Viewpoint Tethering for Remotely Operated Vehicles: Effects on Complex Terrain Navigation and Spatial Awareness

Justin G. Hollands and Matthew Lamb, Defence Research and Development Canada, Toronto, Ontario, Canada

Objective: The effect of viewpoint on the navigation of complex terrain and on spatial awareness was examined with the use of a simulated remotely operated vehicle.

Background: The ability to build terrain models in real time may soon allow remote vehicular control from any viewpoint. A virtual tether couples the viewpoint to the vehicle’s position and orientation, but shows more of the terrain than a fully immersive egocentric display. In this sense, it provides visual momentum by providing a view that incorporates egocentric and exocentric qualities.

Method: For this study, 12 participants navigated a simulated vehicle across complex virtual terrain using five different display viewpoints: egocentric, dynamic tether, rigid tether, 3-D exocentric, and 2-D exocentric. While navigating, participants had to avoid being seen by simulated enemy units. After the navigation task, participants’ spatial awareness was assessed using a recognition task.

Results: The tethered displays minimized the time during which the participant’s vehicle was visible to enemy positions. The egocentric display was more effective than exocentric displays (2-D or 3-D) for navigation, and the exocentric displays were more effective than egocentric for time seen during navigation and the recognition task. The tethered displays produced intermediate results for navigation and recognition.

Conclusion: Viewpoint tethering produced the most effective displays for minimizing time seen, but tethered displays were less effective than egocentric and exocentric displays for navigation and recognition, respectively.

Application: A tethered display is recommended for applications in which it is necessary to understand the relation of nearby locations to one’s own location.

Keywords: navigation, remotely operated vehicles, spatial awareness, spatial cognition, tethered displays, unmanned ground vehicles, virtual environments

INTRODUCTION

Because of advances in remote sensing and geospatial analysis technologies, the integration of data from local and regional sensors will provide remotely operated vehicle (ROV) or uninhabited ground vehicle (UGV) operators with dynamically updated 3-D models of terrain data. These models can potentially be viewed using any conceivable viewpoint on the terrain. In addition to conventional 2-D or 3-D representations, virtual or augmented reality visualizations allow the use of viewpoints that would not normally be feasible in reality. Designers are then faced with the choice of optimal viewpoint parameters that maximize human performance. In addition, there is mounting evidence that aerial viewpoints, as might be obtained from an uninhabited aerial vehicle (UAV), provide useful information for a ground robotics controller searching for a target (Chadwick, 2008; Chen & Clark, 2008). There may be distinct advantages to providing views on terrain different from that obtained “through the windshield” of an ROV.

It is widely accepted that the nature of the task dictates the best viewpoint on geographic terrain (Wickens & Hollands, 2000). Tasks involving shape understanding are best performed with 3-D viewpoints because all three dimensions are integrated into one representation (Hollands, Pavlovic, Enomoto, & Jiang, 2008; St. John, Cowen, Smallman, & Oonk, 2001), whereas 2-D viewpoints are best for tasks judging relative position, given the distortions associated with 3-D viewpoints (St. John et al., 2001; Wickens, Thomas, & Young, 2000). For navigation and wayfinding, local guidance is best performed from an egocentric perspective, whereas global spatial awareness tasks should be carried out with an exocentric, fixed viewpoint that shows most or all of the terrain (Wickens & Hollands, 2000).

Tethered Displays

Although the optimal viewpoint is task dependent, switching between multiple displays can
be disorientating and degrades performance (Hollands et al., 2008). The tethered viewpoint commonly used in computer gaming has been proposed as one solution (Colquhoun & Milgram, 2000; Wickens & Prevett, 1995). A tethered view couples the viewpoint with a vehicle’s position and orientation, but typically is higher up and behind the vehicle, showing more of the terrain. In this sense, it provides visual momentum (Woods, 1984) by providing a view that incorporates both egocentric and exocentric qualities. There is some evidence in support of the tethered concept: Wang and Milgram (2001) found that a tethered display produced better performance than an egocentric display for aerial navigation, and Wickens and Prevett (1995) found advantages with a tether-like display (vs. egocentric) for spatial awareness.

However, a rigid tether violates the principle of motion compatibility. As the operator directs the vehicle to the right, the viewpoint of the virtual camera shifts to the left. The rigid tether also behaves like a compensatory tracking system (Colquhoun & Milgram, 2000) and can produce a jerky motion with a long tether because any small angular change in vehicle heading produces a large displacement of the virtual camera and hence the entire visual scene (Wang & Milgram, 2009).

A dynamic tether (equivalent to a nonrigid, mass-spring-damper system), defined by the force equation \( F = (kd – cv) / m \), creates a display incorporating both compensatory and pursuit tracking attributes. The net effect is a smoother motion as the vehicle moves through the environment, because the nonrigid dynamics mean that abrupt changes in heading produce a slightly lagged motion of the camera that nullifies small abrupt heading changes. In general, a dynamic tether is constrained by several parameters including its length, \( l \); spring constant, \( k \); damping coefficient, \( c \); and mass, \( m \).

The tether’s natural frequency \( \omega \) is defined as \( \sqrt{(k / m)} \), while the damping ratio \( \zeta \) is defined as \( c / 2 \sqrt{mk} \). Wang and Milgram (2009) varied \( \zeta \) and \( \omega_n \) experimentally when using a virtual tether to control the path of a simulated aircraft through sinusoidal tunnels (6 degrees of freedom [DOF] aerial navigation). With the tether natural frequency \( \omega_n = 30 \text{ rad/s} \) close to the task frequency, the root mean square error for their participants was minimized with critically damped tethers (\( \zeta \) near unity, \( k = 900 \text{ kg/s}^2, c = 60 \text{ kg/s}, m = 1 \text{ kg} \)). A critically damped tether will tend to produce an exponential decay toward its final state without significant oscillation. These dynamic tethers were found to be more effective than the rigid tether for this task.

Ground-based navigation differs from aerial navigation in that the operator controls only one rotational DOF, yaw. Yu and Andre (1999) used an arcade-driving simulator to evaluate four viewpoints. Participants drove a vehicle on a simulated race track. A tethered viewpoint slightly removed from the vehicle (“near exocentric”) provided the best driving performance, and a mid-exocentric viewpoint provided the best situational awareness. However, this driving task differed in several fundamental ways from off-road navigation in complex terrain. For example, there was no route choice in their task, and there was no requirement to use or remember the spatial layout beyond the immediate demands of the driving task.

Nielsen, Goodrich, and Ricks (2007) developed what they called an ecological interface in which camera imagery from an ROV was combined with a virtual environment consisting of a map floor and a virtual robot placed on the map, as viewed from behind and above. As the participant drove the ROV, the camera imagery and map were updated in coordinated fashion. This method produced a net display effect similar to a tethered display. In comparisons with a set of 2-D views of the same information, the ecological interface developed by Nielsen et al. was shown to be more effective for navigation through various maze environments. Although they examined a range of navigation tasks, Nielsen et al. did not assess spatial awareness directly. Furthermore, although their maze environments included occluding walls and objects, there was no variation in elevation (i.e., the floors that the robots traversed were flat).

In summary, the existing literature suggests that tether-type displays can offer a useful display perspective and show some advantages for navigation and situation awareness compared with some other display formats. It is not clear, however, how these results would extend to the case of ground-based navigation of an ROV over complex terrain varying in elevation. An experimental evaluation of tethered displays in this situation has not yet been.
explored, perhaps because the use of a tethered display for ROV navigation has only recently become feasible.

3-D Terrain Models in Real Time?

If a 3-D terrain model is built and updated in real time with imagery obtained from the ROV, the model can potentially be viewed from any conceivable viewpoint. A related possibility is that a small UAV is virtually tethered to the UGV and the model is generated or updated from imagery obtained from cameras on both uninhabited devices (in fact with this setup, one could just use imagery from the UAV directly) or that a preexisting model is augmented or integrated with UAV or other sensor imagery, as with the Terra-Sight system (Sarnoff Corporation, 2008). We assume that a reasonably accurate terrain model is available or can be constructed from imagery data in near-real time. At present, 3-D reconstruction algorithms exist that can build such a model in real time at limited resolution with the use of stereoscopic imagery (e.g., Pollefeys et al., 2008; Sizintsev et al., 2010; Sizintsev & Wildes, 2010), but the algorithms are developing at a rapid rate, and the availability of a 3-D model is becoming an increasingly common assumption for robotic control (e.g., Bruemmer, Few, Hunting, Walton, & Nielsen, 2006), UAV systems (e.g., Calhoun, Draper, Abernathy, Delgado, & Patzek, 2005; Hunn, 2005, 2006), and navigation systems (e.g., Cornelis, Leibe, Cornelis, & Van Gool, 2008). An accurate 3-D model offers a wide range of virtual camera positions and associated display options whose effectiveness should be assessed for both navigation and spatial awareness tasks. Thus, our aim here was to select some potentially useful viewpoints for a navigational task involving an ROV. We were also interested in how different viewpoints might affect a user’s spatial awareness of the terrain being navigated, both during navigation and afterward.

Given current trends, in the future it is likely that obstacle avoidance and other low-level driving task components will be automated in ROV navigation (see, e.g., Shimoda, Kuroda, & Iagnemma, 2005; Yoon, Shin, Kim, Park, & Sastry, 2009). Thus, the human operator working with such automation in the future will be doing less actual driving and will be asked to place greater focus on higher-order route choice and planning tasks.

Navigation and Spatial Awareness

For our experiment, we were interested in the navigation of an ROV and an operator’s spatial awareness more so than the actual manipulation and control of the ROV. We did not ask participants to drive a vehicle through a complex environment in which there were many objects (trees, buildings) and in which there was risk of collision (leading to the need to reverse the vehicle, assign penalties to collisions, etc.). Instead, we created an environment that was relatively free of occluding objects other than a few waypoints and the terrain itself—keeping the environment as simple as possible—and then examined the selection of a route through terrain varying in elevation (a navigation task) that would meet certain criteria (concurrent spatial awareness). Our intent was to examine navigation rather than the mechanics of driving per se, given the future likelihood of autonomous driving algorithms for future ROVs. In addition, there are many austere geographic regions that are relatively “featureless” (e.g., desert, arctic, alpine), and such regions are often of military or strategic interest (e.g., Afghanistan).

In particular, we asked our participants to choose a route through a set of waypoints that would minimize the time they could be seen from a set of locations marked on the terrain. We define spatial awareness as a comprehension or understanding of the properties of a 3-D space. The mental representation supporting the awareness can be egocentric or exocentric. During navigation, it is likely that such spatial awareness is maintained in an egocentric reference frame. A good measure should indicate how well a user’s egocentric location is understood relative to other terrain locations—locations that are task relevant because they should influence navigational choices. We refer to spatial awareness during navigation as concurrent spatial awareness.

In contrast, a user should also retain some spatial awareness of an environment after navigating it. In this case, we are interested in the generic layout or global properties of the terrain—what Thorndyke and Hayes-Roth (1982) called “survey knowledge”—rather than detailed information about specific locations. We refer to this type of spatial awareness as retrospective spatial awareness. To the extent that the user maintains a representation of generic layout, he or she should be
able to discriminate the navigated terrain from unfamiliar terrain. Thus, the user should be able to recognize just-navigated terrain among a set of similar terrains shown in an exocentric manner.

Spatial awareness as defined here likely bears some relation to the well-known situation awareness concept (Endsley, 1995), although we did not intend to measure situation awareness per se in our experiment.

**Experimental Manipulations, Predictions, and Measures**

In a simulated environment, we examined the effect of viewpoint on controlling a tank navigating complex terrain and on concurrent and subsequent spatial awareness. Five viewpoints were used (see Figure 1): egocentric, rigid tether, dynamic tether, 3-D exocentric, and 2-D exocentric (God’s-eye view). We predicted that navigational performance would be better in the egocentric view compared with 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 (Wickens & Hollands, 2000). We also predicted that the tethered display would be more effective than the exocentric displays for navigation and as effective as the egocentric display. Finally, we predicted that dynamic tethering would improve navigational performance compared with rigid tethering, by reducing the motion compatibility problem and allowing pursuit tracking.

In general, we adopted those dynamic tether parameters shown to be most effective by Wang and Milgram (2001, 2009). However, we used a relatively long tether length. We found in pilot testing that with shorter tethers, the hills near the tank would occlude objects of interest for the navigation task. Even with long tethers, however, we found that egocentric characteristics were retained more than one would expect because the viewing position shifted with the changing tank position and, less surprisingly, were effective at showing global terrain characteristics as well.

To assess concurrent spatial awareness, while participants navigated between waypoints, they were instructed to avoid being seen by simulated enemy units (these were simple markers in the terrain). The time that the tank was seen from at least one of the enemy positions was recorded. We refer to this measure as *time seen*.

After each trial, to assess retrospective spatial awareness, we asked observers to distinguish the terrain just navigated from a set of distractor terrains. We measured the accuracy of this judgment (*recognition accuracy*) as well as the time taken to decide (*recognition response time* [RT]). We argue that both of these tasks require spatial awareness and a sense of the global characteristics of the terrain. Thus, we predicted the opposite order of effectiveness: Exocentric should be worse than egocentric, and the tethered display should be as effective as the exocentric display. We did not predict any effect for dynamic (vs. rigid) tethering for either spatial awareness task. Table 1 lists each specific dependent measure as a function of the generic task being assessed.

We measured each participant’s map-reading ability using the Map Reading Test (Goerger, 1998). This test was designed to determine whether an individual can read the terrain features on a map and associate them with real-world terrain features. We were interested in determining whether map-reading ability affected performance on the navigation and spatial awareness tasks.

**METHOD**

**Participants**

For this study, 12 (7 female, 5 male) volunteers ages 18 to 50 (mean age = 29) were recruited. Most of the participants were in their 20s or 30s. All had normal or corrected-to-normal vision and were recruited from Defence Research and Development Canada–Toronto and the nearby community. Participants were financially compensated for their participation.

**Stimuli and Apparatus**

We created 11 terrain maps from digital elevation data using Multigen Paradigm Creator software (1 terrain map for familiarization + 10 for experimental trials). Mountainous terrain south of Lake Tahoe and west of Bridgeport, California, was selected for the terrain maps. Digital terrain elevation data (DTED) were obtained from Creator Terrain Studio at 30 m resolution (DTED Level 2). Vector topographical lines were created.
with the use of U.S. Geological Survey (USGS) hypsography data (1:100,000 scale). Each terrain map was one 7.5° latitude × 7.5° longitude quadrangle (approximately 14 km × 11 km). Quadrangles that contained mountainous regions were chosen to provide sufficiently complex terrain for our experimental tasks.

Each terrain map was rendered in a simulated environment with the use of C++ and Open GL on a Windows 2000 workstation with a 3Dlabs Wildcat III 6110 graphics card. The environment had lighting and shading but did not have gravity or collision detection. The tank was placed at a fixed distance above the surface of the terrain; it essentially “floated” on the terrain surface. The environment was constructed so that no collisions could take place and the tank could always move forward. Environments were presented on a liquid crystal display (LCD) measuring 50.8 cm diagonally at 1,024 × 768 resolution. The virtual camera’s geometric field of view was held constant across displays conditions (45.0° vertical, 57.8° horizontal).

The vehicle was a 3-D model of a tank, controlled by a joystick. The tank was not to scale with respect to the terrain. It was approximately 20 times larger than actual size. In effect, the mountains from the DTED data represented hills for our simulation. Given the scale adjustment, we defined graphical display units (GDUs) to refer to distances in our virtual terrain, rather than using kilometers. Each terrain model was approximately 1.42 × 1.14 GDUs.

Two initial tank positions (ITPs) were defined for each terrain. A set of three waypoints was selected for each ITP. ITPs and waypoints were determined randomly with the following constraints: All points were chosen to be within the central two thirds of the terrain. Because it is more effective to hide in low terrain, waypoints and the ITP were chosen to be below the average of all points on the terrain. The ITP and the waypoints were kept at least 0.3 GDUs apart from each other. The distance between consecutive waypoints (including the distance between the ITP and the first waypoint) was less than 0.35 GDUs.

For each ITP, three enemy positions were chosen randomly with the following constraints: Enemy points were restricted to those heights that were above average (to simulate lookout posts) and were chosen so as not to be visible from the waypoints or the initial tank position. The enemy positions had to be greater than 0.3 GDUs from each other, from any of the waypoints, and from the ITP.

Waypoints were represented by blue spheres on posts; cube markers on posts represented enemy positions. These terrain maps were shown with five viewpoints (egocentric, rigid and dynamic tether, 3-D exocentric, and 2-D exocentric). Three status bars representing line of sight from enemy positions were shown in the upper left of the display. A heading indicator was placed at the bottom of the display. Figure 1 shows examples of the five display types with sample terrain.

The exocentric (fixed) display conditions were defined by the distance the virtual camera was placed from the terrain ($l_{\text{fixed}}$) and the zenith angle of the viewpoint from the terrain ($\phi_{\text{fixed}}$). For the 2-D exocentric condition, the virtual camera was located 1.80 GDUs along a normal vector centered on the terrain such that the entire terrain could be seen within the field of view of the display ($l_{\text{fixed}} = 1.80$ GDUs, $\phi_{\text{fixed}} = 90°$). For the 3-D exocentric condition, the virtual camera was located 1.80 GDUs along a vector 45° from normal ($l_{\text{fixed}} = 1.80$ GDUs, $\phi_{\text{fixed}} = 45°$). A schematic representation is shown in Figure 2.
The tethered display conditions were defined by the distance between the virtual camera and the tank (tether length $l_{tether}$) and the zenith angle of the viewpoint from the pitch of the tank ($\phi_{tether}$). The tethered virtual camera had a rest position 1.00 GDU along a vector elevated 30° up and behind the tank roughly halfway between the egocentric and exocentric conditions ($l_{tether} = 1.00$ GDU, $\phi_{tether} = 30^\circ$). For the egocentric condition, the virtual camera was placed within the tank, producing a viewpoint akin to looking out the tank’s “windscreen” ($l_{tether} = 0$ GDUs, $\phi_{tether} = 0^\circ$). The virtual camera for the egocentric condition was fixed to the orientation of the tank and could not be independently adjusted.

Parameters specific to the motion of the dynamic tether include its length ($l$), the mass of the viewpoint ($m$), its translation spring constant ($k_{tr}$), its...
rotation spring constant ($k_{ro}$), its translation damping coefficient ($c_{tr}$) and its rotation damping coefficient ($c_{ro}$). The dynamically tethered viewpoint mass was assigned a value of 1 kg. As described by Wang and Milgram (2002), the translation of the dynamically tethered virtual camera was subject to the force $F_{tr} = \frac{(k_{tr}d_{tr} - c_{tr}v_{tr})}{m}$ and its rotation by $F_{ro} = \frac{(k_{ro}d_{ro} - c_{ro}v_{ro})}{m}$, where $d$ is displacement of the virtual camera from the tank (initially determined by $l_{tether}$ and $\phi_{tether}$) and $v$ is the velocity of the camera ($d$ and $v$ each having linear and angular components).

In general, our parameters (except for length) and tether behavior were the same as those used in the DR4 condition described by Wang and Milgram (2009). The translation and rotation spring constants $k_{tr} = k_{ro} = 900$ kg/s$^2$ and damping coefficients $c_{tr} = c_{ro} = 60$ kg/s were chosen heuristically on the basis of the response of the tether to tank motion over the terrain, given the tether length, and to maintain a critically damped system with damping ratio $\zeta = c / 2\sqrt{mk} = 1$, as found to be optimal by Wang and Milgram. The choice of mass and spring constants gave the dynamic tether the natural frequency $\omega_n = \sqrt{k/m} = 30$ rad/s. A schematic of the egocentric and tether conditions is shown in Figure 2.

**Design and Procedure**

The experiment had a one-way within-subject design with five conditions: egocentric, rigid tether, dynamic tether, 3-D exocentric, and 2-D exocentric.

Each participant completed three 1-hr sessions in a maximum of 4 days. In the first session, participants completed the Map Reading Test and then performed a familiarization task in which they explored the five display conditions, the joystick, and the general environment. Written and oral instructions were provided to the participant.

For the experimental trials, each of the 10 terrains had two sets of enemy locations and
waypoints, or scenarios, leading to 20 trials for each of the five conditions, or 100 trials in total. The 100 trials were run in five blocks of 20. Block 1 was run on the first day, Blocks 2 and 3 on the second, and Blocks 4 and 5 on the third. For each block, 20 trials were randomly selected from this set of 100. This procedure meant that the order of conditions was randomized within each block, and each participant had a different random order.

For the navigation task, the participant was instructed to drive the tank from one waypoint to the next as quickly as possible while remaining out of sight of enemy positions. The shape of the terrain affected the tank’s speed. Traveling on flat terrain, the tank had a constant speed of approximately 92 GDUs per hour. The tank moved faster going downhill and slower going uphill. If the terrain caused the tank to roll, it would also slow the tank’s speed. Speeds ranged from 1 GDU per hour to 205 GDUs per hour with an average speed of approximately 82 GDUs per hour.

Participants moved the joystick left and right to control yaw rotation, the former to rotate the tank clockwise and the latter counterclockwise. When the tank entered the line of sight of an enemy position, the enemy cube marker and the corresponding status bar changed color (from black to red). The heading indicator showed the direction of the next waypoint in the tank’s forward 180° field of view.

The tank’s position was sampled approximately every 100 ms. The task commenced 2 s after the simulation had loaded. The trial ended when the third waypoint had been reached. Path length was correlated with navigation time (longer paths generally took more time to traverse) and is therefore not separately reported.

The amount of time the tank was within the sight of any enemy position was recorded, and this...
time-seen measure was used to assess concurrent spatial awareness. A four-alternative forced-choice recognition task followed the completion of each trial and provided rotating images of four terrains. Participants were required to choose which terrain they had traversed in the navigation task. Recognition accuracy and RT were recorded. Participants were debriefed at the end of the experiment.

RESULTS

Performance Measures

For each performance measure, a mean score was computed for each participant in each display type condition for each terrain. These means were submitted to a two-way 5 × 10 (Display Type × Terrain) within-subject ANOVA in each case. Planned comparisons were computed on those contrasts relevant to predictions. The alpha level was set to .05 for all significance tests. Figure 3 depicts the mean values for each measure and associated within-subject standard error values (Jarmasz & Hollands, 2009).

Navigation Time

Display type affected navigation time, F(4, 44) = 6.94, MSE = 10.4, p < .001. Navigation time was shorter for the egocentric than exocentric displays, F(1, 11) = 8.20, MSE = 234.88, p < .05. Navigation time was also shorter for the egocentric than for the tethered conditions, F(1, 11) = 5.70, MSE = 168.97, p < .05. Although mean navigation time for tethered displays was shorter than for exocentric displays, the difference merely approached conventional significance levels, F(1, 11) = 2.84, MSE = 86.98, p = .12. There was no difference in navigation time between rigid and dynamic tethers, F < 1.

Terrain also affected navigation time, F(9, 99) = 21.31, MSE = 138.00, p < .0001. An examination of the distance between waypoints showed that the longest navigation times occurred for the terrains having the longest distances between waypoints. There was no interaction between display type and terrain, p > .05.

Time Seen

One terrain had to be excluded from the 5 × 10 ANOVA because of a mean value of zero (averaged across participants) in one display condition for that terrain. Display type affected time seen, F(4, 44) = 8.04, MSE = 17.29, p < .0001. Time seen was shorter for exocentric than for egocentric displays, F(1, 11) = 6.44, MSE = 26.99, p < .05. Time seen was shorter for tethered than for egocentric, F(1, 11) = 18.87, MSE = 27.57, p < .005, and exocentric displays, F(1, 11) = 7.65, MSE = 17.58, p < .05. There was no difference in time seen between rigid and dynamic tethers, F < 1.

Terrain affected time seen, F(8, 88) = 42.47, MSE = 35.99, p < .0001, but more importantly, terrain interacted with display type, F(32, 352) = 3.644, MSE = 16.04, p < .0001. The interaction indicated that for some terrains, tethered displays produced a lower time-seen value than did more egocentric or exocentric displays (the display type effect noted earlier), but for other terrains, display type had little effect on time seen.

Recognition Accuracy

Display type affected recognition accuracy, F(4, 44) = 28.57, MSE = 0.13, p < .0001. Accuracy was greater for the exocentric than for egocentric displays, F(1, 11) = 103.52, MSE = 0.138, p < .0001. Accuracy was greater for tethered than for exocentric displays, F(1, 11) = 26.91, MSE = 0.165, p < .0005. Accuracy was greater for exocentric than for tethered displays, F(1, 11) = 25.00, MSE = 0.169, p < .0005.

Terrain had no effect on recognition accuracy, nor was there an interaction between terrain and display type, p > .05 in each case.

Recognition RT

Display type affected recognition RT, F(4, 44) = 2.98, MSE = 12.95, p < .05. Recognition RTs were shorter for exocentric than for egocentric displays, although the difference failed to reach conventional significance levels, F(1, 11) = 1.57, MSE = 32.43, p = .24. There was no difference in recognition RT between tethered and egocentric displays, F < 1. Recognition RTs were shorter for exocentric than for tethered displays, F(1, 11) = 8.42, MSE = 17.02, p < .05. Terrain had no effect on recognition RT, nor did it interact with display type, p > .05 in both cases.

Map-Reading Test and Navigation Training

A map-reading test score was obtained for each participant. These scores were correlated...
Participants were asked to indicate whether they had navigation or orienteering training. If they had, they were asked to indicate how many years of training they had. Of the participants, 8 had no training, 2 had less than a year, 1 had between 1 and 2 years, and 1 had more than 5 years of training. These categories were coded as 0 to 3, respectively, and correlated with each performance measure averaged across all conditions. No correlation reached conventional significance levels ($p > .05$ in all cases).

**DISCUSSION**

The tethered displays were the most effective displays for spatial awareness in regard to the time-seen measure. Not only were the tethered displays more effective than the egocentric...
display, but they were also more effective than the exocentric displays. Use of the tether minimized the time during which the tank was visible to enemy positions. For other measures, 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.

We obtained no effect of tether dynamics in the navigation task. Wang and Milgram (2001) also did not find a performance advantage, although they found that their participants preferred the dynamic to the rigid tether. The display problems associated with the rigid tether may not be evidenced until certain levels of fatigue are reached. Performance benefits for the dynamic tether may therefore be difficult to demonstrate in short experimental sessions.

As predicted, the egocentric display was more effective than exocentric displays (2-D or 3-D) for navigation, and the exocentric displays were more effective than egocentric for spatial awareness, for both time seen during navigation and the recognition task (significantly for accuracy but not significantly for RT). This result is in keeping with a large literature contrasting the relative benefits of egocentric and exocentric displays for navigation and spatial awareness tasks, respectively (e.g., St. John et al., 2001; Wickens & Hollands, 2000).

Why was the tethered display so good for minimizing time seen? It would appear that the user needs to see enough of the terrain to make a good initial route choice between waypoints (a need for exocentricity). Thus, for example, with the tethered display, the user can seek to maximize the distance between enemy positions and the planned route. This task would be more difficult with an egocentric display, because the user likely cannot see all the waypoints or enemy positions. Moreover, the user also needs to be able to adjust the route on the basis of local information as the tank traverses the terrain so that the tank will not be seen (a need for egocentricity). For example, with a tethered display, the user may see that a hill occludes the line of sight between an enemy position and a waypoint. With an exocentric display, it might be difficult to tell whether the hill occludes or not or even whether the hill exists. On the basis of our results, the tether appears to provide the right combination of egocentric and exocentric elements to optimize route choice and route adjustment and thereby to minimize time seen. Terrain affected the time-seen measure and interacted with display type. Some terrains showed very little effect of display type; others showed an advantage for tethered displays compared with exocentric or egocentric displays. We examined specific terrains and waypoint sets to interpret this interaction. We show two terrains in Figure 4. For the terrain on the top, the user must drive through a single valley to reach all the points and minimize time seen. There is not a lot of flexibility in route choice; one cannot choose the high ground, because it will increase time seen, and the user can observe this situation regardless of viewpoint. In contrast, for the second terrain, we see some route choices, given the broad split valley (favoring a more exocentric viewpoint), but two of the points (for both sets) are relatively high and close to nearby hills, and it may not be easy to observe these terrain characteristics with the most exocentric displays. The user needs the right combination of seeing enough to detect all the points to make a good route choice (exocentricity) with enough egocentricity to adjust the route on the basis of local information as the user traverses the terrain so that the tank will not be seen. Similar differences were observed with other terrains. In summary, we suggest that with some terrains, there was not much that could be done to minimize time seen; with others, there was route choice (so that tethered display was more effective than egocentric) and route adjustment based on local terrain (so that tethered display was more effective than exocentric).

Our participants’ map-reading abilities did not affect their performance on these tasks. Participants with more orienteering or navigation training showed no evidence of improved ability to perform on the tasks. Thus it appears that specialized navigation training did not affect the pattern of results observed.

Caveats

One could ask how our results would have differed if features (e.g., trees, buildings) had been
added to the terrain models. One possibility is that the obtained differences would be enhanced. With greater object clutter, the greater occlusion for the immersed egocentric view could make it more difficult to obtain exocentric spatial awareness, and navigation performance could be degraded even further with the exocentric view. We suspect that with greater object clutter, the tethered condition would continue to show its advantage for time seen, because trees and buildings would likely be of small scale relative to the hills. However, given the increased likelihood of collision with the added objects, our tethered displays would likely not be as effective for navigation as the egocentric displays. It is also possible that one might see some navigation advantages for a shorter tether compared with an egocentric display in this situation, given the keyhole-effect problem sometimes observed with narrow fields of view (e.g., Calhoun, Draper, Nelson, & Ruff, 2006; Pazuchanics, 2006; Voshell & Woods, 2005). The use of an adjustable tether might also be considered and needs further exploration.

The argument could be made that a display based on a long virtual tether is not different in kind from an exocentric display. If this were the case, it should have produced results similar to our exocentric displays. However, the tether condition produced different results than the exocentric displays—so if the long tether was similar to an exocentric format, why were the results different? We think that linking the viewpoint to the motion of the vehicle being controlled is a fundamental difference, even when the tether is long. The arguments made earlier highlighting the advantages of a tethered view for route choice and route adjustment also apply here.

The availability of the status bar information could have affected the time-seen measure. For example, a participant could have used the status bar color in isolation to determine whether he or she was being seen. However, to reduce time seen the participant would need to steer to a position where he or she could not be seen. To do so, the participant would need to observe the enemy unit’s location. Moreover, any advantage of the status bar would probably be greatest in the egocentric condition (given that the enemy unit position was least likely to be displayed because of the occlusion from terrain). However, this was the worst condition for the time-seen measure. Hence, we argue that the status bar does not appear to have played a large role in influencing time seen.

Many factors covary with changes in the frame of reference. It is difficult to provide an exocentric view without depicting more terrain with less detail or to provide an egocentric view without providing more detail while showing less of the terrain. In our experiment, although the angular field of view was constant, the distance of the camera from the terrain varied with the display condition. In effect, this meant that linear field of view—the scale or amount of terrain shown—varied with condition. However, the linear field of view cannot be computed deterministically when the viewpoint (and therefore distance to terrain) moves during the trial, as it did in the egocentric and tethered conditions. Generally, there is a high correlation between changing the frame of reference and changing the field of view (see, e.g., the illustration used in Wickens & Hollands, 2000, p. 168). Indeed, it is one of the reasons we use different fields of view when we navigate (e.g., when one canoes through a group of islands within a lake, one refers to a map because it shows the entire lake at once, with all the islands visible, whereas the view from the canoe shows only a subset). Thus, in our experiment, there was a necessary confound between display type and field-of-view size. We argue that had we not varied linear field of view with reference frame, we would have created displays lacking in much practical utility. Thus, a change in frame of reference necessarily involves changing a set of covarying display parameters (location, including scale; azimuth rotation; and vertical viewpoint rotation; Wickens & Hollands, 2000).

One could argue that for our recognition task, the rotating terrains were similar to the exocentric display formats, and therefore, it is not surprising that those formats produced the best recognition performance. First, we note that because the terrains rotated, the participant saw each terrain from multiple viewpoints, not just the same viewpoint he or she would have seen earlier with the 3-D (or 2-D) exocentric display. So the recognition displays were not identical to the exocentric displays used for navigation. Second, to assess spatial awareness, we needed a task that indicated how well an individual could distinguish the
space just traversed from other similar (distractor) spaces. To assess general spatial awareness of the terrain, the whole terrain needed to be shown; otherwise, the participant could not definitively say it was the same terrain (any part of the terrain not shown could be different from that shown during navigation). Although other methods are possible (e.g., having participants “recall” the terrain traversed by sketching it), these methods come with their own challenges (e.g., scoring the accuracy of the sketch is problematic).

**Conclusion**

The tethered view maximized concurrent spatial awareness (minimizing time seen). It did so while providing a viewpoint that was generally “second best” for an egocentric or exocentric task. With further tether adjustments, perhaps it can be made as effective for egocentric or exocentric tasks while still providing situated spatial awareness. Our demonstration of the advantages of an egocentric and exocentric frame of reference for their task counterparts is essentially a replication of earlier findings, including those of Wickens (e.g., Wickens & Prevett, 1995). In a sense, these results serve as a foundation to situate the results obtained with the tethered displays.

In summary, the tether appears to be useful in spatial awareness involving knowledge of locations of interest with respect to one’s own position while navigating (the time-seen measure). In this sense, perhaps our results help to identify a component of navigation whose performance is dissociated from conventional exocentric spatial awareness and egomotion. The tethered display may be the most effective display for this egocentric spatial awareness task.

**ACKNOWLEDGMENTS**

This research was supported by funds from the SIREQ (Soldier Information Requirements) Project, part of Defence Research and Development Canada’s Technology Demonstration Program. We thank Kevin Trinh for conducting analyses and preparing the graphs and Linda Bossi and John Frim for arranging financial support. Preliminary experimental results were presented at the 2005 Annual Meeting of the Human Factors and Ergonomics Society.

**KEY POINTS**

- The ability to build terrain models in real time may soon allow remote vehicular control from any viewpoint.
- A virtual tether couples the viewpoint with the vehicle’s position and orientation but shows more of the terrain than a fully immersive egocentric display.
- For this study, 12 participants navigated a simulated vehicle across complex virtual terrain using five different display viewpoints, including two tethered displays.
- Viewpoint tethering produced the most effective displays for minimizing the time the vehicle could be seen from simulated enemy positions, but tethered displays were less effective than egocentric and exocentric displays for navigation and recognition, respectively.
- We recommend the use of a tethered display for applications in which it is necessary to understand the relation of a vehicle location to nearby locations.

**REFERENCES**


Justin G. Hollands is a defence scientist and senior advisor for the Human Systems Integration Section at Defence Research and Development Canada–Toronto. He received his PhD in psychology from the University of Toronto in 1993.

Matthew Lamb is a research engineer at Defence Research and Development Canada–Toronto. He received his BASc in systems design engineering from the University of Waterloo in 2006.

Date received: May 28, 2010
Date accepted: January 6, 2011