THE NEED FOR INFORMATION VISUALIZATION DESIGN GUIDELINES

Leah Reeves
Kay Stanney, Ph.D.
University of Central Florida
Orlando, FL

LT Jim Patrey, Ph.D.
Robert Breaux, Ph.D.
Naval Air Warfare Center Training Systems Division
Orlando, FL

The current objective of Information Visualization (IV) research is to transform vast amounts of information into decision-supportable knowledge structures and patterns that capitalize on the way humans process information through perception and action. In general, the aim is to assist users in finding appropriate task information by presenting this information in a comprehensible manner through interactive computer graphics displays that present underlying relationships of concrete and/or abstract information in easily identifiable perceptual forms. New trends in IV are emerging, likely driven by the explosive growth of the internet, the computerization of business and defense sectors, the deployment of data warehouses, and use of virtual environment (VE) and advanced distributed learning (ADL) technology for entertainment, educational, and training applications.

Although many commercial tools are now available for creating visualizations for both concrete and abstract information, there are few, if any, theoretically-based or empirically-validated guidelines provided to developers or users regarding which technique(s) is most appropriate for a given domain context, user, or task. The objective of the present research is to identify the need to develop effective visualization design guidelines from the perspective of human information processing, visual display, and problem solving theories. Information processing theories suggest that to achieve comprehension of visual displays they must be developed such that they are readily perceived, interpreted, and acted upon. Visual display theories suggest that information should be organized and displayed in such a way that it is congruent with the methods in which one scans the environment. Research on integrating such theories with how people solve problems can be used to build principle-driven design guidelines that may assist visualization designers in successfully transforming information into the appropriate perceptual form for their users' domain task/goal(s). As a future objective of this research, critical issues to be addressed include determining: how to best characterize the existing knowledge of human perception and presentation design to develop theoretically-based visualization design guidelines; how to categorize and classify domain contexts, users, and tasks/goals; how to extend and augment principles developed for 2D visualizations to 3D; how immersion may enhance visualizations; and whether design principles will be generalizable or domain/task specific. Empirical testing and validation of future proposed design guidelines will be conducted in the shiphandling testbed of the Virtual Environment Training Technology (VETT) project, located at the Naval Air Warfare Center Training Systems Division (NAWCTSD). In this paper, the results of a preliminary study provide empirical evidence for the utility of visualizations in communicating human performance of an underway replenishment (UNREP) task.

Biographical Sketches:

Leah Reeves is pursuing her Ph.D. in Industrial Engineering from the University of Central Florida, with a focus on Human-Computer Interaction. She is sponsored by ONR, and her research is conducted at NAWCTSD in Orlando, Florida. Her undergraduate Honors-in-the-Major, Masters thesis research, and six years of professional experience have all been focused on human-factors issues associated with interactive systems design, including virtual environment research.

Dr. Kay M. Stanney received her B.S. in Industrial Engineering from SUNY Buffalo in 1986, after which time she spent three years working as a manufacturing/quality engineer for Intel Corporation in Santa Clara, California. She received her M.S. and Ph.D. in Industrial Engineering, with a focus on Human Factors Engineering, from Purdue University in 1990 and 1992, respectively. Dr. Stanney has over nine years experience in the design, development,

and evaluation of human-computer interactive systems. She is currently an associate professor in the Industrial Engineering and Management Systems Department at the University of Central Florida.

Robert Breaux earned the Ph.D. from Texas Tech University in Experimental Psychology, and he serves NAWCTSD as leader for the Virtual Training Technology Team.

LT Jim Patrey is a designated Aerospace Experimental Psychologist in the U.S. Navy. He earned his Ph.D. in Cognitive Psychology from the University of Illinois at Chicago.

THE NEED FOR INFORMATION VISUALIZATION DESIGN GUIDELINES

Leah Reeves
Kay Stanney, Ph.D.
University of Central Florida
Orlando, FL

LT Jim Patrey, Ph.D.
Robert Breaux, Ph.D.
Naval Air Warfare Center Training Systems Division
Orlando, FL

INTRODUCTION

In the past two decades, there has been a vast increase in the amount, types, and sources of information capable of being gathered and produced. Consequently, there is a definite need for the identification of information assimilation and management techniques to support users (i.e., the 21st century warfighters and their instructors) in knowledge acquisition and information modality changes using attention alerting mechanisms and intelligent task aids. Visualization, which focuses on the development of such techniques, is a particular area that has become an active research and development field in the last decade. Much of the initial research in visualization has been driven by the scientific community in its efforts to contend with analyzing huge volumes of scientific data (Gershon, Eick, & Card, 1998). New trends in visualization are emerging because of the explosive growth of the internet, the computerization of business and defense sectors, the deployment of data warehouses, and use of virtual environment (VE) and advanced distributed learning (ADL) technology for entertainment, educational, and training applications (Gershon et. al., 1998; Westphal & Blaxton, 1998; van Dam, 2000).

A specific domain that may benefit from the application of effective visualization techniques is military training. For example, improved visualization techniques may facilitate the understanding of patterns in data from large sets of simulation training results. Use of the appropriate visualization techniques will allow the information to be represented and manipulated in ways that are easy for users to understand and interpret. For example, techniques need to be focused on the goal-relevant attributes for the training domain and need to consider the types of information processing that the student (i.e., the future warfighter) or the instructor (i.e., the human analyst) will be required to do when analyzing the training results. Students may therefore identify strengths and weaknesses in their performance, while instructors may readily assess in which areas a student may need more practice. For example, interacting with an effective visualization of a student's shiphandling performance data may aid an instructor in identifying which particular maneuvers the student needs more practice. The instructor could then tailor the student's training regimen accordingly.

Helping students acquire spatial knowledge or how to visualize patterns or relations within non-visual information is another domain in which the improved effectiveness of visualization techniques may benefit military training. For instance, interacting with a VE-based visualization may aid shiphandling students in acquiring the abstract reasoning skills necessary to understand the Venturi effect and other relative motions associated with shiphandling. Trainee attrition rates could also potentially be reduced by utilizing effective visualizations to minimize fatigue and reduce error, which are often associated with information overload from poor visual display designs.

Furthermore, training tools employing principle-driven visualization guidelines may help reduce time and cost associated with military training system development and deployment by identifying a set of common design guidelines that may be applied intelligently and consistently across multiple training system domains. By decreasing the complexity and procedural inconsistencies currently associated with employing visualization techniques, improvements in developing effective military training systems may be realized through design simplification, procedural consistency, efficiency, and job aiding via intelligent human-centered design.

It should be noted that some visualizations may concern the perception and retrieval of multiple categorical elements (Robertson, Card, & Mackinlay, 1993; Stanney & Salvendy, 1994), while others may involve the graphical perception of multiple quantitative

information elements or variables (Bennett & Flack, 1992). Consequently, visualizations are currently being divided into roughly two areas: (1) scientific visualization that focuses mainly on physical data (e.g., the human body, earth, geographical landscape, molecules), and (2) information visualization that focuses on abstract, nonphysical data (e.g., hierarchies, system interdependencies, relative motion) (Mackinlay, 2000; Westphal & Blaxton, 1998). The term 'visualization' will be used herein when referring to both scientific and information visualizations.

THE ISSUE

In recent years, application developers and researchers have developed a diverse range of visualization tools and techniques suitable for various data and information types. Unfortunately, due to a lack of empirical evidence and systematic guidelines, visualization designers often encounter difficulties when choosing the most appropriate techniques for given applications or specific training tasks (Mackinlay, 2000). Such issues may specifically affect the military of the 21st century. For example, completion of time-critical and externally-paced missions in an accurate and timely manner will depend on the warfighter being able to start and launch multiple tasks and mission packages and monitor progress without having to deal with unnecessary information (e.g., manipulate, open, or close hundreds of windows and click on thousands of objects) (S&T Manning Affordability, 2000). Furthermore, many of these tasks may also involve the input and output of information via various sensory modalities (e.g., visual, auditory, tactile). Consequently, design simplification, procedural consistency, efficiency, and job aiding through intelligent design are necessary to the development of successful 'human-in-the-loop' man/machine system interfaces. The implication is that inefficient watchstation design and improper user training may hinder task or mission performance by inducing future warfighters to experience information overload, where one fails to identify or locate information that is vital to the successful completion of a task.

The future objective of the present study is to develop effective visualization design guidelines from the perspective of human information processing, problem solving, and visual display theories. It is proposed that such theories can be used to build principle-driven design guidelines that may assist visualization designers in successfully transforming information into the appropriate perceptual form for their users' application domain and task goals.

THEORY BACKGROUND

Many believe the common adage that "a picture is worth a thousand words," yet not all pictures are equally communicative. Information processing theories suggest that to achieve comprehension of visual displays they must be developed such that they are readily perceived, interpreted, and acted upon. By designing an interface that appropriately and consistently maps perceptual features of a display to the goal-relevant domain properties, the visual representation may support the development of an accurate mental model (Robertson, 1990, 1991). Thus, continued interaction with the visualization should facilitate information comprehension and streamline data access (Thüring, Hannemann, & Haake, 1995). For instance, considering that the cost of visual sampling (i.e., fixating on a particular area and extracting the pertinent information from it) a visual display is inversely proportional to the ease with which the display can be interpreted by the user, a number of factors must be accounted for in order to improve display interpretation by minimizing sampling costs (Ware, 2000), such as: points to be sampled for comparison should be spatially close (e.g., factors influencing the purchase of one house over another); items displayed intermittently as reminders should be done at appropriate intervals (e.g., people are much more reliable at monitoring information presented at one-minute intervals versus 20-minute intervals), and; only goal-relevant information should be displayed. There is a need to identify methods to aid visualization designers in identifying goal-relevant domain properties for particular domains and tasks.

Visual display theories suggest that information should be organized and displayed in such a way that it is congruent with the methods in which one scans the environment. For example, visual primitives such as texture, color, length, and slope are encoded more efficiently than shape, area, orientation, and containment (Lohse, 1997). Visual displays that differentiate by color may thus be more readily perceived than those demarcated via shape. Other examples for improving visual interpretation of computer generated information involves the use of effective graphics and display techniques. For example, the enhancement of edges is an important part of the techniques utilized by some artists to make certain parts of a painting more distinct than others (Ware, 2000). Similarly, this technique may be utilized to enhance key parts of images in a data visualization environment. Further, three-dimensional (3D) representations have been found to be differentially superior to 2D, with benefits in processing found with continuous as opposed to discrete data (Lee &

MacLachian, 1986) and with lateral and altitude tracking accuracy as opposed to airspeed (Haskell & Wickens, 1993). Designing visualization displays according to such theories and findings may facilitate problem solving by streamlining information processing and reducing cognitive demands on users (Woods, 1984).

Since external representations extend and alter the cognitive process (Zhang, 1997), people will solve problems with diagrams differently and possibly more efficiently than they would without diagrams. That is, visualizations may function as memory extensions that enable cognitive operations that would otherwise be impossible. Visualization displays may therefore be considered the conceptualization aids that employ graphics techniques to assist users in finding relevant data, visualizing domain semantics, and restructuring their view of the problem (Woods & Roth, 1988). This iterative process may be considered as a problem-solving feedback loop through which the visualization user, as analyst, forms hypotheses about the presented data or information and refines them via an interactive visualization process (Ware, 2000).

Recent research in identifying display techniques for facilitating problem solving involves the utilization of different display modalities (i.e., visual, audio, tactile) (Turk & Robertson, 2000). Mills and Noyes (1999) suggest that early studies support the notion that the multi-modal aspect of VEs may result in enhanced human performance. For instance, Breaux, Martin, and Jones (1998) suggest that a primary advantage of VE technology is that it enables a training environment to be manipulated by both the instructor and the student to facilitate the integration of knowledge into an appropriate mental model, thereby enhancing learning and improving performance. It is also generally accepted that dynamic or interactive visual displays (e.g., VEs) facilitate a user's understanding of complex relations between represented parameters (Kalawsky, 1993). Moreover, it is theorized that VE technology may be an effective visualization tool for teaching and learning because it makes better use of the human central nervous system's ability to handle a rich set of sensory inputs (e.g., sight, sound, touch) (Doyle & Cruz-Niera, 1999). There is a need to better understand the utility of extending visualizations into multi-modal presentations.

Auralization

Gaver (1997) suggests that data auralization can be used to represent multidimensional data as it exploits the ability of sound to form readily perceivable patterns, communicate distinctive figures one can associate with events, and provide feedback about user actions. Auditory systems consist of technologies that provide aural communication to users (e.g., speech recognition, 3D auditory localization, synthetic speech output) (Kalawsky, 1993). Since sound often plays a major role in real world perception, it may also enhance a user's interaction with a visualization environment. Virtual environment researchers, for example, aim to develop more sophisticated virtual acoustic presentations that may accurately resemble a user's real world aural experiences (Gabbard & Hix, 1997). Such aural feedback may include any sounds resulting from the user's own actions, other's actions, and natural, ambient, or other sounds.

If implemented effectively, sound may be used to enhance perception and improve user performance by increasing a user's physical and spatial awareness in interactive visualization systems like virtual environments. Sound in VEs may also be an effective sensory substitution, such as when no tactile feedback is available (Bowman, Kruijff, LaViola, Poupyrev, 2000). It may be useful in collaborative VEs by providing users with a sense of other's locations in the environment. However, if the human process of auditory spatial localization is not correctly accounted for by the aural system being implemented in VE systems, the concept of a true 3D audio localization system may not be realized (Kalawsky, 1993), potentially hindering user experience and task performance. Aural design guidelines are needed to ensure visualization designers and their subsequent users do not encounter such issues.

Krapichler et. al., (1999) indicate how the combination of visual and acoustic information can be both redundant and complimentary by providing stereo sound as acoustic feedback for different user actions (e.g., opening a 3D control panel, pressing a button, moving a slider). This type of feedback may enhance user performance because specific events can be located acoustically even when they may be hidden visually. Beyond localization, auditory cues could be used to communicate distance and direction, or to provide genre sounds that are associated with given events (Gaver, 1997). The temporal nature of sound could also be used to communicate data streams that vary over time. Aural “textures” have also been used in conjunction with visual textures to assist with the identification and exploration of data spaces (Grinstein & Smith, 1990). A better understanding of what sounds should be presented and how these sounds should be generated is needed. Kaur (1999) notes, for example, how few designers utilize specific strategies to implement sound due to insufficient guidance available to them. Accordingly, more research is needed to

determine the following: when sound may or may not be necessary for particular visualization applications; how to implement needed sounds, and; when and how sound may be used complimentary to other sensory information.

Some visualization applications may require feedback in addition to visual and aural (Bowman, Kruijff, LaViola, & Poupyrev, 2000; Gabbard & Hix, 1997; Kalawsky, 1993, 1999). Concerning VE applications, for example, providing actual force feedback when a virtual object is grasped may be necessary for ensuring successful completion of a visualization task (e.g., virtual surgery, molecule exploration, simulation of the Venturi effect). The next section addresses issues associated with providing tactile stimulation and force feedback (i.e., haptic feedback) within visualization applications.

Haptic Feedback

Gabbard & Hix (1997) and Durlach & Mavor (1995) assert haptic feedback (i.e., force and tactile presentation) may be provided by the following two means: (1) kinesthetic—information sensed through movement and/or force to muscles and joints, and (2) tactile—information received through nerve receptors in the skin (e.g., with a finger pad) which convey shapes and textures. Herein, the term 'haptic' will therefore refer to both tactile and kinesthetic presentation. It is generally accepted that when haptic output is provided both kinesthetic and tactile feedback should be incorporated, thereby enabling both senses to work together simultaneously and compliment one another (Gabbard & Hix, 1997). However, to design a visualization system that effectively incorporates haptic feedback, an understanding of the specifics as to how humans kinesthetically and tactilely perceive information is necessary.

To provide users with the appropriate ‘feel’ for the environment and the task being performed, it may be necessary to provide tactile feedback enabling users to feel resistance (e.g., temperature change, vibrations) when an object is touched (Kalawsky, 1993). Tactile perception is produced through stimulation of receptors in the skin and depends on cutaneous sensitivity, which includes the ability to detect thermal, mechanical, and electrocutaneous stimuli at the skin's surface (Stuart, 1996). With more than 10,000 parallel channels (receptors), a large amount of information may be processed by the tactile system to substitute for or to compliment visually and/or aurally represented domain attributes if the information is effectively presented (Kaczmarek & Bach-y-Rita, 1995). An understanding of the variation in tactile perception over different parts of the body due to varying receptor types and concentrations is necessary to design effective tactile feedback. For example, a human's ability to distinguish a two-point touch from a one-point touch varies drastically from the finger to the thigh (Boff & Lincoln, 1988). The two-point resolution varies from about 1mm at the fingertip to 60mm at the thigh due to these body parts having different types and concentrations of tactile receptors. This has direct implication to the specific body parts that may be utilized to incorporate tactile feedback within a visualization system. That is, if a user needs to be able to distinguish temperature or vibration differences between cues physically located 5 or 10mm apart, utilizing a tactile device for the hand and fingers would be more appropriate than one for the leg to interact with a visualization environment. Further determination of such differences in human tactile perception may lead to the identification of haptic primitives and an understanding as to how they are most readily perceived for specific types of tasks. Certain studies have already begun to discuss the application of tactile display to vigilance and tracking tasks (Sherrick & Cholewiak, 1986) and object identification and depth assessment tasks (White et. al., 1970). While et. al., for instance, found users could identify many common objects and assess their relative positions on a table using tactile display. More empirical research is needed, however, to begin classifying haptic primitives according to their affect on human performance.

Some aspects of a visualization application may require force feedback when a virtual object is grasped. As with tactile feedback, in order to effectively design such a visualization system it is necessary to understand the specifics of the human kinesthetic sense. Although the terms may be used interchangeably, kinesthetic perception (or proprioception) is perception of movement, position, and torque of the limbs and other body parts (e.g., a baseball player's perception of the motion of his arms as he bats or a tennis player's perception of the motion of a forehand stroke), while force perception depends on sensitivity to the muscular opposition that resists mechanical forces (e.g., the perceptions one has when pushing to open a door or grasping to lift a suitcase) (Stuart, 1996). Kinesthetic/force perception may be either passive or active and whole-body or part-body. Passive type refers to when your limbs are moved by an external force (e.g., by a machine or another person), and it involves only afferent information (i.e., the product of sensory input). Active type refers to when the movement is self-induced, and it involves both afferent and efferent information (i.e., the brain's transmission of neural impulses to produce motor movement). Although some of the same receptors may be involved,

different mechanisms are employed to detect the static position of limbs, their movement, and the force developed by a muscle from its respective inducement (Clark & Horch, 1986). For example, human fingers are more sensitive to force and pressure variations than other parts of the body. Shimoga (1993) reports fingers can detect a force step of 0.5 Newton or a pressure variation of 0.2 N/cm2. Such information has direct implication to how hand-based haptic devices may be utilized to incorporate kinesthetic feedback within a visualization system. If a user needs to be able to distinguish differences in the amount of force being applied to different fingers, for instance, the haptic device must provide pressure no less than 0.2 N/cm2 in any finger, with differences in force being at least 0.5 Newton. Understanding such determinants in human kinesthetic/force perception may lead to the identification of haptic primitives and allow classification regarding how these primitives are most readily perceived for specific types of tasks.

An initial step towards identifying haptic primitives may be to decompose haptic percepts into their salient features (Stuart, 1996). Rosenberg (1993), for instance, decomposes the percept of contact with a rigid wall into three stages: the initial dynamic surface contact, the quasi-static interaction with the hard surface, and the final dynamic release of the surface. He further associates the respective qualities of crispness, hardness, and cleanness with these stages, and shows that all three qualities must be effectively conveyed for the user to experience a convincing haptic percept of a wall. More empirical studies are needed to validate and extend Rosenberg's findings to further identify and define qualities of haptic primitives associated with tasks other than contact with a rigid wall. Most researchers speculate that the applications for which haptic displays will prove most valuable are those where a visual display of the data is difficult, complex force fields are important to performance, and the understanding of complex or abstract information is a goal (Stuart, 1996). For instance, the training of shiphandling tasks may be an application that will benefit from haptic display research because shiphandling involves skills which rely on a person's ability to understand complex force effects (e.g., forces generated by a ship's wake, it's hull moving through the water, and the wind and seas).

Experimental research results in the fields of robotics and VEs have shown the presence of haptic feedback increases task efficiency and accuracy in various situations, such as remote and/or virtual manipulation tasks (e.g., remote inspection), complex molecular docking tasks, and applications where other senses may not be usable (e.g., acoustically muted or dark environments) (Gabbard & Hix, 1997). More empirical research is needed to establish guidelines for incorporating haptic feedback into particular visualization applications. An implication is that the effectiveness of any type of multi-modal display must be assessed against knowledge of the full range of human sensory capabilities (Durlach & Mavor, 1995). Thus, advancing the state-of-the-art in multi-modal visualization systems will therefore require multidisciplinary expertise in a variety of areas (e.g., speech and hearing science, perception and vision, linguistics, psychology, signal processing, and pattern recognition) (Oviatt & Cohen, 2000).

A particular military training task that may benefit from presenting performance data via a multi-modal visualization environment is the underway replenishment at sea (UNREP) task, where one ship (i.e., ownship) receives supplies, food, fuel or other essentials from a replenishment ship (i.e., oiler) while both ships are moving alongside each other. This task is essential to military training because it influences the Navy's ability to project sea power over long ocean distances by allowing ships to sustain themselves at sea for long periods without land-based support (Stavridis, 1992), "The UNREP is the most commonly encountered evolution, and the manner in which it is executed is one of the basic yardsticks by which a ship's performance is measured…it can teach more about seamanship and shiphandling than almost any other evolution" (p. 107–108). The UNREP task involves a potentially large number of subjective and objective variables that may be recorded and utilized to assess performance. Unfortunately, so much data is collected during an UNREP training simulation that it may be unclearly presented to instructors, thereby hindering their assessment of a student's performance. To address this issue, an investigation of a preliminary visualization technique has been performed and is discussed next. The findings from this preliminary investigation will be utilized to develop future visualization design guidelines for presenting UNREP performance feedback to both instructors and students.

PRELIMINARY FINDINGS

The UNREP task involves several phases, each of which must be accomplished in a smooth and timely manner to avoid operation delays and/or damage to multi-million dollar ships or injuries to crewmembers. The *approach* phase is when ownship begins to move towards the oiler, and the *alongside* phase is where the ownship matches speed and maintains course with the oiler. During the *alongside* phase, the ownship connects to the oiler via phone and distance lines and tries to maintain a constant lateral separation, the ability

of which is dictated by wind, seas, and the pressure differential created around the hulls of the moving ships (e.g., the Venturi effect). Once the replenishment is complete and all transfer lines are disengaged, the *breakaway* phase begins. Here, the ownship must smoothly move away from the oiler with appropriate increases in speeds and separation angle to avoid crashing the stern into the bow of the oiler. Each of the phases entails the ownship's conning officer (i.e., the one giving commands to maneuver the ship) constantly monitoring relative alignment and separation distances and giving the subsequent appropriate commands to maneuver and/or maintain ownship on the correct course with the oiler. Therefore, being able to assess the shiphandling abilities of a conning-officer-in-training is of paramount importance to the Navy in determining which officers are ready to command ships at sea.

Recent attempts have been made to improve job efficiency and aid shiphandling instructors in evaluating student performance data for the UNREP task by minimizing the amount of information the instructor has to examine to effectively assess a student's performance. For instance, to keep instructors from having to replay a video of an entire UNREP simulation in order to review a student's performance for maintaining an UNREP course, visual diagrams of the student's UNREP track history (i.e., vertical and lateral separation distance between the trainee's ship and the replenishment ship) are utilized.

During the UNREP computer training simulation, performance diagrams are created as the computer plots, as XY coordinates of the trainee's ship (i.e., ownship) position relative to the oiler, the student's UNREP track history over time from start (e.g., waiting station, approach, alongside) to finish (i.e., breakaway). These course plots are then printed out as diagrams (see Figure 1 for a sample of the paper-based course plots), and the instructors evaluate the students' overall performance by rating the performance diagrams on a scale of 0 to 100. As seen in Figure 1, the paper-based visualization of UNREP task performance data incorporates the use of curves to represent the path of the trainee's ship course. While curvature is detected and processed via less efficient serial scanning, the UNREP visualization inherently has redundant distance (i.e., vertical and lateral separation) cues, which reduce the information processing load because they are detected and processed in parallel. Subject matter experts (SMEs) had previously determined the vertical and lateral separation distance between ownship and the oiler to be the most goal-relevant domain attributes for utilizing a single diagram to evaluate a trainee's performance of an UNREP task. Consequently, these attributes were utilized to create the course plots (e.g., Figure 1) from the vast amount of trainee simulation data, thereby reducing time and energy necessary for evaluator's to rate a student's UNREP performance.

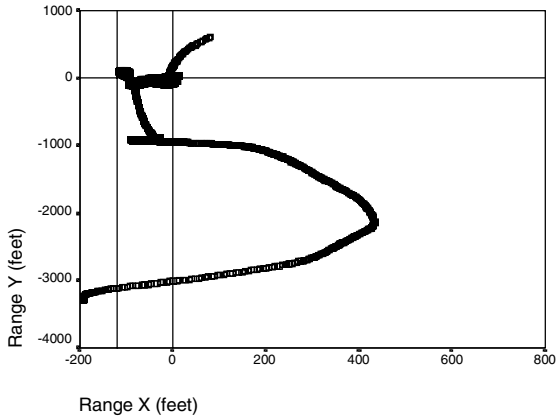

Figure 1. Sample visual diagram of a student's UNREP course performance to be rated by an instructor.

It may be argued that the single course plots can not compare to the richer portrayal given by video replay (see Figure 2). Therefore, to address this issue, NAWCTSD's Virtual Environment Training Technology (VETT) team performed a validation experiment to see how reliably evaluators could rate student's performance from the single paper course plot diagrams as compared to watching a video replay of a student's point of view during the UNREP simulation. The experiment involved a within-subjects, randomized design by having three SMEs each rate thirteen different paper course plots from thirteen different anonymous students and then rate the students' thirteen randomized video performances.

Figure 2. Snapshot from a video replay of a student's UNREP simulation during the *approach* phase.

Validation Results for UNREP Course Visualization

The correlation between the instructors' video and diagram ratings (0.87, 0.85, 0.85) were significant ($p < 0.01$), and the mean intra-rater reliability was 0.85 (SD = 0.01). These results indicated that the paper course plot diagrams are a valid, reliable measure for instructors to assess a student's UNREP performance. Given the excessive time required by instructors to view videotapes of an entire UNREP simulation run, rating diagrams of plotted UNREP course tracks appears to be a fairly robust and suitable visualization method for performance assessment. This preliminary finding provides empirical evidence for the utility of visualizations in communicating human performance of tasks such as the UNREP task.

The results of this validation experiment also provided anecdotal evidence regarding the potential for improving the effectiveness of the UNREP visualization. For example, SME's commented that the course plot diagrams (see Figure 1) did not clearly indicate the ownship's starting and finishing positions, and the UNREP phases (e.g., approach, alongside, breakaway) were not demarcated on the course plots. Other comments included: ship commands given by the students were not available and could therefore not be evaluated, the difficulty level of the UNREP scenario (e.g., sea state, time of day) was not available, and; students should be able to review similar visualizations combined with an instructor's feedback in order to examine in what areas they need to improve their performance. Based on these validation results and preliminary findings future studies will focus on developing and validating principle-driven visualization design guidelines for enhancing the UNREP performance feedback provided for both instructor and student assessment.

DISCUSSION

Some preliminary visual determinants, or goal-relevant domain attributes (i.e., curvature, vertical and lateral separation distances), of human performance have been identified and utilized to present UNREP simulation results to instructors. The success of the preliminary study in presenting a diagram of a student's UNREP track performance without having to show the entire UNREP simulation to the instructor for evaluation indicates much promise regarding transferring these diagrams into interactive computer visualizations. For example, providing a redundant color-coding scheme to represent things like changes in the UNREP phases could enhance this visualization. A visual representation of the angle between ships could also be provided, as both color and angles are readily perceived visual primitives. Moreover, the visualization could be augmented with multi-modal feedback (i.e., auditory, haptic) for improved human performance and performance assessment. For instance, multi-modal determinants of the evaluator's performance (i.e., how accurately he may rate a student's UNREP track performance) could include iconic mappings, such as: aural feedback representing the course commands given by the student at a particular point of the UNREP course, and/or aural feedback representing what the virtual helmsman repeated back to the student after giving a course command. Arbitrary auditory mappings could also be used to represent student command errors. For example, the pitch (or frequency) of a sound cue may be used to represent the severity of a student's command error, whereby the higher pitched tone may represent a more severe error and vice versa (Gaver, 1997). The implication is that such auralisations could compliment the course performance visualizations in presenting the complex training data and thus improve an evaluator's ability to efficiently assess performance. Similarly, to efficiently review their own performance, students may benefit from utilizing similar displays.

From a different training aspect, multi-modal determinants of a student's performance may include providing haptic feedback to a student's hands as he interacts with the oiler and UNREP ships to 'feel' the Venturi effect. For example, each hand could represent a ship and the tactile and kinesthetic/force feedback could be distributed along points of the opposing hands and fingers. Such feedback may facilitate a student's understanding of effects that occur when ships are alongside; hence, the student's UNREP performance may improve due to improved shiphandling techniques acquired from a better understanding of hydrodynamics.

Eventually, design and implementation guidelines are to be developed regarding which computer visualization techniques (e.g., static versus dynamic scenes, verbal versus graphic feedback, color versus gray tones, visual versus auditory versus haptic feedback, and combinations thereof) may be most appropriate for the UNREP task and generalizable to other applications and similar tasks. The efforts of this continued research may transition directly to training systems already in procurement. For instance, the Conning Officer Virtual Environment (COVE) training system is focused on training future conning officers to perform UNREP tasks, as well as basic shiphandling commands and maneuvers. The performance feedback provided to students after completing and UNREP scenario on the COVE system involves the computer’s visual display of a course plot diagram similar to the paper course plot diagrams presented earlier (see Figure 1). Therefore,

being able to identify goal-relevant domain attributes other than separation and alignment for the UNREP task, as well as knowing how to most effectively represent attributes visually and/or aurally, may lead to improved intuitiveness of the COVE system, thereby improving the ability of both students and instructors to review and evaluate performance feedback. This research may also transition to future VE training systems, such as those aimed at improving a shiphandler's ability to command the Landing Craft Air Cushion Vehicle (LCAC), which involves some shiphandling tasks similar to the UNREP task. The implication is that goal-relevant domain attributes identified for the UNREP task may be generalizable to tasks performed in the LCAC. Consequently, similar visualization techniques could be employed for similar tasks performed in these different training systems, hence facilitating design simplification and procedural consistency for military training systems.

Unfortunately, the potential of VE technology as a tool for creating effective training environments may not be fully realized until the following guiding principles are identified from the current and future state of knowledge in the visualization and VE fields: how to identify which attributes are most important for a given domain or user, how the attributes should be represented (e.g., via multi-modalities, 3D versus 2D), which interaction technique(s) is most appropriate, when an immersive or a non-immersive VE may be more effective, or when a traditional desktop environment may be sufficient. Without such knowledge and guidance, designers may produce perceptually ambiguous displays too complex, if not impossible, to interpret, thereby hindering rather than facilitating the user's decision-making process, abstract reasoning capabilities, and overall performance (Kalawsky, 1993). Visualization research, such as that proposed in this paper, is thus needed to exploit capabilities of VEs for presenting performance feedback to students and instructors.

CONCLUSION

A review of the literature has identified the following critical issues in visualization research:

- Principle-driven design guidelines are greatly needed for both visualization input (i.e., interaction techniques) and output (i.e., display representation).
- To begin identifying principle-driven visualization design guidelines, an initial starting point may be to determine how to take many seemingly similar tasks and decompose them into the fundamental lower level attributes that comprise them; then, identify which visual, aural, and/or haptic primitives may be used to efficiently (e.g., to allow users to detect them and organize them in parallel vs. serially) represent these goal-relevant domain attributes and which of these determined representations are generalizable across application domains and which are not.
- Different people may interpret visual images differently due to physiological makeup, previous experiences, and memories; thus, a comprehensive understanding of human perception and cognition is necessary for designing effective visualizations.
- Channels other than the visual (e.g., auditory, haptic, olfactory, heat, etc.) have not been fully explored regarding when one may be more appropriate or substituted for another in particular visualization applications, or how or when combinations of input and output modalities may be successfully integrated.
- Virtual environments may be viewed as a specific sub-set of multi-modal user interfaces and, therefore, visualizations utilizing VEs as a representation and interaction tool require a multidisciplinary approach to their design and implementation.
- Virtual environment technology is an effective yet unexplored tool for visualization.

The above issues must be addressed if principle-driven visualization design guidelines are to be established and if VE technology is to realize its full potential as a tool for creating effective visualizations across multiple domains and for various users. The research could benefit any discipline involved in teaching or training which involves information from either a physical or conceptual realm.

REFERENCES

Bennett, K. B., & Flach, J. M. (1992). Graphical displays: implications for divided attention, focused attention, and problem solving. *Human Factors, 34*(5), 513–533.

Boff, K. R., & Lincoln, J. E. (1988). *Engineering data compendium: Human perception and performance.* Wright Patterson Air Force Base, OH: AAMRL.

Bowman, D., Kruijff, E., LaViola, J., & Poupyrev, I. (2000). *The Art and Science of 3D Interaction.* Tutorial notes from the IEEE International Virtual Reality 2000 conference. New Brunswick, NJ, March 18–22.

Breaux, R. B., Martin, M., & Jones, S. (1998). Cognitive precursors to virtual reality applications. In *Proceedings of SimTecT* (pp. 21-26), Adelaide, S. Australia, March 2–6.

Clark, F. J., & Horch, K. W. (1986). Kinesthesia. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of perception and human performance* (13:1–13:62), Vol. 1. New York: Wiley.

Doyle, P., & Cruz-Neira, C. (1999). Virtual reality and visualization in education. *Syllabus: New Directions in Education Technology, 12*(9), 18–22.

Durlach, N. I., & Mavor, A. S. (1995). *Virtual reality: Scientific and technological challenges*. Washington, DC: Academic Press.

Gabbard, J. L., & Hix, D. (1997). A taxonomy of usability characteristics in virtual environments. http://csgrad.cs.vt.edu/~jgabbard/ve/taxonomy/ (2000, July 12).

Gaver, W. W. (1997). Auditory interfaces. In M. Helander, T. K. Landauer, and P. Prabhu (Eds.), *Handbook of human-computer interaction* (pp. 1003–1041). Amsterdam, Netherlands: North-Holland.

Gershon, N., Eick, S. G., & Card, S. (1998). Information Visualization. *Interactions*, *Mar/April,* 9–15.

Grinstein, G., & Smith, S. (1990). The preceptualization of scientific data. In E. Farrell (Ed.), Extracting meaning from complex data: Processing, display, interaction (pp. 190–199). *Proceedings of the SPIE*, Vol. 1259.

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*(2), 87–109.

Kaczmarek, K. A., & Bach-y-Rita, P. (1995). Tactile Displays. In W. Barfield and T. A. Furness III (Eds.), *Virtual environments and advanced interface design* (pp. 349–414). New York: Oxford University Press.

Kalawsky, R. S. (1993). *The science of virtual reality and virtual environments.* Wokingham, England: Addison-Wesley.

Kalawsky, R. S. (1999). VRUSE—a computerized diagnostic tool: for usability evaluation of virtual/synthetic environments systems. *Applied Ergonomics*, 30, 11–25.

Kaur, K. (1999). Designing virtual environments for usability. Unpublished doctoral dissertation. City University, London.

Krapichler, C., Haubner, M., Lösch, A., Schuhmann, D., Seemann, M., & Englmeier, K.-H. (1999). Physicians in virtual environments—multimodal human-computer interaction. *Interacting with Computers. 11*(4), 427–452.

Lee, J. M., & MacLachian, J. (1986). The effects of 3D imagery on managerial data interpretation. *Management Information Systems Quarterly, 10*(3), 257–269.

Lohse, G. L. (1990). Models of graphical perception. In M. Helander, T. K. Landauer, and P. Prabhu (Eds.), *Handbook of human-computer interaction* (pp. 107-135). Amsterdam, Netherlands: North-Holland.

Mackinlay, J. D. (2000). Opportunities for Information Visualization. *IEEE Computer Graphics and Applications, 20*(1), 22.

Mills, S., & Noyes, J. (1999). Virtual reality: an overview of user-related design issues. *Interacting with computers, 11*(4), 375–386.

Oviatt, S., & Cohen, P. (2000). Multimodal interfaces that process what comes naturally. *Communications of the ACM, 43*(3), 45–53.

Robertson, G. G., Card, S. K., & Mackinlay, J. D. (1993). Information visualization using 3D interactive animation. *Communications of the ACM, 36*(4), 57–71.

Robertson, P. K. (1990). A methodology for scientific data visualization: Choosing representations based on a natural scene paradigm. In *Proceedings of Visualization '90* (pp. 114–123). Los Alamitos, CA: IEEE.

Robertson, P. K. (1991). A methodology for choosing data representation. *IEEE Computer Graphics and Applications, 14,* 56–67.

Rosenberg, L. B. (1993). Perceptual design of virtual haptic sensations. *Proceedings of Virtual Reality Systems Fall '93 Conference.* New York: SIG-Advanced Applications.

S&T Manning Affordability (2000). Multi-modal watchstation. <http://www.manningaffordability.com/S&tweb/index.htm> (2000, February 12).

Sherrick, C. E., & Cholewiak, R. W. (1986). Cutaneous sensitivity. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of perception and human performance* (12:1–12:58), Vol. 1. New York: Wiley.

Shimoga, K. B. (1993). A survey of perceptual feedback issues in dexterous telemanipulation, Part I: Finger force feedback. *VRAIS '93: 1993 Virtual Reality Annual International Symposium* (263–270). Piscataway, NJ: IEEE.

Stanney, K. M., & Salvendy, G. (1994). Effects of diversity in cognitive constructing skills on human-computer performance. *Ergonomics*, *37*(4), 595–609.

Stavridis, J. CDR (1992). *Watch officer's guide: A handbook for all deck watch officers.* Annapolis, MD: United States Naval Institute.

Stuart, R. (1996). *The design of virtual environments.* New York: McGraw-Hill.

Thüring, M., Hannemann, J., & Haake, J. (1995). Hypermedia and cognition: Designing for comprehension. *Communications of the ACM, 38*(8), 67–66.

Turk, M., & Robertson, G. (2000). Perceptual user interfaces. *Communications of the ACM, 43*(3), 33–34.

van Dam, A. (2000). Immersive virtual reality for scientific visualization: A progress report. *Keynote speech at the IEEE Virtual Reality 2000 Conference,* New Brunswick, NJ., March 18–22.

Ware, C. (2000). *Information visualization: Perception for design.* San Diego: Academic Press.

White, B. W., Saunder, F. A., Scadden, L., Bach-y-Rita, P., & Collins, C. C. (1970). Seeing with the skin. *Perception and Psychophysics,* 7, 23–27.

Woods, D. D. (1984). Visual momentum: A concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies,* 21, 229–244.

Woods, D. D., & Roth, E. M. (1988). Cognitive systems engineering. In M. Helander (Ed.*), Handbook of human-computer interaction* (pp. 3–43). Amsterdam: Elsevier.

Westphal, C., & Blaxton, T. (1998). *Data Mining solutions: Methods and tools for solving real world problems.* New York: John Wiley & Sons.

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science, 21*(2), 179–217.