

Limits of Display Realism: Human Factors Issues in Visualizing the Common Operational Picture

Harvey S. Smallman, Ph.D., Mark St. John, Ph.D.

Pacific Science & Engineering Group
9180 Brown Deer Road
San Diego, CA 92121
USA

Michael B. Cowen, Ph.D.

Space and Naval Warfare Systems Center
Code 244210, 53345 Engineer Street
San Diego, CA 92152
USA

E-mail: Smallman@pacific-science.com / Stjohn@pacific-science.com / Mike.Cowen@navy.mil

SUMMARY

*The vision for the US military's future common operational picture (COP) is to create a display that integrates data from all service arms to promote situation awareness (SA) and coordination. This tactical display must visualize the locations of a variety of blue forces and other assets overlaid and integrated with a variety of amplifying information (planning, weather, etc). Display designers are faced with the task of visualizing this complex situation in an intuitive and useful fashion that promotes SA. A popular approach is to maximize the realism of the display, on the assumption that realism will promote natural, intuitive, and easy human interaction. To this end, technological advances have enabled designers to create real-time, realistic 3-D perspective views of the tactical picture for "at a glance" SA. However, in a series of empirical studies conducted over the last five years for the US Navy, we have found that realistic 3-D views are only appropriate for specific tasks and generally do not enhance SA. Rather, they are misperceived and promote errors. This approach to display design of maximizing realism also assumes that realistic, real-time displays will provide adequate support for detecting significant changes to a situation. However, a wealth of psychological studies have documented the tremendous human susceptibility to miss changes in natural scenes. Why would designers create, and users prefer, displays that do not serve them well? Apparently, users harbour a **Naïve Realism** – a misplaced faith in their ability to extract information from natural scenes that translates into a desire for realistic displays. It is paradoxical and worrying that at a time when basic perceptual science is revealing just how flawed and sparse is our visual representation of natural scenes that display designers are striving towards photo-realistic naturalism. In this talk, we review this troubling trend and layout a set of human factors guidelines and display concepts for the COP that, though sometimes counter to Naïve Realism, are likely to promote superior performance.*

1.0 INTRODUCTION

The vision of the future common operational picture (COP) is a display that integrates data from all service arms into a common, augmented tactical picture to promote SA. All forces operating in a defined geographical region would be shown in the same representation. By integrating all blue forces assets in a common representation, the possibility of tragic blue-on-blue incidents could be minimized. By augmenting the picture with various information such as plans, force projections and weather, the impact of these factors on an operation could be more fully understood. The question for COP display designers is how to visualize this complex picture in an intuitive and useful fashion for users.

Smallman, H.S.; St. John, M.; Cowen, M.B. (2005) Limits of Display Realism: Human Factors Issues in Visualizing the Common Operational Picture. In *Visualisation and the Common Operational Picture* (pp. 16-1 – 16-18). Meeting Proceedings RTO-MP-IST-043, Paper 16. Neuilly-sur-Seine, France: RTO. Available from: <http://www.rto.nato.int/abstracts.asp>.

A popular approach to situation awareness display design is to maximize the realism and naturalness of the display. This is done under the assumption that viewing a display that so closely approximates what it is like to literally view the situation with one's own eyes will allow users to bring their natural powers of observation to bear for nearly effortless comprehension of the situation. One manifestation of this approach is to harness three-dimensional (3-D) displays to visualize 3-D situations. In essence, the COP is a display of a three-dimensional (3-D) battlespace. There is, therefore, a natural motivation to portray the situation realistically in 3-D, see Figure 1. The designers of the 3-D prototype display shown on the left of Figure 1 state that it should require "minimal interpretive effort" because it approximates what it is like to see the actual battlespace [Dennehy 1994]. By 3-D display, we mean a display that shows a perspective view of a scene on a flat CRT. The image is two dimensional (2-D), but the oblique viewing angle means that all three dimensions are projected and represented to provide a 3-D perspective. There are various other 'true' 3-D lenticular, holographic and view-sequential display technologies under development [e.g., Travis 1995]. Some of these new displays even closer approximate the experience of viewing the situation because they provide additional depth cues that would be available on real viewing of the scene, but most interest in 3-D displays is still in the flat screen displays like those shown in Figure 1. To further increase realism, the objects in the scene (e.g., the planes and aircraft of Figure 1, left) can be rendered as realistic, shaded 3-D icons instead of the more arbitrary, unrealistic symbols usually found in conventional 2-D displays. The two displays shown in Figure 1 are manifestations of this display philosophy, although only the one on the left employs realistic icons at all view magnifications. Traditional top-down, 2-D displays, on the other hand, are unrealistic in that they are often populated with arbitrary symbols and they require the addition of non-natural design elements in order to represent the missing dimension (typically altitude). For example, on many air traffic control displays, the altitude of an aircraft is represented digitally in a small data field next to the aircraft symbol, and other displays may represent altitude in a separate "side view" or "profile view."

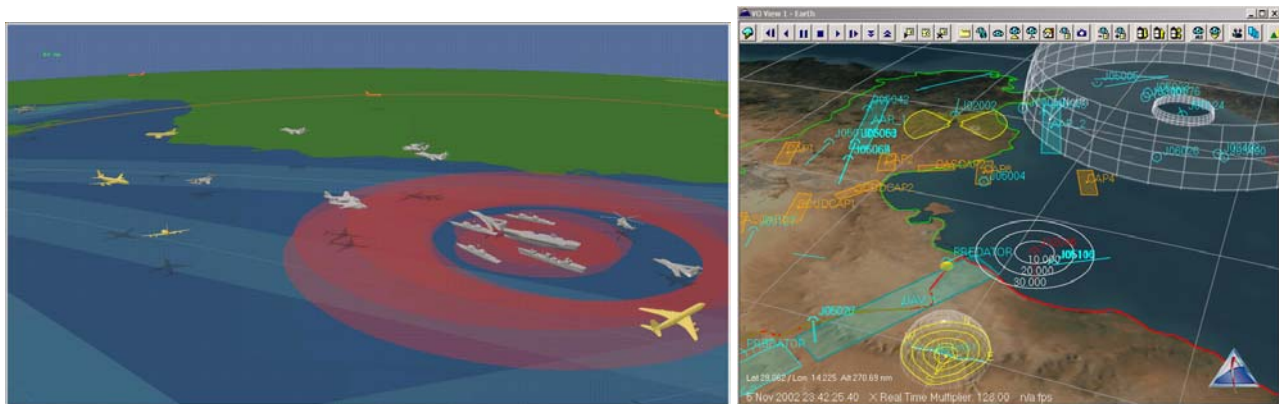


Figure 1: Left, a prototype 3-D tactical air warfare display [Dennehy 1994]. Right, a 3-D prototype COP display (T3DCOP from Northrop Grumman and AGI, 2004, at <http://www.stk.com>).

The touted superiority of 3-D displays over traditional 2-D displays, and users' frequent preferences for 3-D displays does not necessarily correlate with better performance. It is well known that users do not always want what is best for them [Andre 1995]. The literature on evaluating 3-D displays is extensive and tangled, with 3-D displays showing mixed benefits, at best, over 2-D displays [see St. John 2001a]. The reasons for this confusion are many, but they include the fact that a wide variety of different interfaces and tasks have been employed that each, in their own way, control performance outcomes. For example, altitude on 2-D control displays can be shown analogically on a separate side-view display or digitally via a character-readout that may be available on selection either on the main display or on a side window. The depth cues, camera

geometry and rendering of objects as realistic icons (as in Figure 1, left) or unrealistic symbols (as in Figure 1, right) also varies widely among the 3-D experimental displays. Further, the tasks that have been employed vary widely in the information they require users to extract from the displays. Some tasks, for example, require great spatial precision, such as checking lines of sight or predicting collisions, whereas other tasks simply require a gross sense of the layout of all three dimensions of space, such as initial route planning. Some tasks require precise compensatory control in three dimensions, such as flying a final approach, while other tasks require a broad monitoring and identification of objects across an airspace.

In section 2.0, we briefly review an extensive series of empirical and modelling efforts conducted over the last five years for the US Navy that provides a consistent and counterintuitive picture of the human factors of 3-D display use. Our approach has been to break down the question of user performance with 3-D views into a series of tractable sub-questions, each of which could be answered empirically. We carefully manipulated the method of information presentation and measured its effect on performance on a variety of lower level perceptual and cognitive tasks that are essential components of operational tasks (e.g., distance estimation, visual search, track identification). We developed a new theory to predict which categories of tasks are better with which display formats, and we extended this theory into a prescriptive, quantitative model of the misperceptions experienced with different viewing angles shown in perspective.

Our exploration into the mismatch between user preference and performance in the use of 3-D perspective views has led us to new insights into another mismatch. In section 3.0 we briefly review our research in monitoring and SA recovery. While users strongly trust their abilities to monitor complex dynamic situations and detect significant changes, the burgeoning psychological literature on the severe constraints on human visual attention, suggest this trust is misplaced. It is rare that research in one domain parallels that in another, ostensibly unrelated, domain, but perhaps this parallel points to a more general phenomenon. In section 4.0, we introduce the concept of Naïve Realism to unify these two areas and account for the often bizarre persistence of user preference for displays that beguile but that under-perform. In section 5.0, we summarize the implications of this research for the design and visualization of the COP.

2.0 3-D REALISTIC DISPLAYS: PROMOTING SA

3-D realistic displays have been developed to promote rapid perception and comprehension of the identity and layout of objects in 3-D space. As such, 3-D displays are designed to be, what are often colloquially referred to as, “at a glance” displays. In terms of Endsley’s triarchy of the levels of situation awareness (SA), we are studying Level 1 SA – perception of the situation [Endsley 1995]. Level 1 SA is in many ways the most critical element of SA because without veridical perception, higher level processing (Level 2 SA) and projection into the future (Levels 3 SA), will be based on faulty or noisy data. Given the complexities of evaluating a multi-faceted issue as this, we have found it useful to group the research sub-questions into the where, why, what and how of 3-D display use.

2.1 WHERE of 3-D (depicting space)

Where refers to users’ ability to localize objects in space to understand scene layout and to make precise spatial judgments among objects. 3-D views are perspective projections from shallow viewing angles, typically between 20 to 45 degrees above the ground plane. 2-D views, on the other hand, are top-down (90 degrees above the ground plane). In a series of simple experiments, we found that 3-D views are better for shape understanding and for judging the overall layout of a scene, whereas 2-D views are better for precise relative position judgments [St. John 2001a]. We found this pattern of results first with a series of simple,

abstract block-world stimuli and later generalized it to more realistic and operationally relevant, terrain stimuli.

We created a mini-battery of shape understanding tasks that tapped participants' general understanding of the shape of 3-D objects or layout of terrain. The tasks involved such things as mental rotation and rough line of sight (can "A-see-B"?). For example, in one shape understanding task, participants had to imagine standing in the middle of a piece of terrain and match the view they would see from an indicated corner of the terrain (the "4-Corners" task), see Figure 2. The 3-D view led to faster response times and fewer errors. The relative position tasks tapped participants' metric perception of distances and angles in a scene. The relative position tasks included relative position and distance judgments (is "A-higher-than-B"?). For these tasks, it was the 2-D views led to faster response times and fewer errors. In fact, the same cross-over pattern in the results held when the same exact stimuli were used but the task was switched from shape understanding to a relative position one, or vice versa.

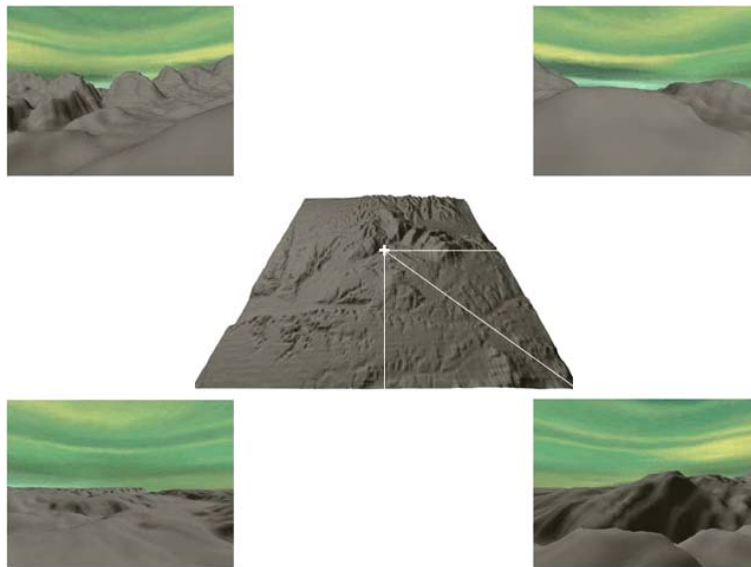


Figure 2: A 3-D View Trial from the 4-Corners Task. Participant had to pick the correct ground-level view they would see from the center of the terrain [St. John 2001a]. The correct answer is top-right. Shape understanding tasks, such as this one, were the only tasks performed better with 3-D.

Why do these results obtain? Three-dimensional views seem to be superior for shape understanding because they integrate all three dimensions into a single display with multiple, reinforcing depth cues. Supporting this notion was the finding that 2-D views could be made to perform almost equivalently to 3-D views for shape understanding by rendering them with shading information, the so called "3-D 90 degree" view. This view consisted of a perspective view, with its depth cues of shading and occlusion, but rendered from directly above, at 90 degrees. This "hybrid" design worked well for terrain, where there were important shapes to render and understand in 3-D, but it would not work well for air traffic, for example. In the next section, we address the complementary question, why are 3-D perspective views inferior for relative position judgments?

2.2 WHY of 3-D (understanding perception)

Why refers to understanding why users misperceive space with 3-D perspective views and predicting the degree of spatial misperception and imprecision with different viewing angles. In a series of experiments, we

discovered a key reason why 3-D views are inferior to 2-D views for judging relative positions [Smallman 2002; 2003a]. This discovery enabled us to create a new perceptual theory of space perception with 3-D perspective views. When perceiving a scene rendered as a 3-D perspective view, users act as if they harbour a naïve misconception about the nature of perspective projection. When rescaling image extents in the perspective views into perceived spatial extents they apply a simplifying assumption in their interpretation of the perspective view that is only a coarse approximation of the actual image geometry of the situation.

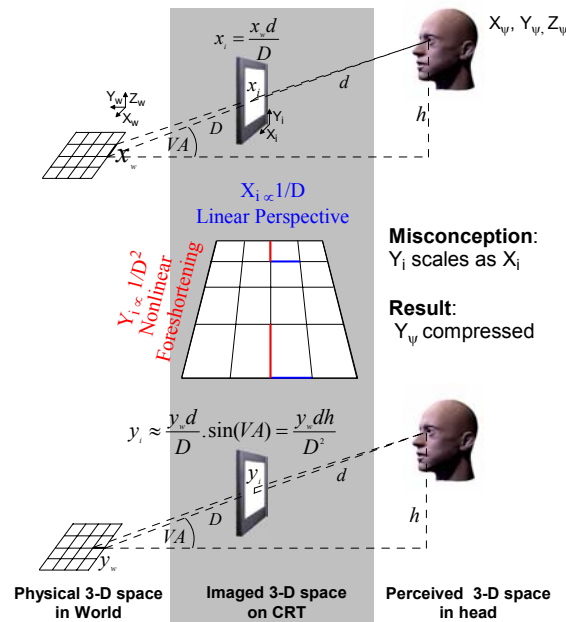


Figure 3: The Cross Scaling Model of 3-D Space (Mis)Perception [Smallman 2002; 2003a]. Humans apparently employ a simplifying assumption when reconstructing realistic perspective views that leads to severe underestimation of distances with shallow, 3-D, viewing angles.

We asked participants to view a scene composed of poles standing or laying on a large textured ground plane. We measured the accuracy of their perceptual reconstruction of the scene from the matches they made of the poles lengths in the cardinal (height, width, depth) directions of visual space when the scene was shown to them from different viewing angles (from 90 deg down to 22.5 deg). We found an unexpected and systematic pattern of errors in their matches with viewing angle [Smallman 2002]. Widths across the scene were always reconstructed (matched) accurately but depths into the scene became increasingly underestimated as the viewing angle grew more shallow. We developed a simple model, the Cross Scaling Model, to explain this underestimation of depth by assuming that participants fail to understand how widths and depths compress differently in perspective and that they instead employ a simplifying assumption that they compress at the same rate – they cross scale depths by width, see Figure 3. Further, since the difference in compression between widths and depths grows increasingly severe at shallower viewing angles, the Cross Scaling Model predicts the observed increasingly worse misperceptions at shallower viewing angles as participants’ naïve assumption increasingly fails to match reality.

The Cross-Scaling model also explains why 2-D views are superior for relative position judgments [section 2.1, above] because a 90 degree (2-D) viewing angle is the only one where the misconception is irrelevant. We subsequently validated the model by showing it could predict both the pattern of errors made by

participants as they attempted to trace out equal distance intervals across 3-D perspective views, and it could predict the geometry of simple line drawings that participants made of 3-D scenes from shallow viewing angles [Smallman 2003a]. Supporting the notion that the misconception is widespread is the fact that even experts appear to make it. The misconception is evident in the diagrams of a perception textbook author [see Gillam 1981].

It was always widely appreciated that 3-D perspective views possessed certain ambiguities but these were thought to be fixable with the addition of certain augmentations. The location of objects along a line of sight in any display is ambiguous [Sedgewick 1986]. This is the reason that altitude is inherently ambiguous in 2-D. The line of sight ambiguity in 3-D displays, on the other hand, extends to all three dimensions meaning that augmentations such as drop-lines or drop-shadows must be added to aircraft symbols in 3-D displays in order to disambiguate their ground-plane locations. For example, in Figure 1, left, drop-shadows are shown under each plane's location. In fact, in an early study in the project we showed that drop lines lead to better precision than drop-shadows, simply because the line's termination on the ground plane is more sharply defined than an icon's shadow [Smallman 2000a]. In addition to line of sight ambiguity, earlier NASA work had shown setting the field of view of the viewing camera too large or too small led to errors perceiving the pitch of an aircraft tethered to the ground-plane via a drop line [McGreevy 1986]. But this error, too, was also fixable with additional augmentations between airtracks and the ground plane. Our work shows that 3-D perspective views may be **inherently** misperceived because of the visual system's use of simplifying heuristics for a flawed, active scene reconstruction process. Interpreting realistic 3-D perspective views is thus anything but effortless or flawless.

Given the pervasive and pernicious nature of the perspective misconception, how can it be combated? It seems unlikely that we will be able to convince the world to abandon 3-D views, see section 4.0 below. Therefore, methods should be developed to limit the misperception, educate users to the problems, and guide them toward more useful viewing angles. Grids on the ground-plane are an obvious augmentation but they are cluttered and space within each grid element is still open to misinterpretation and performance with them is still not error-free [Smallman 2002]. Consequently, grids represent a trade-off and raise the question of when they should be shown and at what granularity. We have recently begun to apply the Cross-Scaling Model to the development of novel display concepts that both alert users to the extent of their likely perceptual error with a given viewing angle and how the error is likely to change with different viewing angles. Figure 4 shows our applied perception technology (APT) 3-D prototype in the domain of littoral strike-route planning. A viewing angle controller is shown, bottom right, that has integrated into it the Cross-Scaling Model's predicted ground plane misperception. A red warning region (a 'Shark's fin') indicates for each viewing angle the difference between the perceived middle of the scene (left side) and the actual middle of the scene (right side). The width of the fin thereby represents the degree of misperception for each viewing angle. The yellow camera indicates the current viewing angle. Note that the misperception applies across the entire scene, but the shark's fin only represents the effect on the middle of the scene. Thus, at 90 degrees, the 2-D top-down view, there is no red area and no misperception while at shallower viewing angles, the misperception grows increasingly severe. At the 22.5 degree viewing angle shown in Figure 4, the middle air defence dome, which actually lies in the middle of the scene, appears shifted closer to the rear dome than to the front one when, in fact, the three domes are equally spaced.

The shark's fin alerts users to the likely misperception at the rear of the display and shows them that by raising the viewing angle, they can experience a more veridical percept. Changes between viewing angles are accomplished by a faint yellow preview camera icon to the new viewing angle. Dragging the preview camera causes the 3-D view to morph smoothly to the new view, since smooth morphing has been shown to be a superior method to promote visual momentum between views [Hollands 2002; 2003]. Although still in the

early stages of development, the APT-3D tool shows promise at tackling the tricky problem of alerting users to their likely misperception of what they are looking at. The APT-3D tool bows to the power of the preference for 3-D views while gracefully informing users that more accurate views are available.

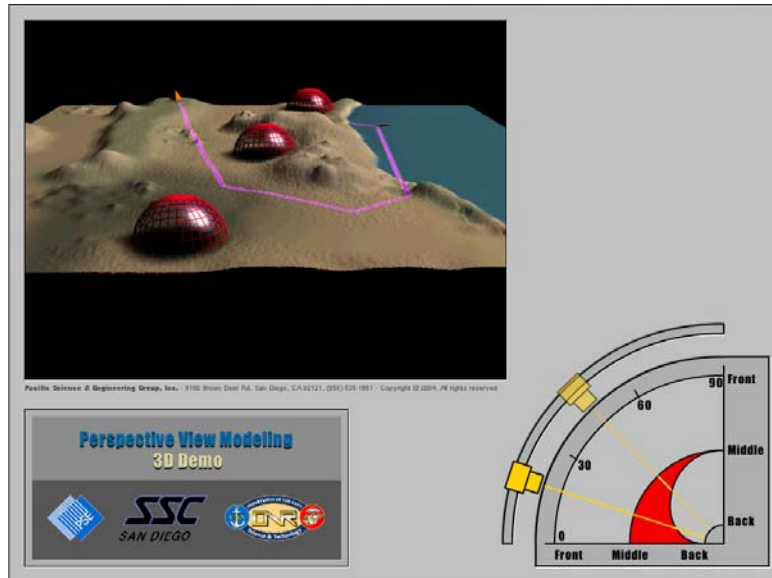


Figure 4: The APT-3D demo incorporates the Cross-Scaling model’s error predictions into a viewing angle controller (bottom right), here in the domain of land strike route planning. Copyright Pacific Science & Engineering Group 2004, all rights reserved.

2.3 WHAT of 3-D (depicting objects)

What refers to users’ ability to identify objects and their attributes from different views. Battlespaces are populated with a variety of force assets and these objects must be visualized somehow, normally either as icons or symbols, see Figure 5. 3-D views are popular partly because they tend to depict force assets as miniature, realistically rendered icons of ships and planes, whereas conventional 2-D displays depict assets as abstract military symbols [Smallman 2000c]. Across a battery of tasks including naming, memorization and visual search we found that although users rated 3D icons as preferable and likely to aid performance, comparable military symbols produced consistently superior identification performance [Smallman 2000a,b,c]. As we found for ‘Where’, above, the beguiling realism of 3-D perspective view displays actually serves to undermine their utility for many tasks. One reason for this problem is that an iconic code retains a visual similarity between the depicted object and its referent. When the set of depicted objects are inherently similar (e.g., many aircraft look somewhat alike, as do many ships) then users will have difficulty discriminating their icons and will misidentify them. Military 2-D symbols, however, are designed to be mutually discriminable.

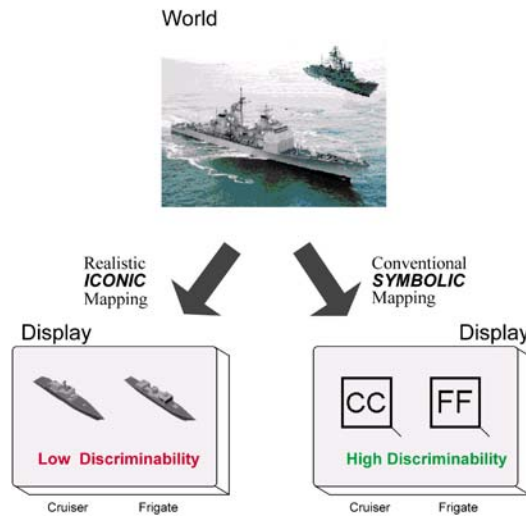


Figure 5: Two ways of mapping real-world military platforms onto a display, left, as realistic 3-D icons, right, as 2-D symbols - from the MIL-STD 2525B symbol set, from [Smallman 2000c].

The problems of icons are further compounded by rendering them realistically in perspective. The length of an altitude dropline is affected by distance, since altitudes compress toward the back of the scene in the same way as widths do, see Figure 6. In order to correctly perceive the altitude, the user must rescale the length of the dropline based on the available depth cues in the display. However, as just discussed in the ‘Why’ section, above, depth cue rescaling is unreliable. Furthermore, a 3-D view confounds the pitch of aircraft with their heading. Judged by their identical course leaders, both the aircraft in Figure 6 are flying in the same direction. However, one is flying level going southeast, while the other is descending going east. In visual search experiments we have found that search times for the heading, pitch and altitude of aircraft shown in perspective are much slower than for those shown in 2-D [Smallman 2001a,b] – this despite the fact that users predicted before the experiments that they would perform better with the 3-D icons. Users either believe they can compensate for these problems or they are oblivious to them. Echoing the picture that emerged from the ‘Where’ section above, users apparently do not understand how hard it is to extract information from realistic views, especially when precision is required.

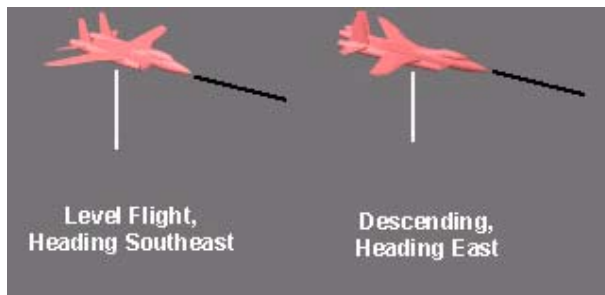


Figure 6: The beauty of 3-D icons is skin deep. Although the leaders and drop lines are almost identical, these similar icons code very different aircraft. Which aircraft is descending? Which aircraft is headed due east? Answers in white text under each icon. From [Smallman 2001b].

Given this discussion of the difficulty of interpreting 3-D icons, it may surprise to hear that an early study used the speed of detecting descending aircraft shown as realistic 3-D icons in perspective to bolster the claim that 3-D displays promoted more rapid SA than 2-D displays [Baumann 1997]. By following-up on this study we found additional clues to why 3-D displays are so liked by users and how their features can be employed to create better 2-D displays. A close look at the method in the Baumann et al. study reveals that the better performance of the 3-D display was in fact due to poor design of the 2-D, control display. In the 3-D display, pitch was coded analogically by the tilt of the realistic 3-D icons that was continuously available in a selection-free way on the display. In the 2-D control display, pitch was coded digitally in a text readout that was only available in a separate display when an aircraft's symbol was selected. When we controlled for these different information access methods, we found that analog coding was faster than digital coding and no-selection was much faster than selection but, when these interface details were equated, 3-D with icons were consistently slower than 2-D with symbols [Smallman 2001b]. Furthermore, the tasks in both studies involved little precision, the tasks simply involved discriminating descending planes from others. Had the tasks involved precise judgments of pitch (e.g., which aircraft are descending at 80°/sec), the 3-D display would probably have fared even worse. In sum, a key feature of 3-D displays that may not be immediately appreciated is that they portray the third dimension in an analog, selection-free, and integrated way. This probably contributes as much to their sustained appeal to users as does their realism. By designing 2-D displays with same advantages, we can create correctly perceived displays with easy information access. For example, although less natural and realistic, the introduction of altitude pegs with arrows to indicate ascent or descent onto symbols in a 2-D display should be highly effective.

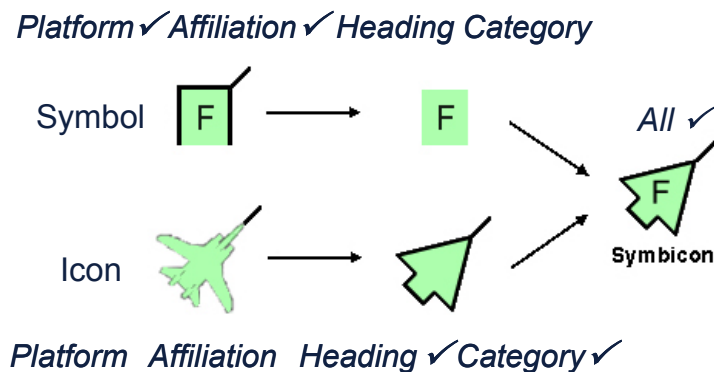


Figure 7: Symbicons are a new hybrid symbology designed to combine the best elements of symbols and icons. Although less realistic than icons, they are superior. From [Smallman 2001c].

Other lessons can be learned from the pattern of user performance with icons and leveraged to improve the design of symbols on tactical displays. When the issue of a 3-D perspective view is put aside and performance of icons and symbols are simply compared in a 2-D format, icons are, in fact, better for some tasks, such as identifying heading and platform category [Smallman 2000c; 2001a]. Icons are better for those attributes because they code them in a more visually conspicuous way than do symbols (e.g., by turning an entire icon's body in the direction of travel and using a leader vs. just using a leader alone). Symbols remain superior for platform identity and for affiliation (e.g. own forces, unknown, hostile). Symbols are better for those attributes because they code them in a more discriminable way than do icons. As discussed above, symbols can be arbitrarily discriminable and affiliation can be double-coded by shape and a uniform swatch of colour). Based these complimentary advantages, we developed a new, hybrid symbology called "Symbicons", designed to combine the best elements of both, see Figure 7. Symbicons use the symbolic coding of platform (e.g. F for

fighter aircraft), the uniform colour swatch for affiliation (e.g. green), the rotation of the entire symbol for heading (e.g. northeast), and the simplified outline for category (e.g. aircraft). Using a visual search task, we showed that this new, hybrid symbology was superior for all attributes [Smallman 2001c]. In follow-on work conducted at DRDC, Canada, Symbicons have been shown to also promote better awareness of heading changes in tactical displays than conventional, MIL-STD 2525B symbols [Keillor 2003].

In sum, users of the COP are likely to need to accurately and rapidly locate and identify force assets. Symbicons offer an attractive new, third way to depict these assets over the previously employed symbols and realistic icons. Symbicons caricature and pair down reality to leave users with a discriminable, conspicuous representation that enables task-relevant attributes be quickly extracted. Again, less realism leads to greater performance.

2.4 HOW of 3-D (displays for tasks)

How refers to how users should use different displays or suites of displays for different tasks. Complex operational tasks are likely to contain task elements that involve both relative position judgments (better in 2-D) and shape understanding (better in 3-D). Hence, the question becomes how can different display formats best be employed over time so that users can take advantages of the view that best fits the task as task needs change. We developed, and then empirically tested, a new human factors display concept called “Orient & Operate” whereby users orient to a display in 3-D for rough layout then switch to 2-D to operate and to make precision judgments [St. John 2001b]. We developed a quasi-operational terrain routing task where participants had to create a chain of antennas across the map so that consecutive antennas could “see” each other while remaining within operating range of each other. We found that antenna placement was performed better with the 2-D plan view, but that initial planning of the antenna route was performed better with the 3-D perspective view.

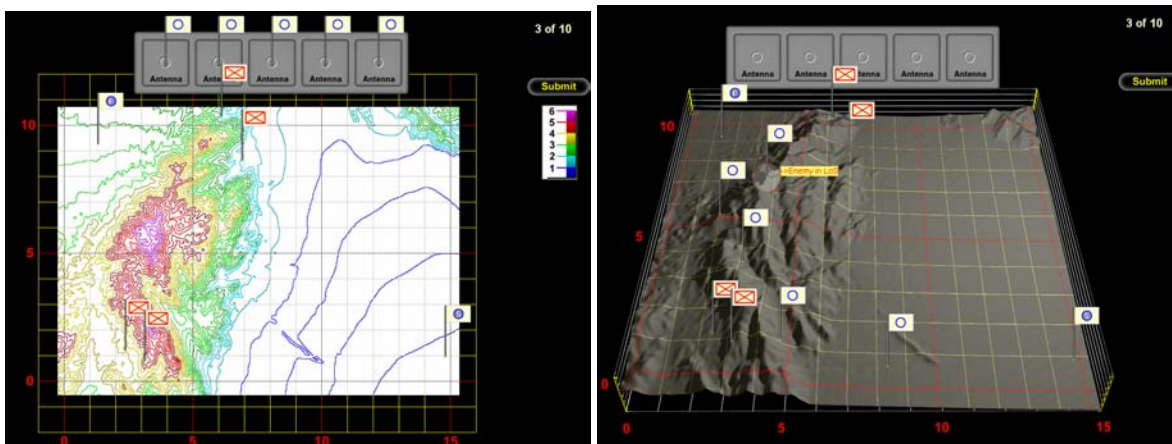


Figure 8: Users performing complex operational tasks are likely to need both 2-D and 3-D displays to Orient to 3-D then Operate in 2-D [St. John 2001b]. Here, the 2-D plan view (left) and 3-D perspective view (right) of the antenna placement experiment are shown side by side. Enemy positions are identified by flags with a red “X”. Antennas are identified by flags with blue circles. Other possible 2D/3D display combinations are discussed in the text.

How can the “Orient & Operate” human factors concept be most effectively embodied in the design of suites of 2-D and 3-D displays? In our original study, we evaluated perhaps the simplest combination which is a

suite of a 2-D and 3-D display of the same piece of terrain side by side, see Figure 8. We found a marginal benefit for the side-by-side condition over either display alone. The superiority may not have been pronounced because the display suite did little to help users solve the problem of determining corresponding locations between the two views. In follow-up, unpublished studies, we pitted several different display combinations against each other to see which might work best. These combinations included mini-orient 3-D maps, smooth morphs between 2-D and 3-D views, portals of one view superimposed on the other and the simple side by side combination again. However, we could not establish a statistically clear advantage for one combination over the others, although the smooth morphs yielded the overall best performance. One reason for the lack of a clear winner related to the complexities of the antenna task. In related work, Hollands and colleagues have concentrated on the smooth morphing combination in the context of our simpler terrain understanding (A-high-B and A-see-B) tasks and have shown that it yields better performance than a discrete change between static 2-D or 3-D displays [Hollands 2002; 2003].

In conclusion, the COP is likely to be used for a variety of complex tasks, some of which will require more shape understanding/gestalt elements performed better in 3-D and some of which will require more relative position/precision elements performed better in 2-D. We have developed and validated a new design principle called Orient & Operate, whereby users initially orient to, and occasionally re-orient to, 3-D displays to get a rough sense of scene layout, but spend the majority of their time making decisions and operating on 2-D displays for precision. Finally, there is evidence that enabling users to smoothly morph between views produces good visual momentum and a useful tool suite for simpler tasks, although it remains an open question whether the other techniques (e.g. mini-orient maps) could be as useful for more complex, operational tasks.

3.0 REAL-TIME REALISTIC DISPLAYS: MAINTAINING & RECOVERING SA

Even though most research has been focused on promoting SA, situation displays must support the maintenance of SA over time and, critically, the swift recovery of SA once it has been lost. In a separate project for the US Navy, we have been investigating users' ability to maintain and recover SA with tactical displays and the discrepancy between the nature of the tools that they desire, and those that they need, to support them in these tasks. The picture that emerges from this research interestingly parallels that on 3-D from section 2.0, above. While the 3-D research highlighted some of the limits of the utility of maintaining realistic spatial continuity with what is being shown, the research we review highlights some of the limits of maintaining realistic temporal continuity with what is shown, as well.

Users monitor situation displays over time primarily to determine if they have changed in any significant way. Unfortunately, several factors conspire to make changes extremely easy to miss. First, situations may be dense and complex, camouflaging changes and making them hard to find. Second, users may be under high workload, continually interrupted, or multi-tasking, and miss changes. Third, numerous cognitive studies have shown that even under near-ideal monitoring conditions, humans are remarkably poor at detecting changes – they exhibit a surprising “change blindness” [Rensink 2002]. Given these cognitive studies, it is surprising that situation displays have been designed with very little support for users in the critical task of reconstructing changes. By maintaining temporal realism and only representing the present state of the system, displays force users to remember and mentally integrate information over time to determine for themselves whether changes have occurred.

Most research into mitigating the effects of interruptions has focused on static task environments [e.g. McFarlane 2002] such as helping users pick up where they left off during a mission planning or problem solving task. This research has primarily focused on prospective interventions, such as negotiating the timing

of interruptions and helping the user prepare cues that can later be used to prompt memory for what users were trying to do before they were interrupted [Trafton 2003]. What makes monitoring potentially so difficult is that the task environment is dynamic and prospective interventions for recalling the interrupted situation are almost useless – the longer the time away, the more dissimilar the situation is from the one that was interrupted. That users need support was recently highlighted in a study that showed that changes to naval air warfare displays are often missed when users are interrupted when the change occurs [DiVita 2004]. In our own study, we confirmed that changes were missed with interruptions, and we also showed that change detection can be near-chance even while users are actively monitoring a display [Smallman 2003b]. In that study, we examined the efficacy of different HCI concepts for detecting changes to a tactical air warfare display. We examined both real-time detection and reconstruction of changes following a one minute interruption. One concept that we developed and evaluated was called CHEX (Change History EXplicit), see Figure 9. The concept of CHEX is to automatically detect and log changes in a table that can be flexibly sorted and filtered by the user to support a variety of tasking. The tool does not clutter the situation display with alerts, rather, users viewed changes by selected rows of the table, which caused all changes by the aircraft to become highlighted, as well as causing the aircraft to become highlighted on the situation display. Conversely, table entries of changes related to an aircraft were also available on selecting the aircraft on the situation display. We showed that CHEX resulted in 80% faster change detection and identification times than a baseline display that only showed the current situation on the map.

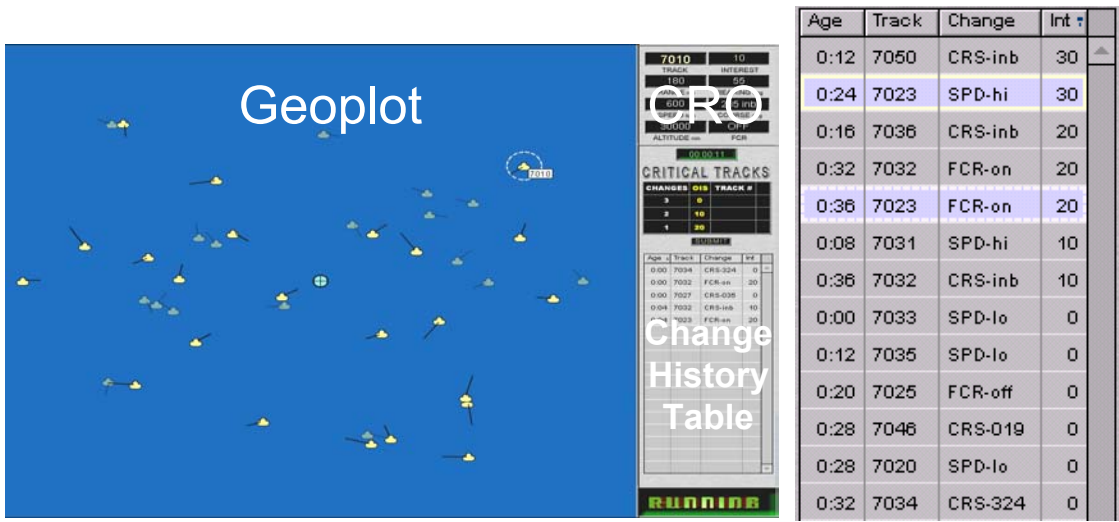


Figure 9: Left, a screenshot of the situation display, CRO (character read-out), and the Change History (CHEX) Table. Right, enlarged view of the CHEX table.

Interestingly, users were overconfident in their ability to spot changes, and they underestimated the extent that CHEX would help them. Others had shown that people are overconfident in their abilities and unaware of how poor they really are at detecting changes in natural scenes [Levin 2000]. In a study we have very recently completed, we compared the efficacy of CHEX to other change awareness tools, including hi-speed visual replay of the air warfare situation display [St. John 2004]. Replay is natural and realistic in that it maintains the temporal integrity of the actual sequence of events. Although intuitive and predicted as useful by colleagues and participants, replay was actually less useful than having no support tool at all, and it was far worse than the less realistic, explicit change information available with CHEX. In the next section, we introduce a theory to account for the bizarre patterns of mismatch between what users believe they want and what actually serves them well.

4.0 NAÏVE REALISM

An intriguing pattern has emerged from our exploration into the human factors of visualizing tactical information in 3-D space and time. Again and again we have found that users naïvely predict superior performance for, and strongly prefer, those displays that mimic and maintain the integrity of realistic scenes over non-realistic ones, in spite of demonstrably worse performance. We term this paradoxical behaviour Naïve Realism. We review its roots, causes and, we believe, very important implications for the way many displays are designed.

Naïve Realism has its origin in perception. The visual system delivers a 3-D perceptual world for us to experience and enjoy that is apparently rich and seamless, and that rich world is made available to us instantly and effortlessly upon opening our eyes. Unfortunately, the richness of this perception is easily mistaken for objectivity and fidelity. It leads us to harbour an inflated view of our ability to extract information from the world and from realistic depictions of it in displays. However, as others have wryly put it, “the mother of all illusions is the illusion of objectivity” [MacLeod 1995]. Back-stage, fully a third of our brain labours to keep up the ‘objective illusion show’ that is visual perception. And what a rickety production that show turns out to be! Rather than rich and seamless, current perceptual theory sees the representation of the visual scene as flawed and Spartan [Cavanagh 1995]. The spate of change blindness and related cognitive studies suggests that little is actually sensed of a scene other than a sample of fixations, but that the brain “fills in” the remainder and gives the viewer the sense of having an accurate representation of the entire scene [O’Regan 1992]. It is this sparse sampling that leads to change blindness and the occasional feeling of surprise at suddenly seeing something that has been “hidden in plain sight.”

Perception is flawed, not out of malice, but out of necessity. Making images of the world is a much easier proposition than the inverse process of interpreting the 3-D world that gave rise to those images. The brain employs a wide range of simplifying assumptions to make tractable the otherwise the tricky process of image interpretation and this “inverse optics.” For example, when interpreting a scene, the brain must disentangle the shape of a surface from the location of a light source even though both sources of information are conflated in the light intensity profiles falling on the retina. To solve the task, the brain simply assumes that the light source is above the surface [Ramachandran 1986]. Analogously, perspective views project three spatial dimensions into two-dimensional images. This projection results in massive ambiguity. There are an infinite number of different 3-D scenes that could give rise to the same 2-D image. Recovering the specific 3-D layout that gave rise to the perspective view requires assumptions analogous to disentangling the light source from the shape of an object. The Cross-Scaling model reviewed in section 2.4, above, reveals one such simplifying assumption used to recover the 3-D scene, the misconception of equal compression in width and depth with distance in perspective. Unfortunately, this assumption, like most other assumptions, creates systematic errors. The illusion of objectivity is that the ubiquity of these errors and the actual sparseness of visual sampling go unobserved, fostering and maintaining Naïve Realism. The brain is a master at concealing its tricks and only occasionally does one get to glimpse the real Wizard of Oz behind the curtain. For example, we laugh off and dismiss as an “illusion” the surprising morph of craters into bumps when we invert a photo of the moon’s surface – a demonstration of the light source assumption. In other cases, the tricks are kept literally out of reach. The absolute errors in perceived distance resulting from the Cross-Scaling misconception only start to reveal themselves at distances greater than arm’s length, and hence remain inconspicuous as we go about our busy lives.

That the sparseness of perception could also remains inconspicuous seems extremely hard to swallow. However, throughout history, that sparseness has been discovered and put to effect by various professions. For example, magicians and card sharks live off the permeability of visual attention. Film editors, too, have discovered that they can get away with dramatic lapses in continuity by simply cutting to a new point of view

to fool an audience. Only recently has cognitive science begun to systematically study these phenomena with change blindness studies etc. Participants in these studies are so convinced by the seamless nature of perception that they dramatically overestimate their change detection ability and are stunned by their inability to do the experimental tasks. People's own understanding of their perceptual abilities is actually becoming a focus of research in its own right [Levin 2003].

The overconfidence in perceptual abilities couples dangerously with a desire for familiarity in the displays people use. Because nothing appears more effortless and natural to users than the task they spend all their waking hours engaged in, seeing, they demand realism in their displays. Display designers are willing to oblige, as they want to satisfy their end-users. Also, the few concrete human factors principles that exist to guide them specifically encourage display designers to employ realism. For example, the principle of "Pictorial Realism" [Roscoe 1968] is engrained in current Human Factors textbooks and guidelines. This principle, and its companion principle, that of the "Moving Part" [Roscoe 1968; 1981], council that displays should maintain spatio-temporal correlation with their real world analogs in order to be intuitive and useful (i.e., if it moves in the world it should move similarly on the display). But displays contain both symbolic as well as spatial information [Ellis 1993], and increasing the realism can actually decrease interpretability by conflating spatial and symbolic information together. Recall, for example, the discussion of the difficulty of perceiving realistic icons in section 2.3 (Figure 6).

Developed to support post-war aviation design at a time when existing displays were obscure and user-unfriendly, Roscoe's principles were undeniably helpful. Now, driven by constant improvements in computer speed and technology, the design principles are being driven to extremes and slavishly followed. Technology is enabling designers to make realistic displays of such sophistication that they ever closer approximate to reality. Naïve Realism leads to the fallacy that the gold standard of displays is one that is almost entirely transparent because it shows the user what it is like to "be there." When presenting our own work on the limits of 3-D displays, we have often been challenged that everything would have turned out differently if we had just been able to add to the 3-D views all the missing depth cues that participants would have had in the real world. Of course, such super-realistic displays just throw users back on their own far from perfect abilities to extract information from natural scenes.

Naïve Realism accounts for a trend seen in a wide range of HCI domains. In the domain of telecommunications, Hollan and Stornetta issued an early rebuke to designers for "imitating the medium rather than facilitating the message" [Hollan 1992]. However, the trend persists. In the development of new virtual collaborative 'groupware spaces' there is a tendency among designers to mimic realistic discourse by rendering humans and workspaces realistically in near real-time 3-D virtual environments [e.g., Benford 2001]. Naïve Realistic expectations are so ingrained that they have led to superfluous research in some domains in order to maximize realism, even though low fidelity tools can offer superior functionality. In animation research, for example, there is a growing literature showing that animation sequences do little to support user comprehension of events over time, compared to static snapshots [Tversky 2002]. Intuitions persist that animation's temporal realism must ultimately show its efficacy and so the literature grows.

Meanwhile, Naïve Realistic expectations have led to chronic under-attention in other domains where high-fidelity displays are thought to provide adequate support even though they actually do not. Our own SA recovery work has highlighted the fact that little research was even deemed necessary to improve on monitoring displays in order to support change detection. It was thought that the temporal sequence of events was sufficient to support change detection, whereas our studies of SA recovery have found that explicit change detection support is actually required for users to achieve good change awareness [Smallman 2003b]. CHEX is effective because it temporally filters the situation display to just extract and explicitly represent

changes and discards the chaff. Nevertheless, in our recent experiment we found that users continued to employ replay tools at a consistent rate throughout the study in spite of the fact that they offered no advantage [St. John 2004].

Good display design is more than slavishly adhering to realism. As Tufte put so succinctly, “design is choice” [Tufte 1983]. Design must be informed both by the information requirements of the tasks for which the displays are used and knowledge of how the mechanisms of visual perception are likely to transform and represent what is shown. Displays should highlight task-relevant information. This process of highlighting inevitably entails pairing down reality, which immediately creates a tension with the Naïve Realism display philosophy. For example, realism means that ephemera are shown in a display (e.g., shaded texture gradients across a 3-D landscape) that serve to mask the extraction of symbolic information that users care about (e.g., the location and identity of an antenna on a hillside). If perception of natural scenes were fully faithful, this would be less problematic. But as we have elaborated here at length, that is far from the case. The problem we have highlighted here is that the zeitgeist of basic perceptual science and display design philosophy are completely misaligned. Users and designers are locked in an unhealthy conspiracy that neither party is guilty or conscious of to create increasingly photo-realistic, real-time displays which beguile but under-perform.

5.0 CONCLUSIONS AND RECOMMENDATIONS

In this paper we have reviewed a series of studies bearing on the human factors of how to visualize the COP. One approach to visualizing the COP is to render it realistically in 3-D perspective view, populate it with icons and update it in real-time. This Naïvely Realistic display would certainly be compelling and popular with naïve users. However, a detailed consideration of the architecture of visual cognition paired with a consideration of task requirements point to a set of guidelines for the COP that, though sometimes counter to naïve intuition, should promote far better performance. Of course there are many more Human Factors guidelines that relate to the design of the COP, but the ones listed below arise directly from the research reviewed above.

- Users should be trained to, and the display should encourage, orienting to 3-D perspective views for scene layout, then switching to operate in 2-D
- Users should be trained to, and the display should encourage, shape judgements be performed in 3-D perspective view, and precise relative position judgements in 2-D plan views
- 3-D views should be available either continuously either as mini-orient views or on-demand from 2-D plan views via a smooth morph to shallower viewing angles
- Depth misperception should be minimized at shallow viewing angles by the use of tools such as the APT-3D “shark fin” viewing angle controller to warn users of the misperception dangers
- The use of realistic icons in 3-D perspective views should be avoided. Symbols, in particular, Symbicons, are recommended
- For those object attributes that are not available in the display’s native format (e.g., altitude and attitude in 2-D), the symbology should continuously make available discriminable, explicit analog representations of those attributes
- Explicit change history information, such as the CHEX tool, should be made available for SA maintenance and recovery
- Naïve Realism is no panacea for creating effective display designs, and there is no substitute for cognitive task analysis and decision-centered display design.

6.0 ACKNOWLEDGEMENTS

We gratefully acknowledge the technical assistance of several developers at PSE over the course of these projects including Tim Bank, Dan Manes, Gene Averett, Geoff Williams and Mike Quinn. This work was sponsored by the Office of Naval Research, program managers, Drs. Helen Gigley and Astrid Schmidt-Nielsen.

7.0 REFERENCES

- [Andre 1995] A.D. Andre, and C.D. Wickens, “When users want what’s not best for them”, *Ergonomics in Design*, pp. 10-14, 1995.
- [Baumann 1997] J.D. Baumann, S.I. Blanksteen, and M. Dennehy, “Recognition of descending aircraft in a perspective naval combat display,” *Journal of Virtual Environments*, 1997. Available, <http://www.hitl.washington.edu/scivw/JOVE/Articles/mdjbsb.html>.
- [Benford 2001] S.D. Benford, C.M. Greenhalgh, T. Rodden, and J. Pycocock,” *Collaborative Virtual Environments*”, *Communications of the ACM*, vol. 44 (7), pp. 79-85, 2001.
- [Cavanagh 1995] P. Cavanagh, “Guest editorial: vision is getting easier every day”, *Perception*, vol. 24, pp. 1227-1232, 1995.
- [Dennehy 1994] M.T. Dennehy, D.W. Nesbitt, and R.A. Sumey, “Real-time three-dimensional graphics display for anti-air warfare command and control,” *Johns Hopkins APL Technical Digest*, vol. 15, pp. 110-119, 1994.
- [DiVita 2004] J. DiVita, R. Obermayer, W. Nugent, and J.M. Linville, “Verification of the change blindness phenomenon while managing critical events on a combat information display”, *Human Factors*, vol. 46(2), pp. 205-218, 2004.
- [Ellis 1993] S.R. Ellis, “Pictorial communication: pictures and the synthetic universe,” in *Pictorial communication in virtual and real environments*, S.R. Ellis, M. Kaiser, and A.J. Grunwald, Editors. London: Taylor and Francis, pp. 22-40, 1993.
- [Endsley 1995] M.R. Endsley, “Toward a theory of Situational awareness in dynamic systems”, *Human Factors*, vol. 37(1), 32-64, 1995.
- [Gillam 1981] B. Gillam, “False perspectives”, *Perception*, vol. 10, pp. 313-318, 1981.
- [Hollan 1992] J.D. Hollan, and S. Stornetta, “Beyond Being There”, *Proceedings of ACM CHI’92 Conference on Human Factors in Computing Systems, Perspectives on the Design of Collaborative Systems*, pp. 119-125, 1992.
- [Hollands 2002] J.G. Hollands, and N. Ivanovic, “Task switching with 2D and 3D displays of geographical terrain: the role of visual momentum”, *Workshop presentation at NATO IST-021RTG-007, Information visualisation in massive military datasets*, Army Logistics and Management College, Halden, Norway, 2002.

[Hollands 2003] J.G. Hollands, N. Ivanovic, and Y. Enomoto, "Visual momentum and task switching with 2D and 3D displays of geographical terrain", Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society, pp. 1620-1624, 2003.

[Keillor 2003] J. Keillor, L. Thompson, H.S. Smallman, and M.B. Cowen "Detection of Heading Changes in Tactical Displays: A comparison of two symbol sets", Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society, pp. 1610-1614, 2003.

[Levin 2000] D.T. Levin, N. Momen, S.B. Drivdahl, and D.J. Simons, "Change blindness blindness: The metacognitive error of overestimating change-detection ability", Visual Cognition, vol. 7, pp. 397-412, 2000.

[Levin 2003] D.T. Levin (Editor), "Thinking and Seeing: Visual metacognition in adults and children" Cambridge, MA: MIT Press.

[MacLeod 1995] D.I.A. MacLeod, and J.D. Willen "Is there a visual space?" In R.D. Luce, M.D. D'Zmura, D.D. Hoffman, G.J. Iverson, and A.K. Romney (Eds.), Geometric representations of perceptual phenomena: Papers in honor of Tarow Indow on his 70th. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 47-60, 1995.

[McFarlane 2002] D.C. McFarlane, "Comparison of four primary methods for coordinating the interruption of people in human-computer interaction", Human-Computer Interaction, vol. 17, pp. 63-139, 2002.

[McGreevy 1986] M.W. McGreevy, and S.R. Ellis, "The effect of perspective geometry on judged direction in spatial info instruments", Human Factors, vol. 28, 439-456, 1986.

[O'Regan 1992] J.K. O'Regan, "Solving the 'real' mysteries of visual perception: The world as an outside memory", Canadian Journal of Psychology, vol. 46, pp. 461-488, 1992.

[Rensink 2002] R.A. Rensink, "Change detection", Annual Review of Psychology, vol. 53, pp. 245-277, 2002.

[Ramachandran 1986] V.S. Ramachandran, "Guest Editorial: The neurobiology of perception", Perception, vol. 14, pp. 97-103, 1986.

[Roscoe 1968] S.N. Roscoe, "Airborne displays for flight and navigation", Human Factors, vol. 18, pp. 321-332, 1968.

[Roscoe 1981] S.N. Roscoe, "Aviation Psychology", Iowa City: University of Iowa Press.

[Sedgewick 1986] H.A. Sedgewick "Space perception", in Handbook of perception and human performance. Vol 1. K.R. Boff and J.P. Thomas, Editors. New York, NY: Wiley Press, pp. 2101-2157, 1986.

[Smallman 2000a] H.S. Smallman, E. Schiller, and M.B. Cowen, "Track location enhancements for perspective view displays", SPAWAR System Center San Diego, CA. Technical Report 1847, 2000.

[Smallman 2000b] H.S. Smallman, E. Schiller, and C.A. Mitchell, "Promoting rapid situation awareness in tactical displays: the role of 3-D perspective views and realistic symbols", Proceedings of the 44th Annual Meeting of the Human Factors and Ergonomics Society, pp. 394-397, 2000.

- [Smallman 2000c] H.S. Smallman, M. St. John, H.M. Oonk, and M.B. Cowen, “When beauty is only skin deep: 3-D realistic icons are harder to identify than conventional 2-D military symbols,” in Proceedings of the 44th Annual Meeting of the Human Factors and Ergonomics Society, pp. 480-483, 2000.
- [Smallman 2001a] H.S. Smallman, H.M. Oonk, M. St. John, and M.B. Cowen, “Searching for tracks imaged as symbols or realistic icons: A comparison between two-dimensional and three-dimensional displays,” SPAWAR Systems Center, San Diego, CA. Technical Report 1854, 2001.
- [Smallman 2001b] H.S. Smallman, M. St. John, H.M. Oonk, and M.B. Cowen “Information availability in 2D and 3D displays”, IEEE: Computer Graphics and Application, vol. 21, pp. 51-57, 2001.
- [Smallman 2001c] H.S. Smallman, M. St. John, H.M. Oonk, and M.B. Cowen, “‘SYMBICONS’: a hybrid symbology that combines the best elements of SYMBols and ICONS”, Proceedings of the 45th Annual Meeting of the Human Factors and Ergonomics Society, pp. 110-114, 2001.
- [Smallman 2002] H.S. Smallman, M. St. John, and M.B. Cowen “Use and misuse of linear perspective in the perceptual reconstruction of 3-D perspective view displays”, Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society, pp. 1560-1564, 2002.
- [Smallman 2003a] H.S. Smallman, D.I. Manes, and M.B. Cowen “Measuring and modelling the misinterpretation of 3-D perspective views”, Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society, pp. 1615-1619, 2003.
- [Smallman 2003b] H.S. Smallman, and M. St. John “CHEX (Change History EXplicit): new HCI concepts for change awareness”, Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society, pp. 528-532, 2003.
- [St. John 2001a] M. St. John, M.B. Cowen, H.S. Smallman, and H.M. Oonk “The use of 2D and 3D displays for shape understanding versus relative position tasks”, Human Factors, vol. 43, pp. 79-98, 2001.
- [St. John 2001b] M. St. John, H.S. Smallman, T.E. Bank, and M.B. Cowen, “Tactical route planning using two-dimensional and three-dimensional views of terrain”, Proceedings of the 45th Annual Meeting of the Human Factors and Ergonomics Society, pp. 1409-1413, 2001.
- [St. John 2004]] M. St. John, H.S. Smallman, and D.I. Manes, “Recovery from interruptions to a dynamic monitoring task: the beguiling utility of instant replay”, manuscript, in preparation.
- [Trafton 2003] J.G. Trafton, E.M. Altmann, D.P. Brock, and F.E. Mintz “Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal”, International Journal of Human-Computer Studies, vol. 58, pp. 583-603, 2003.
- [Travis 1995] A.R.L. Travis, S.R. Lang, J.R. Moore, and N.A. Dodgson, “Time-multiplexed three-dimensional video display,” Journal of the Society for Information Display, vol. 3, pp. 203-205, 1995.
- [Tufte 1983] E.R. Tufte “The visual display of quantitative information”, Cheshire, CT: Graphic Press, 1983.
- [Tversky 2002] B. Tversky, J.B. Morrison, and M. Betrancourt, “Animation: can it facilitate?”, International Journal of Human-Computer Studies, vol. 57, pp. 247–262, 2002.