Chapter 4 – A SUMMARY OF RESEARCH CONDUCTED ON THE COMMAND VISUALISATION TESTBED

J.G. Hollands and Jocelyn M. Keillor
Defence R&D Canada – Toronto, Canada

4.1 INTRODUCTION

4.1.1 Background

NATO Research Task Group IST-021/RTG-007 interpreted visualisation as a human activity supported by technology. Visualisation is therefore a means by which humans make sense of complex data. The RTG considered visualisation technologies, including display devices and techniques in relation to how they help humans to perform their tasks effectively. The RTG emphasized the human use of the computational subsystem in ensuring that the right information is available in the form and at the time needed.

An Action Item was assigned to the IST-021/RTG-007 Canadian Point-of-Contact, requesting a section for the Final Report on progress with the Defence R&D Canada – Toronto (DRDC Toronto) Command Visualisation Testbed (CVT). As originally conceived, existing testbed facilities at DRDC Toronto were to be used to examine several prototype command-and-control related systems involving advanced visualisation concepts developed in the various NATO nations. These included: NORCCIS (Norway), Master Battle Planner (MBP) (UK), and SICCE (Portugal). However, despite repeated attempts it was not possible to obtain working versions of any of these systems. Thus, evaluation of the various national systems could not be performed.

Nonetheless, work relevant to IST-021/RTG-007 was conducted in the DRDC Toronto visualisation laboratory, funded under the auspices of the DRDC Technology Investment Fund (TIF). This work involved the empirical evaluation of a wide range of visualisation display concepts (detailed below). This chapter will describe that work. Implicit in the approach is the assumption that the effectiveness of visualisation display concepts is a question best assessed empirically – that is, by having human observers attempt to make judgments with different display arrangements, and collecting performance data from them as they do so. While this approach is not without its shortcomings, it offers an understanding of the factors that affect human judgment with visualisation systems.

4.1.1.1 Testbed History

Command and control systems need a single, intuitive, common tactical picture that displays the necessary information to increase battlespace awareness [1]. However, tactical displays provide poor support for battlespace visualisation [2] and for the cognitive tasks of commanders [1]. In response, the Canadian Department of National Defence (DND) has identified leading-edge technologies for information management [3] as a focus for DRDC, and information and knowledge management for decision making in a complex environment as one outcome of the research programme.

There is a clear trend towards military systems that handle more and more data, while at the same time there is pressure to reduce the number of personnel in operational systems. Therefore there is a need to produce more effective information management systems. The Naval Command and Control Way Ahead [4] notes that, “Display technologies and interfaces reflect more the technical limitations of the era in which they were designed rather than the true requirements for battlespace visualisation and decision support. Information is presented in a manner that maps inefficiently onto human perception and reasoning processes. … little support is provided for cognitive tasks throughout the decision cycle, including planning, intelligence preparation of the battlespace, course of action development and selection, mission rehearsal and simulation." The report further lists the development of "advanced approaches to battlespace visualisation techniques for the naval tactical environment" as an R&D goal.

Further, to augment communication across various command clusters, commanders have a need for a Common Operational Picture (COP). The COP has been defined as the integrated capability to receive, correlate and display heterogeneous sources of information in order to provide a consistent view of the battlespace. A fundamental problem with the communication process for collaborative working is that of the co-ordination of views on an information space, including geographic terrain.

A common finding in research on graphical displays is that the effectiveness of various graphical formats is task dependent [5,6]. Thus, for the range of tasks involved in command and control situations, an effective display must allow users to transition smoothly from one format to another. Typically, proposals for future battlespace displays show a map or chart background with relevant data superimposed in appropriate geographic locations, either graphically (line, bar graphs, timelines) or in tabular form. Some displays show the scene in geometric perspective (God's eye view) and use techniques and algorithms developed for scientific and information visualisation. However, the underlying human factors of using such displays are not well understood and their effective contribution to military command tasks has not been shown.

4.1.2 Purpose and Scientific/Technical Objectives

In response to these client needs, research scientists in the Human-Computer Interaction (HCI) Group at DRDC Toronto proposed the development of the CVT, so that experiments measuring human performance in representative tasks could be conducted. CVT therefore provides the capability to investigate whether proposed visualisation algorithms, constructs, and display concepts are consistent with human perception and cognition and whether they improve command decision making.

Specific scientific objectives included gaining a better understanding of techniques that reduce disorientation when an observer makes quick shifts from global to local views of the battlespace, developing methods for reducing perceptual bias in magnitude judgments of graphical battlespace elements, examining sets of mental operations used in command visualisation in order to maximize performance efficiency, and developing methods for establishing how much information is available from a display "at a glance".

4.1.3 Approach

The CVT includes state-of-the-art graphics workstations, display hardware, three-dimensional (3D) graphics software, and two eyetrackers. The facility was built with the aim of initiating a multi-experiment research program to determine how to relate human perception and reasoning processes to the elements of a command battlespace display. Using a range of military tasks, the collaborators investigated four general themes.

1) *Frame of reference and visual momentum.* The utility of two-dimensional (2D) and 3D displays to depict terrain information has been extensively investigated and the results indicate that the effectiveness of such displays depends on the judgment task. Since the task changes with context, the commander may need a variety of displays to accomplish various ends, leading to the need to switch displays periodically as tasks shift. Problems include disorientation and the need to mentally perform spatial transformations when transitioning from one format to another. Various visual

momentum [7] techniques are available to help commanders transition between two-dimensional (2D) and three-dimensional (3D) displays. This includes techniques such as smooth rotation and tethering.

2) *Perceptual bias and reference points.* Human judgments of the geometric volumes and areas that are commonly used to depict quantitative values in 3D data representations in statistical graphs and maps are biased [8]. Previous work has shown that reference points can reduce judgement error in graphical displays such as those used in command visualisation systems. Two further questions were investigated:
    a) The effect of response method on perceptual bias, and
    b) Bias engendered by the use of perspective rendering in 3D displays.
3) *Modeling mental operations.* Follette [9] proposed that two factors affect quantitative judgments with graphs:
    1) The number of operations necessary; and
    2) The effectiveness of the perceptual features used as input for the operations.

    Empirical work was conducted to determine the effects of these factors on error in quantitative judgments made with graphical displays.
4) *Visual attention and visual span.* Even when a tactical display accurately depicts all relevant data, the human observer may not attend to all displayed elements. The effectiveness of different symbologies to provide relevant tactical information "at a glance" was demonstrated using a change-blindness paradigm. Using an eyetracker, a gaze-contingent display can be constructed that depicts only that region of the display upon which the observer fixates. By varying the size of that region using a staircase procedure, the result indicate how much display information can be attended "at a glance" by a human observer.

More specifically, each research theme provides an improved understanding of how information processing in the command visualisation context can be improved. Results from Theme 1 (Frame of Reference and Visual Momentum) identified effective techniques for transitioning from specific (immersed) to general (world-centred) reference frames, or vice versa. Theme 2 (Perceptual Bias and Reference Points) results provided improved understanding of the selection of display variables to code different kinds of battlespace and command information, and of the utility of reference points to help reduce judgment error for dimensions used in 2D maps and 3D terrain imagery (e.g. area, volume). Theme 3 results (Modeling Mental Operations) provided better understanding of the nature of information processing by testing a framework of mental operations in the command context. Theme 4 (Visual Attention and Visual Span) demonstrated methods for increasing the speed and detectability of tactical display targets. In sum, the results obtained from the proposed work will generally enhance the efficiency of information processing using command visualisation systems.

4.1.4 Delivery Methods/Preliminary Data

Empirical data collection was conducted with the CVT at DRDC Toronto. Experiments used the subject population available to DRDC Toronto (i.e. military and civilian staff, student research assistants). Performance was quantified through the use of time and accuracy measures. Some subjective measures were also recorded (e.g. ability, workload, and preference data, but are not reported here for brevity). Results from the research have been demonstrated to the client group for potential incorporation into future programs, to researchers at other DRDC Centres, and to scientific and military personnel from various NATO and TTCP nations. Prototypes of visualisation techniques and a general purpose 3D graphics engine were developed. The results have been disseminated through presentations at peer-reviewed conferences and workshops, and submitted for publication in scientific journals. This chapter provides a summary of the work, and treats the results from each research theme in turn.

4.2 THEME 1: FRAME OF REFERENCE AND VISUAL MOMENTUM

4.2.1 Theoretical Background

The intent of the COP is to provide a shared understanding of the battlespace to improve responsiveness and provide decision dominance. Visualisation technology offers a means to establish the COP and should help the commander transition across strategic, operational and tactical levels. The ultimate aim is to obtain an integrated visualisation environment where commander and staff can gain a shared understanding of the changing battlefield situation.

Current technology provides a means to display tremendous quantities of data to the human commander. Geospatial data, sensor data, network data – electronic and human, socio-political data, data on troop status, materiel, data from news media, and so on. Further, the nature of the future operational environment requires a high degree of flexibility and adaptability, due to factors such as: asymmetric threat, enlarged areas of operation; non-contiguous and non-linear operations; requirements for a three-block war, use of complex terrain, and effects-based operations within a maneuverist approach [10].

To address this state of affairs, it is argued that the future commander can only be successful if the command team functions collaboratively. A rigidly hierarchical command structure is too slow and inflexible to respond in a timely manner to the nature of the asymmetric threat and resulting warfighting characteristics. There is a need therefore, to provide an integrated command and control system that supports collaborative working. This includes the design of computer software and displays to facilitate the collaborative working concept.

While co-ordinating information is shared across different echelons, commands, environments, government departments, and nations, the information available varies across such organizations, and is often represented and portrayed in different ways. A fundamental problem with the communication process for collaborative working is that of the co-ordination of views on an information/knowledge space. For instance, if a shared geospatial awareness is required – the platoon commander with troops on the ground looking at a group of buildings versus the company commander examining a 2D plan view (aerial photographs, maps) of the same urban terrain – it can be difficult for one commander to communicate to the other. In one view, task-relevant information may be visible: in the other, invisible. Is the company commander aware of what is visible to the platoon commander? What is left and right in the forward field of view (FFOV) may be reversed with the map depending on orientation. If a shared understanding of network access data is required (e.g. a complex set of intrusion detection data) two analysts in different locations may have access to only limited views on the data, but need a common representational format (e.g. a dynamic 3D graph) to communicate.

Hollands, Lamb, and Keillor [11] argued that such problems are related to the frame of reference concept [12,13,14] and present a framework for that concept that classifies factors that improve or degrade performance when co-ordinating information across views of spatial data. They also considered similar display concepts from information visualisation in human-computer interfaces (depicting file structures, networks, the web, windows and other interface elements; see [14]) and from display design guidelines from other domains (e.g. process control, medical imaging), and note the fundamental similarities. In particular the relevant literatures underline a recurring need for depicting both global context and local content, which leads to the need for multiple displays. Further, since different viewpoints have various advantages and disadvantages for various operational contexts, there is a need to provide methods for improving visual momentum across displays.

Hollands et al. [11] note how such methods can be split into two basic types: those that allow both views to be shown simultaneously (compromise displays) and those that ease the abrupt shift between global and local views by showing the mapping between display elements in different views (transition displays). Hollands et al. [11] also presented a preliminary taxonomy that lists the factors that distinguish egocentric from exocentric displays. This includes the distance and angle of elevation of the viewpoint with respect to a visual scene (Level I), the distance and angle of elevation with respect to a particular object of interest within the scene (Level II), distance and viewing angles with respect to the motion of an object or objects of interest within a scene (Level III), and the distance and angular compatibility between the viewpoint and the motion of a controlled object within the scene (Level IV). In the Level IV situation, the control order dynamics and the rigidity of the link between the viewpoint and controlled object are also of concern.

The Hollands et al. [11] taxonomy can be used to predict performance: as the number of shifts in reference frame increases, increased time and error are predicted. This is because each shift is essentially a transformation in reference frame; each transformation requires time and there is some likelihood of error if the computation required is not performed accurately [15,16,17]. This would be true for communication of shared understanding among collaborating workers, or for shifts in viewpoint over time for a single individual. The classification should lead to the appropriate use of display techniques to help the commander minimize mental effort, maintain good situation awareness, and improve communication across different levels of command, coalition partners, and public or private agencies.

We can also consider how multiple sensors can provide a new interface for controlling and monitoring platforms and their surrounds. Sensor data streamed from a moving vehicle integrated with data from other nearby sensors will provide remotely operated vehicle (ROV) operators and their commanders with reconstructed 2D and 3D visual representations of complex terrain. Virtual or augmented reality visualisations allow the use of viewpoints different from that “out the window”. Designers are then faced with the choice of optimal viewpoint parameters that maximize human performance.

It is widely accepted that the nature of the task dictates the best viewpoint on geographic terrain. Tasks involving shape understanding are best performed with 3D viewpoints because all three dimensions are integrated into one representation. 2D viewpoints are best for precise tasks judging relative position, due to the distortions associated with 3D viewpoints [18]. For navigation and wayfinding, local guidance is best performed from an egocentric perspective, while global spatial awareness tasks should be carried out with an exocentric, fixed viewpoint showing most or all of the terrain [14].

The implication, then, is to provide the appropriate viewpoint for the relevant task. However, as we have noted, switching between multiple displays is disorientating and leads to the need for spatial transformation. Each added transformation requires extra processing time and increased likelihood of error. We describe here two approaches we have taken to this problem. One involves the use of a tethered display, which serves as a compromise display incorporating both egocentric and exocentric elements. The other examines the utility of smooth rotations between display formats, serving as a transition display. Both methods, we argue, provide good visual momentum [7] between egocentric and exocentric display formats.

4.2.2 Tethering

The tethered viewpoint [19,20] is commonly used in computer gaming and couples the viewpoint to the position and orientation of a moving object (or *avatar*). The viewpoint is typically higher than the avatar and behind it, showing more of the terrain than would be seen from the avatar's viewpoint. In this sense, it provides visual momentum by providing a view that incorporates egocentric and exocentric qualities. There is some evidence in support of the tethered concept: Wang and Milgram [21] found that a tethered display produced better performance than an egocentric display for aerial navigation, and Wickens and Prevett [20] found advantages with a tether-like display (versus egocentric) for spatial awareness.

However, a rigid tether violates the principle of motion compatibility. As the operator directs the avatar to the right, the visual scene moves to the left. The rigid tether also behaves like a compensatory tracking system [19]. Modelling the tether as a dynamic mass spring damper system creates a display incorporating both compensatory and pursuit tracking attributes and may reduce the motion compatibility problem. However, Wang and Milgram [21] did not find any advantage to the dynamic tether over a rigid tether for six-degrees-of-freedom (DOF) aerial navigation.

Ground-based navigation differs from aerial navigation in that the operator is only controlling one rotational DOF (yaw). (Although pitch and roll vary when driving, they are typically not under the direct control of the operator). Yu and Andre [22] used an arcade driving simulator to evaluate four different viewpoints. A tethered viewpoint slightly removed from the vehicle provided the best navigation and situational awareness. However, this driving task differs in several fundamental ways from off-road navigation in complex terrain.

Using the DRDC Toronto CVT, Lamb and Hollands [23] examined the effect of viewpoint on controlling a vehicle navigating complex terrain, and on concurrent and subsequent spatial awareness. Five viewpoints were used (see Figure 4-1): egocentric, rigid tether, dynamic tether, 3D exocentric (perspective view on terrain, but viewpoint does not change with vehicular position), and a 2D exocentric map (God's eye view). Lamb and Hollands predicted that navigational performance would be better with the egocentric than either exocentric view because the egocentric display provides well-learned egomotion cues commonly available as one navigates through an environment, either in vehicle or on foot [14]. Lamb and Hollands also predicted that the tethered display would be more effective than the exocentric displays for navigation, and as effective as the egocentric display. Finally, they predicted that navigational performance would be better with dynamic than rigid tethering, reducing the motion compatibility problem and allowing pursuit tracking.

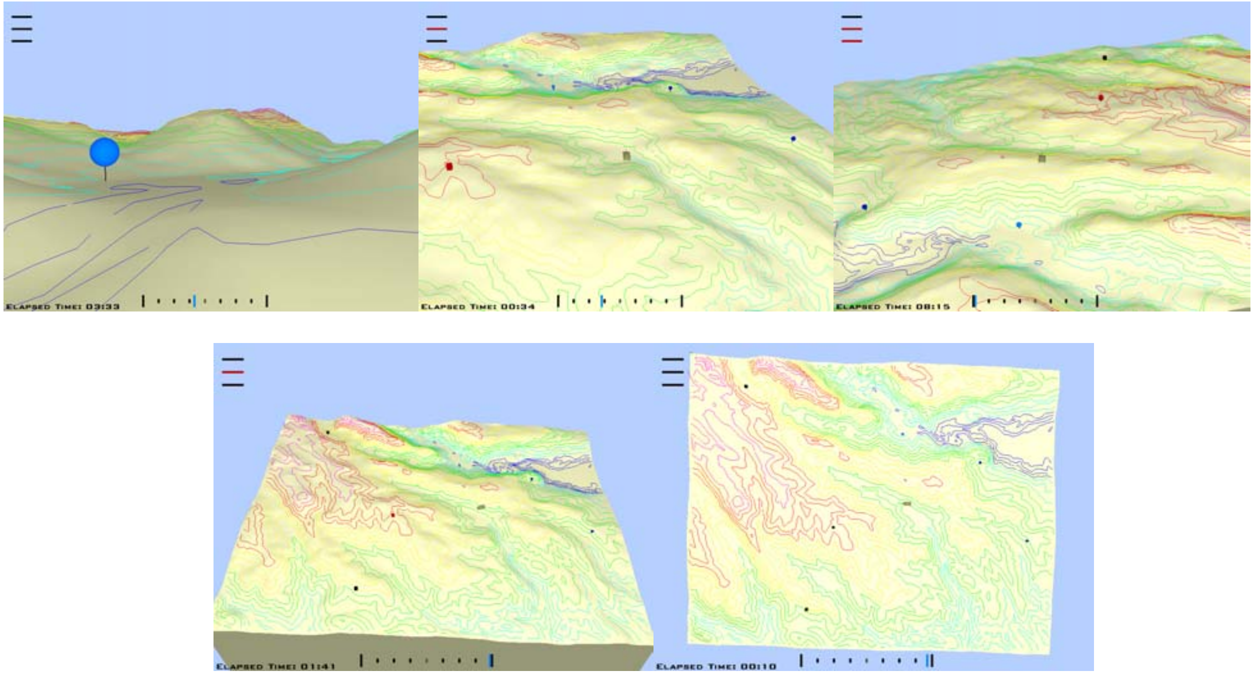

Figure 4-1: Example Views of each Display Type. Top row (left to right): egocentric, dynamic tether, rigid tether. Bottom row: 3D exocentric, 2D exocentric.

In the Lamb and Hollands [23] study, participants were instructed to avoid being seen by enemy units while navigating the simulated vehicle between waypoints. This was done to assess spatial awareness during navigation. The time that the tank was seen from at least one of the enemy positions was recorded. After each trial, observers had to choose the terrain just navigated from a set of distractor terrains. Given that both of these spatial awareness tasks require a sense of the global characteristics of the terrain, Lamb and Hollands predicted the opposite order of effectiveness: exocentric worse than egocentric, with the tethered display as effective as the exocentric display. They did not predict any effect for dynamic (versus rigid) tethering for these tasks.

Consistent with predictions, Lamb and Hollands [23] found that the egocentric display was more effective than exocentric displays (2D or 3D) for navigation, and the exocentric displays were more effective than egocentric for spatial awareness, both for time seen during navigation and the recognition task. The tethered displays generally produced intermediate results. For navigation, they were less effective than the egocentric display and roughly equivalent to the exocentric displays. For spatial awareness recognition the tethered displays were more effective than the egocentric display, but less effective than the exocentric displays.

More importantly, the tethered displays were the most effective displays for spatial awareness for minimizing time seen. Not only were the tethered displays more effective than the egocentric display, they were also more effective than the exocentric displays. Use of the tether minimized the time during which the participant's avatar was visible to enemy positions. There was no effect of tether dynamics in the navigation task.

In summary, the tethered display was useful in spatial awareness involving knowledge of locations of interest with respect to one's own position while navigating. In this sense, perhaps the Lamb and Hollands [23] results identify a navigation task whose performance is dissociated from conventional exocentric spatial awareness and egomotion. The tethered display may be the most effective display for this type of egocentric spatial awareness task.

4.2.3 Visual Momentum and Smooth Rotation

Background. In a series of experiments using the DRDC Toronto CVT, Hollands and co-workers [24-26] have examined the utility of *visual momentum* for the depiction of geographic terrain, a topic that has received relatively little empirical attention despite its clear importance. The concept was first defined as such by Hochberg and Brooks [27], who described techniques used in film to help an audience maintain spatial understanding of a scene across discrete film cuts. Woods [7] extended the visual momentum concept to user-computer interaction, and defined it in that context as the user's ability to extract and integrate data from multiple consecutive display windows. More recently, Wickens and Hollands [14] summarized four basic guidelines for improving visual momentum: consistent representations; graceful transitions; highlighted anchors; and world maps. Various specific approaches to improving visual momentum have been proposed or implemented [28-31], and examined empirically [29,32-36].

Hollands and co-workers were particularly interested in the problem of the depiction of geographic terrain for command and control. This is an important component of battlespace visualisation systems [37]. As described earlier, there is benefit to providing multiple views on terrain, and therefore both 2D and 3D display formats should be made available to the observer. In the command and control context an observer needs to switch tasks frequently while viewing geospatial information, leading to spatial disorientation and the need for spatial transformation. A gradual transition between 2D and 3D perspectives (and vice-versa) incorporating animation of viewpoint during task switching may provide visual momentum and alleviate the problem.

Hollands and co-workers were interested in the question of whether smooth transition aided observers as they switched tasks. To examine this question, they used two tasks developed by St. John et al. [18]. A shape understanding task required the participant to judge whether one ground location was visible from another (*A-See-B* Task), and a relative position task required a judgment of which one of two points was of higher altitude (*A-High-B* Task). St. John et al. found that the A-See-B task was performed better with a 3D display, whereas the A-High-B Task was performed better with a 2D topographic map.

Hollands and co-workers had participants switch tasks across trials to determine whether knowledge of terrain obtained when performing one task affected performance on a different task on a subsequent trial. In a *continuous transition* condition, the display rotated in depth from one display format to the other. The design of the control condition (discrete transition) varied across experiments. In Experiment 1 [24], a blank screen was shown for a duration equal to that used for the continuous transition. Participants were immediately shown the alternate display format in Experiment 2 [25]. In Experiment 3 [26] the 3D display was oriented to be aligned with the A-B vector. For the continuous rotation condition, the terrain was rotated in its azimuth prior to the rotation in depth. In the discrete case, the azimuth rotation was in the opposite direction and then the terrain "snapped" to the final orientation to examine the role of preview during transition.

For all experiments, it was predicted that the continuous transition would improve second-trial performance relative to discrete transition. This is because the smooth rotation should provide improved visual momentum between consecutive displays. We provide only a synopsis of the three experiments here; for details of the experimental methods, the reader should consult the original articles.

In general, terrain models were created from Digital Terrain Elevation Data (DTED) using Creator/TerrainPro [38] modelling tools. Each model represented a 13351 m x 11288 m region of the US state of Wyoming.

2D and 3D displays were constructed to resemble those used by St. John et al.[18]. The Vega visual simulation system [39] was used to render each terrain model as a 3D display, and an example is shown in Figure 4-2. The 3D display depicted the terrain model at a viewing angle of 45 degrees with respect to the ground plane. MICRODEM (Microcomputer Digital Elevation Models) [40] was used to create a 2D display with coloured contour lines (see Figure 4-2 for an example). 2D and 3D displays depicting each of the 8 A-B pairs were constructed, resulting in 16 different displays per terrain model. Each location in a pair was represented by a point superimposed on the map labeled A or B.

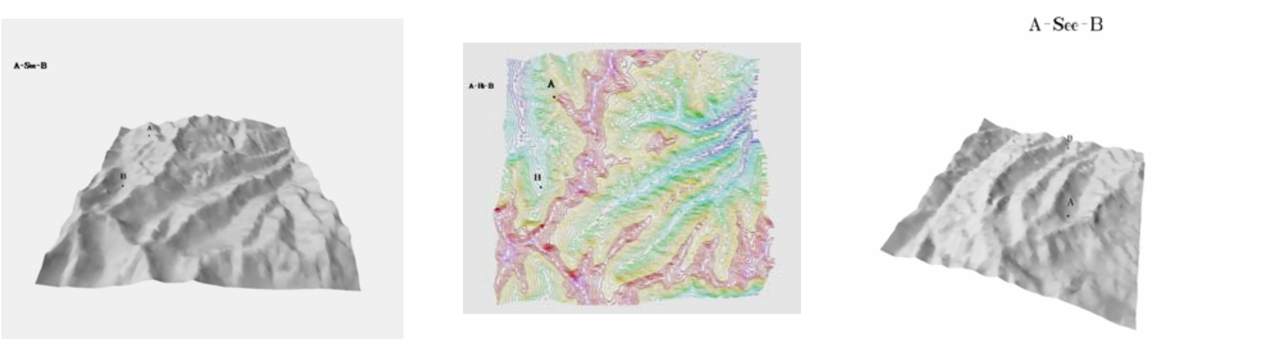

Figure 4-2: Example Views of 3D and 2D Displays Used in the Visual Momentum Experiments [24,25]. The right most 3D display was used in Experiment 3 [26].

Experiment 1. In Experiment 1 [24], participants made judgments about the properties of two points placed on terrain depicted as 2D or 3D displays. They performed the tasks in pairs of trials, switching tasks and displays between trials. On half the trials (continuous transition), the display dynamically rotated in depth from one display format to the other. On the other half (discrete transition), a blank screen was shown for the same duration.

As predicted, the results showed that a continuous transition between the display types improved performance on the trial after transition relative to the discrete condition. Participants were faster and there was a trend toward greater accuracy on the second trial of the pair with the continuous transition. Given the RT advantage without any trade off in accuracy, it appeared that the smooth rotation provided improved visual momentum between consecutive displays.

Experiment 2. Experiment 2 [25] was a replication of Experiment 1 with a key difference: in the discrete transition condition, participants were immediately shown the alternate display format on the second trial of a pair. If the advantage for continuous transition was not obtained in Experiment 2, it would imply that the Experiment 1 results were due to participants forgetting terrain information while the blank screen was shown in the discrete condition. In contrast, if the advantage in Experiment 1 was due to the visual momentum provided by smooth rotation, then the results should still obtain when there is no blank screen in the control condition.

The results of Experiment 2 showed that participants were faster on the second trial of the pair with the continuous transition. As observers switched tasks and display formats, a continuous transition improved performance relative to discrete condition. In the discrete condition of Experiment 2 the second trial was shown immediately following the first, and so this result could not have occurred because participants forgot information gained on the first trial in that condition. Accuracy was generally higher with continuous transition, both before and after transition. This meant that there was no evidence of a speed-accuracy trade off with respect to the transition effect – accuracy was higher after the continuous transition.

Experiment 3. Because the terrain was shown during the smooth transition in the continuous condition, participants in Experiments 1 and 2 may have used this preview to anticipate the second trial of the pair. Was the source of the obtained advantage for the continuous condition due to preview or improved momentum? In Experiment 3 [26] this question was tested by showing the rotating terrain in the control condition for as long as in the continuous transition condition.

In Experiments 1 and 2, the rotation from 2D to 3D (and vice versa) only occurred in depth. In Experiment 3, the terrain was also rotated in the azimuth so that the viewpoint for the 3D display was aligned with an imaginary line connecting points A and B (see Figure 4-2). This was done to make the 3D display more immersive or egocentric [14], and to provide a method for equalizing the rotation time in continuous and discrete conditions. For the discrete transition condition, azimuth rotation was in the direction opposite to that which occurred in the continuous case. For example, if the azimuth rotation to the A-B vector was 120 degrees counter-clockwise in the continuous condition, then it was 120 degrees clockwise in the discrete condition. Upon reaching this position, the display orientation would immediately switch to the azimuth position aligned with the bottom of the 2D map. Then the horizontal translation occurred, followed by rotation in depth to produce the 2D view. The opposite sequence was used to transition from 2D to 3D.

An advantage for the continuous rotation was observed. There was no evidence for a speed-accuracy trade off for this effect – accuracy was not reduced after continuous transition. We argue that continuous transition to the correct position therefore provided improved visual momentum between displays.

Smooth rotation and visual momentum summary. The primary intent of the visual momentum work was to investigate the effect of continuous transition and the results suggest that the source of improved performance is the uninterrupted flow of terrain views. Presumably, this flow provides improved visual momentum. This is not to suggest that there cannot be an advantage to having preview or that displaying task relevant information during the transition will not aid performance. In many real-world contexts, it is probable that such factors will co-occur. Smooth rotation takes time to portray, and it seems likely that the human observer will use this time to prepare for subsequent task demands. The claim is not that these factors will not have an effect, but rather that their presence is not required to produce a performance advantage. The use of dynamic transition is therefore recommended when observers examine multiple views of terrain over time.

4.2.4 Theme 1 Frame of Reference Summary

The research conducted under Theme 1 was concerned with evaluating methods for helping an observer transition between different views of terrain. The results indicate that a tethered view can be a helpful device to monitor one's egocentric position, but still maintain a broader awareness of other relevant locations than is possible with a strictly egocentric view or an exocentric map view. The visual momentum research indicates that smooth rotation provides a flow of terrain views that assists an observer switch between egocentric 3D and exocentric 2D representations. Presumably, these display techniques (tethering and smooth rotation) reduce the amount of spatial transformation required relative to seeing only an egocentric or exocentric view or discrete view switching, respectively.

Future frame of reference work in the CVT will involve conducting empirical studies examining those visual momentum methods useful for switching between different data formats (e.g. tables of materiel for different echelons, vs. maps showing locations vs. graphs and timeline data). These may include linked views between data formats, allowing the user to drag objects from one data format to another [31]. Given the increased interest in collaborative working (ref), we are interested in examining the frame of reference problem and visual momentum techniques as solutions to the communication of information between collaborating partners in a group context.

4.3 THEME 2: PERCEPTUAL BIAS AND REFERENCE POINTS

Human judgments of the geometric volumes and areas that are commonly used to depict quantitative values in 3D data representations used in statistical graphs, maps, and command visualisation systems are biased [8]. However, the use of reference points can reduce judgment error in such judgments. Two research topics were investigated:

a) Bias engendered by the use of perspective rendering in 3D displays; and

b) The effect of response method on perceptual bias.

We first discuss bias in proportion judgments, and then examine each experiment in turn.

4.3.1 Background – Bias in Proportion Judgments

Many real-world tasks require us to estimate the proportion one quantity is of another, larger quantity. For example, we might establish that our gas tank is less than half full by glancing at the fuel gauge while driving. In order for a commander to make an informed decision, it is often necessary for quantities to be compared, leading to the requirement to compute per capita figures from raw data (e.g. number of injured casualties as a function of size of unit). A number of studies have found a consistent pattern of overestimation and underestimation (constant error) in proportion judgments, despite wide variation in task demands and display format: Proportions less than .5 tend to be overestimated, and proportions greater than .5 underestimated [41,42]. Other studies show the pattern cyclically repeating over the range of stimuli tested. For example, Spence and Krizel [43] found a bias pattern whose cycle repeated (over-under, over-under) when subjects judged proportions shown in conventional graphs (e.g. pie chart, divided bar graph). Other studies found a four-cycle pattern (e.g. with angle judgments [44]). Sometimes the pattern reverses (under-over) [45].

Two questions arise when considering these findings. First, what caused the bias to occur? Why, for example, should a small proportion be overestimated, and a large one underestimated? Second, what might account for multiple cycles of bias? Hollands and Dyre [8] proposed a cyclical power model based on Stevens' power law to answer these questions. Stevens' law [46] states that the relationship between perceived magnitude of a stimulus, Φ, and its physical magnitude, Π, is expressed as a power function,

$$\Phi = \alpha\Pi^{\beta}.$$

The coefficient α represents a scaling factor (translating objective to subjective units) and is not that important for current purposes. However, the *Stevens exponent* β indicates the nature of the relationship between physical and perceived magnitude. Response compression occurs when $\beta < 1$; each increase in physical magnitude causes less and less increase in perceived magnitude. (Response expansion, where each increase in physical magnitude causes progressively greater increases in perceived magnitude, occurs when $\beta > 1$). In the magnitude estimation task commonly used by Stevens, the observer is shown a set of stimuli which differ along some physical dimension (e.g. length, area, volume) and is asked to estimate magnitude by assigning a number to each stimulus. Estimates of the exponent of the power function are around 1.0 for length, 0.8 for area, and 0.6 for volume [46]. Reverse patterns [45] can be accounted for using an exponent greater than unity.

Spence [47] proposed a model of proportion judgments based on Stevens' law. Consider two quantities Π and Ω, where $\Pi + \Omega = 1$. The subjective proportion P is computed as

$$P = \alpha\Pi^\beta / [\alpha\Pi^\beta + \alpha\Omega^\beta]$$

$$= \Pi^\beta / [\Pi^\beta + (1 - \Pi)^\beta].$$

This model predicts a one-cycle bias pattern; when $\beta < 1$, small proportions P are overestimated, and large ones underestimated.

Hollands and Dyre [8] proposed a modification of Spence's model [47] to account for multi-cycle bias patterns. Rather than comparing a part to the whole, Hollands and Dyre proposed that the observer can compare the part to intermediate reference points. The cyclical power model proposes that a judgment requires the use of two reference points to define the range of possible responses, one smaller than the to-be-judged proportion, the other larger. That part of quantity P larger than the smaller reference point is compared to the range between the reference points, rescaled as a proportion of the entire stimulus, and added to the smaller reference point. If R_0, ..., R_n is a set of reference points then a proportion may be computed as

$$P(\Pi) = \frac{(\Pi - R_{i-1})^\beta}{(\Pi - R_{i-1})^\beta + (R_i - \Pi)^\beta} \cdot \frac{R_i - R_{i-1}}{R_n} + \frac{R_{i-1}}{R_n} \qquad \text{if } R_{i-1} \le \Pi \le R_i \cdot .$$

The cyclical power model (CPM) predicts one-, two-, and four-cycle bias patterns for P when the number of reference points is varied. When two reference points are used (0, 1) a one-cycle pattern is predicted. When a third reference point is added at .50, a two-cycle pattern results. When reference points are further added at .25 and .75, a four-cycle result occurs. Hollands and Dyre [8] successfully fit the different versions of CPM to one-, two-, and four-cycle data. The experimental data obtained by Erlick [41], Spence and Krizel [43], and Huttenlocher et al. [44] were best fit by one-, two-, and four-cycle versions of the model, respectively.

A key assumption for CPM is that the number of reference points used determines the bias frequency. Placing tick marks in certain locations on a display may affect the choice of reference points. Varying the number of reference points used should affect the bias pattern frequency without affecting the exponent, β. To examine this claim, Hollands and Dyre [48] showed subjects pie charts with tick marks placed at various intervals around the circumference of the pie, on a horizontal response line placed below the pie, or in both locations. The cyclical power model was fit to the data using the procedures described above. Figure 4-3 shows the results with tick marks at halves (0, .50) and at quarters (0, .25, .50, .75, 1), when tick marks were on the graph, response line, or both. The best-fitting version of the cyclical power model is also shown in Figure 4-3. With tick marks at halves the data were best fit by the two-cycle version, as commonly observed with pie charts. In contrast, with tick marks at quarters, a four-cycle version fit best. This occurred only when the tick marks were placed on both graph and response line, or on the response line alone, implying that the selection of reference points may be more related to response production than stimulus perception. Although the frequency of the bias pattern changed, the Stevens' exponent remained constant at about 0.8 as one would expect given that the perceptual continuum did not change. The exponent corresponds well to the exponents obtained from other data involving proportion judgments with pies [47], and is consistent with the typical exponent obtained from magnitude estimates of area. The increase in cyclical frequency generally reduced error, even given a constant Stevens' exponent.

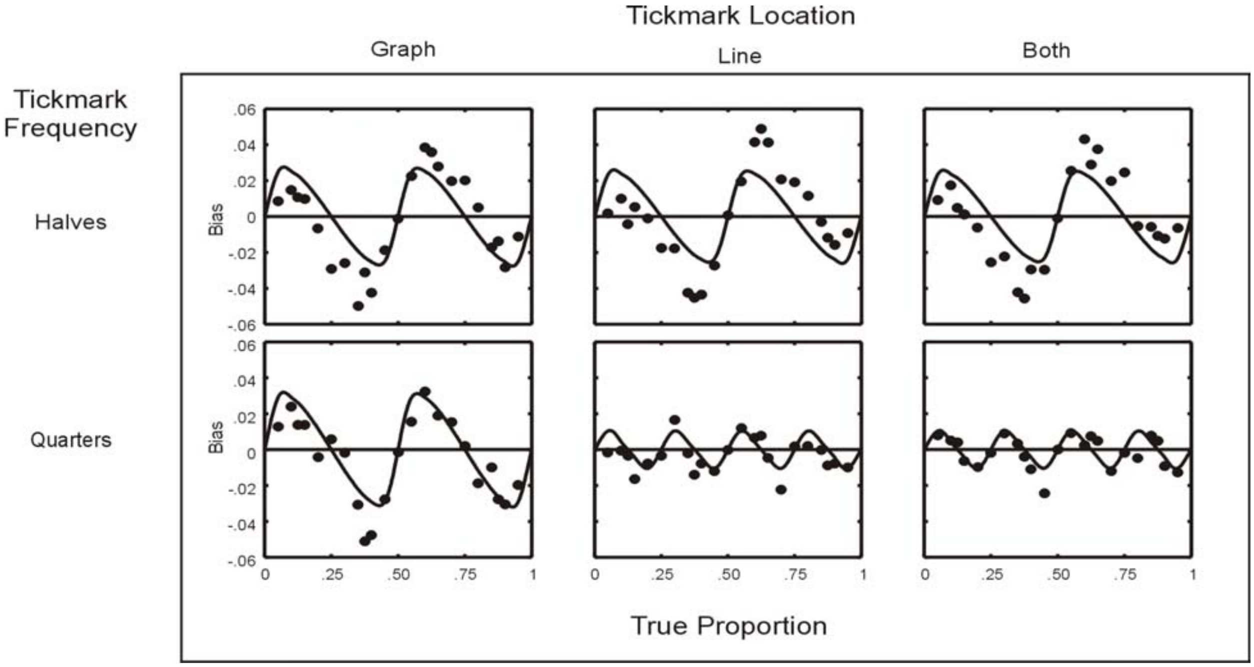

Figure 4-3: Bias (constant error) in Judging Pie Charts as a Function of Tick Mark Location and Frequency [48]. The solid lines represent the predictions of the best fitting version of CPM.

CPM is a general model, and has been shown to account for bias in proportion judgments using a wide range of stimuli [8]. Hence, the model has practical utility for determining how different types of quantitative displays are read. The designer could use the cyclical power model to analyze empirical data comparing various tick mark frequencies and locations to determine if tick marks are being used or not, and also to find the combination producing the smallest bias.

It is important to understand the various types of bias that can affect a perceptual judgment if we are to produce displays that are relatively free of such bias. 3D bar graphs are increasingly used to display complex data of all kinds, including battlespace data [49], and 3D bars are commonly portrayed at different distances from the observer using linear perspective. Does this portrayal at different depths have an effect on the accuracy of proportion judgments? How does it relate to the bias inherent in the judgment of volumes, areas, and lengths discussed above?

Two research topics pertaining to CPM have been investigated using the DRDC Toronto CVT. First, the use of more compatible response methods on the choice of reference points was examined. Second, bias engendered by the use of perspective rendering in 3D displays was investigated.

4.3.1.1 The Effect of Response Method on Perceptual Bias

People make proportion judgments in many situations, and often respond in different ways – by recording a number, by adjusting a control, or by reproducing the position of a display indicator on paper. In the Hollands and Dyre [48] experiment discussed above, participants estimated proportions by dividing a horizontal line into two parts corresponding to the parts of the pie chart. Their results showing that placing tick marks on the response line was the key factor affecting cyclical frequency suggests that something in the nature of the response may affect the choice of reference points utilized. Similar results had been obtained by Taylor [45].

It is a widely accepted human factors design guideline that dial displays are more compatible with rotary controls and linear displays with sliding controls [14,50]. Fitts and Seeger [51] found that such stimulus-response (S-R) compatibility reduced response time and error. Perhaps increased S-R compatibility might also affect the choice of reference points.

Morton and Hollands [52] examined this question by having participants respond using a rotary dial, a horizontal line, or by typing a numeric response. They predicted that greater compatibility between display and response method should increase the number of reference points used and reduce judgment error.

Results showed a two-cycle bias pattern for line and numeric conditions, but a four-cycle pattern for the dial, leading to reduced error for that condition. Response method had no effect on judgment time. Response method did not affect the estimated value of the Stevens exponent (0.83 on average). More compatible S-R relationships between pie chart and dial may have led to the use of higher-frequency reference points, thereby improving judgment accuracy. These results highlight the importance of considering the display-control relationship when attempting to minimize error in proportion judgments.

4.3.2 Biases in Reading 3D Bars

In an attempt to maximize the amount of information available at a glance, modern information visualisation systems make use of display techniques in which 3D graphics are heavily relied upon. In such systems 3D graphs are increasingly used to display complex data of all kinds, and 3D bars are commonly portrayed at different distances from the observer using linear perspective. For example, this occurs when bars are placed at different locations on a map or terrain surface viewed at an oblique angle. It also occurs with 3D bar graphs or histograms where two axes code variables and bar height is used to represent a third variable. However, the use of a 3D representation introduces the potential for error in comparisons of the elements portrayed at different depths.

To give the impression of depth, linear perspective is often used, which results in more distant objects being portrayed at smaller sizes. This mimics the monocular information available when we examine a real 3D scene (e.g. looking out the window with one eye) or when we look at any 2D surface depicting a 3D scene (e.g. a photograph). However, there may be insufficient depth information in an artificial 3D scene to specify object distance and therefore the observer's size-distance scaling may be inaccurate. For 3D displays, depth cues can be added to allow the sizes of objects at different distances to be judged more accurately. Hollands, Parker, and Morton [53] therefore posed the following experimental question: Does size-distance scaling allow accurate judgments of portrayed size in 3D bar graphs given the depth cues typically available (linear perspective, texture, relative height and occlusion)?

Participants were shown two 3D bars placed on a grid, and estimated the proportion the smaller bar at front was of the larger bar, as shown in Figure 4-4. Bars were co-located, placed side by side (near adjacent), or the small bar occluded the large (near or far occluded, respectively).

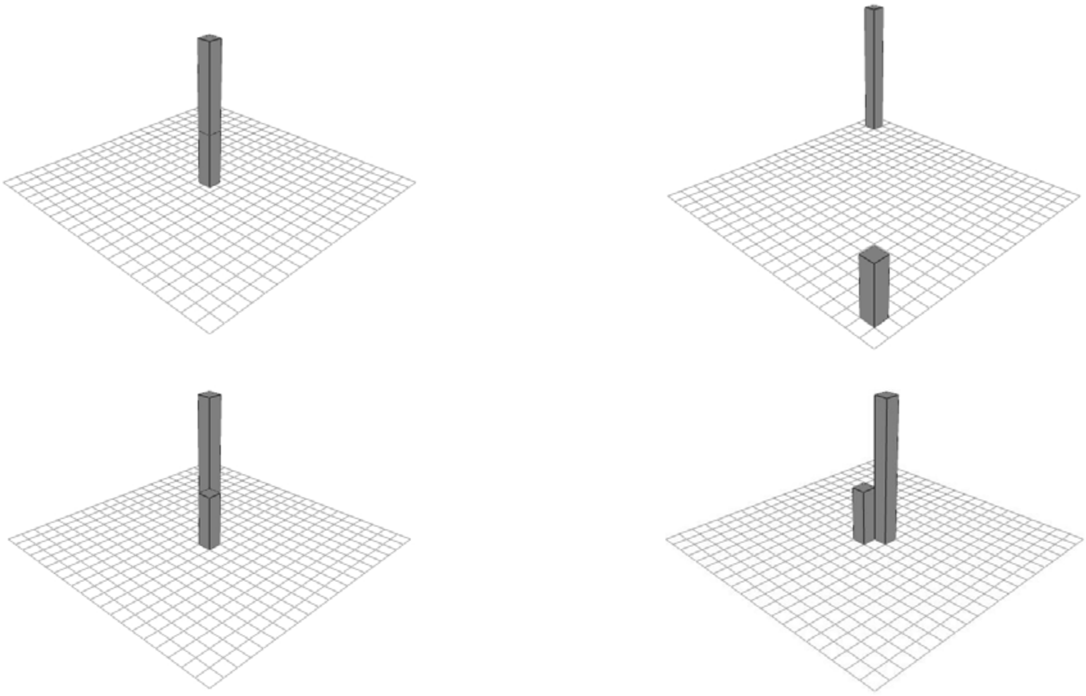

Figure 4-4: Examples of Displays Used in [53] (top: co-located, far-occluded; bottom: near-adjacent, near occluded).

Three types of results are possible if size distance scaling does not hold. If observers judge retinal size (visual angle) then this would lead to general overestimation of the proportion or a positive bias (since the "true" size of the rear bar will be underestimated). If observers perceive the far bar as farther than actually depicted, an overconstancy would result leading to overestimation of the far bar, underestimation of the proportion, and a negative bias. A third possibility is that size estimates simply become less stable, with both overconstancy and underconstancy (either through judgments of retinal size or otherwise) occurring. In this case, there may be no change in net bias, but an increased variability in the scores.

The results showed that judgment error was greater for the far occluded condition and there was greater variability in bias scores across observers. Thus, size estimates in 3D bar graphs were less stable when the bars were placed farther apart, with both overconstancy and underconstancy occurring. Presumably, the available depth information did not accurately specify the relative distances of the two objects, and this poor specification led to imprecise size estimates. In addition, cyclical bias was observed in participants' judgments, with all participants showing best fit with a one- or two-cycle version of CPM.

To account for the observed constant error, Hollands et al. [53] augmented CPM to include an independent bias parameter γ representing size-distance scaling. According to the model, these two types of bias are independently responsible for the observed judgment error. The model was fit to the data with good results. In particular, the absolute value of γ was greater in the far-occluded condition. Inclusion of γ increased R^2 for far- and near-occluded conditions only. In contrast, bar location did not affect cyclical bias.

An implication for graph design is that bars should be portrayed at similar depths in 3D bar graphs if accurate judgments are necessary, or that 2D graphs (where there is no variation in depth) should be considered as an alternative. 3D bars may be augmented by graduated axes scaled to various depths; research is planned to examine the effect of these augmentations on size-distance scaling.

4.3.3 Summary

We draw four conclusions about bias in proportion judgments with commonplace graphical stimuli found in battlespace visualisation systems:

1) Cyclical bias can be expected;
2) The effects of this bias on judgment error can be reduced by adding reference points;
3) More reference points can be added by increasing the frequency of tick marks; and
4) More reference points can be added by increasing the S-R compatibility of the response method and the display arrangement.

4.4 THEME 3: MODELING MENTAL OPERATIONS

4.4.1 Background and Framework

Many studies in graphical perception have shown that the effectiveness of different types of graphs depends on the task [14,54]. For example, Follette and Hollands [55] showed participants bar graphs depicting two quantities A and B. The two bars were either side by side or one was placed above the other (stacked bar). They found that performance on a part-to-whole [A/(A+B)] proportion judgment was better with the stacked bar arrangement, whereas performance on a part-to-part proportion judgment (A/B) was performed better with the side-by-side bar arrangement.

Why does task dependency occur? Presumably, something differs in terms of cognitive processing. Hollands and Spence [56] contend that there must be a difference in mental operations to account for task dependent results. Similar arguments have been made by other researchers [57,58] and models of such processing have been proposed. In general, the models assume that operations are executed in a sequence of steps, and that the time required for a user to complete a task with a graph is linearly related to the number of steps [57]. Each added operation increases the likelihood of error, in keeping with what is generally known about serial processes [59].

Follette [9] extended these ideas by proposing that the quality of a particular perceptual feature to which the operation is applied should also matter. Thus, according to her model, two factors affected graph reading performance:

1) The number of operations necessary given a particular task-graph combination; and
2) The quality or effectiveness of the perceptual features used as input for the operations [60,61].

Cleveland’s ranking [62] (see Figure 4-5) was used to rank the effectiveness of perceptual features. Higher-ranked perceptual features should lead to better performance. Mental operations act on the highest-ranking perceptual feature available within a graph. The set of mental operations proposed by Follette included basic arithmetic operations such as summation and ratio estimation.

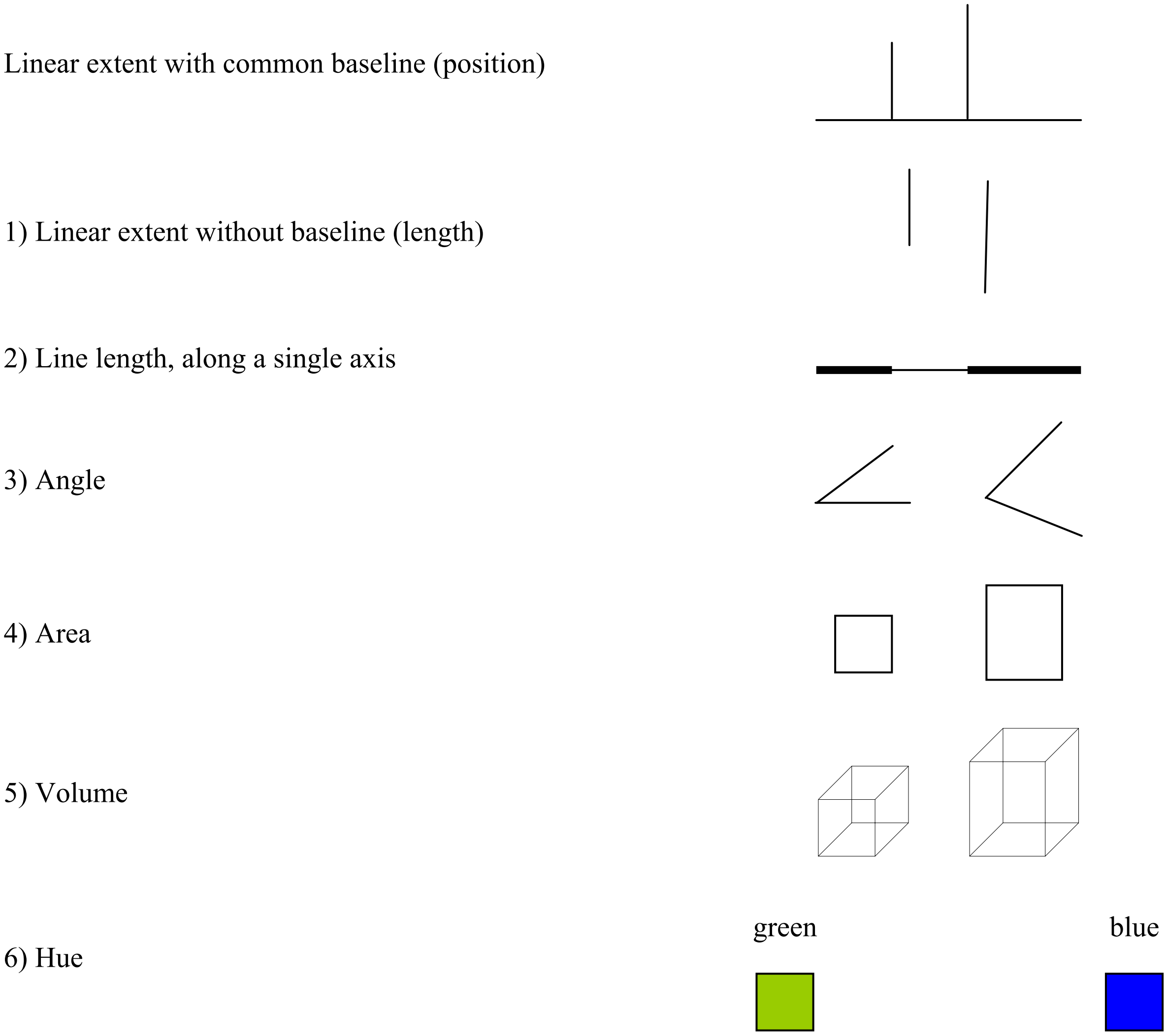

Figure 4-5: Cleveland's [61,62] Ranking of Elementary Graphical–Perception Tasks.

To illustrate the mental operations framework, reconsider the Follette and Hollands [55] results. Table 4-1 (modified from [8,62]) shows the proposed operations for each of the four conditions in their experiment. For example, part-to-part performance with graphical operations was better with bar graphs than stacked bar graphs since the one graphical operation (ratio estimation) had a more highly ranked perceptual feature (position) than did the stacked bar graph (length). Part-to-whole performance was shown to be better with stacked bars than bars since fewer operations were necessary to complete the task (no need to perform the summation operation).

Table 4-1: Follette and Hollands' [55] Results Interpreted in Terms of Mental Operations and Perceptual Features

	PART-TO-PART (A/B)	PART-TO-WHOLE [A/(A+B)]
BAR GRAPH	1) Ratio Estimation (Position); (A/B)	1) Summation (Position); 2) Ratio Estimation (Position); [A/(A/B)]
STACKED BAR GRAPH	1) Ratio Estimation (Length); (A/B)	1) Ratio Estimation (Length); [A/(A/B)]

4.4.2 Object-Based Advantages

Follette [8] examined participants' accuracy when making the two kinds of proportion judgments with several data values. With multiple values it is possible for bars involved in the judgment to be either adjacent or separated by other bars (non-adjacent). The results obtained in Follette's non-adjacent condition are of greater real-world relevance than the adjacent situation, because with more values it becomes less likely that values of interest will be adjacent. Follette found that bar graphs were less accurate than stacked bars for both tasks (although more noticeably in the part-to-part task) when the bars were not adjacent. This is surprising when considering Cleveland's hierarchy [61,62], since one would expect the bar condition to be more accurate than the stacked bar, given that all graphical elements of the bar graph are aligned on one common axis.

St-Cyr and Hollands [63] argued that Follette's [8] results may be explained by object-based theories of attention [64] which generally predict that judgments of object elements are more accurate when the elements belong to a single object rather than different ones. Similarly, in Follette's experiment, the different segments of a stacked bar are part of a single object, producing an object-based advantage relative to separate bars.

To test this hypothesis, St-Cyr and Hollands [63] created a staggered stacked bar arrangement by jittering the bars forming the stack right or left at random so that no pair of bars was contiguous. The vertical alignment of each graphical element was preserved. In their experiment, participants made part-to-part proportion judgments on non-adjacent graphical elements using bars, stacked bars, and staggered stacked bars. They found that judgments of proportions shown in staggered stacked bars were less accurate than judgments of stacked bars.

In addition, bar graphs were less accurate than stacked bar graphs, replicating Follette [8]. An analysis of the effects of distance between elements suggests that the stacked bar advantage was not due to greater distance between bars in the bar graph condition. Thus, a common baseline did not improve participants' judgments, despite the predictions of Cleveland ranking [61,62]. The results also showed that staggered stacked bars were more accurate then bars, an unexpected result. This might be explained by the fact that staggered stacked bar graphs preserve the vertical arrangement of stacked bars, although graphical elements are not part of a single object.

In the St-Cyr and Hollands [63] experiment, participants judged non-adjacent graphical elements. This is a more common situation than adjacent elements given that most graphs depict multiple elements, and the probability of adjacency is reduced with each added data point. Based on our results, this study proposes the following revision of Cleveland's ranking of perceptual tasks [61,62] for graphical perception:

1) Vertical arrangement of graphical elements in a single object;
2) Vertical arrangement of graphical elements; and
3) Graphical elements with common baseline.

Object-based advantages appear to be an important element of graphical perception, although further investigation is necessary.

4.4.3 Conclusions

Follette [8] proposed that two factors affect quantitative judgments with graphs:

1) The number of operations necessary; and
2) The effectiveness of the perceptual features used as input for the operations.

This framework was used to account for task-dependent results obtained by Follette and Hollands [55]. St-Cyr and Hollands [63] extended the framework to include object-based effects by suggesting a revised ordering to Cleveland's hierarchy of graphical elements. Although much empirical work remains to test the model's assumptions and predictions, the model offers an intriguing simple account for how graphically depicted information is processed.

4.5 THEME 4: VISUAL ATTENTION AND VISUAL SPAN

Even when a tactical display accurately depicts all relevant data, the human observer may not attend to all displayed elements. The effectiveness of Symbicon symbology to provide relevant tactical information "at a glance" was demonstrated using a change-blindness paradigm. Using an eyetracker, a gaze-contingent display can be constructed that depicts only that region of the display upon which the observer fixates. By varying the size of that region in small amounts using a staircase procedure, the result indicate how much display information can be attended "at a glance" by a human observer.

4.5.1 Detecting Change in Tactical Displays

Keillor, Thompson, Smallman, and Cowen [65] were interested in determining the degree to which different types of tactical symbology could be selectively monitored for change. The operator of a naval tactical display has multiple tasks, but the primary task is to maintain awareness and understanding of the tactical picture by monitoring changes. During such monitoring tasks, operators are susceptible to *change blindness* [66], which refers to the phenomenon that humans have difficulty in perceiving major changes in objects from one scene to another. Normally, low-level motion transients help observers direct attention toward areas of the display that rapidly change. However, if an observer blinks, glances away from the screen, or orients focal attention toward another portion of the display, the transients fail to produce a shift of attention to the changing location, and the observer is left with no awareness that a change has occurred [67]. A flicker paradigm [68] is typically employed to investigate change blindness. This methodology uses the brief flicker of a blank screen to simulate eye blinks or diversion of attention away from the changing area of the display (Figure 4-6), thereby simulating the attentional shifts that occur during naturalistic monitoring tasks.

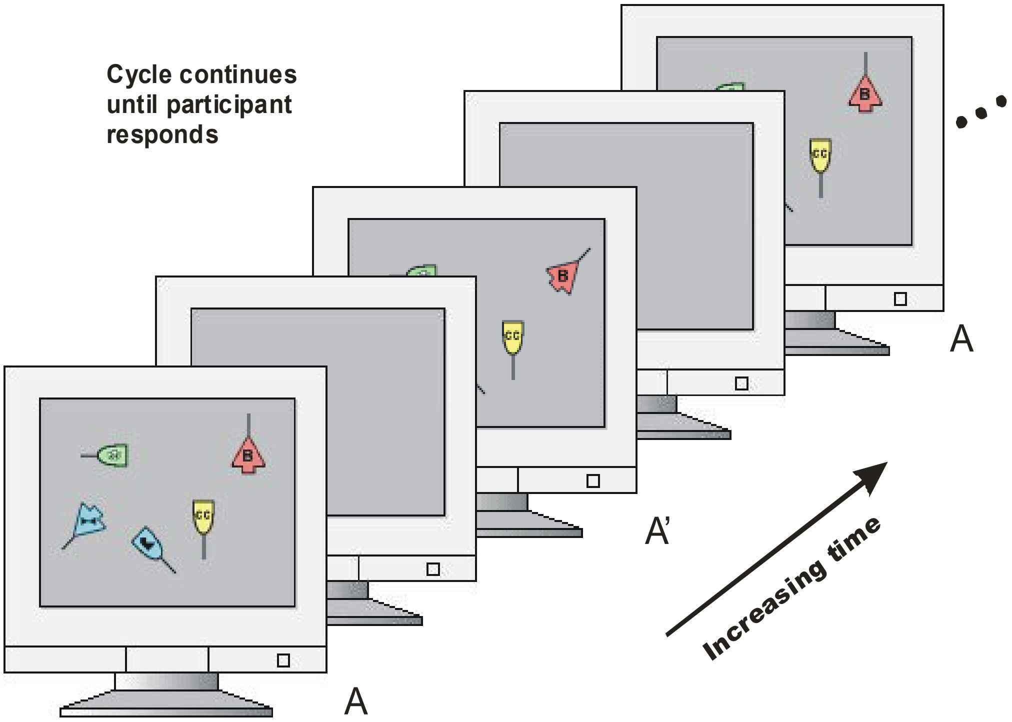

**Figure 4-6: An Illustration of the Change Blindness Paradigm Used in [65].
Note the change in heading for aircraft B from screens A to A′.**

Figure 4-7 illustrates a set of symbols typically used on tactical displays (MIL-STD-2525B) along with a set of symbols called Symbicons recently developed by the United States Navy [69]. Symbicons were designed to incorporate the platform (centre letter/icon) and threat (colour and frame) information of conventional military symbols with the platform classification (frame) and heading (leader line) information of realistic icons in order to promote the rapid appreciation of all depicted track attributes. Thus the main differences between the two sets are the frame shape and the depiction of heading. For MIL-STD-2525B, heading is indicated by the orientation of a leader line, but in the case of the Symbicons, the entire symbol rotates along with the leader line to face the compass direction.

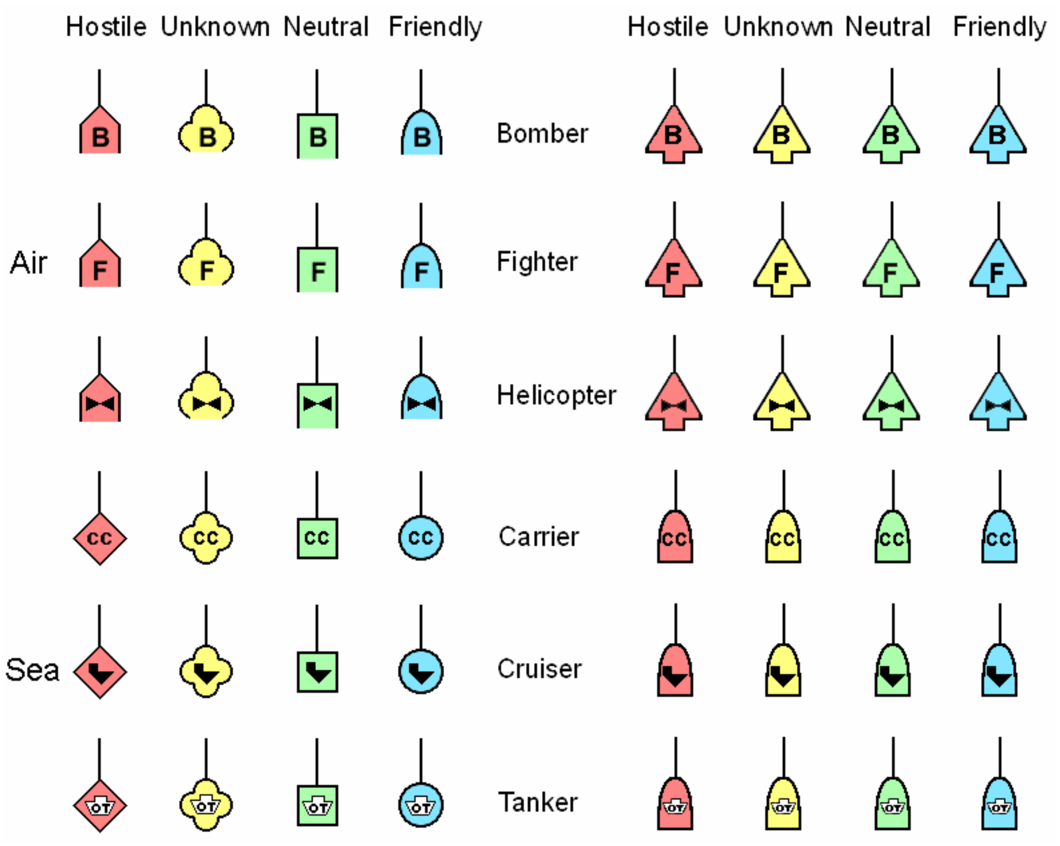

Figure 4-7: Specific Examples of MIL-STD-2525B Symbology (left) and Symbicons (right), for a Variety of Platforms and Threats.

Given the potentially disastrous consequences of change blindness, it would be beneficial if a symbology could be developed that shows greater resistance to change blindness by permitting larger numbers of contacts to be attended at once. The features of Symbicons described above may provide such resistance. Using the DRDC Toronto CVT, Keillor et al. [65] examined how quickly operators detected changes in a simulated tactical display as a function of symbology (MIL-STD-2525B vs. Symbicons).

A change detection task permits an evaluation of the extent to which operators selectively monitor subsets of relevant contacts within a display. Keillor et al. [65] examined the ability of their observers to monitor "sea" and "air" contacts for heading changes with the two symbologies. This task involved selectively attending to the frame shape (platform classification) of the symbols. For each trial, a flicker sequence comprising the original image, a grey screen, the modified image, and a grey screen was presented (Figure 4-6). The sequence continued until the participant responded "change" or "no change" by pressing the right or left button respectively.

Keillor et al. [65] found that change detection was better for Symbicons relative to MIL-STD-2525B, and that this effect was greater when participants were told in advance what type of contact (platform category) could change heading. The advantage for the Symbicons is consistent with groups of similar contacts being simultaneously attended and monitored for change.

4.5.2 Visual Span

As displays become larger or more complex, it becomes important to develop metrics to determine how much of a given display an operator can take in "at a glance". A measure of the spatial extent of visual processing is useful both in terms of optimizing display design, and developing a model for how an operator accomplishes a task when interacting with the display.

The most widely accepted way to measure visual span is by using a gaze-contingent moving window technique. By using eye position information to drive the display it is possible to obscure all information in a given display with the exception of that which is centred on the participant's fixation. By manipulating the size of this moving-window it is possible to determine how much of the display is being used, such that making more of the display visible is of no benefit. Recently, a more efficient psychophysical approach has been adopted by some researchers, in which an iterative algorithm varies the size of the window over successive trials to determine the visual span [70].

Keillor, Desai, Hollands, and Reingold [71] used a variation of the gaze-contingent algorithm [70] to examine the visual span for tactical displays as a function of the task and type of symbology used. Specifically, MIL-STD-2525B and Symbicon symbologies were evaluated. In order to determine whether visual span is affected by the task, participants were required to search for targets within an array based on threat affiliation (coded by colour), heading (coded by colour and orientation), or by the conjunction of threat affiliation and heading (coded by a combination of colour and heading).

Keillor et al. [71] found that the amount of a display processed "at a glance" is affected by the type of search task, even given identical displays. Visual span was smaller for the more complex conjunction search task. The result demonstrates that it is not the display complexity itself that underlies the span of visual processing in a search task, but instead the cognitive demands of the task. Visual span is therefore best thought of as a measure influenced by strategic and cognitive factors as well as a task's perceptual demands, making it an excellent indicator of the ease with which information may be extracted and processed from a given type of display.

4.5.3 Visual Attention and Visual Span Conclusions

In sum, the visual attention work conducted at CVT indicates that the Symbicon symbology which encodes platform classification and heading more iconically rather than symbolically better enabled human observers to detect change. Using an eyetracker, a gaze-contingent display was constructed that depicts only that region of the display upon which the observer fixates. The visual span for an observer conducting visual search was smaller for more complex conjunction searches.

4.6 CONCLUSIONS

IST-021/RTG-007 considered visualisation technologies, including display devices and techniques in relation to how they help humans to perform their tasks effectively. The RTG emphasized the human use of the computational subsystem in ensuring that the right information is available in the form and at the time needed. The experiments conducted on the DRDC Toronto CVT were essentially attempts to empirically assess those factors that make information available in the right form given task constraints and data requirements.

Specifically, a commander's tasks may require both egocentric and exocentric views on a battlespace. We examined the effectiveness of two visual momentum methods to assist the user switch between these diverse views of geospatial data. The first method (a compromise method) involved the use of a tethered display and showed that it was an effective method for depicting one's egocentric position with respect to other key locations in the terrain, which requires a certain exocentricity not available in the completely immersed view. The second method (a transition method) involved depicting smooth rotation from a 2D to a 3D terrain view, and was shown to be more effective than a discrete shift between 2D and 3D views.

We examined the use of reference points to improve the accuracy of quantitative judgments in information displays. First, having more reference points available in the display can improve performance. Second, the particular response method used, which would probably be driven by task demands, turned out to have a clear effect. Third, the use of linear perspective to depict bars at different distances from an observer introduces an source of error in judgment error and should be avoided to the extent possible.

We described work that assessed the utility of a model of mental operations in graphical perception. This model was used to account for a set of data and extended to include object-based effects. From a design perspective, the intent is to shorten sequences of perceptual operations conducted on graphical elements of the highest quality according to Cleveland's [61,62] ranking.

Finally, we described work assessing how a human observer attends to information on a tactical display. This was done in two ways: first by using the change blindness paradigm to assess what display factors affected how quickly an observer could detect a change in a tactical display; second, by using a gaze-contingent display to assess how the operator's visual span is affected by changes in the complexity of an operator's task. Such methods should ultimately allow a determination of those factors leading to more of the display surface being processed by a human observer.

In future work we plan to implement many of these display concepts into prototype systems for navy, army and joint command and control systems.

4.7 REFERENCES

[1] Chief of Maritime Staff (Canada) (2000). The Naval C4I Way Ahead: A Blueprint for 2010. Ottawa: Department of National Defence.

[2] RAdm Buck (1999). Plenary Session, *TTCP Symposium on Co-ordinated Battlespace Management.*

[3] Department of National Defence (Canada) (1999). *Defence Strategy 2020: Shaping the future of the Canadian Forces.* Ottawa: Author.

[4] Naval Command and Control Way Ahead (Chief of Maritime Staff, January 2000).

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

[6] Lewandowsky, S. and Behrens, J.T. (1999). Statistical graphs and maps. In F. Durso et al. (Eds.), *Handbook of Applied Cognition*. W. Sussex, UK: Wiley.

[7] 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.

[8] Hollands, J.G. and Dyre, B.P. (2000). Bias in proportion judgments: The cyclical power model. *Psychological Review, 107,* 500-524.

[9] Follette, L. (2000). *Comparing the effectiveness of bar and stacked bar graphs for judgments of proportion.* Unpublished Master's thesis, University of Idaho, Moscow, Idaho.

[10] Beno, E., Dubé, R., Duval, A., Holmes, W., Pelletier, C., Thériault, G. and Thibault, G. (2004). Concept of operations for collaborative working. Draft Contract Report (v.2.2), Valcartier, Quebec: DRDC Valcartier System of Systems Section.

[11] Hollands, J.G., Lamb, M. and Keillor, J.M. (2004, September). *Modeling 3D frames of reference for the common operational picture*. Workshop presentation, NATO IST-043/RWS-006, Visualisation and the Common Operational Picture [VizCOP], Canadian Forces College, Toronto, Ontario.

[12] Amorim, M.-A. (2003). "What is my avatar seeing?": The co-ordination of "out-of-body" and "embodied" perspectives for scene recognition across views. *Visual Cognition, 10,* 157-199.

[13] Milgram, P. and Colquhoun, H., Jr. (1999). A taxonomy of real and virtual world display integration. In Y.O.A.H. Tamura (Ed.), *Mixed reality – Merging real and virtual worlds* (pp. 1-16). Tokyo and Berlin: Ohmsha & Springer-Verlag.

[14] Wickens, C.D. and Hollands, J.G. (2000). *Engineering psychology and human performance* (3rd Ed.). Upper Saddle River, NJ: Prentice-Hall.

[15] Shepard, R. and Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171,* 701-703.

[16] Trafton, J.G., Trickett, S.B. and Mintz, F.E. (2001). Overlaying images: Spatial transformations of complex visualizations. Model Based Reasoning.

[17] Trafton, J.G., Trickett, S.B. and Mintz, F.E. (in press). Connecting internal and external representations: Spatial transformations of scientific visualizations. *Foundations of Science.*

[18] St. John, M., Cowen, M.B., Smallman, H.S. and Oonk, H.M. (2001). The use of 2D and 3D displays for shape-understanding versus relative-position tasks. *Human Factors, 43,* 79-98.

[19] Colquhoun, H. and Milgram, P. (2000). Dynamic Tethering for Enhanced Remote Control and Navigation. In Proceedings of the Human Factors and Ergonomics Society 44th Annual Meeting (pp. 146-149). Santa Monica, CA: Human Factors and Ergonomics Society.

[20] Wickens, C.D. and Prevett, T.T. (1995). Exploring the dimensions of egocentricity in aircraft navigation displays. *Journal of Experimental Psychology: Applied, 1,* 110-135.

[21] Wang, W. and Milgram, P. (2001). Dynamic viewpoint tethering for navigation in large-scale virtual environments. In *Proceedings of the Human Factors and Ergonomics Society – 45th Annual Meeting* (pp. 1862-1866). Santa Monica, CA: Human Factors and Ergonomics Society.

[22] Yu, P.C. and Andre, A.D., (1999). A simulator evaluation of four perspective scene viewpoints. In *Proceedings of the Silicon Valley Ergonomics Conference and Exposition-ErgoCon 99* (pp. 233-242).

[23] Lamb, M. and Hollands, J.G. (2005) *Viewpoint tethering in complex terrain navigation and awareness.* Submitted to Human Factors and Ergonomics Society 49th Annual Meeting.

[24] Hollands, J.G. and Ivanovic, N. (2004). Task switching with 2D and 3D displays of geographic terrain: The role of visual momentum. In *RTO Meeting Proceedings 50, Massive Military Data Fusion and Visualisation: Users Talk with Developers* [Papers presented at the RTO Information Systems Technology Panel (IST) Workshop held at the Norwegian Defence Logistics and Management College, Halden, Norway 10-13 September 2002, RTO-MP-105, AC/323(IST-036)TP/20], Neuilly-sur-Seine, France: NATO Research and Technology Organisation.

[25] Hollands, J.G., Ivanovic, N. and Enomoto, Y. (2003). Visual momentum and task switching with 2D & 3D displays of geographic terrain. In *Proceedings of the Human Factors and Ergonomics Society – 47th Annual Meeting* (pp. 1620-1624). Santa Monica, CA: Human Factors and Ergonomics Society.

[26] Hollands, J.G., Pavlovic, N.J., Enomoto, Y. and Jiang, H. (2004). Advantage for visual momentum not based on preview. In *Proceedings of the Human Factors and Ergonomics Society – 48th Annual Meeting* (pp. 1803-1807). Santa Monica, CA: Human Factors and Ergonomics Society.

[27] Hochberg, J. and Brooks, V. (1978). Film cutting and visual momentum. In J.W. Senders, D.F. Fisher and R.A. Monty (Eds.), *Eye movements and the higher psychological functions* (pp. 293-313). Hillsdale, NJ: Erlbaum.

[28] Andre, A.D., Wickens, C.D., Moorman, L. and Boschelli, M.M. (1991). Display formatting techniques for improving situation awareness in the aircraft cockpit. *International Journal of Aviation Psychology, 1,* 205-218.

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

[30] Bederson, B.B., Hollan, J.D., Stewart, J., Rogers, D., Vick, D., Ring, L., Grose, E. and Forsythe, C. (1998). A zooming web browser. In C. Forsythe, E. Grose and J. Ratner (Eds.), *Human factors and web development* (pp. 255-266). Mahwah, NJ: Erlbaum.

[31] Roth, S.F., Chuah, M.C., Kerpedjiev, S., Kolojejchick, J. and Lucas, P. (1997). Toward an information visualisation workspace: Combining multiple means of expression. *Human-Computer Interaction, 12,* 131-186.

[32] Neale, D.C. (1996). Spatial perception in desktop virtual environments. In *Proceedings of the Human Factors and Ergonomics Society – 40th Annual Meeting* (pp. 1117-1121). Santa Monica, CA: Human Factors and Ergonomics Society.

[33] Neale, D.C. (1997). Factors influencing spatial awareness and orientation in desktop virtual environments. In *Proceedings of the Human Factors and Ergonomics Society – 41st Annual Meeting* (pp. 1278-1282). Santa Monica, CA: Human Factors and Ergonomics Society.

[34] Olmos, O., Liang, C.C. and Wickens, C.D. (1997). Electronic map evaluation in simulated visual meteorological conditions. *International Journal of Aviation Psychology, 7,* 37-66.

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

[36] St. John, M., Smallman, H.S. and Cowen, M.B. (2002) *Evaluation of schemes for combining 2-D and 3-D views of terrain for a tactical routing task*. Pacific Science & Engineering Group Technical Report.

[37] Barnes, M.J. (2003). *The human dimension of battlespace visualisation: Research and design issues.* Army Research Laboratory Technical Report (ARL-TR-2885). Aberdeen Proving Ground, MD: Army Research Laboratory.

[38] MultiGen-Paradigm (2001a). Creator (v. 2.5). [Computer software]. San Jose, CA: Author.

[39] MultiGen-Paradigm (2001b). Vega (v. 2.5). [Computer software]. San Jose, CA: Author.

[40] Guth, P. (2001). MICRODEM (v. 5.1). [Computer software]. Annapolis, MD: Oceanography Department, U.S. Naval Academy.

[41] Erlick, D.E. (1964) Absolute judgments of discrete quantities randomly distributed over time. *Journal of Experimental Psychology, 67,* 475-482.

[42] Varey, C.A., Mellers, B.A. and Birnbaum, M.H. (1990). Judgments of proportion. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 613-625.

[43] Spence, I. and Krizel, P. (1994). Children's perception of proportion in graphs. *Child Development, 65,* 1193-1213.

[44] Huttenlocher, J., Hedges, L.V. and Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review, 98,* 352-376.

[45] Taylor, M.M. (1961). Effect of anchoring and distance perception on the reproduction of forms. *Perceptual and Motor Skills, 12,* 203-239.

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

[47] Spence, I. (1990). Visual psychophysics of simple graphical elements. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 683-692.

[48] Hollands, J.G. and Dyre, B.P. (1997). Bias in proportion judgments with pie charts: The cyclical power model. In *Proceedings of the Human Factors and Ergonomics Society – 41st Annual Meeting* (pp. 1357-1361). Santa Monica, CA: Human Factors and Ergonomics Society.

[49] Rayson, J.K. (1999). Aggregate towers: Scale sensitive visualization and decluttering of geospatial Data. *Proceedings IEEE Symposium on Information Visualization* (pp. 92-99). San Francisco: IEEE Computer Society.

[50] Ely, J.E., Thomson, R.M. and Orlansky, J. (1963). Design of controls. In C.T. Morgan, J.S. Cook, A. Chapanis and M.W. Lund (Eds.), *Human engineering guide to equipment design* (pp. 247-280). New York: McGraw-Hill.

[51] Fitts, P.M. and Seeger, C.M. (1953). S-R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology, 46,* 199-210.

[52] Morton, A. and Hollands, J.G. (2003). Effects of response method on error and cyclical bias in proportion judgments. In *Proceedings of the Human Factors and Ergonomics Society – 47th Annual Meeting* (pp. 1654-1658). Santa Monica, CA: Human Factors and Ergonomics Society.

[53] Hollands, J.G., Parker, H.A. and Morton, A. (2002). Judgments of 3D bars in depth. In *Proceedings of the Human Factors and Ergonomics Society – 46th Annual Meeting* (pp. 1565-1569). Santa Monica, CA: Human Factors and Ergonomics Society.

[54] Hollands, J.G. and Spence, I. (1992). Judgments of change and proportion in graphical perception. *Human Factors, 35,* 313-334.

[55] Follette, L.J. and Hollands, J.G. (1998, October). Bars versus stacked bars: Effectiveness is task dependent. Poster presented at the Annual Meeting of the Human Factors and Ergonomics Society, Chicago, Illinois. In *Proceedings of the Human Factors and Ergonomics Society – 42nd Annual Meeting* (p. 1652). Santa Monica, CA: Human Factors and Ergonomics Society.

[56] Hollands, J.G. and Spence, I. (1994). Mental operations in graphical perception. In *Proceedings of the 12th Triennial Congress of the International Ergonomics Association* (Vol. 6, pp. 235-237). Mississauga, ON, Canada: Human Factors Association of Canada.

[57] Gillan, D.J. and Lewis, R. (1994). A componential model of human interaction with graphs: 1. Linear regression modeling. *Human Factors, 36,* 419-440.

[58] Lohse, G.L. (1993). A cognitive model for understanding graphical perception. *Human-Computer Interaction, 8,* 353-388.

[59] Swain, A.D. (1990). Human reliability analysis: Need, status, trends and limitations. *Reliability Engineering and System Safety, 29,* 301-313.

[60] Casner, S.M. (1991). A task-analytic approach to the automated design of graphic presentations. *ACM Transactions on Graphics, 10,* 111-151.

[61] Cleveland, W.S. and McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association, 79,* 531-554.

[62] Cleveland, W.S. (1985). *The elements of graphing data.* Pacific Grove, CA: Wadsworth.

[63] St-Cyr, O. and Hollands, J.G. (2003). Judgments of proportion with graphs: Object-based advantages. In *Proceedings of the Human Factors and Ergonomics Society – 47th Annual Meeting* (pp. 1659-1662). Santa Monica, CA: Human Factors and Ergonomics Society.

[64] Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113,* 501-517.

[65] Keillor, J., Thompson, L.K., Smallman, H.S. and Cowen, M.B. (2003) Detection of heading changes in tactical displays: A comparison of two symbol sets. In *Proceedings of the Human Factors and Ergonomics Society – 47th Annual Meeting* (pp. 1610-1614). Santa Monica, CA: Human Factors and Ergonomics Society.

[66] DiVita, J.C. and Nugent, W. (2000). Managing critical events in a multimodal watchstation. *Proceedings of the IEA 2000/HFES 2000 Congress. 6,* 469-472. Santa Monica, CA: Human Factors and Ergonomics Society.

[67] Rensink, R.A., O'Regan, J.K. and Clark, J.J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8,* 368-373.

[68] Rensink, R.A. (2000). Visual search for change: A probe into the nature of attentional processing. *Visual Cognition, 7,* 345-376.

[69] Smallman, H.S., St. John, M., Oonk, H.M. and Cowen, M.B. (2001). 'SYMBICONS': A hybrid symbology that combines the best elements of SYMBols and ICONS. In *Proceedings of the 45th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 110-114). Human Factors and Ergonomics Society, Santa Monica, CA.

[70] Reingold, E.M., Charness, N., Pomplun, M. and Stampe, D.M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science, 12,* 48-55.

[71] Keillor, J.M., Desai, I., Hollands, J.G. and Reingold, E.M. (2004). The span of visual processing in tactical displays. In *Proceedings of the Human Factors and Ergonomics Society – 48th Annual Meeting* (pp. 1798-1802). Santa Monica, CA: Human Factors and Ergonomics Society.