Aviation Human Factors Division
Institute of Aviation

University of Illinois
at Urbana-Champaign
1 Airport Road
Savoy, Illinois 61874

Three-Dimensional Displays as an Effective Visualization Technique for Power Systems Monitoring and Control

Stephan M. Hoppe, Gavin R. Essenberg,
Douglas A. Wiegmann
(Aviation Human Factors Division)

and Thomas J. Overbye
(Electrical and Computer Engineering)

Technical Report
AHFD-04-07/PSERC-04-1

April 2004

Contract PSERC CORNELL WEI

Three-Dimensional Displays as an Effective Visualization Technique
for Power Systems Monitoring And Control

ABSTRACT

Due to recent reconfiguration of the interconnected North American electric power grid, operators must now comprehend a vast and growing amount of multivariate data over more extensive network regions. Power systems engineers and designers have thus begun to develop visualization tools to aid these operators in managing such regions. In this study, we investigated the use of three-dimensionality as a potential display enhancement to support the tasks of fault detection and diagnosis across an integrated one-line diagram. The participants were first required to acknowledge all overloaded transmission lines for a given scenario and then solve the scenario by reducing all line loadings to 100% of capacity or less. Participants performed the tasks using one of three display types: 2-D numerical, 2-D graphical, and 3-D. Although the display type did not have a significant effect on acknowledgment time, a 3-D advantage was apparent for solution time. Accuracy results indicated a significant 3-D advantage for the number of attempted output increases of maxed-out generators. However, no significant effects of display type were found on the number of erroneous generator output adjustments during the solution task. More difficult tasks that require both focused attention and parallel processing may further reveal whether three-dimensional displays can indeed enhance accuracy. Workload was not affected by display dimensionality. Overall, our results support the claim that three-dimensional displays can improve performance in tasks that require monitoring and controlling variables of complex systems.

INTRODUCTION

The interconnected North American electric power grid is “one of the largest and most complex man-made objects ever built” and consists of “billions of individual components, tens of millions of miles of wires and thousands of individual generators” (Overbye, 2000, p. 220). In a massive system such as this, operators must comprehend a vast and growing amount of multivariate data across an extensive network (Overbye, Klump, and Weber, 1999; Rich, Essenberg, Wiegmann, Overbye, and Sun, 2003). Furthermore, deregulation and the subsequent reconfiguration of the power system industry to permit non-utility generators, marketers, brokers, and load aggregators to share the transmission system has caused the size of network regions controlled by individual operators to increase (Overbye, Weber, and Laufenberg, 1999; Overbye, 2000). Power transactions are no longer confined to those between adjacent utilities, and the actions of operators now influence not only the utility companies who employ them but also market participants across the entire continent (Overbye, Sun, Wiegmann, and Rich, 2002). As evidenced by the blackout of August 14, 2003, problems can rapidly propagate across this complex grid and lead to quick information overload for system operators (Mahadev and Christie, 1994a). In order to aid these operators in the acknowledgement, interpretation, and application of this often overwhelming volume of data, new graphical visualization techniques have and must continue to be developed.

Specifically, the role of three-dimensional displays to aid power system operators in detecting potential problems and identifying solutions quickly and accurately has been explored with various visualization platforms (Klump, Schooley, and Overbye, 2002; Overbye, Weber, and Laufenberg, 1999; Overbye, Wiegmann, and Thomas, 2002; Overbye, Klump, and Weber, 1999). Unfortunately, though, few empirical investigations have been conducted to identify the tradeoffs involved with using these and other 3-D visualization tools to analyze power system flow and capacity. Therefore, in this study we examine the potential advantages and disadvantages of providing operators with three-dimensional displays of power systems data and compare the effectiveness of these displays to that of their two-dimensional counterparts.

Before addressing the topics specific to the present investigation, however, the general characteristics of process control and the specific power system operator tasks of monitoring displays and detecting, diagnosing, and resolving faults must be addressed. Following this, descriptions of the displays traditionally used to support these tasks, tabular displays and one-line diagrams, are provided. We then discuss various modern power system displays and enhancements that have been investigated empirically in limited controlled settings. A thorough review of three-dimensional visualization techniques and applications follows, coupled with a description of the proposed study and predicted results.

Process Control Characteristics

Tasks performed by power systems operators are typical of those in the process control domain. Process control tasks are characterized by four general characteristics (Wickens and Hollands, 2000). First, the variables being monitored and controlled by the operators are slow. As a result, the system’s response to a command delivered by an operator may be significantly delayed. Second, operators are often controlling these variables, typically analog and continuous in nature, with discrete adjustments. It is thus imperative for an operator’s mental model of the system to account for this control incongruence (Hollan, Hutchins, and Weitzman, 1984). Third, the quantity of interrelated parameters and variables is large. This complexity, coupled with ineffective data representation, can often lead to information overload (Mahadev and Christie, 1994a) and incorrect system manipulation. Finally, processes in this domain often involve a high level of risk, with serious economic and safety ramifications associated with failures.

The specific tasks performed by power systems operators involve three main task categories: monitoring, fault detection, and fault diagnosis and resolution. These are each discussed in detail below.

Power systems operator tasks: Monitoring. Monitoring in power systems applications, often referred to as a vigilance task (Wickens and Hollands, 2000), involves visual search across multiple displays with the ultimate goal of detecting discontinuous, erratic, and infrequent signals. Studies have shown that the steady-state level of vigilance performance in industrial settings (e.g., power systems control rooms) tends to decline sharply during the first half hour of watch (Harris and Chaney, 1969; Parasuraman, 1986), often referred to as a vigilance decrement. This phenomenon is often more pervasive in visual search tasks that involve long periods of inactivity followed by infrequent yet critical signals that must be detected with high levels of speed and accuracy (Wickens and Hollands, 2000).

Power systems operator tasks: Fault detection. In modeling signal detection theory, Green and Swets (1966) suggest that two stages of information processing take place in relatively close temporal proximity during the initial detection of faults such as voltage violations or line overloads. First, sensory evidence is gathered regarding the presence or absence of a fault. Next, a decision is made to confirm the existence of the fault based on this sensory evidence.

During fault detection, attention can be selective, focused, or divided, depending on the nature of the display design and the requirements of the specific fault detection task. In optimal selective attention, the operator chooses to guide attention strategically to display locations that contain the most task-relevant features. Focused attention then involves concentration on a specifically chosen source of information in the display. According to Yantis (1993), a failure in selective attention involves an intentional but erroneous decision to process suboptimal information, whereas a failure in focused attention results when an operator desires to shut out this suboptimal information but is unable to do so. Divided attention, on the other hand, involves simultaneous processing of multiple system display components (Wickens and Hollands, 2000). Limits in divided attention can stem from an operator's inability to time-share performance of multiple concurrent tasks or from failures in the integration of two or more information sources.

Various factors, both inherent to the display and to the individual operator, can alter attention allocation during fault detection. Visual attention will often be brought to stimuli that are most salient (e.g., large, colorful, bright, dynamic, brought on abruptly) (Remington, Johnston, and Yantis, 1992; Yantis and Hillstrom, 1994; Wickens and Hollands, 2000). Expectancy can also drive operators to fixate on specific regions of the display where faults are anticipated to occur (Wickens and Hollands, 2000). Perceptual grouping techniques (Palmer, 1992), although often performed subconsciously, may cause items to be processed together within a display or across multiple displays. Studies have shown that this integrated processing can be quite disadvantageous when the grouped information sources must be processed independently via focused attention (Wickens and Andre, 1990; Mori and Hayashi, 1995; Wickens and Carswell, 1995), as is often the case in the detection of faults. These and other factors indicate a range of both top-down and bottom-up influences on operator attention during fault detection.

Power systems operator tasks: Fault diagnosis and resolution. According to Wickens and Hollands (2000), the task of fault diagnosis must proceed beyond the perception, recognition, and attention selection stages of fault detection. Although these stages will still occur throughout diagnosis due to its iterative nature, operators must now integrate raw perceptual information from multiple channels, formulate a hypothesis about the state of the system, identify the possible ramifications of such a state, and finally provide a solution to a problem if one indeed exists. Operators will often use expectancies or prior beliefs to guide their decisions. Nested in this process is the requirement to incorporate both long-term memory for background knowledge and working memory for frequent mental model revision based on newly arriving sources of information.

Often, information integration presents a major challenge to human attention during fault diagnosis. Operators are frequently confronted with hundreds of potential cues, some requiring selective attention and others parallel processing. In situations like these, where multiple sources of information must be processed and acted upon, it has been proposed that the design of the diagnostic tools should incorporate the *proximity compatibility principle* (Barnett and Wickens, 1988; Wickens and Andre, 1990; Wickens and Carswell, 1995; Wickens and Hollands, 2000). This principle states that if multiple information sources are used within the same task (i.e., possessing high task proximity), these sources should be placed in close display proximity, either spatially or in terms of object-based properties. If, on the other hand, two information sources possess low task proximity, the information sources should be designed with low display proximity. Examples of tasks with low task proximity may include discrete value readings or simple comparisons, whereas high task proximity tasks will often involve complex comparisons and information synthesis. Of course, operators will be required to accomplish tasks from both classifications throughout fault diagnosis, and displays should thus be designed accordingly.

Also critical in the fault diagnosis phase is the creation of an accurate mental model of the statistical properties of events in the environment. This mental model will incorporate the operator’s expectancies about the frequency and location of faults within the system (Wickens and Hollands, 2000) and will be regularly used to guide his or her visual sampling patterns (Bellenkes, Wickens, and Kramer, 1997). Optimal performance will result only when designers seek to understand and translate users’ mental models into the design of the fault diagnosis tool.

Once a fault has been properly diagnosed, standard procedures are typically followed to then resolve the fault. These procedures, often in the form of explicit checklists, will commonly include sets of tasks that do not require substantial levels of effort or skill (e.g., simple computer clicks and switch adjustments). Fault resolution is assumed to be relatively automatic for experienced power systems operators and does not normally incorporate the complex information integration, mental modeling, or decision making techniques involved with fault diagnosis.

Traditional Power Systems Displays

Displays used by power systems operators to monitor, detect, diagnose, and resolve faults have shown little improvement with respect to traditional tabular displays depicting the statuses of system components combined with large, static, wall-mounted diagrams of the power grid (Overbye, Wiegmann, Rich, and Sun, 2001). By merely presenting operators with lists of violations (i.e., system faults) as text alarms, these displays often fail to elucidate the truly complex and dynamic nature of power systems networks. In addition, a diverse array of alarms and other unimportant or redundant messages within these displays can often hide critical violations, particularly in large networks (Mahadev and Christie, 1994b). This can significantly increase the workload of an operator, especially under stressful, time-constrained conditions. This ineffectiveness is compounded by the growth of networks and increasing responsibilities of power systems operators resulting from industry restructuring.

Traditional one-line diagrams have also been used to display power systems data at the substation level. One-line diagrams are "schematic diagrams presented on computer-generated character graphics displays" (Essenberg et al., 2003, p. 4) that use individual lines to symbolize electrical buses and demonstrate their connections to generators, capacitors, and other buses across the grid. Unfortunately, an accurate understanding of the overall system state has been difficult to obtain from these displays due to the rigid nature of their character graphics (Christie, 1994). To enable operators to gain this high-level understanding, power systems control rooms have, in many cases, turned to *full graphics displays*. This enhancement has allowed designers to construct graphical representations of system components and other data sources from individual pixels and thus build more flexible and intelligible displays of power systems. Regrettably, though, operator performance has not improved dramatically with such displays, likely due to their lack of evolution beyond movable windows incorporating standard character graphics (Christie, 1994).

Figure 1 presents examples of these traditional displays in combination with more modern displays in a typical power systems control room. The tabular and static wall-mounted displays are outlined in red, while a set of contemporary, three-dimensional one-line diagrams is outlined in yellow.

Figure 1. A power system control room with traditional tabular and static wall-mounted displays (outlined in red) and a new, three-dimensional graphical one-line display (outlined in yellow).

As evidenced by the aforementioned shortcomings of traditional displays, operators of electrical power networks have often been poorly supported in the tasks of fault detection, diagnosis, and resolution. Process control displays must therefore be redesigned with new and innovative visualization techniques that take advantage of the capabilities of full graphics displays and directly support each operator task (Christie, 1994).

The Redesign of Power Systems Displays

Many graphical display techniques have been proposed by both power systems designers and researchers with the ultimate goal of improving operator performance. The techniques relevant to the fault detection, diagnosis, and resolution tasks of the current study are presented below.

Integrated one-line diagrams. New graphical representations and dynamic digital values have recently been combined with traditional one-line diagrams through full graphics displays known as integrated one-line diagrams (Overbye, Wiegmann, Rich, and Sun, 2001; Overbye, Sun, Wiegmann, and Rich, 2002) (see Figure 2). In order to capitalize on the predictions of the proximity compatibility principle (Bennett and Malek, 2000; Rich, Wiegmann, and Overbye, 2001b; Wickens and Andre, 1990; Wickens and Carswell, 1995; Wiegmann, Rich, Overbye, and Zhang, 2001), power system component representations and component state information can be placed in close display proximity within these displays. Real time effects on individual components and the entire network can now be interpreted more easily and readily in comparison with traditional displays. In addition, these displays can expose *emergent features,* global display properties not evident when viewing individual display elements in isolation (Wickens and Hollands, 2000). The higher-order information presented by these emergent features is often a key component to an accurate and dynamic understanding of the system state.

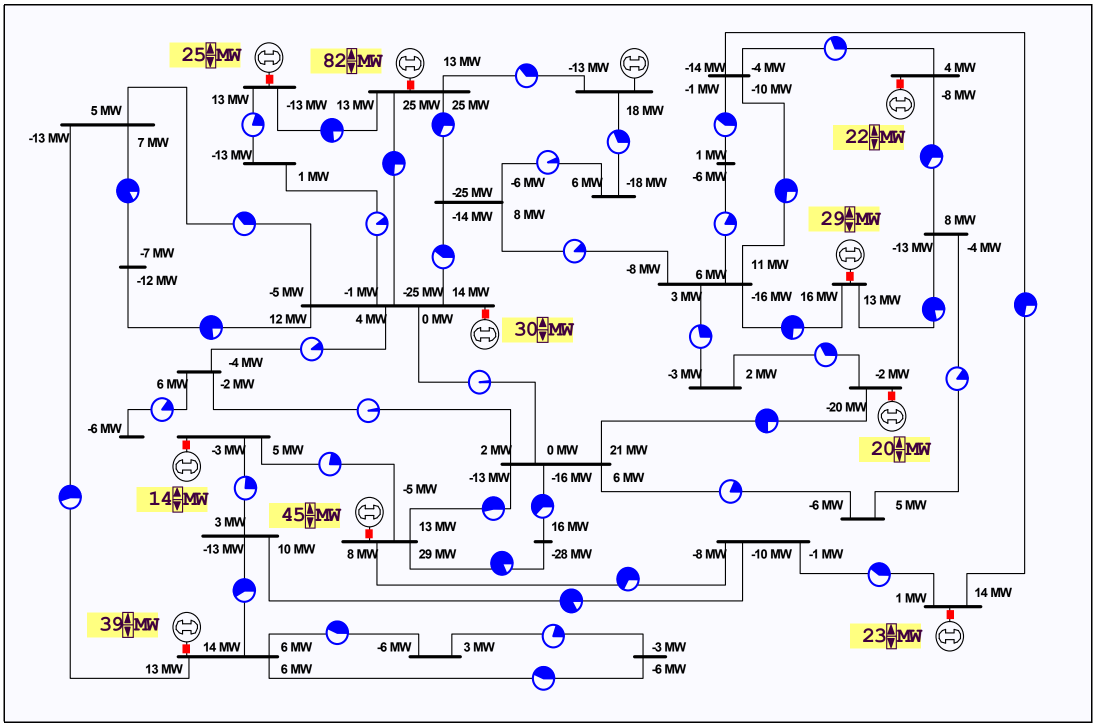

Figure 2. An integrated one-line diagram with power flow on transmission lines indicated by dynamic digital values and percentage loads indicated graphically by pie charts.

Integrated one-line diagrams have been developed to analogically present physical function and flow of information and resources between system components, and can thus be considered animated functional *mimic displays* (Hollan, Hutchins, McCandless, Rosenstein, and Weitzman, 1987; Bennett, 1993; Bennett and Nagy, 1995; Bennett and Malek, 2000, Essenberg et al., 2003). Mimic displays have been shown to "facilitate appropriate mental models and effective causal reasoning skills", "improve the quality of decision support in complex, dynamic domains", and "facilitate the identification of alternative resources required to recover from trouble" (Bennett and Nagy, 1995, p. 2). Such displays capitalize on the principle of pictorial realism by portraying critical systems and subsystems as they exist in the physical network (Wickens and Hollands, 2000; Rasmussen, 1986, p. 17, Essenberg, 2003), thus aiding in the mental interpretation of the network's connections and operating limits. By directly representing the states of internal system components, these displays are capable of making failures of such components explicit and providing valuable support during emergency operations (Kieras, 1992).

Overbye, Sun, Wiegmann, and Rich (2002) compared performance of participants using both traditional displays and integrated one-line diagrams in a simulated 30-bus system. The traditional displays consisted of tabular representations of power system subcomponent information complimented by static paper diagrams mapping out the physical structure of the network. The integrated one-line diagrams were dynamic in nature and represented bus voltages with either digits or color contours. In general, faults were detected faster with the traditional tabular displays, likely due to the relative ease of searching for violations through the neatly formatted columns of these displays compared to the less structured, often disorganized representation of the one-line diagrams. However, because the task of solving the voltage violations required activating capacitors that were “electrically near” the violations, thus necessitating mental integration of the relative locations of violations and capacitors, participants’ solution times were, in general, faster with the integrated one-line displays. This trend was consistent with the predictions of the proximity compatibility principle, as information integration was expedited due to relatively close display proximity of system components and state information.

Graphical enhancements of integrated one-line diagrams. The use of supplemental computer graphics and other scientific visualization techniques (Dyer, 1990; Liu and Qiu, 1998) have been proposed to aid operators in comprehending and assessing power system state (de Azevedo, de Souza, and Feijó, 1996) with integrated one-line diagrams. For example, Mahadev and Christie (1994a) explored the use of color-coded bars to convey bus voltages. Liu and Qui (1998) eliminated the need for alphanumerics by graphically depicting phase angles with pie charts. Mitsui and Christie (1997) used color-coded pixels to show system-level bus voltages. Wiegmann, Essenberg, Overbye, and Rich (2002) applied dynamic arrows to transmission lines with speed proportional to power flow. Klump, Schooley, and Overbye (2002) developed a three-dimensional graphical tool to help identify close-proximity sources of reactive power support. Overbye et al. (2002), Rich et al. (2001b), and Wiegmann et al. (2001) used color contour plots to identify problematic areas of electrical power grids where voltage violations were present. A summary of the visualization tools of color, motion, and three-dimensional graphics will be discussed briefly, followed by a more in-depth analysis of 3-D displays.

Color. The use of color in displays has been shown to have significant performance tradeoffs. On one hand, color-coding in a display can aid rapid location of targets or critical pieces of information through highlighting effects (Christ, 1975) and tie together cognitively integrated display elements that possess low display proximity (Wickens and Andre, 1990; Liu and Wickens, 1992; Wickens and Carswell, 1995). Conversely, color is subject to the limits of absolute judgment and does not naturally define an ordered continuum (Wickens and Hollands, 2000). This, coupled with the possibility that color can be a distracting element that can be processed erroneously due to false population stereotypes, indicates that color must be used sparingly and strategically in complex display design.

In a 2001 study, Wiegmann et al. examined the use of color contours in power systems displays. The experiment simulated a 118-bus system with three distinct integrated one-line diagram classifications: digital-only, contour-only, and digital-plus-contour. In trials containing more than eight initial violations, the subjects were generally fastest in acknowledging the worst bus voltage violations with the contour-only and digital-plus-contour displays. The researchers concluded that a warm color highlighting effect helped direct participants’ attention to the worst violations in these displays. On the other hand, for the task of solving voltage violations by activating capacitors electrically near the violations, subjects were generally quicker and more accurate with the digital-only display. It was concluded that reduced levels of clutter and obscuration of capacitors accounted for the enhancement.

Motion. Although the use of motion in power systems displays is not a novel concept, the empirical evaluation of motion in these displays has received little attention relative to most visualization techniques. In a preliminary study of motion in displays conducted by Wiegmann et al. (2002), fault detection and diagnosis performance was measured for a one-line diagram of a 30-bus power system, with and without the use of motion. Results showed a decrement in detection performance in multiple violation trials for a one-line diagram with moving arrows representing power flow compared to a one-line diagram with only digital values. The distracting effect of arrow size and speed adjustments throughout the display as contingencies occurred was the likely cause of such an outcome. In contrast, there was a trend toward faster solution times with the moving arrow display compared to those with the digital-only display. Apparently, contingencies were solved faster due to the aid provided by the moving arrows in determining power flow directions faster. In an extension of this study, Essenberg et al. (2003) found a clear advantage for motion in geographical displays indicating power transaction flows for a unique buyer-seller selection task. Specifically, subjects were asked to select the buyer and seller in an aggregated electrical power network using displays containing either moving or stationary arrows. Selection times were significantly faster with the moving arrow displays compared to those with the stationary arrow displays, and this response time advantage was enhanced as search difficulty, or buyer-seller separation distance, increased. Accuracy, measured by error rate, was also enhanced with the motion displays and followed a similar pattern to buyer-seller selection time, yielding a greater advantage for motion displays with increasing search difficulty. Hence, a speed-accuracy tradeoff was not apparent.

Motion has been incorporated into a range of other power systems applications as well. For example, the animation of color contours has been proposed to facilitate the detection of evolving contingencies across large power networks (Overbye, Weber, and Laufenberg, 1999). In addition, bus statuses have been animated through time to graphically depict changing values of power flow, voltage, and phase angle and show the fluctuating tendencies of power systems (Kobayashi, Okamoto, Tada, Yamada, and Sekine 1998). Again, the formal empirical evaluation of these and many other motion-based visualization tools has been limited and must continue if the true utility of motion in power systems displays is to be fully understood.

Three-dimensional graphics. In many instances in complex displays, a third dimension is used to convey depth, or the "perceived distance from the observer along an axis perpendicular to the plane of the display" (Wickens and Hollands, 2000, p. 139). In some cases, the three dimensions may correspond to three Euclidean dimensions of space (Wickens and Hollands, 2000), as with, for example, a display intended to help guide a remotely controlled robotic arm operating outside a space shuttle. The third dimension may also be used to graphically convey some other non-distance quantity previously expressed with two-dimensional alphanumerics (Rich, Wiegmann, and Overbye, 2001b; Overbye, Wiegmann, and Thomas, 2002). At the extreme, three-dimensionality may even be used in the creation of virtual reality, the multisensory experience of a set of artificial locations and surrounding objects by means of computer modeling (Carr, 1995). The increased realism in three dimensions compared to two has been identified as an essential element in the design of a truly realistic and interactive virtual world (Furness and Barfield, 1995).

Displays of three dimensions are becoming increasingly common in process control environments and particularly power systems control rooms. For example, Figure 3 shows a sample 3-D integrated one-line diagram of a high-voltage electrical power system. Unfortunately, the empirical evaluation of displays such as this has been relatively nonexistent. To understand the tradeoffs involved with three-dimensional display use, along with the causes of specific systematic distortions when operators use depth information in the natural world, it is first essential to discuss the fundamental characteristics of three-dimensional depth perception. The following section will cover this important issue and provide a brief description of several other key components of three-dimensional display design for complex systems.

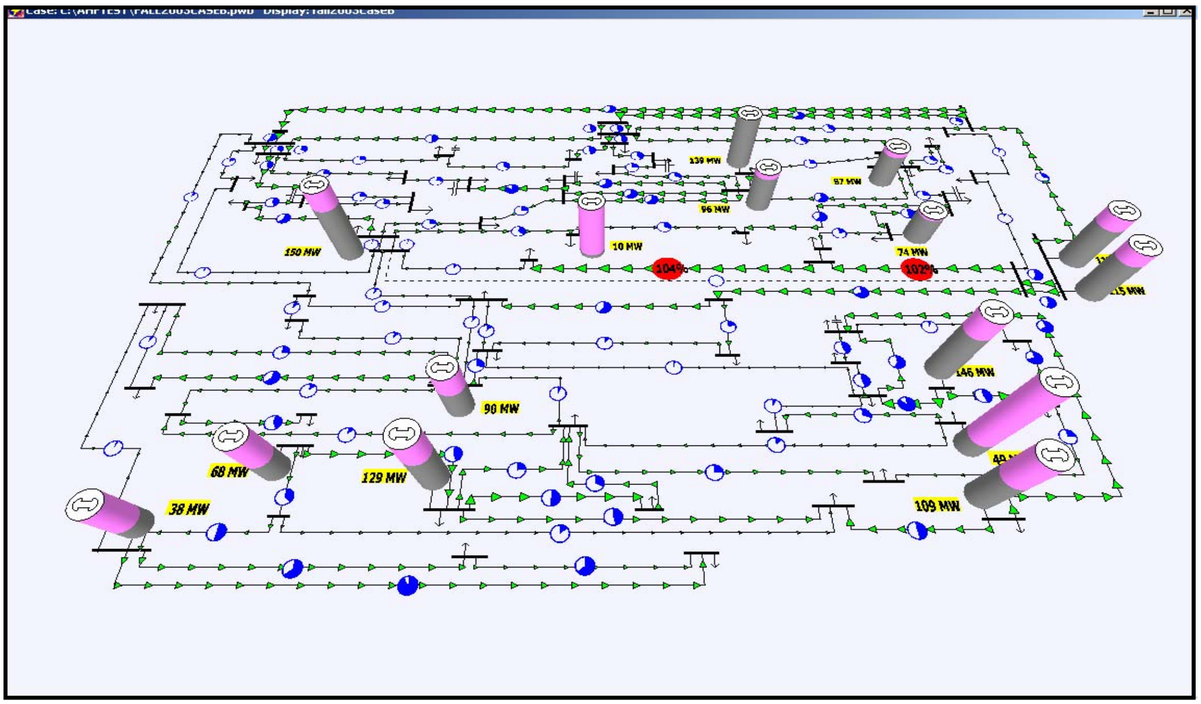

Figure 3. Three-dimensional view of the high voltage electrical power system one-line diagram with cylinders to indicate proportional maximum output, current output, and reserves (3-D group).

Three-Dimensional Display Design

In this section, we first discuss three general categories of reference frames in displays. We then highlight the basic cues for depth perception in three dimensions, including both object-centered cues and observer-centered cues, and the effects of human perceptual limitations in 3-D. Attention allocation in both two- and three-dimensional displays and the effects predicted by the proximity compatibility principle are then discussed. Each of these represents a key element to the present study comparing 2-D and 3-D display use in complex power systems.

Frames of reference. A display's frame of reference refers to the viewpoint from which the graphical information is shown. The challenge to the display designer is to pick the optimal frame of reference to support the various tasks that must be accomplished. Although Wickens

(1999) identifies and compares operator performance across several task-specific ‘camera angle’ viewpoints representing different reference frames, the present discussion will group all frames of reference into three general classifications. The most egocentric, immersed frame of reference (see Figure 4a) is typified by the forward field of view in three-dimensional space (Wickens and Hollands, 2000). Empirical results have shown that this display will improve performance in navigational tasks (Theunissen, 1994; Haskell and Wickens, 1993; Wickens and Prevett, 1995; McCormick, Wickens, Banks, and Yeh, 1998; Olmos, Wickens, and Chudy, 2000) and create a more realistic sense of “presence” in virtual environments (Slater and Usoh, 1993). A less egocentric, although still three-dimensional, tethered view (see Figure 4b) is created when the viewpoint is placed behind and slightly above the viewer’s position, looking toward the viewer’s position (Wickens and Prevett, 1995). While experimental results from Wickens and Prevett indicate that performance in navigation and search tasks can be hindered with this less-immersive viewpoint, the study suggests that situation awareness and global knowledge can indeed improve. Furthermore, a longer “tether length”, resulting in a greater viewable space on an individual display, can enhance one’s ability to integrate spatial information and thus form a consistent mental map (Wickens and Hollands, 2000). The most exocentric frame of reference, commonly referred to as a 2-D plan view (see Figure 4c), results when the viewpoint is located directly above the viewer’s position, looking directly downward (Wickens and Hollands, 2000). Wickens (1999) states that “it is the fully exocentric viewpoint that can most optimally support that aspect of situation awareness related to searching and finding relevant entities” (p. 134). In addition, this highly exocentric viewpoint has consistently been shown to aid tasks requiring a macroscopic knowledge of spatial structure (Aretz, 1991; Barfield and Rosenberg, 1995; Wickens, Liang, Prevett, and Olmos; 1996; Williams, Hutchinson, and Wickens, 1996; McCormick, Wickens, Banks, and Yeh, 1998; Wickens, 1999). Because current applications in power systems control rooms are typically limited to 2-D plan views and 3-D tethered views, the present study will examine performance using only these two reference frames.

Three-dimensional depth cues. According to Wickens and Hollands (2000), a realistic representation of depth and distance can be accomplished by means of various perceptual cues. By creating a compelling and accurate sense of three-dimensionality, these depth cues can signal information to the viewer such as an object’s relative and absolute distance, its 3-D shape, and the orientation of its surfaces (Wickens, Todd, and Seidler, 1989). Some of these cues are properties of our personal visual systems (i.e., “observer-centered” cues), while others are characteristics of the object or environment that we perceive (i.e., “object-centered” cues) (Wickens and Hollands, 2000).

Observer-centered cues (e.g., binocular disparity, convergence, and accommodation) provide information about depth that stem from characteristics of the human visual system (Wickens and Hollands, 2000). Although included in Table 1 for the sake of comprehensiveness, these depth cues will be irrelevant to the current study, as all displays will be fixed in distance to the observer.

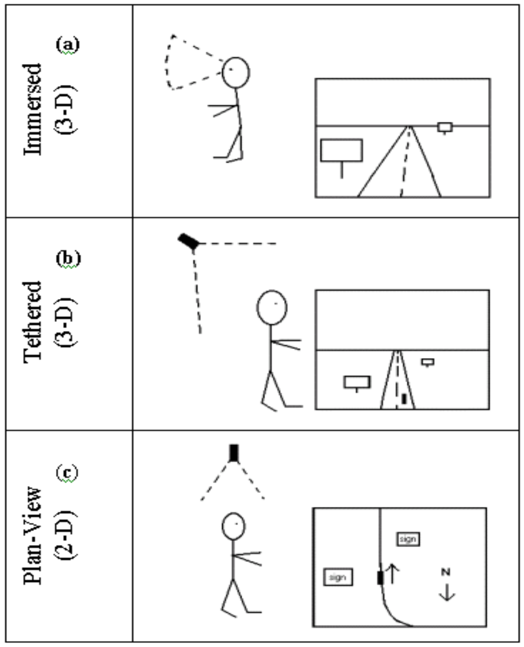

Figure 4. Two-dimensional representation of display frames of reference. Each cell shows a small icon of a viewer, a "camera" positioned relative to the viewer, and a schematic panel showing the contents of the display (Wickens and Hollands, 2000, p. 168).

Table 1. Observer-centered depth cues (Wickens and Hollands, 2000).

Depth cue	**Description**
Binocular disparity	Each eye simultaneously receives a disparate image of an object in view. This level of disparity provides the observer with a basis for assessing distance. *Stereoscopic display systems* refer to computers with accompanying components and software capable of presenting left and right stereo images separately to each eye of the observer (Sollenberger and Milgram, 1993).
Convergence	The level to which an observer becomes "cross-eyed" as an item is brought closer to his or her eyes is a natural indicator of an object's distance.
Accommodation	An image is brought into focus on the retina through the adjustment of the lens shape by the observer's eye muscles. Relative object distance from the observer can be signaled by this degree of adjustment.

Object-centered cues, on the other hand, will be quite relevant to the current study. These cues, frequently referred to as “pictorial” in nature, are commonly built into a two-dimensional picture in order to convey an artificial sense of depth (Wickens and Hollands, 2000). These include linear perspective, interposition, height in the plane, light and shadow, relative/familiar size, textural gradients, proximity-luminance covariance, aerial perspective, and motion parallax. Table 2 provides a detailed description of each of these cues.

Table 2. Object-centered depth cues (Wickens and Hollands, 2000).

Depth cue	**Description**
Linear perspective	If two lines are shown to be converging, the observer assumes that they are relatively parallel lines moving backward in depth.
Interposition	If one object obscures another, the observer assumes that the obscured object is farther away.
Height in the plane	Assuming the typical field of view is from above, the observer assumes that objects higher in the visual field are more distant.
Light and shadow	If light shines on objects from a specified direction, shadows can offer clues about the objects’ relative orientations and shapes.
Relative/familiar size	Assuming objects are identical in size, those presenting a smaller visual image on the retina are presumed to be more distant.
Textural gradients	The grain of a texture, assuming its plane is oriented toward the line of sight, will become finer with distance.
Proximity-luminance covariance	Closer objects are often brighter. Thus, continuously diminishing illumination on objects is assumed to indicate receding distance (Dosher, Sperling, and Wurst, 1986).
Aerial perspective	Objects that are farther from us often tend to be defined with less clarity and contain more of a ‘haze’.
Motion parallax	An observer’s relative movement with respect to a three-dimensional scene causes closer objects to show greater relative motion than those that are farther.

Difficulties in 3-D perception. Although 3-D visualization tools are more realistic and sophisticated than ever, often incorporating many of the aforementioned depth cues, human beings are still inherently limited in ability to accurately judge spatial information in three dimensions. To account for a known discrepancy between perception and reality, observers often use interpretations of the 3-D world to make inferences about relative sizes, depths, and distances of objects. Hence, three-dimensional depth and distance estimations are often governed by *perceptual hypotheses* (Wickens and Hollands, 2000). These hypotheses are often relatively involuntary, habitual, and unconscious, indicating a significant level of top-down processing. As observers interpret the 3-D world, their hypotheses are reinforced by depth cues ideally working together to provide consistent information about the true state of the world.

Unfortunately, perceptual ambiguities of depth, size, and distance will inevitably occur when the 3-D world is graphically depicted on a 2-D display (Gregory, 1970). For example, Hollands, Pierce, and Magee (1998) found that subjects' performance was hindered when estimating the distance between two lines in a three-dimensional cubic display due to interference caused by the depth information. Carswell, Frankenberger, and Bernhard (1991) found a similar performance decrement for participants making relative magnitude estimations with three-dimensional line graphs compared to their two-dimensional counterparts. St. John, Smallman, and Cowen (2000) performed experiments with terrain stimuli and determined that while 3-D perspective views enhanced performance on tasks requiring shape understanding, two-dimensional views were superior with precise judgments of angle, distance, and relative position.

Research has shown that this degradation in depth and distance assessment is exacerbated along the line of sight of the observer (Olmos, Wickens and Chudy, 2000; Wickens, Liang, Prevett, and Olmos, 1996). Significant changes in distance along this axis may produce relatively insignificant changes in displayed distances between objects, leading the observer to underestimate the true magnitude of object separation (Wickens and Hollands, 2000). Perceptual ambiguities are also particularly prevalent in estimations of volume in 3-D graphical formats (Cleveland and McGill, 1985; Carswell, Frankenberger, and Bernhard, 1991; Siegrist, 1996). Volumes have been shown to produce response compression, in which each incremental increase in an object's physical magnitude produces a smaller incremental increase in perceived magnitude (Stevens, 1975; Wickens and Hollands, 2000). Thus, large volumes will tend to be underestimated. For this reason, three-dimensional volumes should be used with caution to code important variables.

Numerous design techniques have been proposed to counter these 3-D perceptual ambiguities. Again, simple artificial depth cues (see Tables 1 and 2) can enhance observers' abilities to accurately assess relative size and position in three-dimensional displays. Of course, some depth cues can offer greater assistance to the observer compared to others, and experiments should be performed to determine the unique combination of cues that optimizes performance in each particular case. In a fusion of several studies of cue dominance, Wickens, Todd, and Seidler (1989) determined that the most governing cues are often interposition, motion parallax, and binocular disparity. In addition, supplementary graphical techniques such as the use of tick marks to subdivide graphical stimuli into subsections can also be strategically incorporated into 3-D displays to reduce perceptual ambiguity (Wickens and Hollands, 2000).

Guiding attention. When addressing the issue of optimal attention allocation in displays, the dimensionality of the display is of prime concern. For example, research has indicated that three-dimensional displays are often ineffective displays for focused attention tasks. Carswell, Frankenberger, and Bernhard (1991) supported this claim by identifying a performance decrement in three dimensions compared to two for a trend classification task involving primarily focused processing. Haskell and Wickens (1993) determined three-dimensional displays to be disadvantageous for the focused-attention task of airspeed control. Liu, Zhang, and Chaffin (1997) concluded that a three-dimensional human form graphic hindered the accuracy of participants' posture assessments when relatively basic, symmetric positions were examined. In each of these cases, depth information in the three-dimensional displays impeded precise judgments of size, distance, and other exact measurements.

In contrast, performance in tasks requiring divided attention, information integration, or mental model development has been shown to improve with three-dimensional displays compared to their two-dimensional counterparts. For example, Carswell, Frankenberger, and Bernhard (1991) found that although accuracy was susceptible, both 3-D line graph and bar chart formats required less time to use compared to 2-D line formats for the estimation of global trends, a task requiring mental integration. Wickens, Merwin, and Lin (1994) discovered that performance with 3-D scatterplot displays exceeded that with 2-D plots for tasks involving information integration. This phenomenon was attributed to the fact that the three-dimensional plots provided superior visual depictions of the intricate shapes of the 3-D surfaces examined. Liu, Zhang, and Chaffin (1997) found that the 3-D human form graphic enhanced observers' abilities to analyze asymmetric postures requiring mental integration of information from all three dimensions of space.

Each of these findings is again consistent with those predicted by the proximity compatibility principle: When information requiring mental integration is indeed integrated into three dimensions, performance on such tasks should improve (Wickens and Carswell, 1995). Three-dimensional, integral, objectlike formats that yield emergent features will allow for more efficient parallel processing, whereas two-dimensional displays will enhance observers' capabilities in focused attention tasks requiring precise judgments.

Strategically presenting cues and signals at different depths in three-dimensional displays can also have a significant effect on how attention is guided. For example, Atchley, Kramer, Andersen, and Theeuwes (1997), through stereoscopic techniques, used cues to identify the relative spatial depths of signals. Participants' signal response times increased when a cue and signal were at contrasting depths compared to instances when both the cue and signal were presented at identical depths. Apparently, a level of focused attention was produced by the cue at its specific depth (Wickens and Hollands, 2000). In two separate studies, Chau and Yeh (1995) and Theeuwes, Atchley, and Kramer (1998) found that by spatially separating targets from distractors by presenting each at different depths, observers were able to locate targets with shorter search times. It can be interpreted from these results that distractor stimuli that would normally prolong search will not do so if presented at varying depths from the target (Wickens and Hollands, 2000). Preserving this depth information may thus be beneficial when distinguishing targets form distractors in three-dimensional displays.

Summary of 3-D costs and benefits. Overall, the advantages of three-dimensional graphic displays over two-dimensional numeric displays are significant. The added dimension and pictorial enhancements will often increase the amount of information (e.g., non-distance quantities) that can be presented on standard display screens, allow for graphic representations of alphanumerics, increase the operator's sense of presence within the display environment, assist in navigation and search activities, aid in tasks requiring information integration through the creation of emergent properties, enable strategic separation of targets and distractors at varying depths, and facilitate more accurate mental modeling of systems being manipulated. Although 2-D displays supplemented with graphics may share some of the aforementioned benefits (e.g., graphical representation of alphanumerics, enhanced information integration), limiting such displays to two dimensions will eradicate many others (e.g., increased presence, depth benefits) and may cause added levels of clutter. Of course, inherent costs of 3-D displays may include increased perceptual ambiguity, limited global awareness, and difficulty in precise judgment of size and distance. Regardless, there exists sufficient evidence that three-dimensional displays can be an effective visualization technique in the domain of power systems monitoring and control.

General Summary

Display design for electrical power systems data has evolved significantly from traditional static map boards and basic one-line diagrams. Graphical representations and dynamic digital values can now be combined with one-line diagrams into functional full graphics displays known as integrated one-line diagrams. The goal now is to design these integrated one-line diagrams to optimize operator performance and empirically demonstrate this performance enhancement. To accomplish this, many human factors issues with important implications for power system applications must be examined. A range of these has been addressed in this study.

Of particular concern in this investigation is the effect of three-dimensional displays on operator performance in power systems monitoring and control. Although various display concepts have been proposed, few, if any, empirical investigations have been conducted exploring the benefits and drawbacks of using such displays to analyze power flow and capacity.

However, to formally evaluate visualization usefulness, at least two critical matters must be addressed (Rich, Wiegmann, and Overbye, 2001b). First, the specific tasks must be determined. Of course, the tasks of a power systems planner, system operator, marketer, and engineer will differ greatly. This particular experiment focused on the traditional process control tasks of fault detection, involving focused attention, and diagnosis, requiring information integration. Next, the degree to which the experimental setup mirrors the actual world must be examined. The objective in this experimental setup was to roughly mimic a utility control environment in which operators must respond to system load disturbances. To perfectly imitate such an environment and realistically test operators' performances with a range of visualization tools would be extremely hard due to the complicated nature of power systems control. Thus, the experimental setup was admittedly of low fidelity compared to that in a true control center, in which skilled operators and engineers monitor and control larger networks with which they are quite familiar (Rich, Wiegmann, and Overbye, 2001b). Nevertheless, the experimental setup used in this study should be feasible for initially evaluating three-dimensional enhancements for operational environments.

Purpose of the Present Study

In the present study, we compared participants' performance in acknowledging and solving (i.e., diagnosing and subsequently resolving) line violations in a simulated electrical power grid for two 2-D display groups (numerical-only and graphical) and one 3-D display group. The power grid, represented by an animated one-line diagram, contained 67 buses, 103 transmission lines and transformers, and 12 generators. The amount of power flow on a given transmission line was controlled indirectly through the changing of relevant generator outputs. Display representations of this system for the 2-D numerical, 2-D graphical, and 3-D groups are shown in Figures 5, 6, and 7, respectively.

The most significant differences across the three display groups were found in the representations of generator power output, capacity, and available reserves. In each 2-D display, a generator was depicted as a circle with a “dog-bone rotor” symbol inside. For the 2-D numerical group, the current power output of each generator was shown numerically in a yellow field close to the generator symbol, with the maximum power the generator could produce shown below it in a magenta field. For the 2-D graphical group, the current power output of each generator was also shown numerically in a yellow field close to the generator symbol, although in these displays gauges were used to graphically indicate proportional maximum power output, current output, and available reserves. The height of each gauge was proportional to the maximum power the generator could produce. The level of the gray shaded portion of the gauge was proportional to the current power output of the generator, and the height of the magenta shaded portion was proportional to the generator’s reserve power. In the 3-D group, each generator was shown as a three-dimensional cylinder whose height was proportional to the maximum power the generator could produce. Similar to the 2-D numerical group, the levels of the gray and magenta shaded portions of each cylinder were proportional to the current power output and reserve power of the generator, respectively. Current power output of each generator in the 3-D display was also shown numerically in a yellow field close to the generator symbol, although this field was on occasion blocked by the generator’s cylinder representation. For all display groups, a line violation symbol consisted of an enlarged red pie chart with the line’s load percentage shown numerically within the pie chart. Transmission lines and paths connecting buses were shown in all displays with thin black lines.

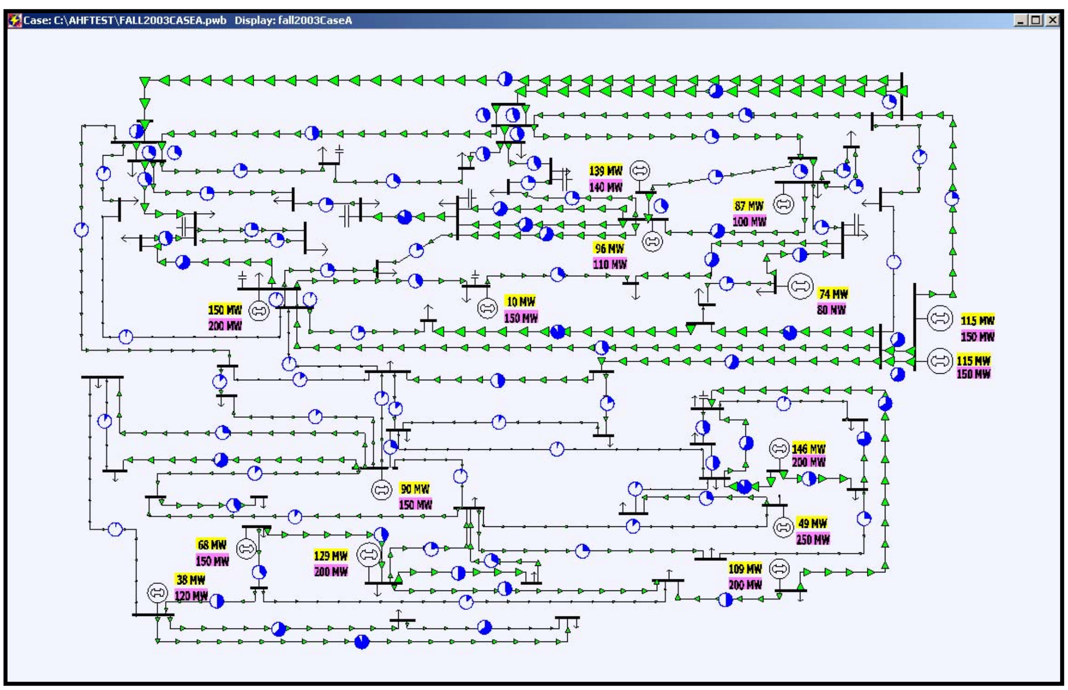

Figure 5. Two-dimensional view of the high voltage electrical power system one-line diagram with maximum generator output shown in a magenta field (2-D numerical group).

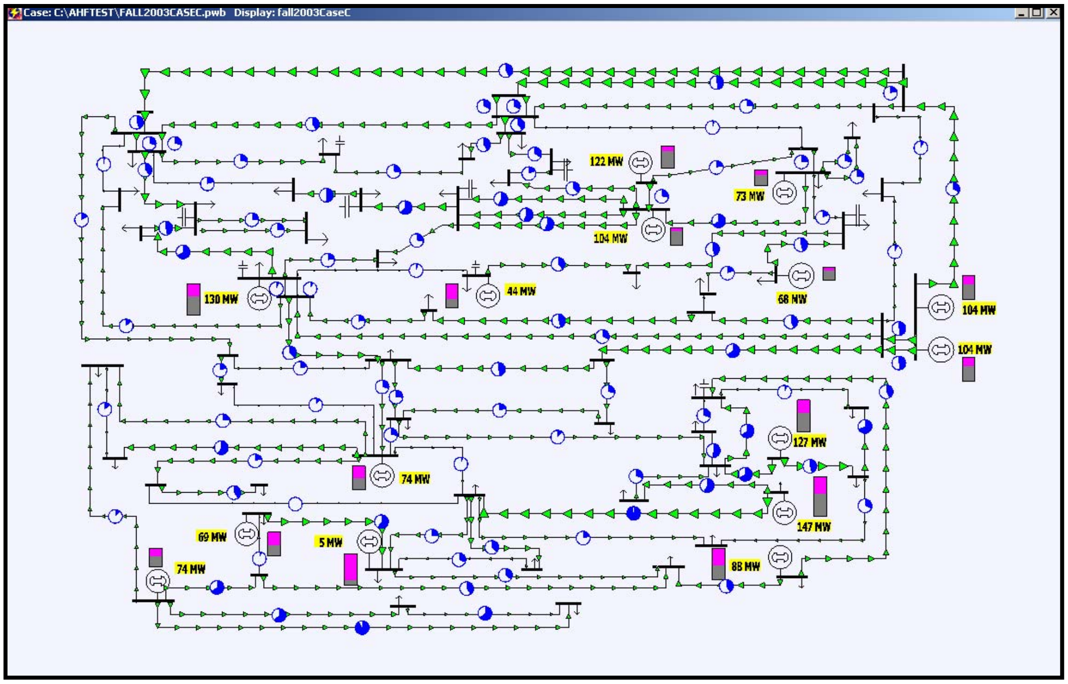

Figure 6. Two-dimensional view of the high voltage electrical power system one-line diagram with gauges to indicate proportional maximum output, current output, and reserves (2-D graphical group).

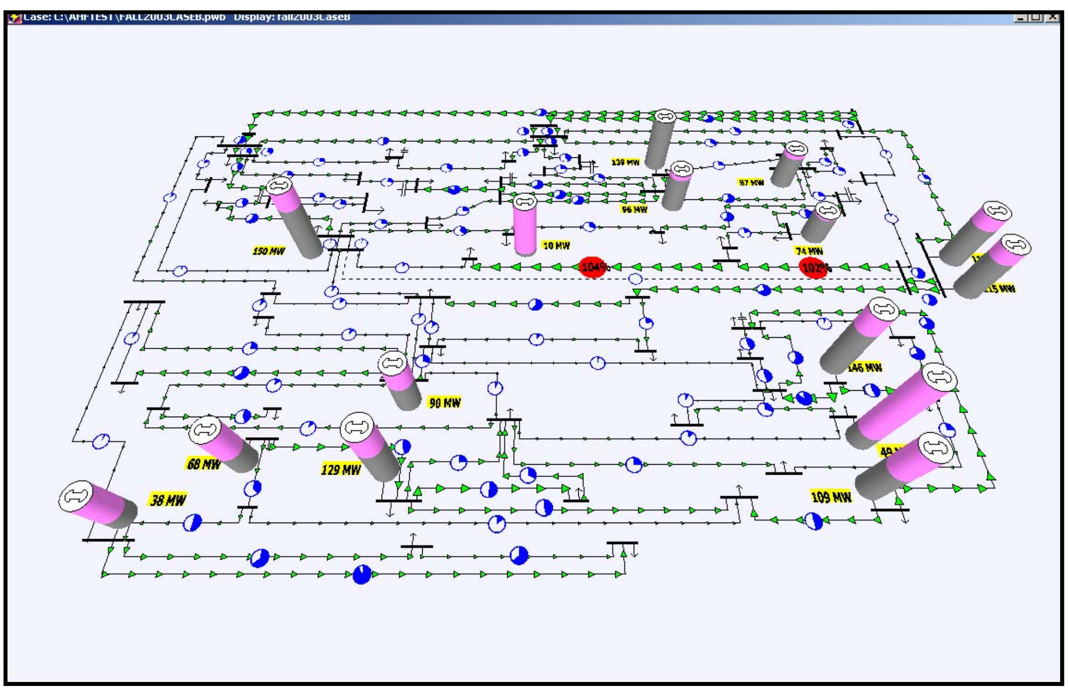

Figure 7. Three-dimensional view of the high voltage electrical power system one-line diagram with cylinders to indicate proportional maximum output, current output, and reserves (3-D group).

At the onset of each trial, all lines were at or below acceptable power flow limits. After some amount of time, a *contingency* occurred, causing either one line to overload for single violation trials or multiple lines to overload for multiple violation trials. Participants were first required to acknowledge all overloaded transmission lines by clicking via mouse on each violation symbol. Once all violations had been acknowledged, subjects were then required to take corrective action by altering the outputs of one or more generators to bring all line loadings to acceptable levels. In some scenarios, the optimal solution required participants to adjust the outputs of multiple generators across the grid.

Hypotheses. Taking into account the many research- and operations-based studies comparing two-dimensional and three-dimensional displays, several hypotheses can be made regarding the current experiment. These hypotheses will be broken into categories of response time, accuracy, and workload, and will differ depending on the level of problem complexity.

The dimensionality effect on response times will depend on the specific task being performed. Due to the relative task simplicity and consistency of violation presentation formats across display types, we do not expect to see significant differences in violation acknowledgment times across display types. Although display compression in the three-dimensional displays will lead to a size reduction of violation symbols, this negative effect will likely be countered by a reduction in travel distances between line violation symbols. On the other hand, we do predict that solution times will decrease for both the 2-D graphical and 3-D groups due to the pictorial enhancements of these displays. Because the solution task will require mental integration of multiple information sources such as current generator output, maximum generator output, and available reserves, we anticipate the integration of these information sources into a single object (i.e., a 2-D gauge or a 3-D cylinder) to tend to reduce solution times, in accordance with the proximity compatibility principle. The non-numeric representation of system variables in the 2-D graphical and 3-D groups will also eliminate the need to perform mental math operations, further reducing solution times.

In addition, the salient illustration of generator reserves in these two groups is expected to expedite fault resolution by quickly drawing operator attention to the generator(s) containing the highest level of power available. Because the cylinders of the 3-D group are larger than the gauges of the 2-D graphical group, and thus more salient, this effect is predicted to be greater for the 3-D group. Furthermore, this size discrepancy will cause dynamic changes in power outputs and available reserves to be more salient for the 3-D group, further aiding operators in assessing system-wide effects of individual generator adjustments. Although display compression and obscuration of system components by the 3-D cylinders will tend to increase solution times, these effects are predicted to be negligible in the present study due to the limited size of the displays. For these reasons, we predict solution times to be fastest for the 3-D group, followed by the 2-D graphical group, and slowest for the 2-D numerical group.

The effects of display dimensionality on solution accuracy are predicted to be limited. We do expect that the graphic enhancements of the 2-D graphical and 3-D groups will aid in preventing unnecessary generator increases when generators are already at their maximum output levels (i.e., "limit errors"), as operators in these display groups will merely need to note when a generator's 2-D gauge or 3-D cylinder representation is completely gray and from this discern that it can yield no additional output. Again, the salience and highlighting effect of the magenta-colored reserve representations should direct an operator's attention towards those with sufficient available power and away from those that are operating at their capacities. However, no other measures of solution accuracy are predicted to be affected by display type or dimensionality, as adjustment direction choice is supported solely by information gained from the moving arrows present in all displays.

Because the temporal and cognitive demands placed on subjects during the violation acknowledgment task are predicted to be negligible, subjective workload should follow a pattern roughly similar to solution time. Specifically, the absence of mental math with the 2-D graphical and 3-D groups should lead to a reduction in frustration level, mental demand, and effort required.

METHODS

Participants

The experiment included 53 participants, 41 men and 12 women, who were recruited from power systems classes in the Department of Electrical and Computer Engineering at the University of Illinois, Urbana-Champaign. Using an entrance questionnaire, we screened out one colorblind male participant from the 3-D group. All participants either had completed or were currently enrolled in one power systems course. The participants' median age was 21, ranging from 19 to 32. They were each paid $12.00 for their voluntary participation.

Computer Software and Apparatus

A modified version of PowerWorld® Simulator software (Overbye, Klump, and Weber, 1999; Rich, Wiegmann, and Overbye, 2001a) was used for the experiment, allowing for configuration and representation of local utilities, equivalent transmission lines, and power transfer in the power system simulation. The software was run on a Dell Dimension XPS B800 computer with an 800-MHz Pentium III processor, 128 MB of RAM, and a 20-inch Dell Ultrascan P1110 monitor. A mouse was used for all input.

Power System Simulation and Tasks

The software simulated a high voltage electrical power system using a one-line diagram. The simulation contained 67 buses, 103 transmission lines and transformers, and 12 generators. As is the case with a real power system, the maximum power limit values for the lines were not identical. In this experiment, they ranged from 16 MW to 200 MW.

2-D numerical group. For the 2-D numerical group, the amount of power being produced by each generator (in MW) was shown numerically in a yellow field close to the generator, with the maximum power the generator could produce shown below it in a magenta field (see Figure 5). The transmission lines were represented as thin lines joining the buses. The power load on each line in proportion to its capacity was indicated using a pie chart on the line. An unfilled pie chart indicated that there was no power flowing down the transmission line, while a pie chart completely filled with blue indicated that the line was loaded at 100%. If the line's loading exceeded 100%, the pie chart increased in size and changed color to red, with the percentage load indicated numerically. Loads, which were present at most buses, were shown using small black arrows. The amount and direction of the power flowing on each transmission line was shown using green moving arrows.

During the simulation, the one-line diagram was animated, with the size of the arrows proportional to the actual amount of power flowing down the transmission line. The amount of power flowing down a transmission line was controlled indirectly by changing the output of the generators. Participants could change generator outputs by left-clicking on the generator's yellow MW field or dog-bone rotor symbol to increase its generation or right-clicking on the generator's yellow MW field or dog-bone rotor symbol to decrease its generation. Each mouse click changed the output of the generator by 2 MW. The change in flow of power on the transmission lines was dynamically reflected on the one-line diagram by changing pie chart values and by the changing size of the green arrows on the transmission lines. Each generator had a minimum output of 0 MW and a maximum output indicated on the one-line diagram by the magenta field. Clicking on the magenta field indicating the maximum output had no effect.

To begin each experimental trial, the line flows all started out below 100%. Then, after a period of 2-12 seconds, a contingency occurred. A contingency consisted of a loss of one or more system devices, such as the opening of a transmission line following a simulated lightning strike. The contingency changed the transmission line loadings, causing one or more lines to overload (i.e., load above 100%). Any lines opened as a result of a contingency were indicated on the one-line diagram with a dashed line.

Once the contingency occurred, participants were to complete two tasks. First, they were required to acknowledge all the transmission lines that were overloaded by left-clicking on the corresponding violation symbols (i.e., enlarged red pie charts). Following each contingency, the computer continued beeping until all overloaded lines were acknowledged. Second, once all violations had been acknowledged, the participants were required to take corrective action by changing the outputs of one or more generators to change the line flows so that all line loadings were at or below 100%. In general, changing the generators that were electrically closest to the overloaded transmission lines was the optimal strategy. Electrical closeness was measured by the number of components (i.e., transmission lines and buses) between the generator and the overloaded line.

Again, the output of each generator varied from a minimum of 0 MW to a maximum of 16-200 MW. Further right or left clicks had no impact on the output of a generator already at its minimum or maximum level, respectively. In some instances, the optimal solution for each scenario required participants to increase or decrease the output of multiple generators. When changing the outputs of generators, it was also possible that participants could cause new violations. Although it was not necessary to acknowledge these new violations, they did need to be corrected.

Once all the line violations had been removed, the trial automatically ended and the subject was prompted to move on to the next trial. If the subject failed to solve the problem within 120 seconds, the trial automatically terminated and the participant was prompted to move on to the next trial. The program indicated when the last trial had been finished.

2-D graphical group. The tasks of the 2-D graphical group were identical to those of the 2-D numerical group. However, for the 2-D graphical group a gauge was shown in close proximity to each generator, with the vertical height of the gauge proportional to the maximum power output of the generator (see Figure 6). The level of the gray shaded portion of the gauge was proportional to the current power output of the generator, and the height of the magenta shaded portion was proportional to reserve power. A generator loaded at its maximum was thus depicted as a completely gray gauge, while an unloaded (0 MW) generator was shown as a completely magenta gauge. The actual amount of power being produced by each generator (in MW) was also shown in a yellow field close to the generator. The locations and sizes of these yellow fields were identical to those in the 2-D numerical group.

During the simulation, participants increased the output of the generators by left-clicking on the generator's yellow MW field, dog-bone rotor symbol, or anywhere on the 2-D gauge. Right-clicking on the generator's yellow MW field, dog-bone rotor symbol, or anywhere on the gauge decreased its power generation. All other experimental protocols were identical to those for the 2-D numerical group.

3-D group. The tasks of the 3-D group were also identical to those of the 2-D numerical group. However, to create the three-dimensional one-line diagram for each scenario with respect to the 2-D one-line diagram, the viewpoint was rotated 30° from the vertical, with the generators depicted as 3-D cylinders (see Figure 7). Each one-line diagram in the 3-D group was thus shown using a perspective view, so that objects farther from the "viewpoint" appeared smaller (i.e., farther away). The height of each cylinder was proportional to the maximum power the generator could produce. The level of the gray shaded portion of the cylinder was proportional to the current power output of the generator, and the height of the magenta shaded portion was proportional to reserve power. Hence, a generator loaded at its maximum was shown as a completely gray cylinder, while an unloaded (0 MW) generator was shown as a fully magenta cylinder. The tops of the generators were represented as circles, each with a dog-bone rotor symbol inside. The actual amount of power being produced by each generator (in MW) was again shown in a yellow field close to the generator, although this field was on occasion partially or completely blocked by the generator's cylinder representation.

Participants in the 3-D group increased the output of the generators by left-clicking on the generator's yellow MW field or anywhere on the cylinder (sides or top). Right-clicking on the generator's yellow MW field or anywhere on the cylinder decreased its power generation. All other experimental protocols were identical to those described for the 2-D numerical group.

Procedure

Phase 1: Group assignment and instructions. Each participant was first randomly assigned to a display group, with the exception that those who had previously used PowerWorld® Simulator or had participated in previous power systems experiments using PowerWorld® Simulator were roughly evenly distributed between the groups. The three display groups were the 2-D numerical group (n = 18), the 2-D graphical group (n = 18), and the 3-D group (n = 17). Participants began the experiment by completing a consent form and providing basic demographic and education information. They were then provided specific written instructions about the display and the tasks to complete during the experiment. Each participant was asked to read the instructions completely and inform the experimenter when ready to proceed or ask questions if any came up.

Phase 2: Trials. Each participant was administered 4 practice trials and 40 experimental trials. The 40 experimental trials consisted of 20 single-violation trials and 20 multiple-violation trials. The experimenter demonstrated successful completion of two practice trials and then the participant began the first practice trial. At the beginning of each trial, the line flows all started out below 100%. After a random interval between 2 and 12 seconds, a contingency occurred, changing the transmission line loadings and causing one or more line violations. The violations on each line were identical across the three display groups in each trial, but varied from trial to trial. After all violations had been acknowledged, the participants were required to correct the violations by altering the outputs of one or more generators. If changes in generator outputs caused new violations, participants were required to resolve them but not required to acknowledge them. The participants were allowed 120 seconds to complete each experimental trial. If the subject failed to solve all violations within the allotted time, he or she was prompted to move on to the next scenario. After each completed or timed-out trial, a dialog box appeared informing the participant of the outcome of the trial. The participant then clicked “OK” when ready to proceed to the next trial or to quit the program in the case of the final trial.

Phase 3: Post-experimental questionnaire and debriefing. After the final trial, the participants completed a post-experimental questionnaire, which included the NASA Task Load Index (TLX) subjective workload assessment. They were then given a debrief sheet explaining the general purpose of the experiment and dismissed.

RESULTS

Dependent Measures

Performance measures consisted of response times, accuracy, and workload. Response time measures included the time from the onset of the initial violations until the acknowledgment of all initial violations (acknowledgment time), the time from the acknowledgment of all initial violations until the reduction of all line loadings to 100% of capacity or less (solution time), and the synthesis of acknowledgment time and solution time (total response time). Solution time was further divided into the time from the acknowledgment of all initial violations until the first generator adjustment (first adjustment time) and the mean time between generator adjustments following the first generator adjustment (mean adjustment interval). Accuracy was measured by the number of generator adjustments in the wrong direction (adjustment errors) and the number of sequences of one or more generator increases or decreases when a generator was already at its maximum or minimum output level, respectively (upper and lower limit error events). Finally, workload was measured with the NASA-TLX subjective workload assessment.

Response Times

Response times were first analyzed using a 3 (Display Type: 2-D numerical, 2-D graphical, 3-D) x 2 (Problem Complexity: single violation, multiple violations) x 2 (Task: acknowledgment, solution) analysis of variance (ANOVA). Post hoc *t* tests were used to determine significant differences between individual data points.

Figure 8 shows response times as functions of display type, problem complexity, and task. The results of the ANOVA showed a significant main effect of display type on total response time, $F(2, 50) = 7.41$, $p = .002$, revealing a 3-D advantage. As illustrated in Figure 8a, participants in the 3-D ($M = 8.47$, $SD = 1.72$) group took less time to complete the combined acknowledgment and solution tasks than those in the 2-D numerical ($M = 11.4$, $SD = 2.68$) or 2-D graphical ($M = 10.4$, $SD = 2.38$) groups. *T* tests showed that the differences in total response times between the 2-D numerical and 3-D groups, $t(33) = 3.85$, $p = .001$, and between the 2-D graphical and 3-D groups, $t(33) = 2.81$, $p = .008$, were significant, but that the difference between the 2-D numerical and 2-D graphical groups was not significant, $t(33) = 1.15$, $p = .257$. There was also a significant main effect of problem complexity on total response time, $F(1, 50) = 535$, $p < .001$, revealing a single violation advantage. Figure 8b shows that participants had faster total response times for single violation trials ($M = 7.05$, $SD = 2.15$) compared to those for multiple violation trials ($M = 13.2$, $SD = 3.25$). In addition, a significant main effect of task, $F(1, 50) = 662$, $p < .001$, was evident. Figure 8c displays much smaller task completion times for the acknowledgment task ($M = 2.11$, $SD = 0.23$) compared to the solution task ($M = 18.2$, $SD = 5.11$).

ANOVA results showed a significant two-way interaction between display type and task, $F(2, 50) = 7.86$, $p = .001$. Figure 9 shows both acknowledgment and solution times as functions of display type, collapsed across problem complexity. The interaction is illustrated by the existence of a significant reduction in solution times but not acknowledgment times for the 3-D display compared to both 2-D displays. In fact, there were no significant differences in acknowledgment times between any of the display types, as indicated by the results of *t* tests, which, with descriptive statistics, are shown in Tables 3, 4, and 5 for the interaction. The 3-D advantage for total response time can thus be attributed solely to the solution task.

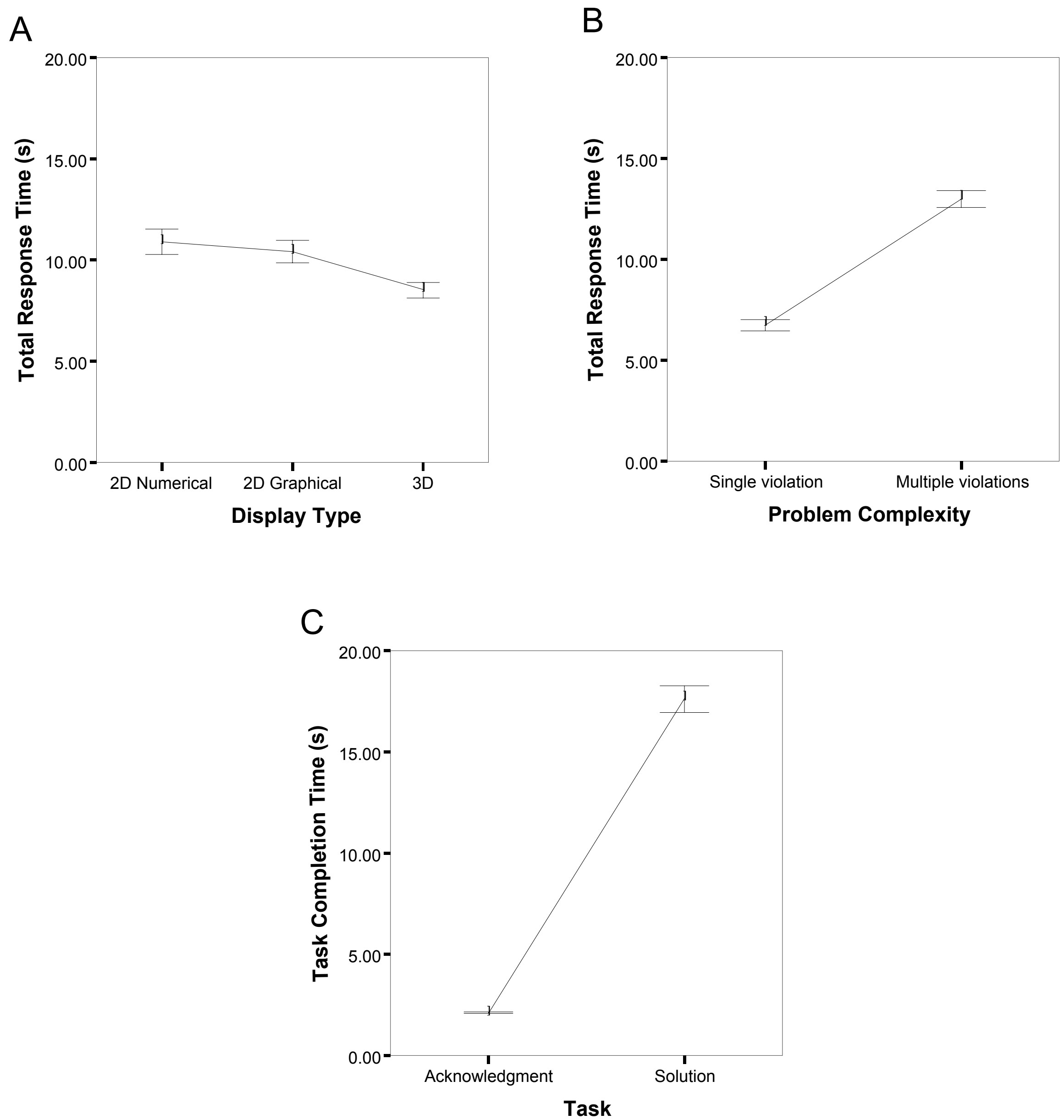

Figure 8. Main effects of (a) display type on total response time, (b) problem complexity on total response time, and (c) task on task completion time.

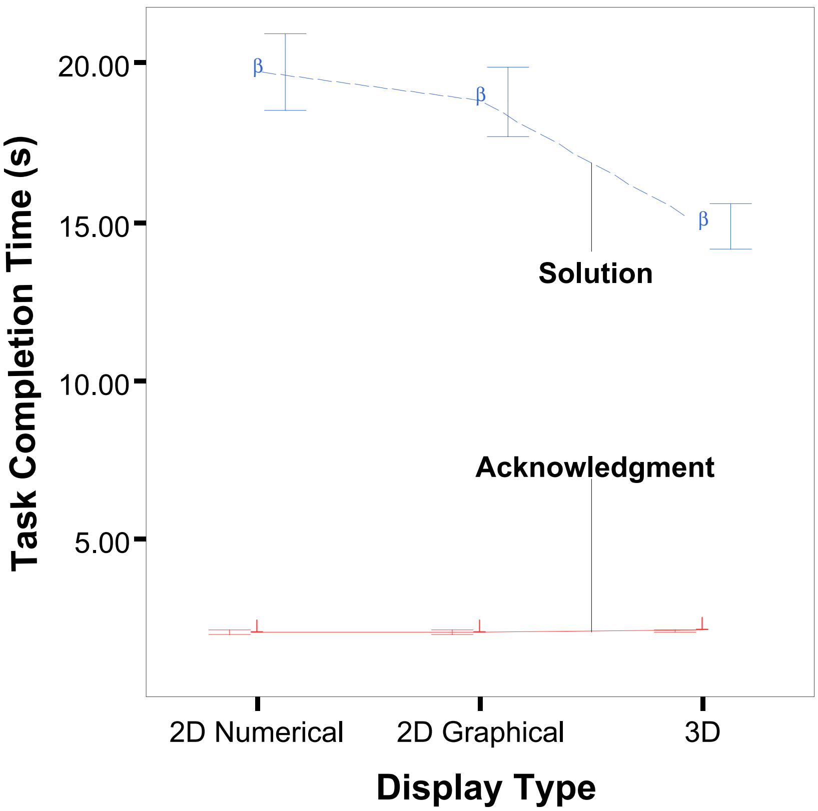

Figure 9. Task completion times as a function of display type. Error bars represent standard error.

Table 3. Descriptive statistics for task completion time by display type, collapsed across problem complexity.

	2-D Numerical		2-D Graphical		3-D	
Task	*M*	*SD*	*M*	*SD*	*M*	*SD*
Acknowledgment	2.09	0.27	2.10	0.25	2.13	0.16
Solution	20.75	5.32	18.79	4.67	14.81	3.43

Note. All values are in seconds.

Table 4. Between-subjects *t* tests of task completion time.

	2-D Numerical – 2-D Graphical		2-D Numerical – 3-D		2-D Graphical - 3-D	
Task	*t*(34)	*p*	*t*(33)	*p*	*t*(33)	*p*
Acknowledgment	-0.15	.880	-0.49	.629	-0.34	.737
Solution	1.17	.249	3.90*	< .001	2.86*	.007

*$p < .05$

Table 5. Within-subjects *t* tests of task completion time by display type.

	2-D Numerical		2-D Graphical		3-D	
	t(17)	*p*	*t*(17)	*p*	*t*(17)	*p*
Sol. – Ack.	14.94*	< .001	15.37*	< .001	15.26*	< .001

*$p < .05$

The solution task was then broken down into first adjustment time and mean adjustment interval in order to identify the solution subtask(s) that contributed to the aforementioned effects. Mean first adjustment time was analyzed using a 3 (Display Type: 2-D numerical, 2-D graphical, 3-D) x 2 (Problem Complexity: single violation, multiple violations) ANOVA. Figure 10 shows the mean first adjustment time as functions of display type and problem complexity. The results of the ANOVA showed a significant main effect of display type, $F(2, 50) = 7.95$, $p = .001$, indicating a 3-D advantage. *T* tests showed that the difference in mean first adjustment time was significant between the 2-D numerical ($M = 2.25$, $SD = 0.85$) and 3-D ($M = 1.26$, $SD = 0.41$) groups, $t(33) = 4.32$, $p <.001$, and between the 2-D graphical ($M = 1.93$, $SD = 0.86$) and 3-D groups, $t(33) = 2.90$, $p = .007$, although the difference between the 2-D numerical and 2-D graphical groups was not significant, $t(34) = 1.11$, $p = .274$. In addition, there was a significant main effect of problem complexity on mean first adjustment time, $F(1, 50) = 12.01$, $p = .001$, with an advantage for multiple violation trials ($M = 1.73$, $SD = 0.77$) over single violation trials ($M = 1.92$, $SD = 0.94$).

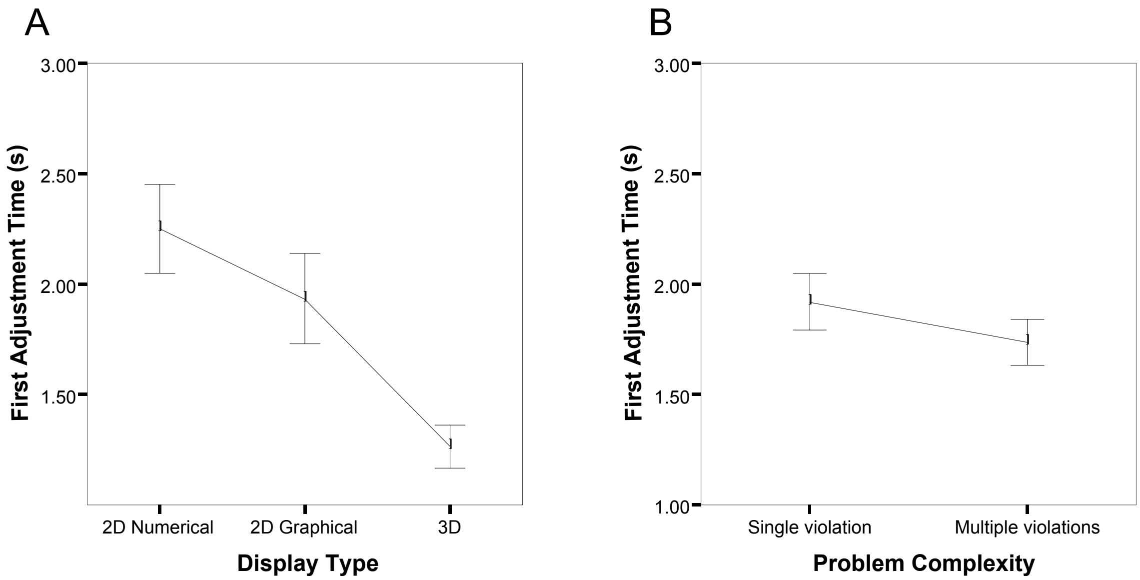

Figure 10. Main effects of (a) display type and (b) problem complexity on mean first adjustment time. Error bars represent standard error.

Figure 11 shows mean first adjustment time as a function of display type expanded across problem complexity. The ANOVA results showed a significant two-way interaction between display type and problem complexity, $F(2, 50) = 4.70$, $p = .013$. The interaction is illustrated by decreasing differences between mean first adjustment times of single and multiple violation trials from the 2-D numerical to 2-D graphical to 3-D display groups. In other words, the 3-D advantage was enhanced for single violation trials. *T* tests revealed that the significant difference between single and multiple violation trials for the 2-D numerical display, the marginally significant difference for the 2-D graphical display, and the non-significant difference for the 3-D display were the main contributing factors to the interaction. In addition, it should be noted that the difference between the 2-D numerical and 2-D graphical displays was not significant for single violation trials. Tables 6, 7, and 8 show descriptive statistics and the results of the *t* tests between all of the data points.

The mean adjustment interval was also analyzed using a 3 (Display Type: 2-D numerical, 2-D graphical, 3-D) x 2 (Problem Complexity: single violation, multiple violations) ANOVA. Figure 12 depicts the mean adjustment interval as functions of display type and problem complexity. Results showed a significant main effect of display type, $F(2, 50) = 3.47$, $p = .039$, indicating a graphical/3-D advantage. *T* tests showed that the difference in mean adjustment interval was significant between the 2-D numerical ($M = 0.35$, $SD = 0.11$) and 3-D ($M = 0.27$, $SD = 0.07$) groups, $t(33) = 2.47$, $p = .019$. However, *t* tests only revealed a trend in the difference between the 2-D numerical and 2-D graphical ($M = 0.30$, $SD = 0.08$) groups, $t(34) = 1.56$, $p = .127$, and no significant difference between the 2-D graphical and 3-D groups, $t(33) = 1.07$, $p = .291$. In addition, there was a significant main effect of problem complexity, $F(1, 50) = 17.41$, $p < .001$, with an advantage for single violation trials ($M = 0.30$, $SD = 0.09$) over multiple violation trials ($M = 0.32$, $SD = 0.10$).

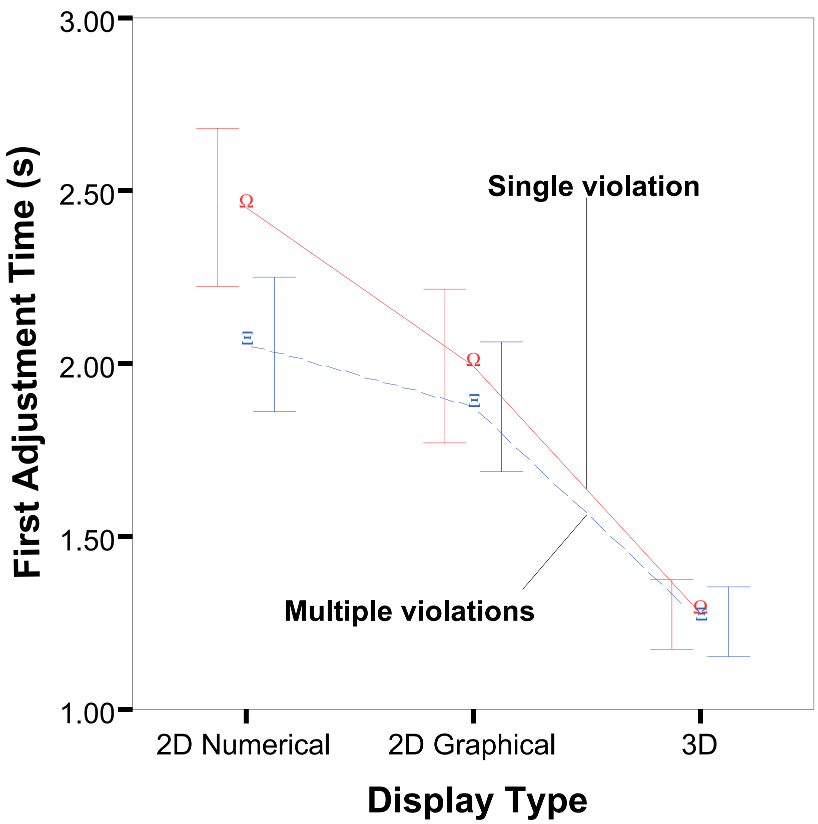

Figure 11. Mean first adjustment times expanded across problem complexity as a function of display type. Error bars represent standard error.

Table 6. Descriptive statistics for mean first adjustment time by display type expanded across problem complexity.

	2-D Numerical		2-D Graphical		3-D	
Problem Complexity	*M*	*SD*	*M*	*SD*	*M*	*SD*
Single violation	2.45	0.97	1.99	0.95	1.27	0.42
Multiple violations	2.05	0.82	1.87	0.79	1.25	0.41

Note. All values are in seconds.

Table 7. Between-subjects *t* tests of mean first adjustment time by problem complexity.

	2-D Numerical – 2-D Graphical		2-D Numerical – 3-D		2-D Graphical - 3-D	
Problem Complexity	*t*(34)	*p*	*t*(33)	*p*	*t*(33)	*p*
Single violation	1.43	.163	4.60*	< .001	2.87*	.007
Multiple violations	0.66	.510	3.60*	.001	2.87*	.007

*$p < .05$

Table 8. Within-subjects *t* tests of mean first adjustment time by display type.

	2-D Numerical		2-D Graphical		3-D	
	t(17)	*p*	*t*(17)	*p*	*t*(16)	*p*
Single - Multiple	2.99*	.008	1.79	.091	0.63	.541

*$p < .05$

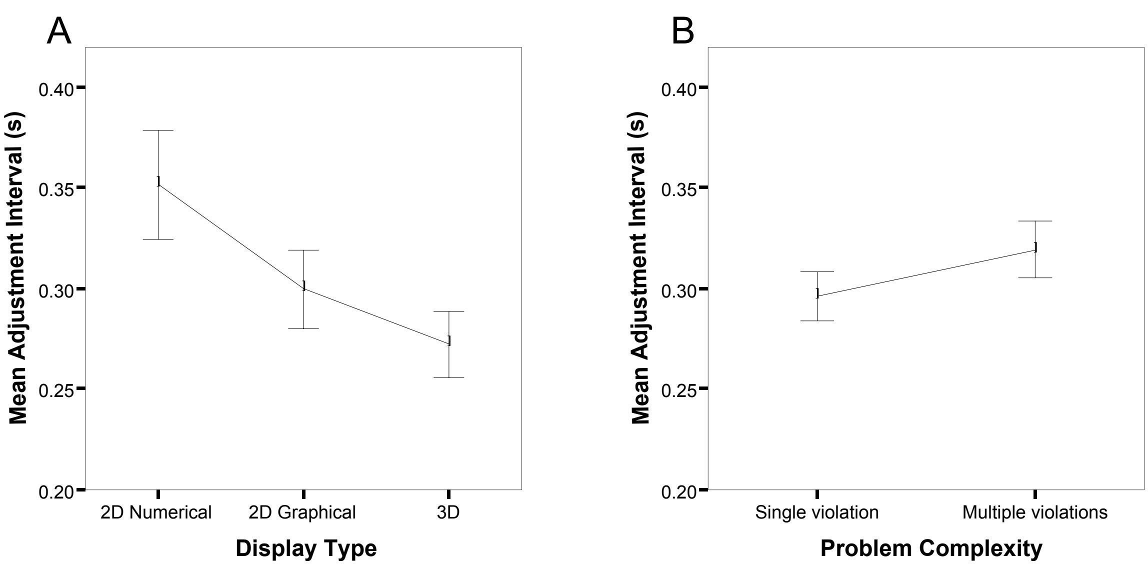

Figure 12. Main effects of (a) display type and (b) problem complexity on mean adjustment interval. Error bars represent standard error.

ANOVA results indicated a significant two-way interaction between display type and problem complexity, $F(2, 50) = 8.54$, $p = .001$. Figure 13 shows the mean adjustment interval as a function of display type expanded across problem complexity. The interaction is illustrated by an enhancement of the graphical/3-D advantage for multiple violation trials. Table 9 shows the descriptive statistics for all of the data points. As indicated in Tables 10 and 11, *t* tests showed that the main contributing factors to the interaction were the significant differences for multiple violation trials between the 2-D numerical and 2-D graphical groups, between the 2-D numerical and 3-D groups, and between single and multiple violation trials for the 2-D numerical group. For single violation trials, the difference between the 2-D numerical and 3-D groups was only marginally significant, while the differences between the 2-D numerical and 2-D graphical groups was not significant. In addition, the difference between single and multiple violation trials was not significant for the 2-D graphical group and constituted only a trend for the 3-D group.

Accuracy

Analysis of adjustment errors and lower limit errors revealed no significant effects. However, a 3 (Display Type: 2-D numerical, 2-D graphical, 3-D) x 2 (Problem Complexity: single violation, multiple violations) ANOVA did reveal a significant main effect of display type, $F(2, 50) = 12.53$, $p < .001$, on the mean number of upper limit error events per trial. Figure 14a shows evidence of a 3-D advantage. *T* tests revealed that the differences in mean upper limit error events per trial were significant between the 2-D numerical ($M = 0.51$, $SD = 0.16$) and 3-D ($M = 0.26$, $SD = 0.13$) groups, $t(33) = 5.07$, $p < .001$, and between the 2-D graphical ($M = 0.50$, $SD = 0.20$) and 3-D groups, $t(33) = 4.21$, $p < .001$, although the difference between the 2-D numerical and 2-D graphical groups was not significant, $t(34) = 0.14$, $p = .891$. Figure 14b shows a main effect of problem complexity, $F(1, 50) = 77.1$, $p < .001$, with an advantage for single violation trials ($M = 0.29$, $SD = 0.18$) over multiple violation trials ($M = 0.56$, $SD = 0.28$).

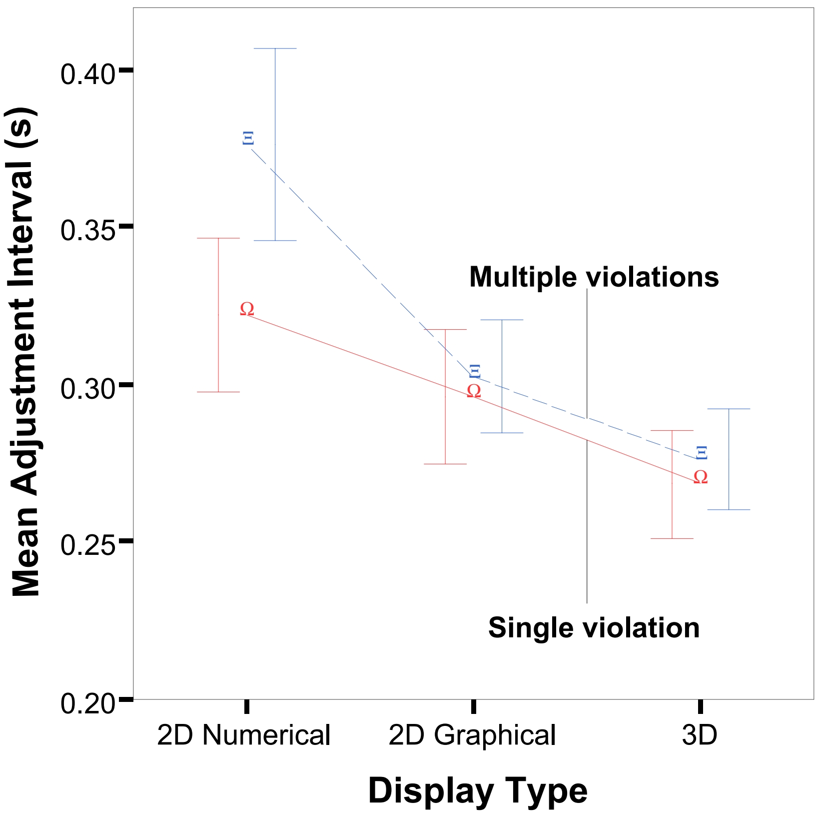

Figure 13. Mean adjustment interval expanded across problem complexity as a function of display type. Error bars represent standard error.

Table 9. Descriptive statistics for mean adjustment interval by display type expanded across problem complexity.

	2-D Numerical		2-D Graphical		3-D	
Problem Complexity	*M*	*SD*	*M*	*SD*	*M*	*SD*
Single violation	0.32	0.10	0.30	0.09	0.27	0.07
Multiple violations	0.38	0.13	0.30	0.08	0.28	0.07

Note. All values are in seconds.

Table 10. Between-subjects *t* tests of mean adjustment interval by problem complexity.

	2-D Numerical – 2-D Graphical		2-D Numerical – 3-D		2-D Graphical - 3-D	
Problem Complexity	*t*(34)	*p*	*t*(33)	*p*	*t*(33)	*p*
Single violation	0.78	.438	1.78	.085	1.02	.315
Multiple violations	2.09*	.045	2.86*	.007	1.08	.289

*$p < .05$

Table 11. Within-subjects *t* tests of mean adjustment interval by display type.

	2-D Numerical		2-D Graphical		3-D	
	t(17)	*p*	*t*(17)	*p*	*t*(16)	*p*
Multiple - Single	4.00*	.001	0.83	.418	1.65	.119

*$p < .05$

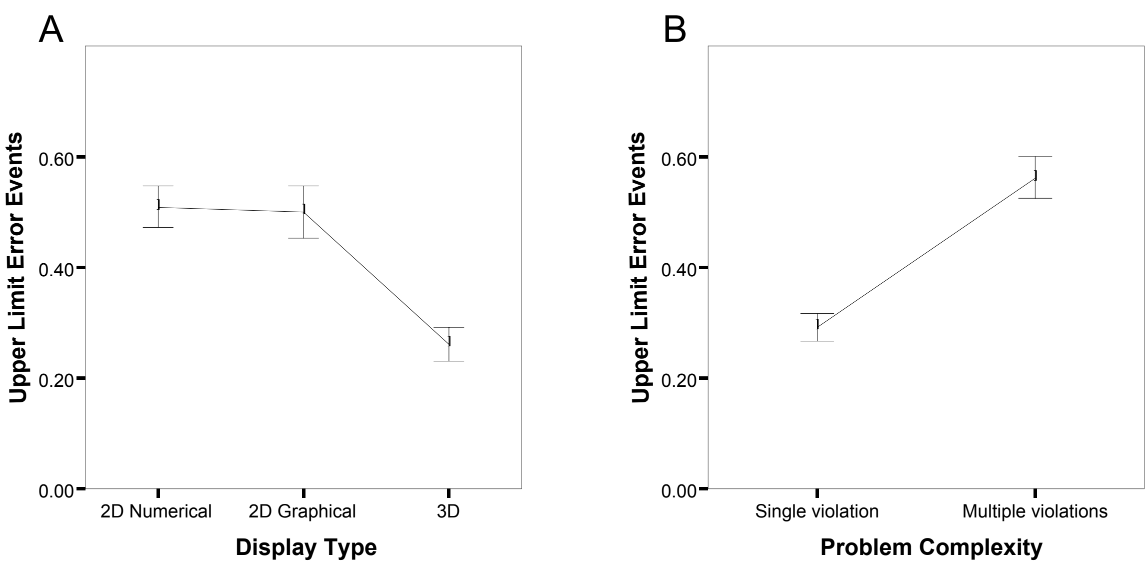

Figure 14. Main effects of (a) display type and (b) problem complexity on the mean number of upper limit error events per trial. Error bars represent standard error.

Figure 15 depicts the mean number of upper limit error events per trial as a function of display type expanded across problem complexity. ANOVA results indicated a significant two-way interaction between display type and problem complexity, $F(2, 50) = 6.36$, $p = .003$. The interaction is illustrated by an enhancement of the 3-D advantage for multiple violation trials; there was a much smaller difference between single and multiple violation trials for the 3-D display group than for the 2-D numerical and 2-D graphical groups. Tables 12, 13, and 14 show descriptive statistics and the results of the *t* tests between all of the data points. The *t* tests indicated that while the difference between the 2-D graphical and 3-D groups was significant for multiple violation trials, it was only marginally significant for single violation trials. In addition, note that the difference between single and multiple violation trials for the 3-D display group *was significant.*

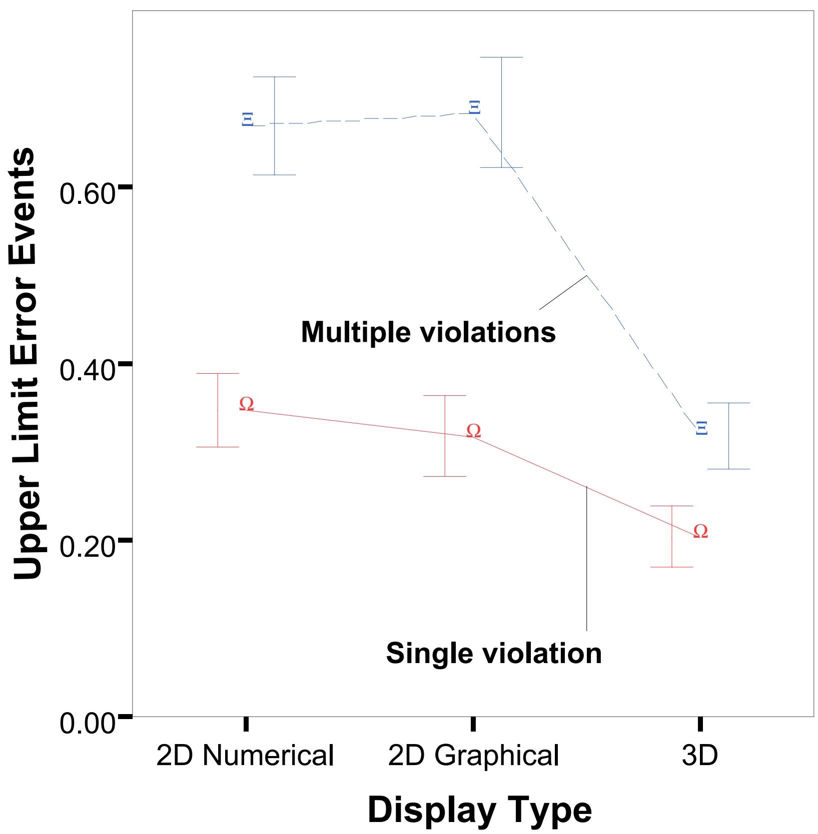

Figure 15. Mean number of upper limit error events per trial expanded across problem complexity as a function of display type. Error bars represent standard error.

Table 12. Descriptive statistics for mean number of upper limit error events per trial by display type expanded across problem complexity.

	2-D Numerical		2-D Graphical		3-D	
Problem Complexity	*M*	*SD*	*M*	*SD*	*M*	*SD*
Single violation	0.35	0.18	0.32	0.19	0.20	0.14
Multiple violations	0.67	0.23	0.68	0.26	0.32	0.16

Note. All values are in seconds.

Table 13. Between-subjects *t* tests of mean number of upper limit error events per trial by problem complexity.

	2-D Numerical – 2-D Graphical		2-D Numerical – 3-D		2-D Graphical - 3-D	
Problem Complexity	*t*(34)	*p*	*t*(33)	*p*	*t*(33)	*p*
Single violation	0.50	.622	2.67*	.012	1.97	.057
Multiple violations	0.17	.868	5.20*	<.001	4.97*	<.001

*$p < .05$

Table 14. Within-subjects *t* tests of mean number of upper limit error events per trial by display type.

	2-D Numerical		2-D Graphical		3-D	
	t(17)	*p*	*t*(17)	*p*	*t*(16)	*p*
Multiple - Single	5.22*	<.001	6.76*	<.001	3.00*	.009

*$p < .05$

Workload

Participants reported mental workload with the NASA-TLX by rating mental demand, physical demand, temporal demand, performance, effort, and frustration level on a scale from 0 to 100 separately for trials 1-10, 11-20, 21-30, and 31-40. Post-analysis revealed it to be unnecessary to group subjects' responses into these trial blocks, and averages were collected over all trials. The participants circled one of 21 tick marks on each of six Likert scales corresponding to each TLX dimension, each tick mark representing an increment of 5 points.

The results of the NASA-TLX are shown in Figures 16 and 17. A 3 (Display Group) x 6 (TLX Dimension: mental demand, physical demand, temporal demand, performance, effort, frustration level) ANOVA revealed a significant main effect of TLX dimension, $F(5, 250) = 12.9$, $p < .001$, but indicated neither a significant main effect of display type, $F(2, 50) = 0.15$, $p = .858$, nor a significant interaction between display type and TLX dimension, $F(10, 250) = 1.05$, $p = .404$.

The descriptive statistics and *t*-test results for the six TLX dimension scores are shown in Tables 15 and 16, respectively. The *t* tests indicated that performance, temporal demand, mental demand, and effort were the most significant contributors to overall workload, in order of decreasing mean score.

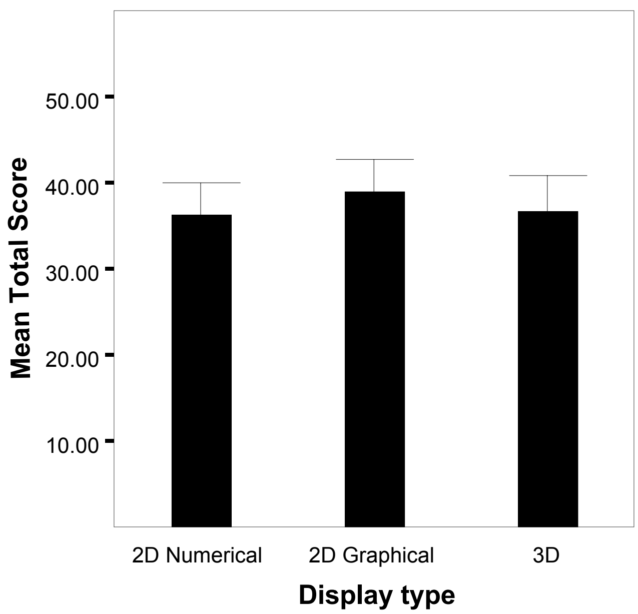

Figure 16. Mean NASA-TLX workload scores as a function of display type. Error bars represent standard error.

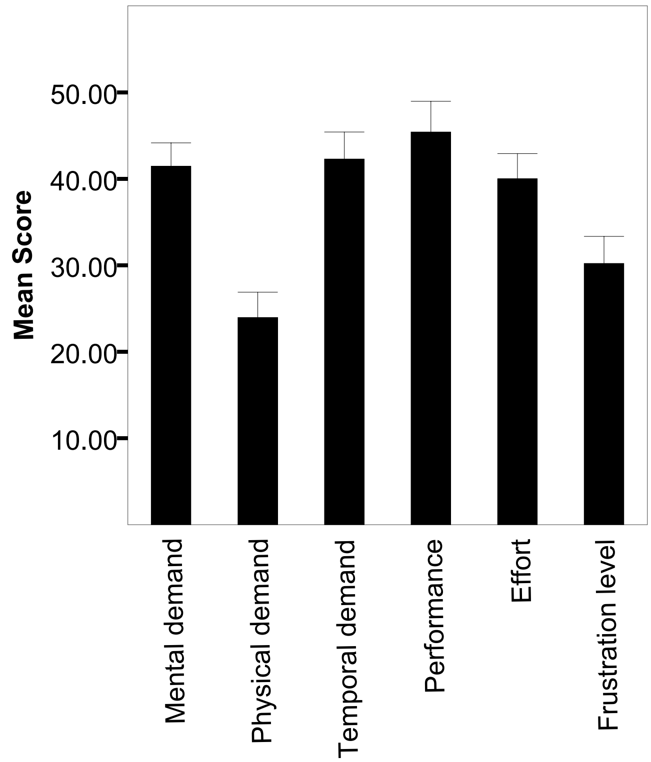

Figure 17. Mean NASA-TLX workload scores as a function of TLX dimension. Error bars represent standard error.

Table 15. Descriptive statistics for NASA-TLX dimension scores.

TLX Dimension	*M*	*SD*
Mental Demand	41.6	18.8
Physical Demand	24.0	21.4
Temporal Demand	42.4	23.2
Performance	45.5	25.1
Effort	40.0	21.4
Frustration Level	30.2	22.6

Table 16. Results of *t* tests between each pair of NASA-TLX dimension scores.

	Physical Demand		Temporal Demand		Performance		Effort		Frustration Level	
TLX Dimension	*t*(52)	*p*	*t*(52)	*p*	*t*(52)	*p*	*t*(52)	*p*	*t*(52)	*p*
Mental Demand	5.39*	< .001	-0.27	.787	-1.09	.281	0.62	.536	4.42*	< .001
Physical Demand			-4.85*	< .001	-5.57*	< .001	-4.67*	< .001	-1.84*	.071
Temporal Demand					-0.81	.420	0.93	.356	4.02*	< .001
Performance							1.57	.123	4.09*	< .001
Effort									3.57*	.001

*$p < .05$

DISCUSSION

Overview of Discussion

In the present study, we investigated the use of three-dimensionality as a potential display enhancement to facilitate the tasks of acknowledging and solving voltage violations across an integrated one-line diagram. For each scenario, the participants were required to acknowledge all overloaded transmission lines and subsequently solve the scenario by reducing all line loadings to 100% of capacity or less. Participants performed the tasks using one of three display types: 2-D numerical, 2-D graphical, and 3-D. In this section we provide a review of our hypotheses regarding the dependent measures of response time, accuracy, and workload and explain our relevant findings regarding such measures. We then touch on the possible limitations and applications of such findings and close with our overall conclusions of the study.

Findings Related to Hypotheses

Response times. Overall, response time results closely followed our initial hypotheses. As predicted, we did not find a significant effect of display type on acknowledgment time. The consistency and uniformity of violation presentation formats, coupled with the simplicity of the task, were likely the primary reasons for the lack of significant differences among the displays. We might have seen more significant effects had the violation symbols varied substantially in size or salience, or had the task simultaneously required the use of display navigation.

The hypothesis that solution times would decrease for both the 2-D graphical and 3-D groups was confirmed by our results. There are numerous reasons why such displays could potentially support the solution task better than the 2-D numerical display. Performance in such a task, requiring divided attention, parallel processing, and mental model development, is often best supported when integrated information is presented within graphical objects and integral formats (Carswell, Frankenberger, and Bernhard, 1991; Wickens, Merwin, and Lin, 1994; Wickens and Carswell, 1995; Liu, Zhang, and Chaffin, 1997), as was the case with both the 2-D graphical and 3-D displays. This finding is consistent with the proximity compatibility principle (Bennett and Malek, 2000; Rich, Wiegmann, and Overbye, 2001b; Wickens and Andre, 1990; Wickens and Carswell, 1995; Wiegmann, Rich, Overbye, and Zhang, 2001). During the solution task, multiple information sources such as current generator output, maximum generator output, and available reserves required parallel processing, and the integration of these information sources into a two-dimensional gauge or a three-dimensional cylinder allowed for this parallel processing to take place efficiently. Because precise judgments were not explicitly required to solve the voltage violations, the 2-D numerical display, which would normally improve performance in this type of task, was of lesser value.

Furthermore, the solution time advantage was, as predicted, enhanced for the 3-D display over the 2-D graphical display, indicating more than a mere graphical advantage. The increased size and salience of the cylindrical generator representations in comparison with the two-dimensional gauges, coupled with a reduced level of clutter in the 3-D displays, allowed for a clearer and more explicit presentation of relevant generator information. In addition, the larger representation of generator cylinders in the 3-D group resulted in more salient dynamic changes in power outputs and reserves, enabling operators to have an even greater understanding of the effects of individual generator adjustments on the entire system.

The development of emergent properties within the 2-D graphical and 3-D displays could have also led to lower solution times. For example, a *high reserve region*, a portion of the one-line diagram where substantial reserves are available, will appear in the 2-D graphical and 3-D displays due to the presence of multiple high-proximity generator gauges or cylinders, respectively, with large magenta components. Such an emergent feature could likely facilitate the solution task by drawing attention to the generator(s) containing adequate levels of available power (Remington, Johnston, and Yantis, 1992; Yantis and Hillstrom, 1994; Wickens and Hollands, 2000). This effect was again likely enhanced in the 3-D display due to the increased size of the generator symbols compared to the gauges in the 2-D graphical display.

Breaking the solution task into subcomponents of first adjustment time and mean adjustment interval revealed that first adjustment time was a clear contributing component to the aforementioned 3-D advantage in the solution task. This 3-D advantage for first adjustment time, an effect that was enhanced for single violation trials, can likely be attributed to a range of factors. First, presenting the one-line diagram in a 3-D pictorial format likely facilitated a more rapid and accurate mental model creation (Hollan, Hutchins, and Weitzman, 1984; Wickens and Hollands, 2000), as dynamic changes to the three-dimensional grid were graphically depicted and continuously updated as they occurred. This enhanced mental model could have allowed for less time "thinking" about the system and the correspondingly appropriate actions to take following the acknowledgment task and before the first generator adjustment. Also, the added salience of current power outputs and available reserves in the 3-D display could have further reduced first adjustment times by quickly drawing attention to high reserve regions. Finally, it is likely that the three-dimensional format was less overwhelming due to a reduction in the display clutter that was prevalent in both the 2-D numerical and 2-D graphical displays. This reduction in perceived clutter may have been further enhanced as operators in the 3-D display group interpreted the generator cylinders as coming out of the one-line diagram, rather than being embedded within it.

On the other hand, it was not clear if there existed a graphical advantage, a 3-D advantage, or a combination of both for the mean adjustment interval component of the solution task. Regardless of its classification, this effect was characterized by a significant advantage for the 3-D group over the 2-D numerical group and a trend in advantage of the 2-D graphical group over the 2-D numerical group, and was enhanced for multiple violation trials.

It is worth noting that, contrary to intuition, complex (multiple violation) trials resulted in faster first adjustment times overall. One possible explanation for this phenomenon is that more violations meant a greater proportion of generators that could help solve the problem, thus resulting in less search time to find an appropriate generator to adjust. Single violation trials meant a smaller proportion of generators that could help solve the problem, thus resulting in more search time to find the appropriate generator(s) to adjust. Higher complexity trials, however, did lead to increased adjustment intervals and overall solution times. Hence, although participants acted quicker following acknowledgment to begin adjusting generators in complex trials, solution times were dominated in such trials by participants spending, on average, more time between generator adjustments.

Accuracy. As predicted, the effects of display dimensionality on solution accuracy were limited. The results confirmed our expectation of a 3-D advantage in reducing the number of upper limit error events, as the number of sequences of one or more generator increases when the generator was already at its maximum was significantly less for the 3-D display in comparison with the 2-D graphical or 2-D numerical display. However, these results may be difficult to generalize to an actual control room setting due to the relatively small screen size of the current experiment. Being limited to a 20-inch single display likely prevented the graphical capabilities that wall-size one-line diagrams (see Figure 1) would enable. For example, such a limited screen size likely hindered participants' abilities to accurately assess how close a generator's actual output was to its maximum level. This limitation was likely enhanced in the 2-D graphical display due to the relatively small size of the generator gauges compared to the 3-D cylinders. However, even with this limitation in mind, one cannot ignore the significant 3-D advantage over the 2-D numerical display, which was enhanced for multiple violation trials.

It is difficult to point to a clear reason for the absence of a significant effect of display type on the number of lower limit error events. Although one would expect a graphical and/or 3-D advantage resulting from the increased salience of fully magenta gauges and cylinders to highlight generators with no output, this effect was negligible over all trials due to very few generators actually operating at or near minimum output levels at the beginning of each scenario.

As we predicted, no other measures of solution accuracy were affected by display type or dimensionality. This was likely because the choice of generator adjustment direction was supported solely by information gained from the moving arrows that were consistently presented across all displays. We may have seen notable improvements in the number of adjustment errors or other measures of accuracy had the 3-D displays strategically incorporated additional line loading information into the third dimension to enhance the operator's mental model of the system and the effects of individual generator adjustments on its overall state.

Workload. Subjective workload was predicted to follow a pattern similar to solution time, as the temporal and cognitive demands placed on subjects during the violation acknowledgment task were assumed to be negligible. Specifically, we expected that the absence of required mental math with the 2-D graphical and 3-D groups would lead to reduced levels of frustration, mental demand, and effort required. Contrary to our predictions, however, workload was not significantly different across display types. A possible reason for this is that the mental exercises (e.g., simple math) required for the solution tasks of this experiment were not sufficiently difficult in the 2-D numerical display to create significant differences in workload.

Limitations

As one would expect, there were limitations to the present investigation. As previously mentioned, a notable limitation was the size of the displays presented to the participants. In an actual power systems control room such as that in Figure 1, operators will often be provided with a combination of multiple personal computer displays and wall-size depictions of each one-line diagram. In the present setup, however, participants were limited to a single 20-inch monitor. Hence, display compression that could potentially hinder performance for 3-D displays in true control rooms was likely negligible in this experimental setup. In addition, the external environment of the operator did not closely mimic that which he or she would face in an actual power systems control room. Participants in the current study were able to, perhaps unrealistically, pay close attention to the tasks at hand with little distraction from their surroundings. Of course, external distractions (e.g., phone calls, interruptions, external noise) are difficult to mimic and monitor in a controlled setting such as that in the present experiment. Finally, an inherent expectancy effect likely resulted in the experiment, as participants had a clear understanding that violations were going to occur at relatively predictable times for each trial. Hence, long periods of vigilance while waiting for violations, common in true power systems environments (Harris and Chaney, 1969; Parasuraman, 1986; Wickens and Hollands, 2000), were absent in this study. Regardless of these limitations, however, the experimental setup used in this study was deemed feasible for initially evaluating three-dimensional enhancements for operational environments.

Possible Applications

Keeping in mind the aforementioned limitations, it is still clear that the advantages of three-dimensional displays over their two-dimensional counterparts can be quite significant. As this study has indicated, the added dimension can allow for non-distance information that was previously confined to alphanumerics to be presented graphically on standard display screens, aid in tasks requiring information integration through the creation of emergent features, facilitate a more accurate mental model of the system being manipulated, and allow for a better understanding of the interconnected nature and structure of a complex system such as an electrical power grid. However, our results do not indicate a clear advantage of three-dimensional displays for tasks that are extremely simple (e.g., the acknowledgment task). In addition, there may be no advantage of three-dimensional displays in terms of accuracy. Although our results indicated an advantage in reducing upper limit error events for the 3-D displays over both 2-D displays, this was the only accuracy measurement that was significant across display types and is a rather crude estimation of solution accuracy. We do, however, predict that with more elaborate future studies involving more complex and relevant accuracy measurements, we will indeed see an overall improvement with three-dimensional displays.

Also, contrary to our predictions, results did not indicate that workload could improve with three-dimensional displays. However, we note that workload was measured rather simplistically through the subjective rating system, and an objective method of assessing workload may be necessary for future experiments. We predict that such experiments will indeed indicate that workload can be reduced through the use of 3-D displays.

Overall, three-dimensional displays can be valuable tools in the process control tasks of monitoring and controlling sets of interrelated variables and the relationships between such variables. Specifically, high-level judgments of current operating levels in relation to upper and lower limits can be facilitated by depicting those levels as objects rising out of a two-dimensional display. Designers of complex systems can use this technique to explicitly present information in ways that were previously impossible with two-dimensional formats, leading to greater operator support and improved performance in such systems.

Conclusions

Operators today are being forced to identify and understand a massive and growing amount of multivariate data over extensive network regions. Visualization tools must thus be designed to aid these operators in managing such regions. This study indicated that three-dimensional displays can be used to support the tasks of fault detection and diagnosis across a subsection of the network, an integrated one-line diagram. The most significant 3-D advantage was a reduction in the time required to solve voltage violations. However, 3-D displays did not seem to aid in acknowledging these violations, nor did they appear to substantially improve accuracy or workload. However, we recognize that experiments involving more sophisticated and realistic displays, combined with more difficult tasks that require both focused attention and parallel processing, may provide more generalizable results regarding these dependent measures. Overall, our results did support the claim that three-dimensional displays can improve performance in tasks that require monitoring and controlling variables of complex systems, and we expect that future studies will provide additional evidence to confirm this claim.

ACKNOWLEDGEMENTS

This material is based on work supported by the National Science Foundation (NSF) through grant NSF DMI 00-60329, the Power Systems Engineering Research Council (PSERC), and the U.S. Department of Energy through its Consortium for Electric Reliability Technology Solutions (CERTS) program. Any opinions, findings, conclusions, or recommendations expressed are those of the author and do not necessarily reflect the views of NSF, PSERC, or the Department of Energy.

REFERENCES

Aretz , A. J. (1991). The design of electronic map displays. *Human Factors, 33,* 85-101.

Atchley, P., Kramer, A. F., Andersen, G. J., & Theeuwes, J. (1997). Spatial cuing in a stereoscopic display: Evidence for a "depth-aware" attentional focus. *Psychonomic Bulletin and Review, 4,* 524-529.

Barfield, W. & Rosenberg, C. (1995). Judgments of azimuth and elevation as a function of monoscopic and binocular depth cues using a perspective display. *Human Factors, 37,* 173-181.

Barnett, B. J. & Wickens, C. D. (1988). Display proximity in multicue information integration: The benefits of boxes. *Human Factors, 30,* 15-24.

Bellenkes, A. H., Wickens, C. D. & Kramer, A. F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation, Space, and Environmental Medicine, 68,* 569-579.

Bennett, K. B. (1993). Encoding apparent motion in animated mimic displays. *Human Factors*, *35*, 673–691.

Bennett, K. B. & Malek, D. A. (2000). Evaluation of alternative waveforms for animated mimic displays. *Human Factors*, *42*, 432–450.

Bennett, K. B. & Nagy, A. L. (1995). Spatial and temporal frequency in animated mimic displays. *Displays, 17,* 1-14.

Carr, K. (1995). Introduction. In K. Carr & England, R. (Eds.), *Simulated and virtual realities: Elements of perception* (pp. 1-9). London: Taylor and Francis.

Carswell, C. M., Frankenberger, S. & Bernhard, D. (1991). Graphing in depth: Perspectives on the use of three-dimensional graphs to represent lower-dimensional data. *Behaviour and Information Technology, 10*(6), 459-474.

Chau, A. W. & Yeh, Y. Y. (1995). Segregation by color and stereoscopic depth in three-dimensional visual space. *Perception and Psychophysics, 57,* 1032-1044.

Christ, R. E. (1975). Review and analysis of color coding research for visual displays. *Human Factors, 17,* 542-570.

Christie, R. D. (1994). Towards a higher level of user interaction in the energy management task. *Proceedings of the 1994 IEEE Conference on Systems, Man, and Cybernetics* (pp. 1086-1091). Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Cleveland, W. S. & McGill, R. (1985). Graphical perception and graphical methods for analyzing scientific data. *Science, 229,* 828-833.

de Azevedo, G. P., de Souza, C. S. & Feijó, B. (1996). Enhancing the human-computer interface of power system applications. *IEEE Transactions on Power Systems*, *11*(2), 646–653.

Dosher, B. A., Sperling, G. & Wurst, S. A. (1986). Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3-D structure. *Vision Research, 26*, 973-990.

Dyer, D. S. (1990). A dataflow toolkit for visualization. *IEEE Computer Graphics and Applications, 10*(4), 60-69.

Essenberg, G. R., Wiegmann, D. A., Overbye, T. J. & Sun, Y. (2003). *Motion as an effective flow visualization technique for power systems monitoring and control* (Tech. Report ARL-03-10/PSERC-03-1). Savoy: University of Illinois, Aviation Research Lab.

Furness, T. A., & Barfield, W. (1995). Introduction to virtual environments and advanced interface design. In W. Barfield & Furness, T. A., III (Eds.), *Virtual environments and advanced interface design* (pp. 3-13). New York: Oxford University Press.

Green, D. M. & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley (Reprinted in 1988, Los Altos, CA: Peninsula).

Gregory, R. L. (1970). *The intelligent eye.* Great Britain: McGraw-Hill.

Harris, D. H. & Chaney, F. D. (1969). *Human factors in quality assurance*. New York: Wiley.

Haskell, I. D. & Wickens, C. D. (1993). Two- and three-dimensional displays for aviation: A theoretical and empirical comparison. *International Journal of Aviation Psychology, 3,* 87-109.

Hollan, J. D., Hutchins, E. L., McCandless, T. P., Rosenstein, M. & Weitzman, L. (1987). Graphic interfaces for simulation. *Advances in Man-Machine Systems Research*, *3*, 129-163.

Hollan, J. D., Hutchins, E. L. & Weitzman, L. (1984). STEAMER: An interactive inspectable simulation-based training system. *The AI Magazine, 5*(2), 15-27.

Hollands, J. G., Pierce, B. J. & Magee, L. (1998). Displaying information in two and three dimensions. *International Journal of Cognitive Ergonomics, 2,* 307-320.

Kieras, D. (1992). Diagrammatic displays for engineered systems: Effects on human performance in interacting with malfunctioning systems. *International Journal of Man-Machine Studies*, *36*, 861–895.

Klump, R., Schooley, D. & Overbye, T. (2002). An advanced visualization platform for real-time power system operations. *Proceedings of the 14th Power Systems Computation Conference.* Zurich, Switzerland: Executive Board of the Fourteenth Power Systems Computation Conference.

Kobayashi, N., Okamoto, H., Tada, Y., Yamada, T. & Sekine, Y. (1998). New methods for visualization of dynamic behavior of power systems. *IEEE Transactions on Power Systems*, *13*(3), 999–1005.

Liu, Y. & Qiu, J. (1998). Visualization of power system static security assessment based on GIS. *Proceedings of the 1998 International Conference on Power System Technology* (pp. 1266–1270). Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Liu, Y. & Wickens, C. D. (1992). Use of computer graphics and cluster analysis in aiding relational judgment. *Human Factors*, *34*, 165–178.

Mahadev, P. M. & Christie, R. D. (1994a). Case Study: Visualization of an electrical power transmission system. *Proceedings of the IEEE Conference on Visualization* (pp. 379-381). Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Mahadev, P. M. & Christie, R. D. (1994b). Envisioning power system data: Vulnerability and severity representations for static security assessment. *IEEE Transactions on Power Systems*, *9*(4), 1915–1920.

McCormick, E., Wickens, C. D., Banks, R. & Yeh, M. (1998). Frame of reference effects on scientific visualization subtasks. *Human Factors, 40,* 443-451.

Mitsui, H. & Christie, R. D. (1997). Visualizing voltage profiles for large scale power systems. *IEEE Computer Applications in Power*, *10*(3), 32–37.

Mori, H. & Hayashi, Y. (1995). Visual interference with users' tasks on multiwindow systems. *International Journal of Human-Computer Interaction, 7,* 329-340.

Olmos, O., Wickens, C. D. & Chudy, A. (2000). Tactical displays for combat awareness: An examination of dimensionality and frame of reference concepts and the application of cognitive engineering. *The International Journal of Aviation Psychology, 10*(3), 247-271.

Overbye, T. J. (2000). Reengineering the electric grid. *American Scientist, 88*(3), 220-229.

Overbye, T. J., Klump, R. P. & Weber, J. D. (1999). A virtual environment for interactive visualization of power system economic and security information. *Proceedings IEEE PES 1999 Summer Meeting* (pp. 682-687). Edmonton, Canada.

Overbye, T. J., Sun, Y., Wiegmann, D. A. & Rich, A. M. (2002). Human factors aspects of power system visualizations: An empirical investigation. *Electric Power Components and Systems*, *30*, 877–888.

Overbye, T. J., Weber, J. D. & Laufenberg, M. (1999). Visualization of flows and transfer capability in electric networks. *Proceedings of the 13th Power Systems Computation Conference* (pp. 420-426). Zürich, Switzerland: Executive Board of the Thirteenth Power Systems Computation Conference.

Overbye, T. J., Wiegmann, D. A., Rich, A. M. & Sun, Y. (2001). Human factors analysis of power system visualizations. *Proceedings of the 34th Hawaii International Conference on System Sciences* (pp 647–652). Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Overbye, T. J., Wiegmann, D. A. & Thomas, R. J. (2002). *Visualization of power systems* (PSERC Report 02-36). Ithaca, NY: Power Systems Engineering Research Center.

Palmer, S. E. (1992). Common region: A new principle of perceptual grouping. *Cognitive Psychology, 24,* 436-447.

Parasuraman, R. (1986). Vigilance, monitoring, and search. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance. Vol. 2: Cognitive processes and performance* (pp. 43.1-43.39). New York: Wiley.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering.* New York: Elsevier.

Remington, R. W., Johnston, J. C. & Yantis, S. (1992). Involuntary attentional capture by abrupt onsets. *Perception and Psychophysics, 51,* 279-290.

Rich, A. M., Wiegmann, D. A. & Overbye, T. J. (2001a). *Usability evaluation of power system visualization software* (Tech. Report ARL-01-3/NSF-01-1). Savoy: University of Illinois, Aviation Research Lab.

Rich, A. M., Wiegmann, D. A. & Overbye, T. J. (2001b). *Visualization of power systems data: A human factors analysis* (Tech. Report ARL-01-5/PSERC-01-1). Savoy: University of Illinois, Aviation Research Lab.

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

Slater, M. & Usoh, M. (1993). Presence in immersive virtual environments. *IEEE Virtual Reality International Symposium* (pp. 90-96). Seattle, WA: IEEE Neural Networks Council.

Sollenberger, R. L. & Milgram, P. (1993). Effects of stereoscopic and rotational displays in three-dimensional path-tracing task. *Human Factors, 35*(3), 483-499.

Stevens, S. S. (1975). *Psychophysics.* New York: Wiley.

St. John, M., Smallman, H. S. & Cowen, M. B. (2000). Designing for the task: Sometimes 2-D is just plane better. *Proceedings of the IEA 2000/HFES 2000 Congress* (pp. 407-410). Santa Monica, CA: Human Factors and Ergonomics Society.

Theeuwes, J., Atchley, P. & Kramer, A. F. (1998). Attentional control within 3-D space. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1476-1485.

Theunissen, E. (1994). Factors influencing the design of perspective flight path displays for guidance and navigation. *Displays, 15,* 241-254.

Wickens, C. D. (1999). Frames of reference for navigation. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII* (pp. 113-144). Cambridge, MA: Bradford Book.

Wickens, C. D. & Andre, A. D. (1990). Proximity compatibility and information display: Effects of color, space, and objectness on information integration. *Human Factors, 32*(1), 61-77.

Wickens, C. D. & Carswell, C. 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* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.

Wickens, C. D., Liang, C. C., Prevett, T. T. & Olmos, O. (1996). Electronic maps for terminal area navigation: Effects of frame of reference and dimensionality. *International Journal of Aviation Psychology, 6*(3)*,* 241-271.

Wickens, C. D. & Prevett, T. T. (1995). Exploring the dimensions of egocentricity in aircraft navigation displays: Influences on local guidance and global situation awareness. *Journal of Experimental Psychology: Applied, 1,* 110-135.

Wickens, C. D., Todd, S. & Seidler, K. (1989). *Three-dimensional displays: Perception, implementation, and applications* (CSERIAC SOAR-89-01). Dayton, OH: Wright-Patterson AFB, Armstrong Aerospace Medical Research Laboratory.

Wiegmann, D. A., Essenberg, G. R., Overbye, T. J. & Rich, A. M. (2002). Motion in mimic displays: Effects on the detection and diagnosis of electrical power system failures. *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (pp. 1-5). Santa Monica, CA: Human Factors and Ergonomics Society.

Wiegmann, D. A., Rich, A. M., Overbye, T. J. & Zhang, H. (2001). *Effects of color contouring on the detection and diagnosis of voltage violations within electrical power systems* (Tech. Report ARL-01-10/SBC 01-263-DMI006032). Savoy: University of Illinois, Aviation Research Lab.

Williams, H., Hutchinson, S. & Wickens, C. D. (1996). A comparison of methods for promoting geographical knowledge in simulated aircraft navigation. *Human Factors, 38*(1), 50-64.

Yantis, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science, 2*, 156-161.

Yantis, S. & Hillstrom, A. P. (1994). Stimulus-driven attentional capture: Evidence from equiluminant visual objects. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 95-107.