewer immune to the effects of graph format—if anything, these papers indicate larger effects of graph format for more knowledgeable individuals, perhaps because knowledge and familiarity is necessary to take full advantage of specific graph formats.

The papers by Smallman and Cook and by Fisher address current challenges in display design raised by the recent explosion in availability of complex data sets and new technologies for visualizing and interacting with these data. Smallman and Cook argue against a naïve view that progress in graphic design is made by packing more and more information into increasingly realistic displays. They point to limitations both in the perception of visual displays and users' metacognition about displays, as well as describing new methodologies for measuring users' intuitions. Specifically, they show that the optimal displays for different terrain understanding tasks are not the most detailed or realistic, and that people generally prefer to use more realistic displays than are optimal for task performance. In conclusion, they question the wisdom of giving users the power to configure their own displays without guidance, given that intuitions about effective displays are not always calibrated to their actual effectiveness.

Fisher's paper highlights developments in the new field of visual analytics, which is premised on the availability of complex data sets, new technologies for interacting with these data sets, and new data mining problems in such areas as scientific research and emergency

management. There is currently a disconnect between the relatively simple visual displays, data sets, and tasks that have been studied by cognitive scientists and the complex data sets, interactive visualizations, and high-level data mining activities carried out in the field of visual analytics. Because of this disconnect, interactive systems for finding patterns in data are currently being designed without insights from cognitive science. Fisher proposes a “translational cognitive science” in which cognitive scientists study how interactive visual displays are used for the types of truly complex tasks being carried out in the real world. He proposes that this effort will require the adoption of new methods for studying how people interact, perceive, and reason with visual displays, as well as a closer connection between research and practice.

## 8. Conclusions

Although cognitive scientists have made important contributions that inform the design of visual displays, many challenges remain. We can currently say more about the design of data graphs than about other types of visual displays, because graphs have received more attention by cognitive scientists. While cognitive scientists have had good success in characterizing the cognitive processes and improving the design of displays for focused, well-defined tasks, it is not clear how these methods will scale up to more complex and ill-defined tasks such as data exploration and more complex displays being developed by the information visualization and scientific visualization communities. There are open questions about when we should design different optimal displays for specific tasks and when we should design a single display that can support many tasks. We know very little about whether and how people use the interactive functions provided by new visualization technologies. While people often have strong intuitions about the effectiveness of displays, these are not always in line with actual effectiveness, so there is a clear need to evaluate display designs empirically and continue to develop cognitive models that will allow us to predict the effectiveness of displays a priori. We hope that this topic inspires both better display design and more research on cognitive science of visual displays to address the many remaining challenges of research on visual-spatial displays.

## References

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*, 183–198.
- Bennett, K. B., & Flach, J. M. (1992). Graphical displays: Implications for divided attention, focused attention, and problem solving. *Human Factors, 34*, 513–533.
- Bertin, J. (1983). *Semiology of graphics: Diagrams networks maps* (W. Berg, Trans.). Madison: University of Wisconsin Press.
- Breslow, L. A., Trafton, G. J., & Ratwani, R. M. (2009). A perceptual process approach to selecting color scales for complex visualizations. *Journal of Experimental Psychology: Applied, 15*, 25–34.
- Brown, L. (1979). *The story of maps*. New York: Dover.

- Card, S. K., Mackinlay, J. D., & Schneiderman, B. (1999). *Readings in information visualization: Using vision to think*. San Francisco: Morgan Kaufmann Publishers.
- 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, M. (1992). Choosing specifiers: An evaluation of the basic tasks model of graphical perception. *Human Factors*, 34, 535–554.
- Cheng, P. C.-H. (2002). Electrifying diagrams for learning: Principles for effective representational systems. *Cognitive Science*, 26, 685–736.
- Cheng, P. C.-H., & Barone, R. (2004). Representing rosters: Conceptual integration counteracts visual complexity. In A. Blackwell, K. Mariott, & A. Shimojima (Eds.), *Diagrammatic representation and inference: Third international diagrams conference* (pp. 385–387). Berlin: Springer-Verlag.
- Cheng, P. C.-H., & Barone, R. (2007). Representing complex problems: A representational epistemic approach. In D. H. Johanssen (Ed.), *Learning to solve complex scientific problems* (pp. 97–130). Mahwah, NJ: Erlbaum.
- Cleveland, W. S. (1984). Graphs in scientific publications. *American Statistician*, 38, 261–269.
- Cleveland, W. S. (1985). *The elements of graphing data*. Monterey, CA: Wadsworth.
- Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation and application to the development of graphical methods. *Journal of the American Statistical Association*, 79, 531–554.
- Dent, B. D. (1999). *Cartography: Thematic map design*. Boston, MA: McGraw-Hill.
- DiSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction*, 22, 293–331.
- Duncker, K. (1945). On problem solving. *Psychological Monographs*, 58, (Whole No 270).
- Durding, B. M., Becker, C. A., & Gould, J. D. (1977). Data organization. *Human Factors*, 19, 1–14.
- Fabrikant, S. I., & Boughman, A. (2006). Communicating data quality through realism. *Proceedings of GIScience*, Münster, Germany, September 20–23.
- Fabrikant, S. I., & Lobben, A. (2009). Guest editor's introduction. Special issue: Cognitive issues in geographic information visualization. *Cartographica*, 44, 129–144.
- Fabrikant, S. I., Rebich-Hespanha, S., & Hegarty, M. (2010). Cognitively inspired and perceptually salient graphic displays for efficient inference making. *Annals of the Association of American Geographers*, 100, 13–29.
- Ferguson, E. S. (2001). *Engineering and the mind's eye*. Cambridge, MA: MIT Press.
- Freedman, E. G., & Shah, P. (2002). Toward a model of knowledge-based graph comprehension. In M. Hegarty, B. Meyer, & N. H. Naryanan (Eds.), *Diagrammatic representation and inference* (pp. 8–31). Berlin: Springer-Verlag.
- Furnas, G. (1986). Generalized fisheye views. In M. Mantei & P. Orbeton (Eds.), *Proceedings of the 1986 SIGCHI Conference on Human Factors in Computing Systems* (pp. 16–23). New York: Association for Computing Machinery.
- Gattis, M., & Holyoak, K. J. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 22, 231–239.
- Gillan, D. J., & Callahan, A. B. (2000). A componential model of human interaction with graphs: VI. Cognitive engineering of pie graphs. *Human Factors*, 42, 566–591.
- Gillan, D. J., & Lewis, R. (1994). A componential model of human interaction with graphs: 1. Linear regression modeling. *Human Factors*, 36, 419–440.
- 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.
- van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, 20, 95–99.
- Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science*, 14(5), 462–466.

- Hegarty, M., Canham, M., & Fabrikant, S. I. (2010). Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *36*, 37–53.
- Henderson, J. M., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision and action* (pp. 1–58). New York: Psychology Press.
- Hollands, J. G., & Spence, I. (1998). Judging proportion with graphs: The summation model. *Applied Cognitive Psychology*, *12*, 173–190.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506.
- Keehner, M., Hegarty, M., Cohen, C. A., Khooshabeh, P., & Montello, D. R. (2008). Spatial reasoning with external visualizations: What matters is what you see, not whether you interact. *Cognitive Science*, *32*, 1099–1132.
- Kirsh, D. (1997). Interactivity and multimedia interfaces. *Instructional Science*, *25*, 76–96.
- Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*, 513–549.
- Kosslyn, S. M. (1989). Understanding charts & graphs. *Applied Cognitive Psychology*, *3*, 185–226.
- Kosslyn, S. M. (1994). *Elements of graph design*. New York: W. H. Freeman.
- Kosslyn, S. M. (2006). *Graph design for the eye and mind*. New York: Oxford University Press.
- Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies*, *65*, 911–930.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.
- Larkin, J., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*, 65–99.
- Lewandowsky, S., & Behrens, J. T. (1999). Statistical graphs and maps. In F. T. Durso, R. S. Nickerson, R. W. Schvaneveldt, S. T. Dumais, D. S. Lindsay, & M. T. H. Chi (Eds.), *Handbook of applied cognition* (pp. 513–549). Hoboken, NJ: Wiley.
- Liben, L. S. (2001). Thinking through maps. In M. Gattis (Ed.), *Spatial schemas and abstract thought* (pp. 45–77). Cambridge, MA: MIT Press.
- Lohrenz, M. C., Trafton, J. G., Beck, M. R., & Gendron, M. L. (2009). A model of clutter for complex, multivariate, geospatial displays. *Human Factors*, *51*, 90–101.
- Lohse, G. L. (1993). A cognitive model for understanding graphical perception. *Human Computer Interaction*, *8*, 353–388.
- Loomis, J. M., & Knapp, J. M. (2003). Visual perception of egocentric distance in real and virtual environments. In L. J. Hettinger & M. W. Hass (Eds.), *Virtual and adaptive environments* (pp. 21–46). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lowe, R. K. (1996). Background knowledge and the construction of a situational representation from a diagram. *European Journal of Psychology of Education*, *11*, 377–397.
- MacEachren, A. (1995). *How maps work. Representation, visualization and design*. New York: The Guilford Press.
- MacEachren, A. M., & Kraak, M.-J. (2000). Research challenges in geovisualization. *Cartography and Geographic Information Science*, *28*, 3–12.
- Mackinlay, J. D. (1986). Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics*, *5*, 110–141.
- Mayer, R. E., Hegarty, M., Mayer, S. Y., & Campbell, J. (2005). When passive media promote active learning: Static diagrams versus animation in multimedia instruction. *Journal of Experimental Psychology: Applied*, *11*, 256–265.
- McCabe, D. P., & Castel, A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition*, *107*, 343–352.
- McCormick, B. H., DeFanti, T., & Brown, M. D. (1987). *Visualization and scientific computing*. Arlington, VA: National Science Foundation.

- Montello, D. R. (1993). Scale and multiple psychologies of space. In A. U. Frank & I. Campari (Eds.), *Spatial information theory: A theoretical basis for GIS* (pp. 312–321). Proceedings of COSIT '93. Berlin: Springer-Verlag, Lecture notes in Computer Science 716.
- Myles-Worsley, M., Johnston, W. A., & Simons, M. A. (1988). The influence of expertise on X-ray image processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 553–557.
- Norman, D. (1994). *Things that make us smart*. New York: Addison-Wesley.
- Novick, L. R., & Catley, K. M. (2007). Understanding phylogenies in biology: The influence of a Gestalt perceptual principle. *Journal of Experimental Psychology: Applied*, 13, 197–223.
- Novick, L. R., & Hurley, S. M. (2001). To matrix, network, or hierarchy: That is the question. *Cognitive Psychology*, 42, 158–216.
- Palmer, S. E. (1978). Fundamental aspects of cognitive representations. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 259–303). Mahwah, NJ: Lawrence Erlbaum Associates.
- Panofsky, E. (1960). *Renaissance and renaissances in western art*. Stockholm: Almqvist & Wiksell.
- Peebles, D., & Cheng, P. C.-H. (2003). Modeling the effect of task and graphical representation on response latency in a graph-reading task. *Human Factors*, 45, 28–46.
- Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp. 73–126). Hillsdale, NJ: Erlbaum.
- Pomerantz, J. R., & Pristach, E. A. (1989). Emergent features, attention, and perceptual glue in visual form perception. *Journal of Experimental Psychology: Human Perception and Performance*, 15(4), 635–649.
- Pylyshyn, Z. W. (2003). *Seeing and visualizing: It's not what you think*. Cambridge, MA: MIT Press.
- Ratwani, R. M., & Trafton, J. G. (2008). Shedding light on the graph schema: Perceptual features versus invariant structure. *Psychonomic Bulletin & Review*, 15, 757–762.
- Ratwani, R. M., Trafton, J. G., & Boehm-Davis, D. A. (2008). Thinking graphically: Connecting vision and cognition during graph comprehension. *Journal of Experimental Psychology: Applied*, 14, 36–49.
- Robertson, G., Czerwinski, M., Fisher, D., & Lee, B. (2009). Selected human factors issues in information visualization. *Reviews of Human Factors and Engineering*, 5, 41–81.
- Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision*, 7, 1–22.
- Sanfey, A., & Hastie, R. (1998). Does evidence presentation format affect judgment? An experimental evaluation of displays of data for judgments. *Psychological Science*, 9, 99–103.
- Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies*, 45, 115–143.
- Schneiderman, B. (1994). Dynamic queries for visual information seeking. *IEEE Software*, 11, 70–77.
- Schwartz, D. L. (1995). Reasoning about the referent of a picture versus reasoning about the picture as the referent: An effect of visual realism. *Memory & Cognition*, 23, 709–722.
- Shah, P., & Carpenter, P. (1995). Conceptual limitations in comprehending line graphs. *Journal of Experimental Psychology: General*, 124, 337–370.
- 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 visual spatial thinking* (pp. 426–476). New York: Cambridge University Press.
- Simkin, D. K., & Hastie, R. (1987). An information-processing analysis of graph perception. *Journal of the American Statistical Association*, 82, 454–465.
- Smallman, H. S., & St. John, M. (2005). Naïve realism: Misplaced faith in realistic displays. *Ergonomics in Design*, 13, 14–19.
- Smith, P. J., Bennett, K. B., & Stone, B. R. (2006). Representation aiding to support performance on problem solving tasks. *Reviews of Human Factors and Ergonomics*, 2, 74–108.
- Smith, L. D., Best, L. A., Stubbs, A. D., Archibald, A. B., & Roberson-Nay, R. (2002). Constructing knowledge: The role of graphs and tables in hard and soft psychology. *American Psychologist*, 57, 749–761.
- Spence, I., & Lewandowsky, S. (1991). Displaying proportions and percentages. *Applied Cognitive Psychology*, 5, 61–77.

- Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science*, 19, 97–140.
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103, 677–680.
- Thomas, J. J., & Cook, K. A. (2005). *Illuminating the path: Research and development agenda for visual analytics*. Richland, WA: IEEE Press.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46, 234–240.
- Trafton, J. G., Kirschenbaum, S. S., Tsui, T. L., Miyamoto, R. T., Ballas, J. A., & Raymond, P. D. (2000). Turning pictures into numbers: Extracting and generating information from complex visualizations. *International Journal of Human-Computer Studies*, 53, 827–850.
- Tufte, E. T. (2001). *The visual display of quantitative information* (2nd ed.). Cheshire, CT: Graphics Press.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.
- Tversky, B. (2001). Spatial schemas in depictions. In M. Gattis (Ed.), *Spatial schemas and abstract thought* (pp. 79–112). Cambridge, MA: MIT Press.
- Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology*, 23, 515–557.
- Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies*, 57, 247–262.
- Vicente, K. (2002). Ecological interface design: Progress and challenges. *Human Factors*, 44, 62–78.
- Wainer, H. (2005). *Graphic discovery: A trout in the milk and other visual adventures*. Princeton, NJ: Princeton University Press.
- White, B. Y. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction*, 10, 1–100.
- Wickens, C. D., & Carswell, M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors*, 37, 473–494.
- Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance*. Upper Saddle River, NJ: Prentice Hall Inc.
- Woods, D. D. (1984). Visual momentum: A concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies*, 21, 229–244.
- Wu, H.-K., Krajcik, J. S., & Soloway, E. (2001). Promoting understanding of chemical representations: Students' use of a visualization tool in the classroom. *Journal of Research in Science Teaching*, 38(7), 821–842.
- Yeh, M., & Wickens, C. D. (2001). Attentional filtering in the design of electronic map displays: A comparison of color coding, intensity coding, and decluttering techniques. *Human Factors*, 43, 543–562.
- Zhang, J. (1996). A representational analysis of relational information displays. *International Journal of Human Computer Studies*, 45, 59–74.
- Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, 18, 87–122.