Overestimation of heights in virtual reality is influenced more by perceived distal size than by the 2-D versus 3-D dimensionality of the display

Melissa W Dixon*, Dennis R Proffitt
Department of Psychology, University of Virginia, 102 Gilmer Hall, Charlottesville, VA 22903-2477, USA
Received 29 February 2001, in revised form 5 September 2001

Abstract. One important aspect of the pictorial representation of a scene is the depiction of object proportions. Yang, Dixon, and Proffitt (1999 Perception 28 445–467) recently reported that the magnitude of the vertical–horizontal illusion was greater for vertical extents presented in three-dimensional (3-D) environments compared to two-dimensional (2-D) displays. However, because all of the 3-D environments were large and all of the 2-D displays were small, the question remains whether the observed magnitude differences were due solely to the dimensionality of the displays (2-D versus 3-D) or to the perceived distal size of the extents (small versus large). We investigated this question by comparing observers' judgments of vertical relative to horizontal extents on a large but 2-D display compared to the large 3-D and the small 2-D displays used by Yang et al (1999). The results confirmed that the magnitude differences for vertical overestimation between display media are influenced more by the perceived distal object size rather than by the dimensionality of the display.

1 Introduction
Research on picture perception often focuses on observers' accuracy in perceiving pictorial space relative to some physical metric space, as it exists in the world. However, this may be misleading, as inaccuracies in perceiving the spatial layout of real scenes have been reported as well (e.g. Wagner 1985; Proffitt et al 1995). As Rogers (1995, page 123) has emphasized:

"The question of veridicality properly concerns the match between perception of spatial layout in pictures and the perception of spatial layout in real-scene controls and not the match between picture perception and the objective measurements of the original scene layout."

Thus, more research that directly compares our perception of three-dimensional (3-D) scenes and their pictorial counterparts is necessary to determine how our perceptual system processes information differently in these two instances and how pictures can be altered so as to compensate for such differences.

One important aspect of the pictorial representation of a scene is the depiction of object proportions, particularly in engineering or architectural renderings in which the veridical perception of a building's proportions may be the focus. Yang et al (1999) recently reported that the magnitude of the vertical–horizontal illusion (VHI) was greater for vertical extents presented in 3-D environments compared to two-dimensional (2-D) displays, even when the visual angles subtended were held constant across display media. However, because all of the 3-D environments were large (e.g. an outdoor environment) and all of the 2-D displays were small (e.g. a monitor screen), the question remains whether the observed magnitude differences were due solely to the dimensionality of the display (2-D versus 3-D) or to the perceived distal size of the extents (small

* Present address and author to whom all correspondence and requests for reprints should be sent: Federal Aviation Administration, Aviation Security Research and Development, AAR-510, William J Hughes Technical Center, Building 315, Atlantic City International Airport, NJ 08405, USA; e-mail: melissa.dixon@faa.gov
versus large). We explored this question by comparing observers’ judgments of vertical relative to horizontal extents on a large 2-D display to the large 3-D and the small 2-D displays used by Yang et al. (1999). The displays in all three conditions in the current experiment were virtual reality (VR) displays.

The vertical–horizontal illusion occurs when a physical vertical extent is overestimated in length relative to a comparable physical horizontal extent. In a series of studies, Yang et al. (1999) systematically compared the magnitude of the VHI in 3-D environments versus 2-D pictorial displays. The 2-D displays were designed to represent the more traditional, visually impoverished studies of the VHI in which a small, approximately 5%, illusion has been reported (e.g., Avery and Day 1969; Künnapas 1955a, 1955b, 1957, 1958). The 3-D environments were designed to represent the more visually-rich studies of the VHI in which a large illusion has been reported (Chapanis and Mankin 1967; Higashiyama and Ueyama 1988; Higashiyama 1996).

In all of the Yang et al. (1999) studies, except experiment 4, observers matched a frontal horizontal extent to the height of a vertical target pole by instructing an experimenter to move a marker out horizontally from the base of the pole (real or pictured), perpendicular to the target pole and to the observers’ line of sight. In experiment 1, observers viewed real objects (a door, a light pole, and two buildings) and the results replicated the findings of Chapanis and Mankin (1967), Chapanis and Ueyama (1988), and Higashiyama (1996) that a greater VHI is found for real objects compared to the traditional ‘L’ figures in pictorial displays. Furthermore, proportional overestimation increased with object height and the magnitude of the illusion was as great as 18%.

In experiment 2, Yang et al. (1999) experimentally manipulated pole height to demonstrate that vertical overestimation increases with object height and that vertical overestimation is greater with real objects than with pictures or line drawings matched for visual angle. Observers viewed either poles (0.61 m to 5.96 m) in a large outdoor field, pictures of the poles in the outdoor field, or pictures of white lines on a black background. Both the pictures and the line drawings of the poles were viewed on a computer monitor and the visual angles subtended by the depicted poles matched the visual angles subtended by the real poles outdoors. The results showed that even though the visual angles were equalized across conditions, a greater VHI was found for the poles viewed outdoors (12%) compared to the poles viewed as 2-D pictures (2%) or line drawings (3%). Furthermore, the magnitude of the VHI increased as a function of the height of the pole in all conditions.

Although the visual angles in experiment 2 were matched, the monitor in the pictures and line drawings conditions restricted the field of view. To test the influence of a small field of view on the magnitude of the VHI, Yang et al. (1999) matched the field of view of the outdoors condition to the pictures and line drawings conditions by having observers view the outdoor poles through a viewing box in experiment 3. The results did not differ from experiment 2; the magnitudes of the VHI for the frame (14%) and the outdoors conditions (12%) were both larger than for the pictures and line drawings conditions. This study thus demonstrated that the reduced overestimation in the pictures conditions was not due to the framing of the visual scene by the computer monitor.

In experiment 5, Yang et al. (1999) demonstrated that an immersive VR head-mounted display (HMD) evokes the same magnitude of overestimation found outdoors. The data from the outdoors and pictures conditions of experiment 2 were compared with two new conditions, poles viewed in a virtual field or pictures of the poles in the virtual field viewed on a computer monitor. Although the virtual field was rendered

---

(1) See Yang et al. (1999) for a thorough discussion of the theories that have been proposed to explain the vertical–horizontal illusion.
as computer graphics images, observers experienced being surrounded by a 3-D world with objects projected at their true distal sizes. Again, the visual angles were matched across conditions. The results showed that the observers in the virtual field produced the same VHI as outdoors observers (12%), and the observers who viewed pictures of the virtual field on a monitor produced a VHI similar to the observers who viewed pictures of the outdoor poles (4% and 2%, respectively).

The results of experiment 5 provided substantial evidence that the magnitude of the VHI is greater in 3-D environments compared to 2-D displays. However, experiment 6 confirmed that the results of the virtual-field condition were not simply a result of the slight optical distortion present in the lenses of the HMD. The data from the virtual-field and pictures-of-the-virtual-field conditions were compared to the data from one new condition, pictures of the poles in the virtual field viewed on a virtual monitor. With this control, not only were the visual angles the same between the virtual-field and virtual-monitor conditions, but it was also the case that exactly the same images were projected into the eyes. The results showed that observers who viewed the pictures of the virtual poles on the virtual monitor produced a similar VHI as the observers who viewed the pictures of the virtual poles on a real computer monitor (3% and 4%, respectively).

To summarize, the results presented by Yang et al (1999) suggest that the perception of object proportions varies as a function of the display medium in which the object is presented. They found that objects presented in small 2-D displays were perceived to be somewhat compressed compared to the same objects presented in large 3-D displays, even when the retinal visual angles were held constant. This research implies that both the perceived distal object size and the dimensionality of the display might play important roles in the VHI and the perception of object proportions. However, none of the large-scale displays used by Yang et al was projected onto a flat 2-D surface, whereas all of the small-scale displays were displayed in this manner. The question thus remains whether the dimensionality of the display (2-D versus 3-D) or the apparent distal object size (small versus large display) was the most important factor in the Yang et al studies.

2 Experiment: Differentiating the influence of perceived display dimensionality and apparent distal object size on the vertical–horizontal illusion

A large, 2-D virtual movie screen was used in this experiment to display the Yang et al (1999) scene of the poles in the virtual field. This condition was compared to the virtual outdoor field and the virtual desktop display conditions that Yang et al used. If the perceived distal object size is the important variable, then observers should produce overestimations similar to those previously found with large, 3-D displays, because a movie screen presents the poles in their true large-scale form. On the other hand, if the dimensionality of the display is the important variable, then observers should produce smaller overestimations like those previously found with small, 2-D displays. The results showed that the perceived distal object size and not the dimensionality of the display most influences the magnitude of the VHI.

2.1 Method

2.1.1 Participants. Seventy-two undergraduates from the University of Virginia participant pool participated in experiment 1 for course credit; twenty-four undergraduates participated in the virtual-movie-screen condition and forty-eight undergraduates participated in the virtual-field and virtual-monitor conditions of Yang et al (1999). All had normal or corrected-to-normal vision with contact lenses. Participants who wore glasses could not participate because the HMD does not accommodate glasses.
2.1.2 Design. The data from the virtual-field and virtual-monitor conditions from Yang et al (1999) were compared with the data from experiment 1 to create a 3 (Viewing Condition) × 10 (Pole Length) mixed experimental design. Viewing condition (virtual field, virtual monitor, and virtual movie screen) was a between-subjects variable and Pole Length (2 ft to 20 ft, in increments of 2 ft) was a within-subjects variable. Each observer participated in 10 trials, one for each pole height. There were twenty-four observers in each of the three conditions.

2.1.3 Stimuli and apparatus. Observers viewed a computer graphics rendering of the virtual field, the virtual monitor, or the virtual movie screen through a VR HMD. All virtual environments were designed and created with the use of the program Alice, version 0.9711 (beta version), a 3-D computer graphics authoring software. A Gateway 2000 computer with a 233 MHz Intel Pentium II processor, the Microsoft Windows 95 operating system, and 256 MB RAM controlled the execution of the program, as well as graphics rendering and head-tracking. A Monster 3-D PCI Video Multimedia graphics card processed the graphics for the virtual-field and the virtual-monitor conditions and the graphics for the virtual-movie-screen condition were processed by a 3-Dfx Interactive, Inc., VooDoo graphics card.

The screen refreshed at a rate of 60 Hz. The 6 degrees of freedom obtained from the HMD—yaw, pitch, and roll—were registered by an Ascension SpacePad magnetic tracking system in the virtual-field and virtual-monitor conditions and with a Polhemus InsideTrak magnetic tracking system in the virtual-movie-screen condition. The computer used the position and orientation coordinates to update the scene appropriately as observers moved their heads. The end-to-end latency of the tracking system, which is the time the system senses the HMD position and orientation change to the time the scene is updated, was approximately 100 ms.

Observers viewed the virtual environment through a Virtuality Visette Pro HMD, with two active-matrix color LCDs, which operated in a pseudo VGA video format. The resolution of each display screen was 640 pixels (horizontal) × 480 pixels (vertical), per color pixel. The field-of-view per eye was 60 deg (horizontal) × 46.8 deg (vertical). The HMD presented a bi-ocular display, not stereo; therefore, the same image was projected to each eye rather than two slightly different images. These images were viewed through collimating lenses that allowed the observer's eyes to focus at optical infinity.

2.1.4 Virtual field. The 360 deg virtual field, which is shown in figure 1, was rendered by Yang et al (1999) to approximate the outdoors condition. Observers viewed this scene through an HMD. The scene contained a green textured ground plane, trees, sky and clouds, and a virtual experimenter, all of which were arrayed appropriately in 3-D. The target poles were modeled from ten white PVC poles, 4.8 cm in diameter, with the following heights: 0.61, 1.22, 1.83, 2.44, 3.04, 3.54, 4.15, 4.76, 5.37, and 5.97 m. The virtual experimenter, in three different postures to hold poles of different heights, was created by cropping digital images of the experimenter holding each of the poles, scaling the images to the correct virtual size, and texture-mapping them onto a transparent graphical object. The scene also contained a white vertical marker pole (1.5 virtual m tall), but it did not contain a second virtual experimenter.

2.1.5 Virtual monitor. As shown in figure 2, the virtual environment was a 360 deg scene rendered to resemble the main features of our laboratory, including counters, cabinets, a chair, a chin-rest, and a computer monitor.

The stimuli were ten static snapshots of the virtual-field scene with the virtual experimenter holding each of the ten poles. The snapshots were created with the field-of-view parameter set to replicate a virtual-field scene photographed with a 35 mm lens.
These images were presented on a virtual monitor, where the size of the picture on the screen was set to the same size as the desktop displays used by Yang et al. (1999): 20.3 cm (tall) × 31.2 cm (wide). The lengths of the poles as measured on the monitor screen were 1.31, 2.60, 3.92, 5.29, 6.63, 7.69, 9.07, 10.37, 11.75, and 12.77 cm.

When viewed from the center of projection, at 34 cm from the monitor screen, the angular subtense of each pictured pole matched the angular subtense of the corresponding poles in the virtual-field condition. A white vertical marker line (3.4 cm tall), but not a second virtual experimenter, was present in the simulation to the right of the target pole. The marker pole subtended the same visual angle as the marker pole in the outdoors condition.
2.1.6 Virtual movie screen. A 360 deg virtual scene was rendered and viewed through the HMD. The virtual scene was composed of a green textured ground plane, a nighttime sky and clouds, numerous buildings, an amphitheater, a popcorn stand, and a movie screen (figure 3). Snapshots of the virtual-field scene were projected onto the virtual movie screen, whereby the length of the poles as measured on the movie screen (distal image size) were: 0.61, 1.22, 1.83, 2.44, 3.04, 3.54, 4.15, 4.76, 5.37, and 5.97 m. The angular subtense of each of the ten target poles projected onto the virtual movie screen matched the angular subtense of the corresponding poles in the virtual-field and virtual-monitor conditions.

![Figure 3. A snapshot of the virtual-movie-screen environment.](image)

A female movie usher was positioned on the right-hand side of the movie screen stage to provide an additional cue for the scale of the movie screen. The usher was removed from the stage when the experiment began to avoid distracting the observers.

Observers stood on a wooden platform, with a 0.91 m × 0.91 m base and a 0.91 m high × 0.91 m wide railing on the front. This platform was rendered on the upper level of the amphitheater facing the movie screen and its construction and position corresponded with a real platform on which observers stood. A 3-D model of Alice from Alice in Wonderland was also present in the scene.

2.2 General procedure
Although observers followed similar procedures in all three conditions when they produced the horizontal extent that matched the vertical extent, each condition had subtle differences relevant to the display medium. The general procedure common to all conditions is first discussed.

In all three conditions, the observers put on the HMD, tightened it until comfortable, and adjusted the focus and interocular settings until the image was clear. Observers were encouraged to look around the virtual scene before starting the experiment to familiarize themselves with the surroundings and to achieve maximum possible immersion in the virtual scene.

At the beginning of each trial, a virtual marker pole was positioned adjacent to the target pole. Although a virtual experimenter held the target pole, there was no second experimenter in the virtual scene to hold the marker pole. The target pole was
displaced one observer eye height to the left to prevent differential foreshortening of horizontal and vertical extents. The visual angles subtended by the virtual poles were held constant across all three conditions.

The experimenter explained to the observer that the task was to instruct the experimenter to move a target pole laterally to make the horizontal distance between the marker pole and the target pole equal to the height of the target pole. When the observer was ready to start, the experimenter moved the virtual marker away from the target pole to the right, using the keyboard. The experimenter moved the marker pole out at a steady pace until the participant verbally instructed the experimenter to stop. The observers used as much time as needed to make the adjustment and they could change their initial response by asking the experimenter to move the marker pole left or right. The computer recorded each of the observers’ responses. After every other trial, observers were asked to look around the virtual scene to remind them of their surroundings and to maintain immersion in the virtual world. These steps were repeated for each of the ten poles. Each session required about 30 min to complete.

2.3 Condition-specific procedures

2.3.1 Virtual field. When the observer put on the HMD in the virtual-field condition, they were in the middle of a virtual field, with a tree-lined horizon. The virtual poles were positioned at a virtual distance from the observer equal to the distance of the real poles in the outdoors condition of Yang et al (1999), experiment 1.

2.3.2 Virtual monitor. Before putting on the HMD in the virtual-monitor condition, the observer sat down in a chair in front of the monitor, which had a chin-rest in front of it. The observer then put on the HMD to view a virtual environment that simulated the laboratory room and its furnishings. Wearing the HMD simulated entering a computer graphics version of the real laboratory. In the virtual room there was, among other objects, a virtual monitor on which the snapshots of the virtual-field scene were displayed.

To enhance the compellingness of the virtual environment and to calibrate the observer’s position in front of the virtual monitor, a magnetic tracker was attached to the back of the real chair in which the observer sat. When the observer moved the real chair, the virtual chair viewed through the HMD moved in correspondence with the real chair. In addition, the experimenter encouraged the observer to reach out and touch the virtual monitor and the virtual chin-rest. Because the position and size of these virtual objects were simulated to correspond to their real counterparts, when the observer reached out to where he or she perceived the virtual monitor and virtual chin-rest, the observer’s hand touched the real monitor and chin-rest in the real laboratory.

Observers moved their heads towards the virtual monitor and positioned their heads on the chin-rest. When observers looked at the virtual monitor, a cross hair with a vertical line drawn to the chin-rest was visible to help observers position themselves at the center of projection with the virtual picture. Observers moved towards the cross hair until they just passed through it and the cross-hair disappeared. At this point they viewed a static snapshot of the scene used in the virtual-field condition on the virtual monitor. After every other trial, observers were asked to remove their chins from the chin-rest and to look around the virtual laboratory by rotating in the chair. This was done to remind observers of their surroundings.

2.3.3 Virtual movie screen. Before putting on the HMD in the virtual-movie-screen condition, the observer stepped onto a wooden platform and stood in front of the platform railing. The observer then put on the HMD to view a virtual environment that simulated a virtual town with an amphitheater, numerous buildings, a wooden platform to which a magnetic tracker was attached, a popcorn stand, and a movie screen on which the snapshots of the virtual-field scene were displayed.
A virtual female movie usher was positioned on the right-hand side of the virtual movie screen stage to provide an additional cue for the scale of the movie screen. The usher was removed from the stage when the experiment began to avoid distracting the observers. After every other trial, observers were asked to perform a visual search for Alice, who could appear in one of twelve locations in the virtual scene. This was done to remind observers of their surroundings.

3 Results and discussion

The results supported the hypothesis that the VHI illusion is influenced more by the perceived distal size than by the dimensionality of the display. That is, larger proportional overestimations were obtained in the large-scale virtual-field and virtual-movie-screen conditions compared to the small-scale virtual-monitor condition, despite the fact that the virtual movie screen was 2-D.

Figure 4 illustrates the two main findings: (i) vertical overestimation was greater for the virtual-field and virtual-movie-screen conditions than for the virtual-monitor condition, and (ii) vertical overestimation increased with the size of the virtual or pictured object.

![Figure 4](image_url)

**Figure 4.** Mean proportional overestimation of height as a function of viewing condition.

Collapsed across pole length, the mean proportional overestimations for the virtual field, virtual monitor, and virtual movie screen were, respectively, 1.12, 1.03, and 1.15. The mean proportional overestimation for individual pole lengths ranged from 1.08 to 1.17 (virtual field), 1.02 to 1.06 (virtual monitor), and 1.07 to 1.21 (virtual movie screen).

A repeated-measures general linear model analysis of variance, with Pole Length as a within-subjects variable and Viewing Condition as a between-subjects variable, revealed a reliable difference between the means for Viewing Condition: $F_{2,647} = 6.88$, $p = 0.002$. Bonferroni a posteriori tests showed that virtual movie screen was reliably different from virtual monitor: $t_{48} = 5.13$, $p < 0.01$, and the virtual field was reliably different from virtual monitor: $t_{47} = 2.64$, $p < 0.05$, but virtual movie screen was not reliably different from virtual field: $t_{48} = 0.88$, $p > 0.05$. Overall, proportional overestimation increased with the height of the virtual object: $F_{1,647} = 21.35$, $p < 0.0001$.

These results provide strong evidence that the VHI is influenced more by the perceived distal object or image size, regardless of the dimensionality of the display.
Although the movie screen was 2-D, the poles presented on it were perceived to have the same distal size as the poles presented in the 3-D virtual field, which was much larger than the distal size of the poles presented in pictures on the virtual monitor screen. These data imply that objects presented in a small-scale display may be perceived as somewhat compressed in the vertical dimension relative to when the same objects are presented in their true large-scale form.

4 Conclusion
The purpose of the present investigation was to assess whether the magnitude differences for the VHI reported by Yang et al (1999) between display media were more influenced by the dimensionality of the display or by the perceived distal object size. Observers matched a frontal vertical extent to a frontal horizontal extent either on a large-scale 2-D virtual movie screen, in a large-scale 3-D virtual field, or on a small-scale computer monitor, with visual angles held constant.

The results showed that observers' judgments in the virtual-movie-screen condition were not significantly different from observers' judgments in the virtual-field condition, both of which were large-scale displays but with different dimensionalities. However, observers in both of the large-scale display conditions produced significantly greater estimations than observers in the small virtual-monitor condition.

With regard to the VHI, this experiment suggests that the magnitude differences obtained by Yang et al (1999) indeed can be explained by the apparent distal object or image size rather than by the dimensionality of the display. The virtual movie screen was perceived to be 2-D, but the distal sizes of the images presented on it were as large as the 3-D poles in the virtual-field condition.

In conclusion, while the small 2-D virtual monitor presented a dual reality between the distal sizes of the images on the display surface and the true sizes of the objects that the images represented, there was no difference between the distal sizes in the display and the true sizes of the objects on the large, 2-D virtual movie screen. The larger a 2-D representation, the more likely the visual system is to achieve a natural perception of large depicted objects. Moreover, because the proportions of large objects may appear compressed when presented on a small-scale display, such images may need to be stretched to preserve the natural perception of their proportions.

Acknowledgements. This research was supported by CMU/DARPA grant 539689-52273, ONR grant N0001401010060, NASA grant NCC2925, and NIMH grant MH52640, to the second author. This experiment was one of several experiments conducted for the dissertation of the first author, in partial fulfillment of the PhD in Psychology at the University of Virginia. The authors wish to acknowledge Brett Lider, an undergraduate research assistant who helped to run participants in this study.

References


Wagner M. 1985 "The metric of visual space" *Perception & Psychophysics* **38** 483–495