The Use of 2D and 3D Displays for Shape-Understanding versus Relative-Position Tasks

Mark St. John, Pacific Science and Engineering Group, San Diego, California, Michael B. Cowen, Space and Naval Warfare System Center, San Diego, California, and Harvey S. Smallman and Heather M. Oonk, Pacific Science and Engineering Group, San Diego, California

Research on when and how to use three-dimensional (3D) perspective views on flat screens for operational tasks such as air traffic control is complex. We propose a functional distinction between tasks: those that require shape understanding versus those that require precise judgments of relative position. The distortions inherent in 3D displays hamper judging relative positions, whereas the integration of dimensions in 3D displays facilitates shape understanding. We confirmed these hypotheses with two initial experiments involving simple block shapes. The shape-understanding tasks were identification or mental rotation. The relative-position tasks were locating shadows and determining directions and distances between objects. We then extended the results to four experiments involving complex natural terrain. We compare our distinction with the integral/separable task distinction of Haskel and Wickens (1993). Applications for this research include displays for air traffic control, geoplots for military command and control, and potentially, any display of 3D information.

INTRODUCTION

Many operational tasks require the comprehension of three-dimensional (3D) objects and environments. For example, air traffic controllers need to understand a complex 3D environment populated with aircraft, air routes, and no-fly zones. Similarly, military officers need to understand 3D environments such as ground and underwater terrain, air routes, and radar zones. 3D displays (or displays with 3D views) seem to provide a natural, and increasingly affordable, solution to these requirements.

What we mean by a 3D view is actually a perspective or oblique view of an object or scene displayed on a computer monitor. The image is two-dimensional (2D), but the viewing angle provides a 3D perspective. For example, rather than displaying an environment from directly above (a planar or bird’s-eye view), perspective view technologies generally display the environment from a 30° or 45° angle. Holographic and other true 3D technologies are being developed, but most interest in 3D displays concerns 3D perspective views.

Many potential users who see 3D perspective views are enthusiastic about them. This appeal of 3D views most likely stems from their capability to convey the shape of complex objects in a natural and integrated way. Within the scientific visualization community, 3D perspective views have been used extensively to view complex objects such as molecules and even abstract objects such as the semantic structure of document collections (e.g., Card, Mackinlay, & Shneiderman, 1999).

However, Andre and Wickens (1995) caution system designers that sometimes “users want what’s not best for them,” preferring systems that hinder rather than enhance performance. Andre and Wickens’ review of studies on input devices, display interfaces, color, and 3D displays provides evidence to support this cautionary note. Although it might be believed...
naively that more display dimensions are always better, there are three sources of problems with 3D views that stem from perspective projection.

First, without other depth cues available in the 3D view, the location of objects is ambiguous along lines of sight into the viewing plane. This problem has been termed **projective ambiguity** by Sedgwick (1986) and **line-of-sight ambiguity** by Boyer and Wickens (1994). Second, in addition to the linear affine transformations in a 2D view (magnification, translation, or both), space is nonlinearly distorted in a 3D view. Specifically, depth into the scene orthogonal to the picture plane is scaled nonlinearly (approximately quadratically) by depth, whereas widths and heights parallel to the picture plane are scaled linearly by depth (see Gillam, 1995). This asymmetric compression of space in a 3D view results in the distortion of distances and angles. Third, the projection of objects tilted toward the line of sight in a 3D view is compressed. Imagine watching someone lower a telephone pole in your direction. The height dimension of the pole slowly shortens until it disappears entirely, and all you can see is the round top of the pole. This effect is known as **foreshortening** (see Sedgwick, 1986).

Interestingly, these distortions can act as cues to depth and help establish the layout of the scene (see Gillam, 1995), though perhaps not in a very detailed way. Others have experimented with adding more depth cues (stereo, texture, etc.) and even artificial enhancements (drop lines and drop shadows) to improve localization (see Wickens, Todd, & Seidler, 1989, for an early review). However, it remains questionable whether these cues could ever make up for the inherent spatial distortion in a 3D perspective view. Perhaps metric judgments of distance and relative position will always be problematic in a 3D view irrespective of how many cues are available.

Further, the empirical evidence for the utility of 3D views is decidedly mixed. Across an array of tasks, a number of studies have found benefits for 3D perspective over 2D (Bemis, Leeds, & Winer, 1988; Ellis, McGreevy, & Hitchcock, 1987; Hickox & Wickens, 1999; Liter, Tjan, Bulthoff, & Kohnen, 1997; Naikar, Skinner, Leung, & Pearce, 1998). Other studies have found rough parity or different results on different measures or tasks (Andre, Wickens, Moorman, & Boschelli, 1991; Baumann, Blanksteens, & Dennehy, 1997; Burnett & Barfield, 1991; Haskell & Wickens, 1993; Van Breda & Veltman, 1998; Wickens, Liang, Prevett, & Olmos, 1996; Wickens & Prevett, 1995), and still other studies have found 2D superior to 3D (Boyer, Campbell, May, Merwin, & Wickens, 1995; Boyer & Wickens, 1994; O’Brien & Wickens, 1997; Wickens, Campbell, Liang, & Merwin, 1995; Wickens & May, 1994; Wickens, Miller, & Tham, 1996). The details of tasks and interfaces vary widely in these studies. For example, altitude is sometimes represented in digital form and sometimes in analog form. In addition, it is sometimes continuously visible on the display, and at other times it must be engaged by selecting an aircraft icon. Consequently, it seems likely that some results depend more on details of the interface implementation than on the nature of the display formats themselves.

Aside from interface differences, perhaps the inconsistency of results stems from the fit between tasks and representations. Several distinctions among tasks or among displays have been drawn to help explain the empirical results. Wickens and Prevett (1995) proposed a distinction between local guidance tasks and situation awareness tasks, and between egocentric displays and exocentric displays. Haskell and Wickens (1995) proposed that 3D perspective-view displays lead to better performance “whenever the tasks to be performed using the display are integrated three-dimensionally or whenever the method of performing the task with the display bears a strong resemblance to a similar task performed without a display. For flight displays, this includes flight control and identifying and making integrated judgments regarding other aircraft. In these cases, the similarity of the display representation to the view in visual-contact flight overcomes possible disadvantages of the three-dimensional display format. However, for tasks that require focused attention and that do not have a visual analog in flight, it may be advantageous to create separate planar displays” (p. 104–105).

This regard for integration relates to the proximity compatibility principle elucidated by Wick-
ens and Carswell (1995). It states that all the information that facilitates task completion should be integrated as much as possible into a coherent display. Unfortunately, it is sometimes difficult to predict which tasks require focused attention and which require integration across dimensions (Haskell & Wickens, 1993, p. 90). For example, if a task can be broken down into separate dimensions, does it fail the criteria for integration?

Here we propose a distinction that makes somewhat similar predictions but that focuses more closely on task demands. A 3D view is most useful for tasks that require understanding the general shape of 3D objects or the layout of scenes – what is termed a structural description in the vision literature (e.g., Frisby, 1980, chap. 5, pp. 106–122). A 3D view is useful for shape understanding because it integrates the three dimensions into a single view and provides natural depth cues such as perspective, shading, and occlusion to help solve the task. Three-dimensional views, however, impair perception of the relative positions of objects because of projective ambiguities and distortions. Consequently, 2D views are more useful for tasks that require judging the precise distances and angles between objects. We tested this shape-understanding versus relative-position judgment distinction with a battery of perceptual tasks, first in the domain of simple, generic blocks and then in the domain of complex terrain. In the general discussion, we compare the predictions of Haskell and Wickens’s (1993) separated/integrated distinction and our shape-understanding/relative-position distinction across a number of published studies.

**EXPERIMENT 1: SHAPE UNDERSTANDING**

In Experiment 1 we tested the hypothesis that a 3D perspective view leads to better shape understanding than a 2D view. Our goal was to make the stimuli simple and generic in the hope that our results would apply to a wide variety of tasks and content domains. Consequently, we created simple 3D block shapes composed of 10 to 16 cubes. These block shapes were rendered as a 3D perspective view or as a set of 2D views (see Figure 1).

Shape understanding was defined by using four identification and mental rotation type tasks. Mental rotation (Shepard & Metzler, 1971) of an object requires shape understanding because all of an object’s parts must be coordinated in the rotation. However, mental rotation is an inherently harder task than identification. Together the relatively easy identification tasks and the relatively hard mental rotation tasks provided a good range of shape-understanding tasks.

**Method**

**Participants.** The participants were 32 Navy and civilian personnel. All were employed at the Fleet Anti-Submarine Warfare Training Center, San Diego, California.

**Stimuli.** The study stimuli were 10 2D and 3D renderings of simple block shapes created using Extreme 3D (1996) and presented on a 15-inch color liquid crystal display (LCD) panel. Each object was composed of 10–16 cubes arranged into a 3D shape. Care was taken to avoid recognizable shapes. For the 3D renderings, a camera was positioned 30° above the horizontal plane of the object and at such an angle that three faces of the object were visible. In choosing the viewing angle, care was taken to ensure that all prominent features of the object were visible.

An omni light and an ambient light illuminated the objects, and shading on the blocks was created by placing a single spotlight source above and 90° to the right of the camera. Shadows were not rendered. An orthographic perspective, rather than a vanishing-point perspective, was used. This perspective made the blocks look like small toys viewed from close range rather than large, looming buildings.

There were three multiple-choice answer objects for each study object. For the Identify 3D task, the answer objects were rendered in 3D. One answer object was the study object, and the other two answer objects were similar to the study object except for one structural change, such as rotating one leg of a block 90° or shifting a leg over by one cube.

It can be argued, however, that the 3D renderings in the multiple-choice answer set provided an unfair advantage to the 3D condition because in that condition, the correct answer
looked exactly the same as the study block. To address this issue, we created a second task, the Identify Real task, in which the answer sets were composed of real wooden cubes (see Figure 1, bottom). Participants still studied either 2D or 3D renderings, but they picked the correct answer from among three real blocks.

The third shape-understanding task was Rotate Yaw. This task required participants to study one block, rendered in either 2D or 3D, mentally rotate the block 90° in a designated direction around the vertical axis, and then identify the correct rotated block from among a set of minimally structurally different alternatives rendered in 3D, akin to the classic mental rotation task of Shepard and Metzler (1971).

The fourth task was Rotate Pitch. This task required participants to mentally rotate the study block 90° in a designated direction around the horizontal axis and then identify the correct block from among three alternatives.

Procedure. For the Identify 3D, Rotate Yaw, and Rotate Pitch tasks, a trial consisted of first viewing a single study object rendered in either 2D or 3D in the upper half of the screen. After 10 s, three multiple-choice answer objects appeared in the lower half of the screen, while the study object remained visible. The participant used a mouse to choose an answer object. Reaction times to complete each trial were recorded, as well as errors. For each task there were 3 practice trials (using simpler objects) followed by 10 test trials. For the Rotate Yaw and Rotate Pitch tasks, a circular arrow next to each study object indicated which direction the study object should be mentally rotated.

Figure 1. The top half of the figure shows 2D and 3D renderings of a typical block shape used in Experiments 1 and 2. The ball and red colored cube (shown as dark gray in Figure 1) appeared only in Experiment 2. Note the ambiguous location of the ball in the 3D view. The bottom half of the figure shows the multiple choice answers to the Identify Real task. Note the structural changes of the incorrect choices.
In the Identify Real task, the study object was displayed on the computer screen in the same place as in the other tasks, but no objects were shown in the lower half of the screen. Instead the participant was shown wooden blocks on a table to the left of the computer screen. For each trial, after a study object was shown for 10 s, a tone sounded and a screen was lifted to reveal the wooden answer blocks. The blocks were labeled \(a\), \(b\), and \(c\). Participants chose the correct block by using the mouse to select an \(a\), \(b\), or \(c\) button on the computer screen.

Each of the tasks was randomly assigned to eight participants: Identify 3D, Identify Real, Rotate Yaw, and Rotate Pitch. Each participant received both the 2D and 3D condition for his or her assigned task. The condition order was counterbalanced across participants. This experimental design allowed us to compare the 2D and 3D versions of each task within participants.

Before receiving the 2D conditions of the tasks, participants were shown a video loop on the screen explaining how to interpret the 2D renderings. The video showed a 3D rendering of an object with the three 2D views wrapped around the top, front, and side of the object. As the video continued to play, the three views unwrapped, flattened out, and moved away from the object until the 3D rendering and the 2D rendering were side by side. Care was taken to point out the orientation of the three views. The video ran forward and backward for as long as the participant desired.

**Results**

For each participant, response times (RT) for correct trials were averaged, and a proportion correct (PC) score was calculated. For each task \((n=8)\), performance in the 2D condition was compared with performance in the 3D condition (see Figure 2). Participants were faster and more accurate with the 3D views on the Identify 3D task; RT: \(t(7) = 7.3, p < .001\); PC: \(t(7) = 4.6, p = .003\), and the Identify Real task; RT: \(t(7) = 3.5, p = .01\); PC: \(t(7) = 2.2, p = .07\).

Note that the mean differences between 2D versus 3D were about the same for both tasks.

![Figure 2](https://example.com/figure2.png)

*Figure 2.* Performance on the shape-understanding tasks. Error bars show one standard error.
This similarity suggests that the 3D advantage in the Identify 3D task was not a result of the graphical similarity of the study and test renderings, because the same advantage was found in the Identify Real task. Instead the advantage must have been a result of the fact that 3D renderings are easier to understand. Participants were also faster and more accurate with the 3D views on the Rotate Yaw task; RT: \( t(7) = 7.5, p < .001 \); PC: \( t(7) = 3.5, p = .01 \) and on the Rotate Pitch task; RT: \( t(7) = 7.3, p < .001 \); PC: \( t(7) = 2.4, p = .05 \).

Discussion

The 3D views were clearly better for understanding the shapes of simple blocks. Reaction times were two to three times faster, and proportion correct scores were 10% to 20% better, using the 3D views. The obvious weakness of the 2D views for any of the shape-understanding tasks is that the top-down, front, and side views must be integrated into a single object, which takes time and mental effort. In the 3D perspective renderings, by comparison, the views are already integrated into a single object, which makes shape understanding fast and easy. This reasoning assumes that all the important features of the object are visible in the perspective view, which was the case in this experiment.

EXPERIMENT 2: RELATIVE POSITION

In Experiment 2 we investigated the effectiveness of 2D and 3D views for determining the relative position of two objects. Relative-position understanding was defined by three tasks: the Over-Same task, the Over-Different task, and the Navigation task.

In the Over-Same task, participants were presented with a 2D or a 3D view, which consisted of a simple block shape (as in Experiment 1) and a ball the size of a single cube that did not cast a shadow. For each view of the block and ball rendered in either 2D or 3D, the participants were asked to determine which cube lies directly underneath the ball. Participants could compare the views to determine the height of the ball and which cube lay underneath it.

One might argue that this task is artificially easy in the 2D condition because participants only need to look at the top view to determine which cube lies directly underneath the ball. However, we believe it is simply the nature of 2D views to make relative-position tasks like this one simple and obvious. Nevertheless, we developed a related task, the Over-Different task, to address this issue. In the Over-Different task, participants were presented with 2D or 3D views of the block and ball just as in the Over-Same task and were again asked to determine which cube was underneath the ball. Simultaneously, participants were shown a 3D view of the block with the ball removed. They used this 3D view of the block to report their answer by clicking on the appropriate cube.

In the third task, Navigation, participants determined how to move from a designated cube (shown in red) in the block shape to reach the ball. Participants indicated the number of cubes necessary to move in each direction to get from the red cube to the ball. In the example shown in Figure 1, the answer is Up 3, East 2, and North 1.

Method

Participants. The participants were 24 Navy and civilian personnel. All were employed at the Fleet Anti-Submarine Warfare Training Center, San Diego, California.

Stimuli. The stimuli were the same 10 block shapes used in Experiment 1, with the addition of a ball displayed somewhere above the block. The diameter of the ball was identical to the width of one of the cubes making up the block. The ball was located in empty space from one to three cube-lengths above the block. The ball did not cast a shadow.

For the 3D condition, two views of the block and ball were rendered. For each view the camera was positioned 30° above the horizontal plane of the block and 45° to the left or right of the front view.

Procedure. On each trial, participants saw a block and ball rendered in either 2D or 3D. They first performed the Over-Same task by...
using the mouse to click on the cube directly underneath the ball. In the 2D version, participants were informed that they could click on cubes in any of the three views but that using the top view would be easiest. In the 3D version, which contained two views, participants were instructed to click on a cube in the block shape rendered on the right side of the screen. Immediate feedback was provided, and the participant could not continue until the correct cube was identified. Total time to find the correct cube and first-choice errors were recorded.

On each trial, immediately after answering in the Over-Same task, participants performed the Navigation task for that stimulus. Following a correct answer in the Over-Same task, one cube of the block turned red, a compass point icon appeared next to the block, and direction menus appeared below the block. There were three separate direction menus: North/South, East/West, and Up/Down. When participants selected a direction menu, a pop-up list of distances appeared on the screen. For example, when the North/South menu was selected, the list showed North 4, North 3, North 2, North 1, 0, South 1 . . . South 4. After participants chose a distance, the menu closed and displayed the chosen distance for that dimension. After choosing distances along all three dimensions, each participant clicked a “submit” button. The answer was evaluated, and if correct, the participant continued to the next trial. After three incorrect tries, the program moved on to the next trial and recorded a failure. Total time to select a correct answer was recorded (trials with three failures were discarded). Additionally, we recorded first-try errors.

All participants received 6 practice trials on the Over-Same/Navigation tasks followed by 10 experimental trials. Half of them received the 2D condition followed by the 3D condition, and half received the reverse order. Next, half the participants were shown the 2D condition of the Over-Different task, and half were shown the 3D condition of the Over-Different task. Participants did not receive both conditions of the Over-Different task in an effort to keep them from becoming too familiar with the stimuli.

Results

Mean response times and proportion correct scores are shown in Figure 3. For the Over-Same task \( (n = 24) \), participants were faster and more accurate using the 2D views than the 3D views; RT: \( t(25) = 7.7, p < .0001; \) PC: \( t(25) = 6.0, p < .0001 \). The 2D views were found to be much more effective, but as previously mentioned, the top view of the block shape in the 2D condition made the task very easy.

The Over-Different task was designed to remedy this complaint by requiring participants to respond by pointing to cubes in a separate 3D view. Participants were still faster, though not reliably more accurate, using the 2D rather than the 3D views; RT: \( t(22) = 2.9, p = .008; \) PC: \( t(22) = 1.5, p = .21 \). For the Navigation task, participants again performed faster and more accurately using the 2D views than the 3D views; RT: \( t(23) = 7.3, p < .0001; \) PC: \( t(23) = 7.1, p < .0001 \).

Discussion

 Unlike the shape-understanding tasks, for the three relative-position tasks, the 2D displays were clearly better. For the Over tasks, the culprit was ambiguity in the 3D perspective view. It is impossible to determine which cube lies directly underneath the ball in a single view. A second view from another angle, which is equally ambiguous, had to be consulted, and then the two views had to be compared to find the correct location. The 2D views from the front and side are also ambiguous. However, the top view is entirely unambiguous for determining which cube is underneath the ball.

For the Navigation task, the culprit for the 3D views is not the line-of-sight ambiguity because the ambiguity is resolved in the Over-Same task. Instead, the problem seems to be the distortion in the 3D perspective view. Participants reported that height was especially difficult to judge because it required estimating distances across empty space. In the 2D views there is no distortion of angles or distances; instead, each dimension is represented faithfully. Consequently, it is easy to judge and move specific directions and distances along each dimension. To move in the dimension that is not represented, one simply has to turn to another view in which that dimension is represented faithfully. To view another dimension requires a shift of the eyes and a reorientation.
to the object or scene, but this perceptual shift does not seem to hinder performance as much as does dealing with the distortions in the 3D views.

EXPERIMENT 3: FOUR CORNERS

The block stimuli were chosen for their simplicity and generality so as to test our hypothesis as cleanly as possible. Because the blocks were composed of cubes, all the angles were right angles, and all the lengths were cube units. These characteristics could be used to compensate for distortions in the 3D view, and they might have facilitated understanding of the 3D renderings. How might the results generalize to more complex and natural stimuli such as terrain?

Before proceeding, it is worth discussing how we chose the 2D and 3D representations that we used in the terrain experiments. There is a fair amount of variety among representations of 3D space and terrain reported in the literature (for reviews, see Aretz, 1991; Mc Cleary, Jenks, & Ellis, 1993). In fact, it is probably best to think of 2D and 3D displays as natural concepts (Rosch, 1973; Wittgenstein, 1958) that have no defining necessary or sufficient features. Instead, they share a set of family resemblances. For 3D or perspective view displays, the primary resemblances that are shared by most representations are oblique views (views from angles that show all three dimensions of an object) and shading. For 2D views, the primary resemblance is the use of normal viewing angles that are orthogonal to a cardinal axis of the space. However, it is possible to have what we would clearly call a 3D view that contains a normal viewing angle.

Figure 5 (left side) shows a 90° view that we consider to be a 3D view. To us, the natural representation of shape and the use of shading make the view 3D despite the 90° viewing angle. In fact as we will demonstrate, it performs much more like the standard 45° 3D view than it does like a 2D topographic view.

In Experiment 3 we tested the hypothesis

![Figure 3. Performance on the relative-position tasks.](image-url)
that 3D views are better for understanding the shape of terrain. Participants viewed terrain rendered in 3D from a 45° angle, in 3D from a 90° angle, or in 2D as a topographic map. Their task was to imagine standing in the center of the map, look to a specified corner of the map, and then choose the correct ground-level view from among four alternatives (the Four-Corners task). Because this task required an understanding of the shapes of terrain features and their layout in the scene, we predicted that the 45° view (see Figure 4) would produce the best performance.

Method

Participants. The participants were 18 civilian personnel. They were either employed at a local engineering company or were students from local universities who were paid for their participation.

Stimuli. A total of 26 stimuli were created from 13 7 × 9-mile U.S. Geological Survey Digital Elevation models. This size terrain is similar to that found on standard 1:50,000 scale military maps. These models were processed through MICRODEM (Guth, 2000) to create elevation bitmaps (latitude was compressed slightly to create a square map, and altitude was exaggerated on some maps to make the terrain more dramatic). Separate renderings were made for a 3D 45° view, a 3D 90° view, and a topographic view. The topographic maps were created by drawing iso-altitude contour lines on the map. Unlike typical topographic maps, which use numbers to indicate altitude, we color-coded the contour lines to indicate altitude to make map interpretation easier. The program assigned dark blue for the lowest altitude on the map, ran through the color spectrum for intermediate altitudes, and assigned magenta for the highest altitude on the map (Figure 5).

The 3D 45° and 90° views were created by importing the elevation bitmaps into 3D Studio Max (1999). The camera had a fairly wide 90°
horizontal field of view and a wide-angle 18-mm lens. An omni light source (the sun) was placed directly west of the center of the map and at 50° above ground level from the center of the map. For the 3D 90° views, the camera was suspended 3.5 miles directly over the center of the map so that the entire map was visible while maintaining a 45° angle between ground level and the line of sight from the camera to the center of the map (Figure 5). For the 3D 45° views, the camera was moved to the south of the map so that the entire map was visible while maintaining a 45° angle between ground level and the line of sight from the camera to the center of the map (Figure 4).

The ground-level views were created in 3D Studio Max (1999) to represent what one would see while standing at the center of the map and looking out to each of the four corners of the map. The ground views were placed in a random order around the map. Two stimuli were created for each terrain map by placing a white cross at the center of the terrain and designating different corners to visualize.

Procedure. The stimuli were presented one at a time on a 17-inch color computer monitor. Participants viewed the central map and imagined the ground-level view for the designated direction. They then chose one of the four ground-level views by using a mouse to click the appropriate on-screen button. Reaction times for correct trials and errors were recorded, and correct/incorrect feedback was provided. Stimuli were blocked by conditions, and participants viewed all three conditions (e.g., 45°, 90°, topographic) in a counterbalanced order.

Results and Discussion

Response times for correct trials were averaged, and proportion correct scores were calculated (see Figure 6). This task proved to be very difficult for all three views. Consequently the response times to correct trials should be viewed with some skepticism because they represent only 40% to 50% of the trials. Nonetheless, the views produced statistically different response times, $F(2, 34) = 6.6, p = .004$. Post hoc analyses indicated that both the 45° view and the 90° view produced faster response times than the topographic view: 45° vs. topo, $t(17) = 2.6, p = .02$; 90° vs. topo, $t(17) = 3.4, p = .003$. The proportion correct scores were not statistically different from each other, $F(2, 34) = 1.7, p = .18$.

These results support our hypothesis that 3D views are better for understanding the shape and layout of terrain, presumably because the 3D depth cues are quick and easy to interpret. The symbolic representation of altitude used in the 2D topographic views, on the other hand, required more time to interpret into 3D shapes.

Interestingly, performance using the 90° view was as good as performance using the 45° view. Apparently, for understanding the shape of terrain, 3D depth cues such as shading are more important than a perspective view.

EXPERIMENT 4: A-SEE-B

In Experiment 4 we again tested the hypothesis...
that 3D views are better than 2D views for understanding the shape of terrain. The task was to judge whether one ground location was visible from another ground location or obstructed by an intervening hill (the A-See-B task). Because this task involves the relationship between two points, one might suspect it to be a relative-position task. However, the task was designed to hinge on assessing the general shape and layout of the terrain between the two points: Would the view of the second point be clearly visible, or would a major terrain feature block the view?

**Method**

**Participants.** The participants were 27 university students who were paid for their participation.

**Stimuli.** The stimuli were created from the same terrain that was used in Experiment 3. A total of 26 stimuli were created, 2 from each of the 13 models of terrain. To create a stimulus, we plotted two points on the terrain. The two points were chosen randomly from different quadrants of the terrain and labeled A and B (see Figure 7). Each stimulus was rendered in 3D from a 45° angle, in 3D from a 90° angle, and as a 2D topographic map, as in Experiment 3.

**Procedure.** The stimuli were presented one at a time on a 15-inch color LCD panel. Participants were asked to determine whether they could see point B if they were standing at point A. Participants responded by clicking an on-screen button for True (A can see B) or False. Response times were recorded and feedback was provided. For this experiment, participants performed in only one of the three conditions. The experiment was easy and fast, so we were concerned that participants would remember the stimuli and their previous responses if they performed in all three conditions.

**Results and Discussion**

The A-See-B task proved to be less difficult than the Four-Corners task, having a mean response time of 5.8 s and a mean accuracy of 81% (see Figure 6). Overall the three viewing conditions were similar, with no significant differences in response time or proportion correct. However, there was a trend towards higher accuracy in the 3D conditions compared to the 2D condition, suggesting that 3D displays may provide an advantage for this type of task.
conditions produced statistically different response times, $F(2, 21) = 4.1, p = .03$. As predicted, the 3D 45° view was faster than the 2D topographic view, $t(14) = 2.7, p = .02$. The 3D 90° view was not reliably faster than the 2D topographic view, $t(14) = 1.8, p = .09$. The proportion correct scores did not differ reliably.

In sum, the shape-understanding tasks, which required extracting information concerning a rough sense of the layout of the terrain and determining whether a view would be obscured, were easier to perform using the 3D 45° view than the 2D topographic view. Performance using the 90° 3D view was mixed. Interestingly, we found the 45° 3D view advantage for both a relatively difficult task and for a relatively easy task, despite the fact that the colored contour lines in the 2D topographic maps made these topographic maps easy to use.

**EXPERIMENT 5: A-HIGH-B**

In Experiment 5, we tested the hypothesis that 2D views are better for judging the relative position of two terrain locations. Participants viewed terrain and judged which of two points was higher (the A-High-B task).

**Method**

**Participants.** The participants were 27 civilian government personnel employed at a local military research and development center.

**Stimuli.** As in Experiment 4, the stimuli were 26 stimuli with A and B points. As before, they were rendered in 3D from a 45° angle, in 3D from a 90° angle, and as a 2D topographic map.

**Procedure.** The stimuli were shown one at a time on a 15-inch color LCD panel. Participants were asked to view each map in turn and indicate which of the two points, A or B, was higher. Response times for correct trials and errors were recorded. Participants performed in only one of the three conditions.

**Results and Discussion**

Response times were not statistically different from each other, $F(2, 24) = 1.8, p = .19$. Accuracy scores, however, were statistically different from each other, $F(2, 24) = 8.5, p = .002$. Post hoc analyses indicated that the topographic view produced more accurate answers than the 45° view, $t(16) = 5.1, p = .007$, and the 90° view, $t(16) = 5.43, p < .0001$ (see Figure 6).

The greater accuracy achieved with the topographic view supports our hypothesis that relative-position judgments are easier to perform with 2D top-down views than with 3D perspective views. In this case, relative height was easier to determine in the topographic view, presumably because elevation was explicitly coded by the color of contour lines. Participants needed to determine only the color of the nearest contour line for each point and then compare colors.

In the 45° view, on the other hand, elevation is conveyed by shadows and rendered height. Rendered height is distorted in the sense that large vertical distances in the background appear very small. Additionally, shadows are an
unreliable height cue because flat ground may
in fact have a subtle but significant gradient.
Therefore, hills that appear to rise equally high
above a plain may have significantly different
elevations. The 90° 3D view does not suffer
from the distortions of distance that are inher-
ent in the 45° view, but it still falls prey to the
problem of using shadows to convey height.

The A-High-B task is a moderately imprecise
and easy relative-position task. This may account
for the small 2D accuracy advantage. The final
experiment tests another relative-position task
that requires more precise judgments of rela-
tive height, as well as relative latitude and lon-
gitude.

EXPERIMENT 6: A-TO-B

For the final terrain experiment, we devel-
oped a terrain navigation task similar to the
block navigation task in Experiment 2. Particip-
ants were asked to estimate the distance in lati-
tude, longitude, and altitude between two points
on terrain. Participants were given multiple-
choice answers that were within a half mile of
the correct answer for latitude and longitude and
within one color change of the correct an-
swer in the six-color topographic map for alti-
tude. Therefore the A-To-B task required fairly
precise judgments of relative position to select
the appropriate answers.

Because of these task demands for precise
distance judgments, we predicted that the 2D
topographic views of the terrain would lead to
better performance than the 3D views. We also
predicted that the 45° 3D views would lead to poor
performance on all three dimensions whereas the 90° 3D views would lead to poor
performance primarily on the altitude dimen-
sion, because latitude and longitude would be
represented faithfully.

Of secondary interest, we were curious
whether we could improve the 3D display by
adding artificial cues for distance and altitude.
To this end we created two new sets of views by
rendering grid lines and colored contour lines
onto the terrain in the 3D views (see Figure 8). Our
hope was that the superimposed grid and contour lines would facilitate relative-position
judgments, whereas the 3D perspective view
would continue to facilitate shape understand-
ing. To be fair, we created one more set of maps
by rendering the topographic maps with super-
imposed grid lines. We expected to find that
the grid and contour lines raised performance
on all three dimensions of the 3D views to the
level of performance of the topographic views.

Method

Participants. The participants were 24 stu-
dents from local universities. They were paid
for their participation.

Stimuli. The stimuli were the same 26 stimuli
with A and B points used in Experiments 4 and
5. Each stimulus was rendered six ways: a 45°
view with no grid or contour lines, a 90° view
with no grid or contour lines, a topographic
map with no grid, a 45° view with grid and con-
tour lines superimposed, a 90° view with grid
and contour lines superimposed, and a topo-
graphic map with grid lines superimposed.

In the No Grid conditions, one-mile marks
were arranged along the south and west sides
of all views, but grid lines did not appear on
the maps. Altitude was indicated on the topo-
graphic maps by showing a color scale to the
right of the map. Altitude was indicated on the
3D views by telling the participants that on a
scale of 1 to 6, the lowest point on the map was
a 1 and the highest point on the map was a 6.

To the left of the maps, three sets of five but-
tons were arranged vertically (see Figure 8). The
buttons in each set displayed multiple-choice
answers for each dimension. The choices for lati-
tude and longitude were arranged in half-mile
increments, and the choices for altitude were
arranged in one-color step increments.

Procedure. The stimuli were displayed one at
a time on a 15-inch color LCD panel. The par-
ticipants’ task was to view the terrain and select
the correct distances between points A and B. If
the answer on any dimension was incorrect, the
computer would sound a buzzer, highlight the
incorrect dimension(s), and gray out the incor-
rect answer(s). We provided this feedback to
ease frustration because this task proved to be
very difficult. After three incorrect answers, the
computer advanced to the next trial. Response
times and errors were recorded for each answer.

The 26 stimuli for each type of view (45°,
90°, and topographic) were grouped into blocks,
and the order of the blocks was counterbalanced.
among participants. Half the participants (n = 12) saw all three views in the No Grid condition, and half (n = 12) saw all three views in the Grid + Topo condition.

Results and Discussion

Proportion correct scores are shown in Figure 9. The columns represent the proportion of trials for which individual dimensions were answered correctly on the first try, and the horizontal bars represent the proportion of trials for which all three dimensions were answered correctly on the first try. For proportion correct scores, the three-way interaction of type of view (45, 90, and topo) by dimension (latitude, longitude, and altitude) and by augmentation (No Grid and Grid + Topo) proved to be statistically reliable, $F(4, 88) = 15.4, p < .0001$. To understand this interaction, we must break it down.

We begin with the No Grid condition. The task proved to be very difficult for most participants: The average proportion correct score was .38. Proportion correct scores for the three views were statistically different, $F(2, 22) = 61.7, p < .0001$. Post hoc tests revealed that the 45° views were worst, the 90° views were better, and the topographic views were best, $ps < .0005$. Total response times for the three views were also statistically different, $F(2, 22) = 13.5, p = .0002$. Post hoc tests revealed that the 45° views (53 s) were slower than the 90° and topographic views (32 s and 36 s, respectively).

As expected, all three dimensions were judged poorly in the 45° view condition, and only the altitude dimension was judged poorly in the 90° 3D view condition. All three dimensions were judged well in the topographic view condition. The interaction of view and dimension was statistically significant, $F(4, 44) = 22.8, p < .0001$. Interestingly, after we added the grid and contour lines, altitude remained the worst dimension overall: only 65% correct for altitude compared with 88% correct for longitude and latitude, $F(2, 22) = 50.1, p < .0001$. Apparently, determining the color of the contour line closest to each position and then counting the number of color steps between them was more difficult than counting the number of tic marks between the two points. The relative difficulty depends on the precision.
required when choosing among the multiple-choice answers.

Augmenting the views with the grid and contour lines substantially facilitated overall accuracy, which nearly doubled to .69, $F (1, 22) = 71.8, p < .0001$. The interaction of augmentation condition and view type was also statistically significant, $F(2, 44) = 17.4, p < .0001$. The interaction indicates that the majority of the improvement from adding the grid and contour lines came from the 45° view condition, whose proportion correct score increased from .13 to .56. Nonetheless, the 45° condition continued to have the worst accuracy performance among the Grid + Topo conditions, $F(2, 22) = 10.3, p = .0007$. Meanwhile, the 90° 3D views with the grid and contour lines reached the level of accuracy of the 2D topographic views.

Adding the grid and contour lines did not reliably improve response times, $F(1, 22) = 3.3, p = .08$. However, the interaction of augmentation condition and view type was significant, $F(2, 44) = 5.27, p = .05$, with all of the improvement coming from the 45° condition. Despite this improvement, among the Grid + Topo conditions, the 45° views (39 s) were answered more slowly than the 90° views (30 s) and the topographic views (35 s), $F(2, 22) = 4.8, p = .02$.

In sum, our hope that adding grid and contour lines to the 3D views would increase their usability for relative-position tasks was only partly borne out. Although the additions did significantly improve performance with the 45° views, the performance still did not match that of the 2D topographic maps. The augmentations did bring the performance on 90° views to the level of the 2D topographic views. Looking across experiments, the 90° views appear very promising for both shape-understanding and relative-position judgment tasks, especially with the addition of grid and contour lines. However, more work must be done to test this possibility.

**GENERAL DISCUSSION**

Three-dimensional displays are compelling, integrated, and natural, yet this natural representation is fraught with ambiguity and distortion. Two-dimensional displays are not as integrated and seem less compelling, but they faithfully represent space. Our goal was to find a useful distinction among cognitive tasks that matched the complementary characteristics of the views. Such a distinction should be valuable to display designers to help them maximize the usability of their displays.

To summarize our findings, a single 3D perspective view was superior to three 2D views for understanding the shape of simple blocks (Experiment 1) and natural terrain (Four-Corners and A-See-B). However, the 2D views were superior to 3D views for understanding the relative positions of two objects (Experiment 2) and two terrain locations (A-High-B and A-To-B).

Three-dimensional views are useful for shape and layout understanding because they (a) integrate all three dimensions into a single rendering, (b) are receptive to supplementary depth cues, and (c) allow features of an object
to be depicted that would be invisible in a normal 2D view. With 2D views, no single view can provide information about all three dimensions of an object; to present a third dimension, a separate view must be added. Information about the shape of an object or scene must then be combined mentally, which is both difficult and time-consuming. A perspective view is easier to use because it integrates the dimensions into a single view. Supporting this notion, others have found shaded perspective views of household objects faster to recognize than silhouettes (L iter et al., 1997).

The second advantage of 3D views is that extra depth cues can be added, such as shadows, object scaling (i.e., distant object features are drawn smaller), texture gradients, and shading, making 3D shape immediately apparent. Others have found that foreshortened views of household objects can be recognized more easily when consistent perspective depth cues are added to the scene (Humphrey & Jolicoeur, 1993). Each additional depth cue adds to the portrayal of depth. In fact, the benefits of each cue appear to be independent and additive (Bruno & Cutting, 1988). With 2D views, few depth cues, apart from occlusion, are applicable.

The third advantage of 3D views is that they allow for the illustration of object features that could be hidden in a 2D view. For example, dashed lines are used in 2D engineering models to designate hidden features of an object.

Two-dimensional views, on the other hand, are useful for judging relative positions because the normal viewing angles (e.g., top-down, side, front) minimize distortion: What is visible is represented faithfully. Ambiguity is confined to a single dimension such as altitude in a top-down view. This confinement of ambiguity to the dimension that is not represented provides better opportunities to deal with the ambiguity. For instance, a user can easily switch among a set of 2D views to obtain exact information about each dimension of interest. In contrast, each dimension of a 3D view is confounded with ambiguity spread across all three dimensions. This ambiguity and distortion make relative-position judgments of any precision difficult.

The integration of dimensions in a 3D view and their separation in 2D views is the crux of the integrated/separated hypothesis proposed by Haskell and Wickens (1993). Tasks that require understanding all three dimensions at once are best performed using a 3D view. Our opinion is that there are counterpoising task requirements for precise metric judgments of relative position that can outweigh the benefits of integration.

For comparison we list in Table 1 a number of published studies that compared 2D and 3D displays. The first column lists the study and the task performed by the participants. The second column lists the point of view (egocentric and exocentric) of the displays. The third column lists the outcome of the experiment – which display type produced better performance. The fourth column lists whether the task predominantly involved shape-understanding or relative-position judgments, and the fifth column scores the match between the result and the hypothesis’ predicted result (2D better for relative-position judgments, 3D better for shape understanding). The sixth column lists whether the task predominantly involved integrated spatial dimensions or separated dimensions. The seventh column scores the match between the result and this hypothesis’ predicted result (2D better for separated dimension tasks, 3D better for integrated dimension tasks). The final column lists the method by which altitude information (the third dimension) was accessed by the participant and displayed.

As mentioned at the outset, there is a great deal of complexity in the literature. Tasks, measures, display details, results, and effect sizes vary enormously. This variety is good for sampling the experiment space but difficult for making comparisons. The goal for Table 1, therefore, was not to test either hypothesis with a formal meta-analysis but to marshal a set of studies, find some coherence, and launch a deeper discussion of theoretical explanations and related issues.

We attempted to limit the table to published studies, but two technical reports were added to increase the number of authors and laboratories represented. Studies often reported more than one experiment, and each experiment often included more than one task or dependent measure. If the separate experiments and measures painted a common picture, the results were summarized in a single row of the table. If the tasks or measures differed in theoretically important ways, or their results were substantially different, multiple rows of the table were used.
## TABLE 1: Empirical Tests of 2D and 3D Displays

<table>
<thead>
<tr>
<th>Study and Task</th>
<th>View</th>
<th>Result</th>
<th>su/ rp</th>
<th>Score</th>
<th>int/ sep</th>
<th>Score</th>
<th>Altitude Access</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boyer et al. (1995) confliction, re-vectoring</td>
<td>exo</td>
<td>2D</td>
<td>rp</td>
<td>1</td>
<td>int</td>
<td>0</td>
<td>various</td>
</tr>
<tr>
<td>Haskell &amp; Wickens (1993) altitude, distance</td>
<td>ego</td>
<td>same</td>
<td>rp</td>
<td>–</td>
<td>sep</td>
<td>–</td>
<td>side view</td>
</tr>
<tr>
<td>Haskell &amp; Wickens (1993) closest point of approach</td>
<td>ego</td>
<td>same</td>
<td>rp</td>
<td>–</td>
<td>int</td>
<td>–</td>
<td>side view</td>
</tr>
<tr>
<td>Wickens et al. (1996) altitude</td>
<td>ego, exo (2D)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>rp</td>
<td>(1)</td>
<td>sep</td>
<td>(1)</td>
<td>side view</td>
<td></td>
</tr>
<tr>
<td>Wickens et al. (1996) relative direction</td>
<td>ego</td>
<td>2D</td>
<td>rp</td>
<td>1</td>
<td>sep?</td>
<td>1</td>
<td>side view</td>
</tr>
<tr>
<td>Wickens et al. (1996) absolute direction</td>
<td>ego, exo 3D</td>
<td>rp</td>
<td>0</td>
<td>sep?</td>
<td>0</td>
<td>side view</td>
<td></td>
</tr>
<tr>
<td>Bemis et al. (1988) closest interceptor</td>
<td>exo</td>
<td>3D</td>
<td>rp</td>
<td>0</td>
<td>int</td>
<td>1</td>
<td>hooking</td>
</tr>
<tr>
<td>Burnett &amp; Barfield (1991) confliction</td>
<td>exo</td>
<td>same</td>
<td>rp</td>
<td>–</td>
<td>in</td>
<td>–</td>
<td>digital</td>
</tr>
<tr>
<td>Burnett &amp; Barfield (1991) altitude</td>
<td>exo</td>
<td>same</td>
<td>rp</td>
<td>–</td>
<td>sep</td>
<td>–</td>
<td>digital</td>
</tr>
<tr>
<td>Ellis et al. (1987) confliction</td>
<td>ego</td>
<td>3D</td>
<td>rp</td>
<td>0</td>
<td>int</td>
<td>1</td>
<td>digital</td>
</tr>
<tr>
<td>Naikar et al. (1998) closest interceptor</td>
<td>exo (2D)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>rp</td>
<td>(1)</td>
<td>int</td>
<td>(0)</td>
<td>digital</td>
<td></td>
</tr>
<tr>
<td>Wickens, Miller &amp; Tham (1996) confliction</td>
<td>exo (2D)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>rp</td>
<td>(1)</td>
<td>int</td>
<td>(0)</td>
<td>digital</td>
<td></td>
</tr>
<tr>
<td>Andre et al. (1991) flight recovery</td>
<td>ego, exo 2D exo&lt;sup&gt;d&lt;/sup&gt;</td>
<td>rp</td>
<td>1</td>
<td>int</td>
<td>0</td>
<td>digital</td>
<td></td>
</tr>
<tr>
<td>O’Brien &amp; Wickens (1997) flight avoidance</td>
<td>exo</td>
<td>2D&lt;sup&gt;e&lt;/sup&gt;</td>
<td>rp</td>
<td>1</td>
<td>int</td>
<td>0</td>
<td>forward view</td>
</tr>
<tr>
<td>Hickox &amp; Wickens (1999) terrain matching</td>
<td>ego</td>
<td>3D</td>
<td>su</td>
<td>1</td>
<td>int</td>
<td>1</td>
<td>none</td>
</tr>
<tr>
<td>Liter et al. (1997) identification time</td>
<td>ego</td>
<td>3D</td>
<td>su</td>
<td>1</td>
<td>int</td>
<td>1</td>
<td>none</td>
</tr>
<tr>
<td>Overall</td>
<td>rp</td>
<td>7 for</td>
<td>int</td>
<td>4 for</td>
<td>3 against</td>
<td>5 against</td>
<td>2 tied</td>
</tr>
<tr>
<td></td>
<td>su</td>
<td>2 for</td>
<td>sep</td>
<td>2 for</td>
<td>0 against</td>
<td>1 against</td>
<td>0 tied</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Parentheses denote results that are partial. The su/rp column indicates whether the task was classified as shape-understanding or relative-position judgment. The following column indicates whether the result matched the prediction of the hypothesis. The int/sep column indicates whether the task was classified as integrated or separated, and the following column indicates the result/hypothesis match.

<sup>a</sup> Speed accuracy trade-off; 2D was more accurate. <sup>b</sup> Conflicting results, but planar views were better for the high-priority alerts. <sup>c</sup> The most difficult condition, changes in heading and altitude illustrated by spatial vectors, was harder in 3D than in 2D. All other conditions were equivalent. <sup>d</sup> However, the 3D display promoted more integrated (multidimensional) recovery maneuvers. <sup>e</sup> Conversely, the 2D display promoted more integrated maneuvers here.
The first section of the table lists studies using air traffic control (ATC) tasks. Participants viewed an airspace and performed several related tasks. Confliction involves predicting whether two aircraft will collide on their current paths or whether an aircraft will collide with terrain or weather. Revectoring involves giving the pilot a new direction to fly, to avoid a collision. Closest interceptor involves identifying the interceptor aircraft that is closest to the target aircraft. We categorized confliction, revectoring, and closest interceptor as integrated tasks because they involve analyzing all three spatial dimensions simultaneously, as do the real-world tasks. We categorized judging relative altitude, relative and absolute direction, and distance as separable tasks because they involve analyzing only one or two dimensions at a time. These ATC tasks are relative-position judgment tasks because they involve judging the relative positions, directions, and distances between two or more objects. Fortunately, categorizing the ATC tasks for both theories seems clear: The waters get murkier for other tasks.

Recovery flight involves flying an airplane, having the displays go blank while a random turn and dive or climb is executed by the simulation, and then having to recover the original flight path as quickly as possible once the displays reappear. Recovery is different from standard flight control because a single large revectoring is required, rather than incremental small changes. Recovery also requires a greater sense of situation awareness than flight control (Andre et al., 1991). But recovery does not require the degree of precision judgment that is required for collision; a relatively gross movement in the correct direction will suffice. Avoidance flight involves steering one’s own aircraft away from an imminent conflict. It shares similarities with revectoring because it involves a single large change, but it does not require much precision. Both recovery flight and avoidance flight are integrated tasks because they involve simultaneous evaluation and change of all three dimensions.

One task that we did not include in Table 1 is flight control. Flight control is difficult to categorize because it requires only gross relative-position judgments. Constant small corrections can be made to keep the aircraft in the path, so no great precision is required. It is also difficult to determine which aspects of flight control are integrated or separable. For example, Haskell and Wickens (1993) argued that lateral and altitude changes are integrated, but air speed changes are separable.

In fact, neither distinction is immune to categorization problems. Line-of-sight judgments offer an interesting case for the shape-understanding/relative-position judgment distinction. We consider the line-of-sight judgments in the A-See-B task to involve shape understanding because they hinge on assessing the rough shape of the terrain between the two points – whether or not a mountain or range of hills lies between the points. However, we consider the line-of-sight judgments in ATC confliction tasks to involve relative-position judgments because confliction predictions require precise judgments of angles and distances, and small changes can make the difference between collision and safety. In fact it should be possible to titrate a shape-understanding line-of-sight task, such as the A-See-B task, to the point at which very fine judgments would be required to determine whether a line of sight was interrupted. At this point the task would have to be reclassified as a relative-position task because of the requirement for precise judgments about the heights and positions of the two points and potential obstructions.

The third section of the table lists the few studies we could find that have characteristics both of requiring shape understanding and of pitting 2D views against 3D views. The proxy for shape understanding in the Liter et al. (1997) study is object recognition. Because shape is a 3D construct, Haskell and Wickens (1993) would predict that object recognition would be easier with integrated 3D views, as indeed it is in the case of Liter et al. Another line of shape-understanding work that also relies on recognition, but in the domain of flight, is that by Hickox and Wickens (1999) and colleagues. They assessed the ability of a pilot to match an out-of-cockpit view of terrain to that on a display console. Three-dimensional views that have the same viewing angle as the forward field of view of the pilot lead to better performance than top-down views.

The complexity of the literature is substantiated in the table: Neither theory has a perfect
prediction record, and results vary from experiment to experiment. Overall, as shown in the bottom of Table 1, the score for the relative-position hypothesis is somewhat better than for the integration hypothesis, but the small sample sizes for these two hypotheses limit any firm conclusions. Our interpretation is that the integrated/separator hypothesis is a valuable factor for predicting and explaining the results in Table 1, but that an equal, if not larger, factor is the task requirement for shape understanding or precise relative-position information.

Much of the variability in the results shown in Table 1 presumably stems from variability in the design details of the displays used in the experiments. For example, one key design detail that surely accounts for some of the variability is how altitude is represented in the 2D views (see the rightmost column of Table 1). Some studies used a single top-down planar view with numbers in a data block to indicate altitudes. In some studies these data blocks were constantly visible, but in other studies individual aircraft had to be “hooked” with a mouse click to reveal the data block. Other studies used a top-down view plus a separate side view, and still others used a single top-down view with topographic lines to indicate the altitude of terrain. Each way of representing altitude requires different user actions and mental processes. Each study, therefore, tests not just 2D versus 3D but 2D with a particular representation of altitude versus 3D.

For complex tasks that involve aspects of shape-understanding as well as relative-position judgments, both types of view may be useful. We are exploring several methods of combining judgments, both types of view may be useful. We are exploring several methods of combining judgments, both types of view may be useful.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the programming efforts of Tim Bank and Gene Averett of Pacific Science and Engineering Group. This work was supported by the Office of Naval Research through the Space and Naval Warfare System Center, San Diego.

REFERENCES


Mark St. John is director of cognitive systems at Pacific Science and Engineering Group, Inc. in San Diego. He received his Ph.D. in cognitive psychology from Carnegie Mellon University in 1990.

Michael B. Cowen is a scientist at the Space and Naval Warfare System Center in San Diego. He received his Ph.D. in cognitive psychology from Claremont Graduate University in 1991.

Harvey S. Smallman is a senior scientist at Pacific Science and Engineering Group, Inc. in San Diego. He received his Ph.D. in experimental psychology from the University of California, San Diego, in 1993.

Heather M. Oonk is a senior scientist at Pacific Science and Engineering Group, Inc. in San Diego. She received her Ph.D. in cognitive psychology from Washington University in 2000.

Date received: April 9, 1999
Date accepted: May 15, 2000