Journal of Experimental Psychology:
Human Perception and Performance
2001, Vol. 27, No. 6, 1314–1334

0096-1523/01/$5.00 DOI: 10.1037//0096-1523.27.6.1314

Accommodation, Occlusion, and Disparity Matching Are Used to Guide Reaching: A Comparison of Actual Versus Virtual Environments

Geoffrey P. Bingham, Arthur Bradley, Michael Bailey, and Roy Vinner
Indiana University Bloomington

The authors used a virtual environment to investigate visual control of reaching and monocular and binocular perception of egocentric distance, size, and shape. With binocular vision, the results suggested use of disparity matching. This was tested and confirmed in the virtual environment by eliminating other information about contact of hand and target. Elimination of occlusion of hand by target destabilized monocular but not binocular performance. Because the virtual environment entails accommodation of an image beyond reach, the authors predicted overestimation of egocentric distances in the virtual relative to actual environment. This was confirmed. The authors used −2 diopter glasses to reduce the focal distance in the virtual environment. Overestimates were reduced by half. The authors conclude that calibration of perception is required for accurate feedforward reaching and that disparity matching is optimal visual information for calibration.

Space perception studies focus on the relative contributions of different sources of information to the perception of spatial properties like distance, size, or shape. Until the last decade or so, reaching was not used systematically as a measure in space perception studies. The only exceptions were a study by Foley and Held (1972), in which reaching was used as a measure of distance perception, and a substantial number of studies on direction perception (Bingham & Romack, 1999; see Welch, 1978, for review). Jeannerod and colleagues (see Jeannerod, 1988, for review) performed studies investigating the contributions of location versus size perception to the control of reaching versus grasping, respectively.

More recently, a number of studies have been published showing dissociations between judgments and action measures of perception (e.g., Aglioti, DeSouza, & Goodale, 1995; Bridgeman, Kirch, & Sperling, 1981; Bridgeman, Peery, & Anand, 1997; Gentillucci & Negrotti, 1994; Haffenden & Goodale, 1998; Loomis, DaSilva, Fujita, & Fukusima, 1992; Marotta, DeSouza, Haffenden, & Goodale, 1998; Pagano & Bingham, 1998; Post & Welch, 1996). Various hypotheses have been formulated to account for these dissociations, for instance, that the dissociations are a function of task specificity (Rieser, Pick, Ashmead, & Garing, 1995), of egocentric versus exocentric distance perception (Loomis et al., 1992), or of relative versus absolute distance perception (Vishton, Rea, Nunez, & Cutting, 1999). However, the most influential hypothesis has been that of Goodale and Milner (1992).

Goodale and Milner (1992) suggested that there are two anatomically distinct visual systems (Goodale, Jakobson, & Servos, 1996; Milner & Goodale, 1995). One (called *perception*) is hypothesized to be for recognition and to include the ability to be aware of and to describe what one is seeing. The other (called *perception–action*) is hypothesized to be for perceptually guided action and to entail a lack of awareness or a lack of an ability to describe. Bridgeman suggested that the perception–action channel is accurate, whereas the perception channel yields distorted perceptions (Bridgeman et al., 1997). This suggestion seemed to be consistent with Milner and Goodale's (1995) conception. Indeed, perceptual judgments have long been known to be subject to various illusions (e.g., Gregory, 1970; Hochberg, 1978). More to the point and as reviewed by Todd, Tittle, and Norman (1995), distortion in the visual perception of egocentric distance and shape has been found in many studies using passive judgments (see also, e.g., Baird & Biersdorf, 1967; Fermuller, Cheong, & Aloimonos, 1997; Gilinsky, 1951; Johnston, 1991; Loomis et al., 1992; Norman & Todd, 1993; Norman, Todd, Perotti, & Tittle, 1996; Parker, Cumming, Johnston, & Hurlbert, 1995; Perotti, Todd, & Norman, 1996; Phillips & Todd, 1996; Tittle & Braunstein, 1993; Tittle & Perotti, 1997; Tittle, Todd, Perotti, & Norman, 1995; Todd & Bressan, 1990; Toye, 1986; Wagner, 1985).

In the context of these results, the question arose about whether such distortions would appear in actions and, in particular, in visually guided reaches. Reaches were found to be free of illusions in some studies (e.g., Aglioti et al., 1995; Haffenden & Goodale, 1998; Marotta et al., 1998; Vishton et al., 1999) but subject to illusions in other studies (e.g., Franz, Gegenfurtner, Bülthoff, & Fahle, 2000; Pavani, Boscagli, Benvenuti, Rabufetti, & Farne, 1999; Post & Welch, 1996). Using targeted reaching as a measure of space perception, Bingham and coinvestigators (Bingham & Pagano, 1998; Bingham, Zaal, Robin, & Shull, 2000; Pagano & Bingham, 1998; Wickelgren, McConnell, & Bingham, 2000) found distortions of egocentric distance, size, and shape akin to those found in previous judgment studies. They concluded that action measures do not necessarily entail accurate perception.

Geoffrey P. Bingham, Michael Bailey, and Roy Vinner, Department of Psychology, Indiana University Bloomington; Arthur Bradley, School of Optometry, Indiana University Bloomington.

This work was supported by Grant RO1 EY11741-01A1 from the National Eye Institute.

Correspondence concerning this article should be addressed to Geoffrey P. Bingham, Department of Psychology, 10th and Grove Streets, Indiana University, Bloomington, Indiana 47405. Electronic mail may be sent to gbingham@indiana.edu.

Normal, everyday actions do require accurate performance to avoid ramming hands into doorknobs or knocking over cups of coffee. The existence of a separate channel would not address the fundamental question that remains for studies on space perception, namely, "what information is available and used to enable accurate reaching performance?" A feature intrinsic to perception–action is that actions yield feedback information that can be used to calibrate visual information about the surrounding space (Bingham & Pagano, 1998). Reaches, for instance, almost always result in contact of hand and object, yielding both visual and haptic feedback about the visually perceived distance, size, and shape.

Reaches typically entail the use of both feedforward and on-line feedback information. Before the arm begins to move, visual information about target distance, size, and shape is used to initiate a reach. At this juncture, definite (Bingham, 1993b) or absolute scale information is required and is used, as shown by the fact that a subsequent ballistic action (i.e., an action without continuous visual guidance) is typically accurate within a tolerance of a few centimeters. As shown by Bingham et al. (2000), the tolerance is improved when feedback can be used to calibrate this information.

Bingham et al. (2000) studied the use of haptic feedback to calibrate the visual perception of definite distance, size, and shape. In the current article, we investigate sources of visual information that might be used for calibration. The first source, *occlusion,* is available to monocular vision as well as binocular. We investigated reaching both with and without occlusion of the hand by the target. The second source of information, *disparity matching,* is only available to binocular vision. We investigated reaching when only disparity matching was available to judge when the hand was at the target. To control and manipulate these sources of information in the context of active perception and reaching, we used a virtual environment.

Comparison of Virtual and Actual Environments

A *virtual environment lab* is potentially a very powerful tool for the study of perception and action. It allows one to control and manipulate both visual and haptic information while preserving the natural coupling of self-motion and optic flow and allowing the measurement of relevant actions like reaching. However, such control comes at the cost of accessory perturbations of perception. As a computer graphics based display system, the virtual environment lab involves the viewing of images. Of course, vision is not normally mediated by a display or a viewed image. The problem in a virtual environment is that one is looking at an image, but one sees a virtual object. The image and the object are not typically at the same place. On the one hand, the observer must focus the image, so he or she must be accommodated to the focal distance of the display optics. In a *head-mounted display* (HMD), a virtual image of each display is viewed through a lens that places the image at a comfortable, fixed viewing distance. Thus, the observer is accommodated to a fixed virtual image distance while, on the other hand, the observer's eyes must converge to the distance of a virtual object, and the latter distance varies as the observer moves around. Normally, accommodation and convergence are coupled and coincident (Howard & Rogers, 1995), but in any virtual environment, they must be decoupled and divergent (Wann, Rushton, & Mon-Williams, 1995). There are other inevitable perturbations that are common to nearly all displays. Although rarely discussed, displays always involve image distortions due to the shape of the screen or to the display optics (Kocian & Task, 1995). HMDs entail a restricted field of view as well as limited spatial and temporal resolution. Finally, virtual environments always entail some phase lag between self-motion and corresponding optical transformations (Barfield & Furness, 1995; Held & Durlach, 1991; Held, Efstathiou, & Greene, 1966; Liang, Shaw, & Green, 1991; Smith & Bowen, 1980; Tharp & Liu, 1992).

Because of such perturbations, it is important to compare performance in actual and virtual environments. We compared performance in visually guided reaching in the two environments. We used tasks previously studied in an actual environment by Bingham et al. (2000). First, we attempted to replicate these results in an actual environment, and then we examined performance in the same tasks in the virtual environment. Two tasks could be performed equally in either environment, namely, normal visually guided reaching and feedforward reaching. In the former task, participants viewed the targets both before and during the reaches, whereas in the latter task, participants viewed the targets only before each reach and then reached without vision. The former allows use of relative distance information, whereas the latter entails use of only definite distance information.

Two features of virtual environments are largely responsible for the experience of a 3-D layout. One is stereo viewing; the other is optic flow or structure from motion, generated by and coupled to self-motion. We tested these under two conditions. The first was intended to be representative of normal performance conditions that entail both binocular vision and a moving observer. Participants used binocular vision and moved their heads from side to side while viewing a target before each reach. This should have maximized the available information about distance, size, and shape. The second condition isolated the use of structure from self-motion. The latter is notably different from standard structure from motion because the optic flow is coupled with somatosensory information about the motion. The somatosensory information could be used, in principle, to enable perception of definite sizes and distances (Bingham & Stassen, 1994; Nakayama & Loomis, 1974). Participants used monocular vision and moved their heads from side to side while viewing targets before each reach. In Experiment 1, we tested visually guided reaching using either binocular or monocular vision. In Experiment 2, we tested feedforward reaching using either binocular or monocular vision. In Experiment 3, we directly tested the effect of a fixed accommodative distance in the virtual environment. Performance in the virtual environment was tested both without and with -2 diopter glasses that reduced the accommodative distance. In Experiment 4, the use of binocular disparity matching was isolated and tested in the virtual environment to confirm our interpretation of results in Experiment 1. Disparity matching was isolated by eliminating occlusion of the hand by the target. Monocular performance was also tested with and without occlusion to determine the role of occlusion in that case.

Experiment 1

We compared performance in actual and virtual environments in a visually guided reaching task. This task allowed the participant to see the target together with a handheld stylus during a reach, and thus relative distance information could be used both to calibrate perception of definite distance and to position the stylus during each reach. Relative distance information can be used to calibrate other definite or absolute distance information when relative distance is 0 (that is, when the hand contacts the target). At this point, the relative distance information itself becomes definite. Without physical contact, there are three potential sources of information about the relative distances of the handheld stylus and the target. The first is the relative image size of the stylus versus the target (assuming the relative actual sizes are known). The second is occlusion of the target by the stylus or of the stylus by the target. (In the virtual environment, when the target distance is exceeded by the stylus, the stylus penetrates the target surface and is occluded by it.) The third is relative disparity. This last source of information is available only to binocular vision.[1] The ability to use relative disparity may yield superior performance in targeted reaching. It is often assumed that the guidance of reaching is the ultimate function of binocular vision (see, for instance, Parker et al., 1995). Indeed, Bingham et al. (2000) found that reaching was more accurate when guided using binocular vision. Their results indicated that both shape and egocentric distance were perceived accurately. In contrast, when participants used monocular vision, shape was expanded in depth and egocentric distance was overestimated by ≈5%.

Method

Participants

Seven adults between the ages of 20 and 44 years participated in the experiment; 5 were men and 2 were women. Five were naive to the goals of the experiment and one was an author (Geoffrey P. Bingham). Six of the participants were paid $5/hr. All participants had normal or corrected-to-normal vision (using contacts) and normal motor abilities. All were right-handed.

Apparatus

Below, we describe the apparatus used in the actual and virtual environments, respectively.

Actual environment. In the actual environment, the target was a Styrofoam sphere that was 7 cm in diameter. The target was painted flat black and randomly covered with phosphorescent dots, each about 0.5 cm in diameter. As shown in the middle panel of Figure 1, the target was held in position by a rigid framework anchored to an optical bench on the floor. One end of the optical bench was positioned directly below the right eye of the seated participant. The target was mounted on the framework so as to appear unsupported in front of the participant at eye level. The framework and procedures for positioning and calibrating the apparatus are described in detail in Bingham et al. (2000). An infrared emitting diode (IRED) attached to the support structure was used to measure the position of the target relative to the hand.

An IRED also was glued to the side at the end of a cylindrical plastic stylus that was 18.5 cm long and 1 cm in diameter. The stylus was painted with phosphorescent paint. The participant held the stylus so that the end with the IRED extended 3 cm beyond the closed fist with the thumb on the stylus. A launch platform (a 7-cm cube) was located to the right of the seated participant's hip. Each trial began with the back end of the stylus inserted in a hole in the launch platform.

The target was viewed in complete darkness. The framework supporting the target was wrapped in black cloth and black curtains hung behind the target from the observer's perspective. Only the stylus and the dots on the target could be seen by the participant.

The positions of the IREDs were sampled at 100 Hz with a resolution of 0.1 cm by a two-camera WATSMART kinematic measurement system (Northern Digital Inc., Waterloo, Ontario, Canada). A WATSCOPE connected to the WATSMART recorded signals from the launch platform. A gauge figure was used to check that the measurement system itself was isotropic (Bingham et al., 2000).

Virtual environment. The virtual environment lab consisted of an SGI Octane graphics computer, a Flock of Birds (FOB; Ascension Technology Corporation, Burlington, VT) motion measurement system with two markers (for head and hand), and a Virtual Research V6 stereo HMD. We developed software that enabled us to produce displays in the HMD portraying a virtual target sphere and handheld stylus. The FOB emitter yielded a measurement volume with a 122-cm radius. The emitter was positioned at a height of 20 cm above the head of the seated participant and at a horizontal distance midway between the head and the handheld stylus at maximum reach. One marker was placed on the V6 HMD and the other on a Plexiglas stylus held in the participant's hand. The stylus was a Lucite dowel 10 cm in length and 1 cm in diameter. The 7-cm diameter virtual target sphere was dark with green phosphorescent-like dots, and it appeared against a dark background so that only the green dots could be seen. The stylus and marker were modeled precisely and appeared as a gray virtual stylus with a blue and red marker at its bottom. The hand was not modeled, so participants only saw the virtual stylus floating in the dark space. Its position and motion were the same as the actual stylus. There were no shadows cast on the target by the stylus or by the target on the stylus.

The HMD displays subtended a 60° field diagonally with complete overlap of the left and right fields. The resolution was 640 × 480, and the frame rate was 60 Hz. The weight of the helmet was 0.82 kg. The sampling rate of the FOB was 120 Hz. As described in the Appendix, we measured the focal distance to the virtual image, the image distortion, the phase lag, and the spatial calibration. The virtual image was at about a 1 m distance from the eyes. The characteristic pincushion image distortion increased in proportion to the distance from the center of the image, reaching a maximum of about 7% at the edges. The phase lag was 80 ms. The spatial calibration yielded a resolution of about 2 mm.

Procedure

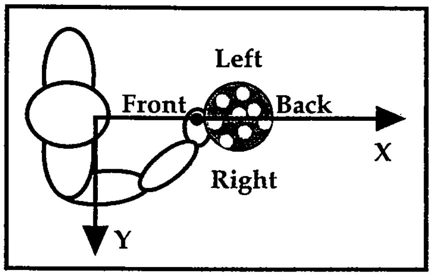
Below, we describe the procedures used in the actual and the virtual environments, respectively.

Actual environment. The procedure for testing in the actual environment was the same as described in Bingham et al.'s (2000) Experiment 3. The target was positioned at eye height. The participant's maximum reach distance was measured and target distance was computed as .70 of the participant's maximum reach. The task and procedure were explained to participants. Participants were instructed to reach to place the tip of the stylus at one of four locations relative to the surface of the target sphere as shown in the top panel of Figure 1. They reached to place the stylus at a distance of 1 cm from the surface to either the front, right, left, or back of the sphere. Participants were instructed not to touch the target and that if they were to do so inadvertently, they should report it at the end of the trial.

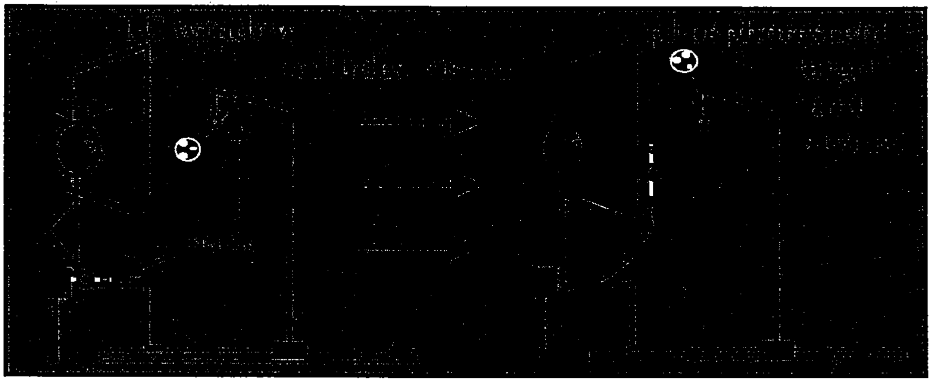
Each trial began with the participant sitting with his or her eyes closed and grasping the stylus in the launch platform. The experimenter announced the reach location, for instance, "front." The participant signaled he or she was ready by echoing this instruction. The experimenter said "start" and started WATSMART sampling. The participant opened his or her eyes and moved his or her head through about 10 cm side to side, three times at a preferred rate, while viewing the target. The participant was

---

[1] Disparity information would also be available to monocular vision if head movements were performed during the reach. This is difficult and awkward to do and is not representative of normal reaching. Nevertheless, we do consider this in Experiment 4.



In Actual Environment



In Virtual Environment

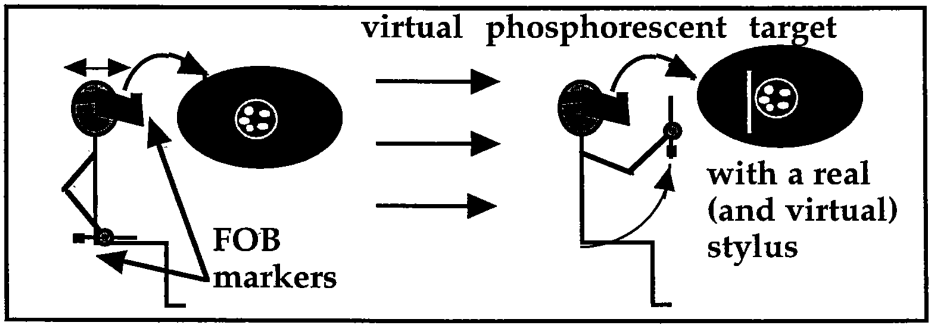


Figure 1. Top: The reaching task illustrated together with the horizontal coordinates used to measure and analyze target and reach positions. Middle: The experimental arrangement and procedure in the actual environment. The liquid crystal (LC) window was used in Experiment 2 and not in Experiment 1. Bottom: The experimental arrangement and procedure in the virtual environment. IRED = infrared emitting diode. FOB = Flock of Birds motion measurement system (Ascension Technology Corporation, Burlington, VT).

instructed to look at the targeted location on the sphere and to assess where he or she was going to reach. The participant then reached. Once the participant had completed the reach and placed the stylus, he or she said "OK," and WATSMART sampling was terminated. The participant then closed his or her eyes and placed the stylus back into the launch platform and prepared for the next trial. The phosphorescent dots were energized with a bright light every 10 trials while the participant sat with his or her eyes closed. Throughout the experiment, the participants were never able to see anything other than the stylus and the dots on the target surface.

Reaches were tested in two viewing conditions, with binocular and monocular vision, respectively. Participants wore a patch over their left eye during monocular viewing. Trials were blocked by viewing condition. Ten blocks of reaches were performed in each condition with the four locations on the target visited in a random order within each block. The order of conditions was counterbalanced across participants. The 80 trials (4 locations × 10 blocks × 2 viewing conditions) were performed in a single session. Participants took a 10–15-min break between viewing conditions during which they walked around the lab. The last 3 participants only performed three blocks of trials in the actual environment because the data from the previous participants showed no drift.

Virtual environment. In the virtual environment, participants also sat in a wooden chair. The experimenter first measured the participant's interpupillary distance using a ruler and entered the value into the software. Participants then placed the HMD on their head using two knobs to adjust the fit. The placement of the lenses in front of the eyes was then adjusted by the participant so that each lens was immediately in front of an eye. Participants then gripped the actual stylus and moved it around in front of their faces so that they were able to see the virtual stylus. We found that some participants had difficulty fusing the two stereo images if they were only allowed to view the target at this point. However, if they were allowed to view the virtual stylus as they moved it by hand, fusion occurred spontaneously and immediately. Participants were then allowed a few minutes to move their head and hand and to explore and acclimate to the virtual environment. Following this, the maximum reach distance and eyeheight were measured by having participants sit in a chair while wearing the HMD, gripping the stylus in their right fists, and holding it out as far as possible in front of their faces. The software used the measured values to position the 7-cm virtual sphere at eye height and at a distance equal to .70 of the maximum reach.

The task was explained to the participants. It was the same as that performed in the actual environment with the following exceptions (see the bottom panel of Figure 1). Between trials, participants sat with their eyes closed and holding the stylus in their lap. At the beginning of each trial, the computer announced to the participant the location to be touched on the target (e.g., front, back, left, or right). The participant then opened his or her eyes and moved his or her head and torso 10 cm side to side, 2–3 times at preferred rates, while counterrotating the head to keep the target centered in the display and to look at the targeted locus on the surface. Participants were instructed to perform this action the same way as in the actual environment. Participants then reached at preferred rates. Once the participant had reached the target, he or she said "OK," and the 3-D coordinates of the top of the stylus were recorded. The participant then closed his or her eyes, placed the stylus back in his or her lap, and the next trial was begun. Ten blocks of trials were performed as in the actual environment. Participants removed the HMD and walked around the department for 10–15 min between the monocular and binocular sessions. The order of these conditions was counterbalanced across participants. Also, the order in which the actual and virtual environments were tested was counterbalanced across participants. Actual and virtual environments were tested on different days, which were separated by a number of weeks for 2 of the participants and by a couple of days for the remaining 5 participants.

Dependent Measures

The method allowed us to evaluate a number of perceptual properties concurrently and to determine the extent to which they covary. Five dependent measures were computed for each block of four reaches. As shown in the top panel of Figure 1, we used Cartesian coordinates such that depth varied along the x-axis, and the y-axis lay in a frontoparallel plane. We computed the *egocentric distance* as the x-centroid of the four reaches. This distance was reported as a proportion of target distance (e.g., reach distance/target distance). Size, as usually studied, is an extent in the frontoparallel plane. The difference in y between reaches to the left and right yielded *width,* which was equivalent to standard measures of size. Exocentric distance or *depth* was computed as the difference in x between front and back (or twice the difference between front and the mean x of left and right). Both depth and width were reported as a proportion of the sum of target diameter and stylus diameter. *Shape* was computed as the aspect ratio of width to depth. The product of width and depth yielded 3-D size.

Design

The independent variables were environment (actual and virtual), viewing (monocular and binocular), and block (1–10). The dependent variables were egocentric distance, width, depth, shape, and 3-D size. All variables were tested within subject.

Results and Discussion

Egocentric Distance

We computed overall mean egocentric distances for each viewing condition and environment. These are shown in the top panel of Figure 2 together with standard error bars representing between-subjects variability. When reaching to actual targets, participants exhibited accuracy comparable to that found in Bingham et al. (2000). Distance was overestimated by only 1–2% when participants used binocular vision. When they used monocular vision, distance was overestimated by 4–5%. Results were somewhat different in the virtual environment. When participants used binocular vision, target distance was underestimated by 2–3%. When participants used monocular vision, distance was overestimated more than with actual targets, by about 8%. We performed a multiple regression on the egocentric distances using, as independent variables, block number and coded vectors (±1) for environment (virtual vs. actual) and viewing (monocular vs. binocular), together with vectors representing the three 2-way interactions and the single 3-way interaction. The results were significant, F(7,

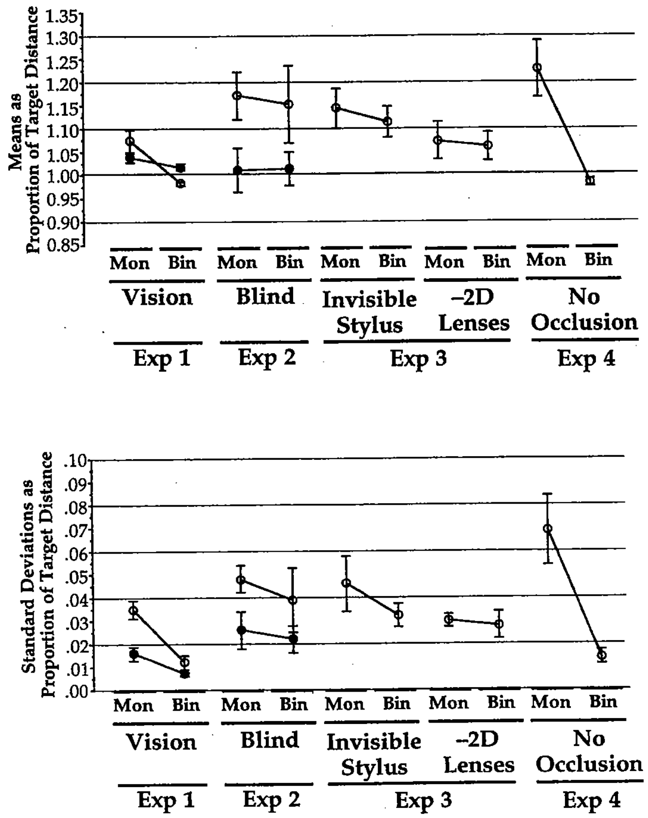


Figure 2. Top: Overall mean egocentric distances plotted for each experiment (Experiment 1: with vision; Experiment 2: blind; Experiment 3: invisible stylus and −2 diopter [−2D] lenses; Experiment 4: no occlusion), viewing condition (monocular or binocular), and environment (actual: filled circles; virtual: open circles). Standard error bars represent between-subjects variability. Bottom: Mean within-subject standard deviations in egocentric distances plotted for each experiment (Experiment 1: with vision; Experiment 2: blind; Experiment 3: invisible stylus and −2D lenses; Experiment 4: no occlusion), viewing condition (monocular or binocular), and environment (actual: filled circles; virtual: open circles). Standard error bars represent between-subjects variability. Mon = monocular; Bin = binocular; Exp = experiment.

230) = 26.4, $R^2 = .44$, $p < .001$.[2] To obtain the best measure of the contributing factors, we isolated them using a procedure described in Pedhazur (1982). We removed all nonsignificant factors in order of smallest partial F values until only significant factors remained. The resulting regression was significant, $F(2, 235) = 91.9$, $R^2 = .44$, $p < .001$. Viewing was significant, partial $F = 116.4$, $\beta = .54$, $p < .001$. Egocentric reach distances were 6% of the target distance farther with monocular vision than with binocular vision. The Environment × Viewing interaction was significant, partial $F = 38.2$, $\beta = .31$, $p < .001$. When guided using monocular vision, reaches were 4% farther in the virtual compared with actual environments. When guided using binocular vision, they were 4% less far in the virtual compared with actual environment.

Mean within-subject variability is shown in the bottom panel of Figure 2. These means reflect differences in the stability of the egocentric distance of reaches. The stability was much greater with the use of binocular vision in both the actual and the virtual environments. When participants used monocular vision, reaches were twice as variable on average in the virtual as compared with the actual environment. However, analyses in which we regressed block number on egocentric distances revealed no systematic drift in reach distances in any of these conditions.

Depth

Next, we analyzed the depths and found a pattern of results similar to that for the egocentric distances, as shown by a plot of the overall means in Figure 3. We performed a multiple regression on the depths using the same design as in the analysis on egocentric distances. The result was significant, $F(7, 227) = 27.0$, $R^2 = .45$, $p < .001$. After removal of nonsignificant factors, the result was $F(2, 232) = 93.7$, $R^2 = .45$, $p < .001$, and only viewing, partial $F = 119.2$, $\beta = .54$, $p < .001$, and the Environment × Viewing interaction, partial $F = 36.2$, $\beta = .30$, $p < .001$, were significant. Depths were overestimated by 67% more using monocular as compared with binocular vision. Using monocular vision, the depth of target objects was overestimated on average in both the actual (by almost 40%) and virtual (by 90%) environments, 50% more so in the virtual environment. In contrast, when participants used binocular vision, the depth of actual objects was only overestimated by about 8%, whereas that of virtual objects was underestimated by about 15%, yielding a 23% difference in depth between the actual and virtual environments.

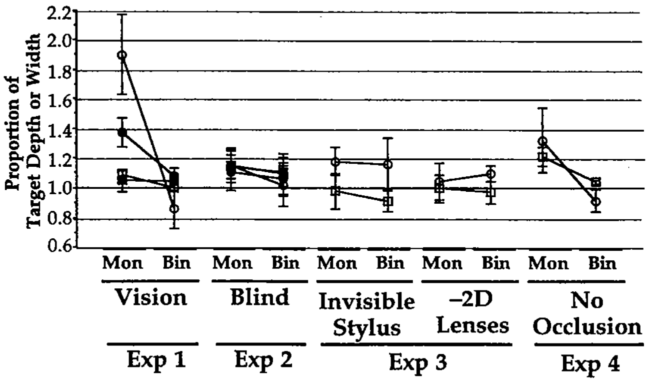


Figure 3. Overall mean depth (circles) and width (squares) plotted for each experiment (Experiment 1: with vision; Experiment 2: blind; Experiment 3: invisible stylus and −2 diopter [−2D] lenses; Experiment 4: no occlusion), viewing condition (monocular or binocular), and environment (actual: filled symbols; virtual: open symbols). Standard error bars represent between-subjects variability. Mon = monocular; Bin = binocular; Exp = experiment.

Width

Next, we examined the widths. The overall means plotted in Figure 3 revealed that widths were overestimated by 10% using monocular vision, but only by 4% or less using binocular vision. The multiple regression performed on widths was significant, $F(7, 226) = 4.4$, $R^2 = .12$, $p < .001$. After nonsignificant factors were removed, the result was $F(2, 231) = 12.3$, $R^2 = .10$, $p < .001$, and both the Block × Environment, partial $F = 7.5$, $\beta = -.17$, $p < .01$, and Block × Viewing, partial $F = 17.7$, $\beta = .26$, $p < .001$, interactions were significant. Overall, variations were on the order of 3% and were small compared with variation in depths. Drift was greatest with the monocular condition in the actual environment.

3-D Size

A multiple regression performed on 3-D size essentially reproduced the pattern of results for depths as shown in Figure 4. The analysis was significant, $F(7, 223) = 23.5$, $R^2 = .42$, $p < .001$. After removal of nonsignificant factors, the result was $F(2, 228) = 82.1$, $R^2 = .42$, $p < .001$, and both viewing, partial $F = 104.8$, $\beta = .53$, $p < .001$, and the Environment × Viewing interaction, partial $F = 28.7$, $\beta = .28$, $p < .001$, were significant. When participants used monocular vision, 3-D sizes were estimated as 72% larger than when they used binocular vision. With monocular vision, virtual objects were estimated as 60% larger than actual, whereas with binocular vision, virtual objects were estimated as 27% smaller than actual.

Shape

Finally, we examined the width:depth aspect ratios as measures of shape. The overall means plotted in the bottom panel of Figure 4 yielded a pattern expected from the analyses of depth and width. The multiple regression performed on the width:depth ratio was significant, $F(7, 223) = 15.7$, $R^2 = .33$, $p < .001$. After removal of nonsignificant factors, the result was $F(3, 227) = 37.1$, $R^2 = .33$, $p < .001$. Environment was significant, partial $F = 9.2$, $\beta = .16$, $p < .005$. Viewing was significant, partial $F = 64.0$, $\beta = -.44$, $p < .001$, as was the Environment × Viewing interaction, partial $F = 19.0$, $\beta = -.24$, $p < .001$. When viewed using monocular vision, shape was expanded in depth in both the actual (≈20%) and virtual (≈30%) environments, more so in the virtual environment. When participants used binocular vision, shape was accurate in the actual environment, but in the virtual environment, shape was compressed in depth by 31%.

---

[2] Data were lost as a function of reflections or occlusion problems with the IREDs because the target spheres were present in the measurement volume.

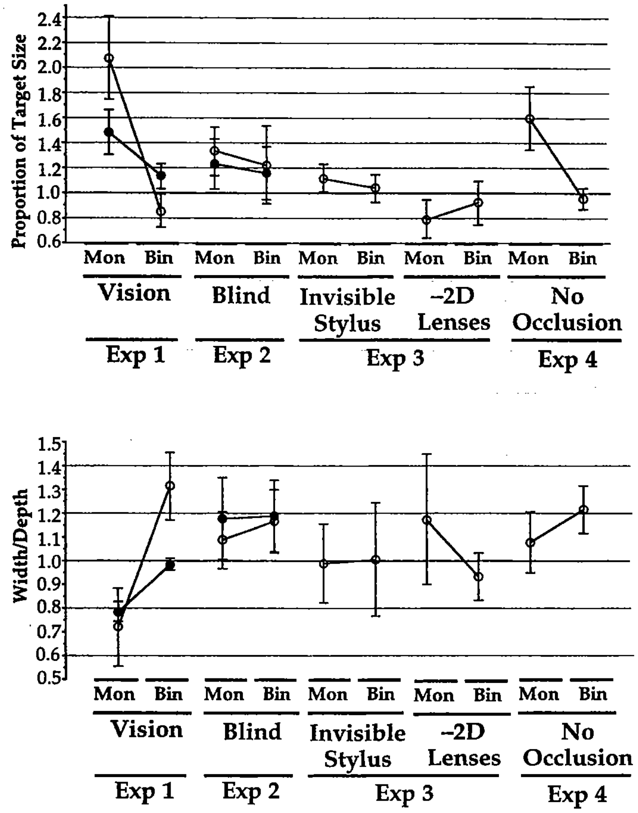


Figure 4. Top: Overall mean 3-D size plotted for each experiment (Experiment 1: with vision; Experiment 2: blind; Experiment 3: invisible stylus and −2 diopter [−2D] lenses; Experiment 4: no occlusion), viewing condition (monocular or binocular), and environment (actual: filled circles; virtual: open circles). Standard error bars represent between-subjects variability. Bottom: Overall mean shape aspect ratios plotted for each experiment (Experiment 1: with vision; Experiment 2: blind; Experiment 3: invisible stylus and −2D lenses; Experiment 4: no occlusion), viewing condition (monocular or binocular), and environment (actual: filled circles; virtual: open circles). Standard error bars represent between-subjects variability. Mon = monocular; Bin = binocular; Exp = experiment.

The finding of accurate shape when participants used binocular vision to reach to actual objects replicated the results of Bingham et al. (2000). However, we found a different result in the virtual environment. The pattern of the raw position data is illustrated in the top left panel of Figure 5 for a representative participant in the binocular, virtual environment condition. As with actual targets, reaches to virtual targets were accurate to the front, left, and right sides. However, reaches to the back of virtual targets placed the stylus inside the back half of the target, somewhat behind the center of the target. Participants were not aware that they were doing this. In debriefing, they stated that they felt their reaching was accurate when they were allowed to use binocular vision during the reach and equally so in the actual and virtual environments. The pattern of results is consistent with the use of disparity matching, that is, placing the stylus in the same depth plane as the target surface so as to be able to look at the stylus and see a single image of the target or vice versa. In reaches to the back of the target, disparity matching would use the visible occluding contour extending over the top of the spherical target with the result that the stylus would be located in the plane of this contour, that is, a little beyond the center of the sphere. This would produce underestimation of depth and egocentric distance, accurate estimation of width, underestimation of 3-D size, and compression of shape in depth. This was exactly the pattern of results. In the actual environment, the physical presence of the target would prevent this from happening.

As can be seen in Figures 2, 3, and 4, there was a relation between egocentric distance and depth and shape results. We computed means of each variable for each participant, in each environment and viewing condition, and regressed the depths, widths, 3-D sizes, and width to depth ratios on egocentric distances, each in separate simple regression including the data from the combined conditions. The regression using the depths was significant and accounted for 66% of the variance in egocentric distances, $F(1, 26) = 49.5$, $p < .001$. Similarly, the regression using the 3-D sizes accounted for 67% of the variance, $F(1, 26) = 53.4$, $p < .001$. The regression using the width to depth ratios accounted for 48% of the variance, $F(1, 26) = 23.5$, $p < .001$. The regression using the widths was significant but only accounted for 16% of the variance in egocentric distances, $F(1, 26) = 5.0$, $p < .05$. The covariation of egocentric distance and shape (and 3-D size or depth) in these conditions is consistent with the findings in Bingham et al. (2000). As shown in Figures 2 and 4, the effect was amplified in the virtual environment simply because egocentric distances and depths were overestimated more when participants used monocular vision, whereas when they used binocular vision, depths and thus egocentric distances were compressed. Reaches to the front of the target were accurate in all cases, that is, when participants used binocular or monocular vision in the actual or the virtual environment. However, difficulties arose in positioning the stylus to the back of the target. When participants used monocular vision, the distance at the back was overestimated with the result, given the accuracy to the front, that both egocentric distance and depth (and shape and 3-D size) were overestimated. This distance was more strongly overestimated in the virtual environment compared with the actual.

In summary, we directly compared performance in virtual and actual environments. We found two differences. First, when participants used binocular vision, we found that shape was accurate in the actual environment but compressed in depth in the virtual environment. We attributed this difference to the use of disparity matching without the constraint imposed by the physical presence of the target. Second, when participants used monocular vision, we found that shape was expanded in depth in both environments, but more strongly so in the virtual environment. Next, we considered why this overestimation was greater in the virtual environment.

Experiment 2

We performed the comparison between environments using a feedforward reaching task. The significance of this task is that information about the relative distances of hand and target can no longer be used to guide the hand progressively to the target. Instead, information about the definite distance, size, and shape of the target must be obtained before a reach and used to guide it. The problem is that purely optical information is angular and temporal and therefore cannot specify definite metric distances and sizes

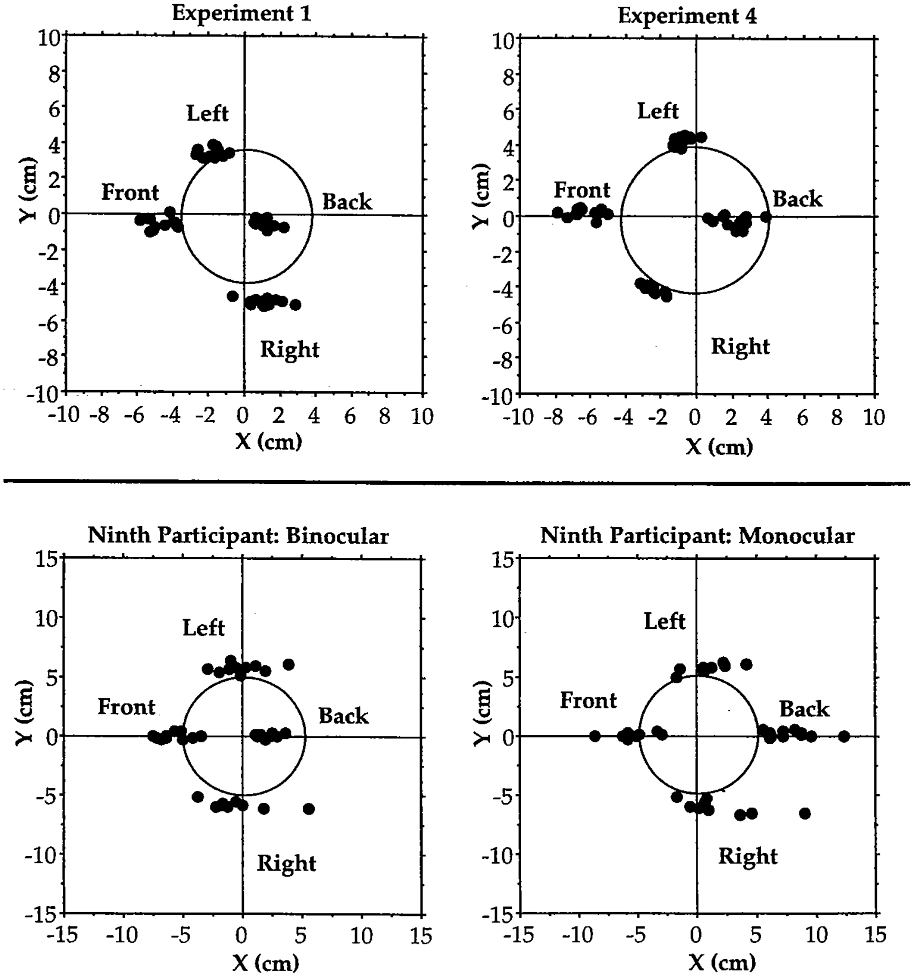


Figure 5. Representative position data from the binocular condition of Experiment 1 (top left) and Experiment 4 (top right). The binocular (bottom left) and monocular (bottom right) position data for the 9th participant in Experiment 4 are shown for comparison. All plots show positions of the stylus in the horizontal *x*–*y* plane. The circle in the center of each plot represents the 7-cm diameter target sphere.

(e.g., Faugeras, 1993; Ullman, 1979). By themselves, binocular disparity, optical flow (including motion parallax), and texture gradients can only specify relative distances and sizes. However, optical information can specify definite metric distance and shape when coupled with extraoptical information, at least in theory (Bingham & Pagano, 1998; Bingham & Stassen, 1994; Eriksson, 1974; Gogel, 1977; Gogel & Tietz, 1979; Johansson, 1973; Ono & Steinbach, 1990; Rogers & Collett, 1989; Rogers & Graham, 1979; Sedgwick, 1986; Steinbach & Ono, 1991; Wickelgren et al., 2000). In principle, definite distance is obtainable from vergence in binocular vision using the stable distance between the two eyes (Howard & Rogers, 1995; Rogers & Rogers, 1992). Observers have been found to use vergence as information about egocentric distance within reaching distance (Brenner & Van Damme, 1998; Owens & Liebowitz, 1980; Swenson, 1932; Tresilian, Mon-Williams, & Kelly, 1999), although results indicate that vergence specifies distance both relative to an adjustable resting vergence level (Owens & Liebowitz, 1980) and relative to the distance of the target last fixated (Brenner & Van Damme, 1998). As distance increases, vergence becomes increasingly unreliable as distance information, and it is likely that vergence information is less effective with increasing distance (Tresilian et al., 1999). Accommodation has been tested monocularly and has been found to yield only ordinal information about reachable distances, that is, whether the distance is greater or less than that of the target just accommodated (Mon-Williams & Tresilian, 1999a). Comparison of averaged judgments to actual distances yields very low slopes (Fisher & Ciuffreda, 1988). It is likely that accommodation is effective through its influence on vergence even in monocular conditions. When accommodation and vergence distances are separated, the normal coupling of accommodation and vergence acts to pull vergence in the direction of the accommodation, with over- or underestimations of distance accordingly (Mon-Williams & Tresilian, 1999b; Peli, 1999; Swenson, 1932; Wann et al., 1995).

A virtual environment entails decoupling of accommodation and vergence. As described in the Appendix, we performed a series of studies to reveal the specific properties of our virtual environment lab. We measured the image distortion, the phase lag, the spatial calibration, and the focal distance to the virtual image. We found that the virtual image in the HMD is at about a 1-m distance from the eyes. As such, it is somewhat beyond reachable targets. In the virtual environment, therefore, we expect overestimation of distance. Participants have to accommodate to the focal distance of the image at 1 m while attempting to converge at the distance of the virtual target sphere at a distance (≈30–40 cm) that is significantly less than 1 m. This should produce convergence to a distance somewhat beyond the target and result in overestimation of distance. This may well have been the source of the greater overestimation of distance found when participants used monocular vision in the virtual environment in Experiment 1.

However, we designed our experiments to be representative of normal, visually guided reaching by requiring observers to move their heads while viewing the targets prior to each reach. Monocular optic flow (i.e., motion parallax and radial expansion) from head movement can, in principle, yield absolute distance and shape when combined with efferent or kinesthetic information about the amplitude or velocity of self-motion (Bingham & Stassen, 1994; Nakayama & Loomis, 1974). Nevertheless, performance levels have once again been found to be less than optimal (Bingham & Pagano, 1998; Bingham et al., 2000; Gogel & Tietz, 1979; Ono & Steinbach, 1990; Wickelgren et al., 2000). Slopes are less than 1 (≈.7) when judged or reached distances are plotted as a function of actual target distances. As in Experiment 1, we studied in Experiment 2 the use of both monocular and binocular information generated by moving observers. In the virtual environment, overreaching might be expected as a result of the need to accommodate to an image beyond reach space. The question was whether overestimation would occur in the context of a representative configuration of information about distance (including motion parallax generated by voluntary self-motion).

Method

The participants were the same as in Experiment 1. The data for Experiment 2 were collected before those for Experiment 1 in sessions that occurred in different days. Participants received no feedback about their performance in Experiment 2 as they did in Experiment 1. The methods in Experiment 2 were the same as in Experiment 1 with the following exceptions. First, in the actual environment, we added a liquid crystal (LC) window to the apparatus as shown in the middle panel of Figure 1. Previous studies have shown that a delay of 4–5 s between looking and blind reaching can increase errors (Elliott & Madalena, 1987; Graham, Bradshaw, & Davis, 1997a, 1997b; 1998). To control this and to minimize the time between occlusion of vision and measurement of the endpoint of each reach, we hung a large 60 cm high × 150 cm wide LC window in front of the participant's face at a distance of 4–5 cm. The window became opaque 5 ms after the back of the stylus left a launch platform at the hip. Thus, only about 1 s elapsed between loss of vision of the target and measurement of the reach endpoint.

Participants were instructed to reach to place the stylus at the locus of the target surface either to front, back, left, or right. As in Experiment 1, Experiment 2 was performed in the dark. As shown in the middle panel of Figure 1, each trial began with the participant grasping the stylus in the launch platform and the target lowered into position for viewing, but with the LC window opaque and occluding the target. The experimenter announced the reach location, for instance, "front." The participant signaled that he or she was ready by echoing this instruction. The experimenter said "start," started WATSMART sampling, and caused the LC window to become transparent. The participant moved his or her head through about 10 cm side to side along the window surface, three times at a preferred rate, while viewing the target. The participant then reached. Removal of the stylus from the launch platform caused the LC window to become opaque and lit a light that signaled a second experimenter to pull a string that raised the target out of the way. Once the participant had completed the reach and placed the stylus, he or she said "OK," and WATSMART sampling was terminated. The participant then placed the stylus back into the launch platform and prepared for the next trial. The phosphorescent dots were energized with a bright light every 10 trials while the participant sat with his or her eyes closed. Throughout the experiment, the participants were never able to see anything other than the dots on the target surface.

In the virtual environment, the only difference in the procedure was that the experimenter turned off the display as the reach was initiated, that is, as the participant began to raise his or her hand and the stylus from his or her lap. The display was turned on again at the beginning of each trial.

Results and Discussion

Egocentric Distance

We computed overall mean egocentric distances for each viewing condition and environment. These are shown in the top panel of Figure 2 together with standard error bars representing between-subjects variability. Participants were more accurate on average when reaching to actual targets than reported in Bingham et al. (2000), in which they were found to overestimate distance by 5–10% on average. However, the results were the same as those obtained by Wickelgren et al. (2000), who also reported substantial individual differences. The error bars shown in the top panel of Figure 2 are quite large. To illustrate the pattern of performance, we computed a mean egocentric distance for each participant in each viewing condition and environment and plotted these together with standard error bars in Figure 6. Some of the participants undershot the target, some overshot, and some were fairly accurate. Despite these variations, the relative overestimation of distance by about 12% in the virtual as compared with the actual

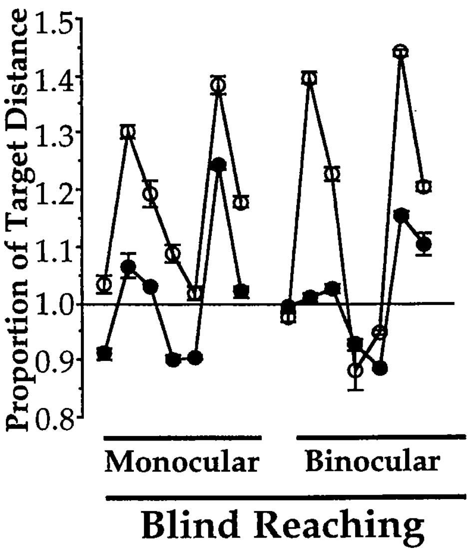


Figure 6. Mean egocentric distances (and standard error bars) plotted for each participant in Experiment 2 (blind reaching) by viewing conditions (monocular or binocular) and environment (actual: filled circles; virtual: open circles).

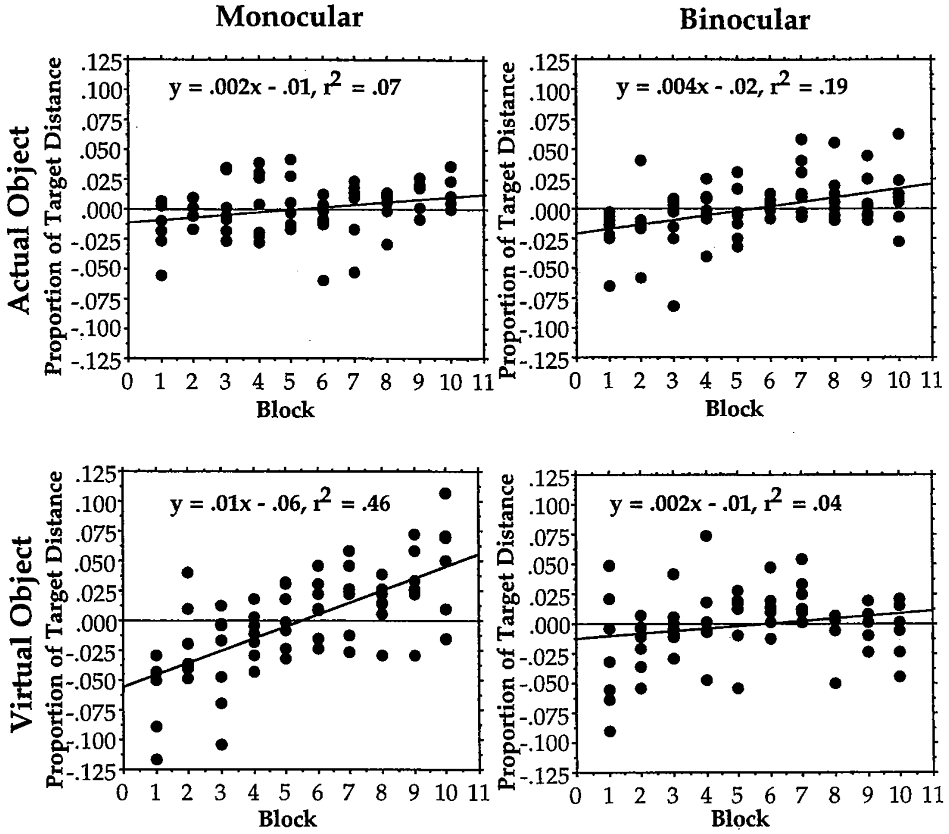


Figure 7. Normalized egocentric distances for 7 participants plotted by block of trials in each of two viewing conditions and two environment conditions. Lines fitted by least squares regression are also shown together with the corresponding equations and r^2 values.

environment was fairly consistent across participants and viewing conditions. We again performed a multiple regression on the egocentric distances using block number and coded vectors (±1) for environment and viewing, together with vectors representing the three 2-way interactions and the single 3-way interaction. The result was significant, $F(7, 272) = 9.9$, $R^2 = .20$, $p < .001$, but only the environment factor was significant, partial $F = 10.6$, $p < .002$. The mean difference in egocentric reach distance as a proportion of target distance was 12.4% between the actual and virtual environments.

Next, we examined the stability of performance over blocks. As shown in the bottom panel of Figure 2, within-subject variability was higher in Experiment 2 than in Experiment 1. Because of the strong individual differences, we normalized the data for each participant. We subtracted each participant's mean from his or her data and then performed simple regressions regressing block number on normed egocentric distances, as shown in Figure 7, for each of the environments and viewing conditions. A general tendency to drift outward in depth over blocks was apparent. Significant drift was found in three of the four cases,[3] with a trend ($p = .08$) present in the fourth case, that is, binocular viewing in the virtual environment. This was consistent with previous results (Bingham et al., 2000; Vindras & Viviani, 1998). Without feedback and the ability to calibrate, perception of egocentric distance is unstable. However, when participants used monocular vision, the instability was greater in the virtual environment as compared with the actual. Reaches drifted a total of 10% of the target distance over the 10 blocks in the virtual as compared with 2% in the actual environment. Drift in the virtual environment was less when participants used binocular vision.

Next, we investigated whether the differences found in egocentric distance would be found in the other measures as well. Do size and shape covary with distance?

Depth and Width

We performed multiple regressions on depths and then on widths, in each case using the same design as we did when analyzing the egocentric distances. Neither result was significant: depth, $F(7, 269) = 1.5$, $R^2 = .04$, $p > .05$; width, $F(7, 269) = 0.7$, $R^2 = .02$, $p > .05$. As shown in Figure 3, both depth and width were overestimated by about 12% on average in all conditions.

3-D Size

We performed a multiple regression on 3-D size and obtained similar results, as shown in the top panel of Figure 4. The regres-

---

[3] Nine outliers (at a distance of two standard deviations from the mean) were removed before these analyses were performed.

sion failed to reach significance, $F(7, 267) = 1.1$, $R^2 = .03$, $p > .05$. In all conditions, 3-D size was overestimated by 20–30% on average (as expected given $1.12^2 = 1.25$).

Shape

Finally, we performed a multiple regression on the width to depth aspect ratios, and the results were also not significant, $F(7, 262) = 0.8$, $R^2 = .02$, $p > .05$. On average, as shown in the bottom panel of Figure 4, shape was compressed in depth (or expanded in width) by 20% in the actual environment and 13% in the virtual environment. However, there was large variability in both cases, so that one-tailed, one-sample *t* tests performed on the subject means, comparing them to 1, showed only a trend in both cases, $t(13) = 1.7$, $p < .06$, and $t(13) = 1.5$, $p < .08$, for the actual and virtual environments, respectively.

Overall, as shown in Figures 2, 3, and 4, the finding was of no difference between performance in the actual and virtual environments except in the respect that had been predicted, namely, overestimation of egocentric distances in the virtual as compared with the actual environment. Accompanying this latter result, egocentric distances were less stable and subject to greater drift in the virtual environment when viewed with dynamic monocular vision as shown in the bottom panel of Figure 2 and Figure 7.

Width (size), depth (exocentric distance), 3-D size, and shape did not follow the pattern of results for egocentric distance, implying that these properties vary independently. This is consistent with the findings of Bingham et al. (2000). We computed means of each variable for each participant in each environment and viewing condition and regressed the depths, widths, 3-D sizes, and width to depth ratios on egocentric distances, each in separate simple regression including the data from the combined conditions, and then together in a multiple regression. None of these regressions reached significance ($p > .05$ in all cases).

Experiment 3

As we had predicted, we found in Experiments 1 and 2 that performance in the virtual environment as compared with an actual environment yielded relative overestimation of egocentric distances. We had hypothesized that this would occur as a result of the decoupling of accommodation and convergence and accommodation to a virtual image that lay at a distance beyond reach space. Next, we tested this hypothesis directly. We used −2 diopter (−2D) lenses to reduce the accommodative distance from 1 m to .33 m in the virtual environment. The resulting focal distance was approximately equal to the virtual target distance in each case. Accordingly, the predicted result was that the amount by which egocentric distances are overestimated should be strongly reduced and comparable with results in the actual environment. We did not predict that reaches would be entirely accurate on average because we had found that egocentric distances were overestimated in actual environments.

We normally would have used feedforward reaching to perform this test. However, the experiment was to be performed entirely in the virtual environment, which presented another possibility. We had found performance with feedforward reaching to be quite variable. The variability might be reduced if participants were able to see the target throughout each reach. However, vision of the stylus would have to be prevented to exclude disparity matching or other forms of feedback control using information about contact with the target. The virtual environment allowed us to do this. Participants were able to see the target sphere both before and throughout each reach, but they did not see a virtual stylus. The stylus simply was not drawn in the display. Participants performed reaches while viewing the target both with and without a pair of −2D glasses. Performance was compared between the two conditions. Performance was also compared with that in Experiment 2.

Method

Participants

Six adults between the ages of 20 and 44 years participated in the experiment. The 44-year-old was pre-presbyopic. Four were men and 2 were women. Five were naive about the goals of the experiment and had not participated in any of the earlier experiments. These participants were paid $5/hr. One participant was an author (Geoffrey P. Bingham). All participants had normal or corrected-to-normal vision (using contacts) and normal motor abilities. All were right-handed.

Procedure

The methods were the same as in the virtual environment portion of Experiment 1 with the following changes. As in Experiment 1, participants were able to see the target sphere both before and during each reach. However, unlike in Experiment 1, participants did not see a virtual stylus. Rather, they reached with an invisible stylus and hand. Ten blocks of reaches were performed by each participant in each of two viewing conditions, monocular and binocular. Both viewing conditions were tested in two additional conditions. In one condition, called *invisible stylus,* the target was viewed simply in the HMD. In the second condition, called *−2D lenses,* the participant wore −2D glasses inside the HMD. (The lenses inside the HMD were adjustable to allow room for the user to wear glasses.) The order of all conditions was counterbalanced across participants.

Results and Discussion

The overall mean egocentric distances are shown in the top panel of Figure 2. Without the glasses, the binocular mean was 1.11 and the monocular mean was 1.14. With −2D glasses, the overshoot in each case was reduced by half. The binocular mean was 1.06 and the monocular mean was 1.07. We performed a multiple regression on the egocentric distances using block number and coded vectors (±1) for viewing (binocular and monocular) and glasses (without and with), together with vectors representing the three 2-way interactions and the single 3-way interaction. The results were significant, $F(7, 232) = 4.4$, $R^2 = .12$, $p < .001$, but the only significant factor was glasses, partial $F = 6.9$, $p < .01$. When all nonsignificant factors were removed, only glasses remained, $F(1, 238) = 26.6$, $R^2 = .10$, $p < .001$.

Multiple regressions performed on depth ($R^2 = .03$), width ($R^2 = .02$), and on the width to depth aspect ratio ($R^2 = .02$) failed to reach significance. A multiple regression on 3-D size was significant, $F(7, 227) = 2.7$, $R^2 = .08$, $p < .02$. With nonsignificant factors removed, only glasses remained, $F(1, 233) = 13.7$, $R^2 = .06$, $p < .001$. Without glasses, the size was overestimated by 11%, and with glasses, it was underestimated by 11%, as shown in the top panel of Figure 4.

We performed a multiple regression comparing egocentric distances in the virtual environment condition of Experiment 2 (feedforward reaching) with those in the current experiment (invisible stylus but visible target) without the glasses. The result failed to reach significance, $R^2 < .01$, $p > .05$. We computed means and standard deviations for each participant in the two experiments and tested the differences using two-tailed, unpaired *t* tests. These tests failed to reach significance ($p > .05$). The overall means are comparable, as shown in the top panel of Figure 2. We performed multiple regressions comparing depth, width, 3-D size, and the width to depth aspect ratio. Only the analysis on width, $F(7, 251) = 3.2$, $R^2 = .08$, $p < .001$, was significant. After removal of nonsignificant factors ($R^2 = .08$), only the experiment condition was significant. With feedforward reaching, the width was overestimated by 10–15%, whereas with vision of the virtual target (but not the virtual stylus), width was fairly accurate.

In summary, the only difference found between reaches performed without vision during the reach versus with vision of only the virtual target was that the width of the target was approximated more accurately. Apparently, the ability to look at the target while positioning the stylus at the occluding contours facilitated performance.

We predicted that when participants wore −2D glasses, the overestimation of target distances would be reduced, and indeed, it was reduced by half. The result supported the hypothesis that the relative overestimation of egocentric distance found in Experiment 2 was produced by the effect of accommodation to the virtual image in the HMD. We infer from the results in the literature (Mon-Williams & Tresilian, 1999b; Swenson, 1932; Wann et al., 1995) that accommodation pulled the vergence in the direction of the virtual image yielding increases in perceived egocentric distance. The result was the same when participants used binocular or monocular vision. The latter was also expected given past findings on the role of vergence in monocular distance perception, known as *vergence micropsia* (Leibowitz, 1966, 1972), or in our case, *macropsia.*

Experiment 4

Next, we used the virtual environment lab to investigate the relatively accurate reaching performance found in Experiment 1 when participants used binocular vision. The question was whether this performance might be attributed to the use of disparity matching as we suggested. Mon-Williams and Dijkerman (1999) have speculated that such a disparity nulling strategy might be used when guiding a reach with binocular vision. We have found that binocular vision does not yield superior reaching accuracy under conditions in which relative distance information is unavailable and disparity matching is not possible. Performance was equivalent in Experiments 2 and 3 and in the comparable studies in Bingham et al. (2000) when participants used monocular or binocular vision. Performance was different and superior when participants used binocular vision in Experiment 1 in which disparity matching was possible. However, other relative distance information was also available, namely, occlusion of the stylus by the target surface. In Experiment 4, we eliminated this as a potential source of information while preserving the possible use of disparity matching. The relative size of the visible stylus (which was the same as the actual stylus held in the hand) was also available in both monocular and binocular viewing conditions. The monocular form of disparity matching, that is, parallax matching, was not available with monocular viewing because participants moved their heads to obtain optic flow information only before and not during each reach. Moving the head while attempting to perform an accurate reach is very awkward and unnatural. It perturbs the torso, which is the base for purposes of postural control, and thus, it tends to perturb the hand position. Nevertheless, for comparison, in the present experiment we tested a single participant in the monocular condition who was instructed to move his head while reaching.

Note that this is an experiment that could only be performed in the virtual environment. We know from Experiments 1–3 that egocentric distances are relatively overestimated because of the need to accommodate to the virtual image distance in the virtual environment. However, we have seen that visual feedback control overcame this perturbation and yielded fairly accurate performance, comparable with that found in an actual environment under similar conditions. The question we investigate here is the extent to which that performance can be attributed to disparity matching in the case of binocular vision and to visual information about contact versus relative distance information in the case of monocular vision.

Method

Participants

Eight adults between the ages of 20 and 44 years participated in the experiment. Four were men and 4 were women. Seven were naive about the goals of the experiment and one was an author (Geoffrey P. Bingham). Four of the participants (including the author) had participated in Experiments 1 and 2, whereas 4 had not participated in the previous experiments. A 9th male participant (30 years old) performed in the monocular condition with head movement during the reach. He had not participated in the previous experiments. Eight of the participants (excluding the author) were paid $5/hr. All participants had normal or corrected-to-normal vision (using contacts) and normal motor abilities. All were right-handed.

Procedure

The procedure was the same as the virtual environment condition of Experiment 1 with a single exception. In reaches to the front, left, or right of the target, the virtual stylus could not be occluded by the target sphere. It remained visible, occluding the sphere even when held centered on the sphere from the perspective of the participant and far beyond the distance of the sphere. The fact that this would occur was explained to the participants, as was the fact that the sphere could occlude the stylus during reaches targeted to the back of the sphere.

The ninth participant was given instructions to move his head while making the reach as well as before the reach in the monocular viewing condition. Head movements were performed before the reach by moving the head and torso about the waist just as done by all participants in these experiments. The head movement during the reach was performed by only moving the head. This occurred as the natural preference of the participant, not by instruction. No head movements were performed by this participant during the reaches in the binocular viewing condition.

Results and Discussion

Egocentric Distance

As shown in the top panel of Figure 2, for 8 participants, reaches in the monocular condition overshot the target distance by 23% on average, whereas in the binocular condition, reaches undershot the target by 2%. The former result did not replicate the result for monocular vision in Experiment 1, and this shows the importance of occlusion as visual information about surface contact in that case. The latter result did replicate the result for binocular vision in Experiment 1 and shows that participants had indeed used disparity matching to achieve that result. We performed a multiple regression on the egocentric distances using block number and a coded vector (± 1) for viewing (binocular vs. monocular) and the two-way interaction.[4] The results were significant, $F(3, 135) = 50.0$, $R^2 = .53$, $p < .001$, but only viewing was significant, partial $F = 17.0$, $p < .001$.

Because the individual differences were large, we normalized the egocentric distances by computing a mean for each participant in each viewing condition and subtracting the mean from each participant's data. We performed a multiple regression on the normalized egocentric distances. The results were significant, $F(3, 135) = 10.5$, $R^2 = .19$, $p < .001$. Block was significant, partial $F = 16.7$, $\beta = .32$, $p < .001$, as was the Viewing × Block interaction, partial $F = 14.2$, $\beta = -.63$, $p < .001$. As shown in the top left panel of Figure 8 reaches drifted away over blocks in the monocular condition at a rate of 1.2% per block, yielding a total drift equal to 12% of the target distance over the 10 blocks of the experimental session. As shown in the top right panel of Figure 8, in contrast, reaches in the binocular condition were absolutely stable. This difference in stability is also reflected in the mean within-subject standard deviations plotted in the bottom panel of Figure 2. Examination of the top and bottom panels of Figure 2 shows that performance using binocular vision was comparable in Experiments 1 and 4, whereas performance using monocular vision was relatively destabilized in Experiment 4.

Depth

We performed a multiple regression on depths and the result was significant, $F(3, 135) = 10.3$, $R^2 = .19$, $p < .001$. The only significant factor was viewing, partial $F = 13.8$, $\beta = -.62$, $p < .001$. As shown in Figure 3, depths were on average overestimated by 32% in the monocular condition and underestimated by 9% in the binocular condition. A multiple regression on normed depths was significant, $F(3, 135) = 5.1$, $R^2 = .10$, $p < .001$. Both block, partial $F = 5.9$, $\beta = -.20$, $p < .02$, and Viewing × Block, partial $F = 9.0$, $\beta = .53$, $p < .01$, were significant. As shown in the middle panel of Figure 8, depth decreased over blocks by 4% per block in the monocular condition, but remained steady in the binocular condition.

Width

The results were similar for width. The multiple regression was significant, $F(3, 135) = 20.3$, $R^2 = .31$, $p < .001$, but only the viewing factor was significant, partial $F = 5.0$, $\beta = -.34$, $p < .03$. As shown in Figure 3, widths were overestimated by 22% on average in the monocular condition, but only by 4% in the binocular condition. A multiple regression on normed width was significant, $F(3, 135) = 3.3$, $R^2 = .07$, $p < .03$, and both block, partial $F = 4.9$, $\beta = .19$, $p < .02$, and Viewing × Block, partial $F = 4.7$, $\beta = -.39$, $p < .04$, were significant. As shown in the bottom panel of Figure 8, width increased over blocks by a total of 10% in the monocular condition but values were steady in the binocular condition.

Shape

A multiple regression performed on the width to depth aspect ratio was significant, $F(3, 134) = 20.3$, $R^2 = .10$, $p < .005$. Both viewing, partial $F = 11.5$, $p < .001$, and the Block × Viewing interaction, partial $F = 6.9$, $p < .01$, were significant. As shown in the bottom panel of Figure 4, shape was compressed in depth by only 7% in the monocular condition and by 22% in the binocular condition. Shape in the monocular condition was drifting into greater compression at the rate of 4.5% per block, whereas in the binocular condition, the drift was toward less compression at the rate of 1.5% per block.

3-D Size

Finally, a multiple regression performed on the 3-D size ratio was significant, $F(3, 135) = 15.2$, $R^2 = .25$, $p < .001$. Only viewing, partial $F = 14.7$, $p < .001$, was significant. As shown in the top panel of Figure 4, size was overestimated in the monocular condition by 60%, on average, whereas in the binocular condition, size was underestimated by 5%.

The most obvious difference in performance between the monocular and binocular conditions in this experiment was in stability. Binocular performance was very stable, whereas monocular reaches were highly unstable. The difference is captured in the bottom panel of Figure 2 and in Figure 8. All five variables drifted over blocks in the monocular condition. Egocentric distance increased; so did width. Depth decreased. Shape compression increased; size increased. Egocentric distance and size were strongly overestimated, and shape was weakly compressed in depth. Removing monocular information about contact with the target completely destabilized performance when participants used monocular vision, as revealed by the comparison between this experiment and Experiment 1.

In contrast, the performance in the binocular condition was the same in the two experiments. All five variables were stable and fairly accurate. Both egocentric distance and size were slightly underestimated, whereas shape was strongly compressed in depth. These results were related. The shape compression was produced by compression of depth with accurate width. The pattern of the raw data is shown for a representative participant in the top right panel of Figure 5, together with a representative participant from Experiment 1 shown in the top left panel. The pattern is the same, and it reflects disparity matching. The shape compression was produced by matching the stylus to the occluding contour of the sphere when reaching to the back. This shape compression aside, the binocular condition is the only one in which we see performance that is consistently accurate, stable, and precise, both in the actual and in the virtual environment. The implication is that disparity matching is a very important component of visually guided reaching and constitutes the true advantage of binocular vision.

---

[4] The data of 1 participant were excluded from these analyses and are discussed subsequently.

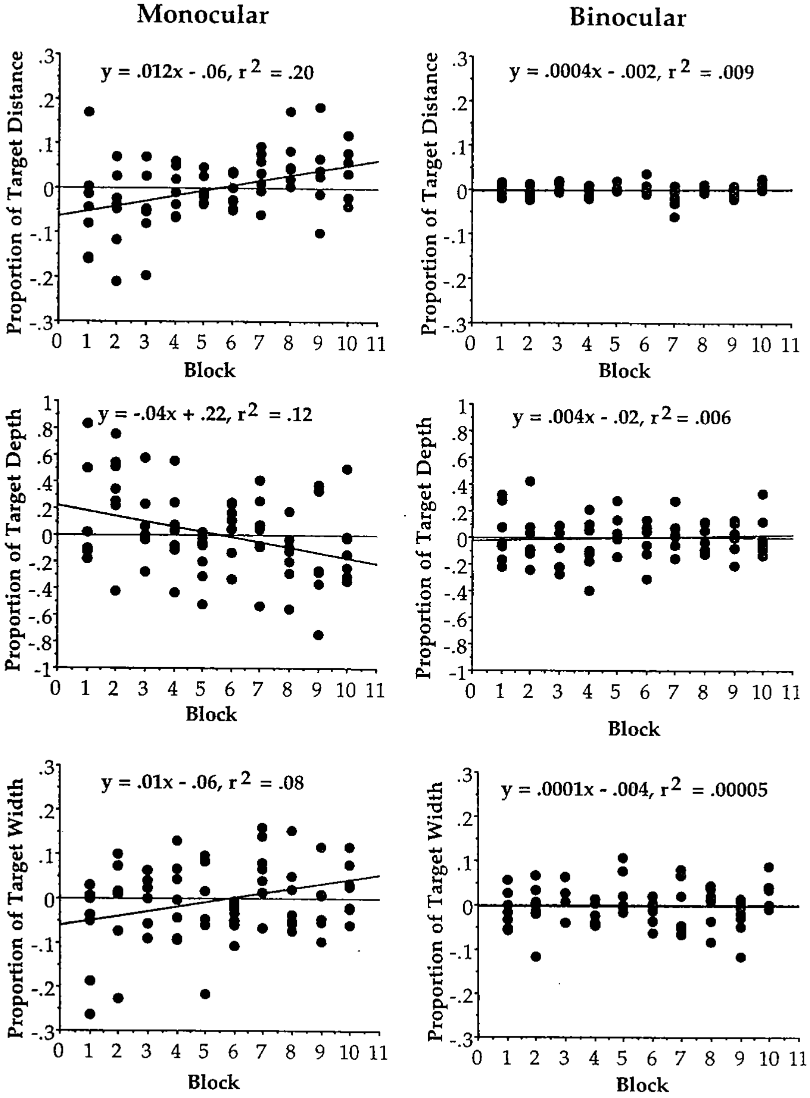


Figure 8. Normalized egocentric distances, depths, and widths for 7 participants plotted by block of trials in each of two viewing conditions in Experiment 4: no occlusion. Lines fitted by least squares regression are also shown together with the corresponding equations and r^2 values.

In contrast, participants were much more inaccurate when they used monocular vision. They overestimated the target distance and strongly expanded the shape in depth. However, we tested a 9th participant who performed the same task in the binocular condition as had the other participants, but who performed head movements during the reach in the monocular condition, unlike the other participants. The raw position data are shown for this participant in the lower panel of Figure 5, where it can be seen that his performance in the binocular condition was the same as found in Experiment 1 and for the remaining participants in the current experiment. His performance in the monocular condition was quite different, however. Initially he overreached, just as the other participants, but then his reaches quickly adjusted to become accurate. In fact, his performance in the monocular condition was more accurate than in the binocular condition because he failed to match to the occluding contour when reaching to the back.

General Discussion

We investigated a variety of sources of visual information about both definite and relative scale of egocentric distance, size, and shape. We compared performance in an actual environment with that in a virtual environment.

Definite Distance, Size, and Shape Perception

We measured feedforward reaching to investigate perception of definite distance, size, and shape. In Experiment 2, first we tested perception in an actual environment. We found overestimation of egocentric distance on average (only ≈3%) with large individual differences. Participants varied from underestimation by 10% (≈4 cm) to overestimation by 20% (≈8 cm). We also found shape to be compressed in depth by ≈20% on average, but with individual differences. Four participants exhibited compression and three exhibited expansion. Finally, our results replicated the instability found by Bingham et al. (2000), in that egocentric distances of reaches drifted away over blocks at a constant rate. Reaches drifted outward at a steady rate, moving a total of ≈3% of the target distance (≈2 cm). These results confirm the hypothesis of Bingham and Pagano (1998) that calibration is required for the accurate and stable perception of definite distance.

To isolate and manipulate calibration information, we used a virtual environment lab. However, this lab entailed other accessory perturbations of vision, so we performed a direct comparison of performance in the actual and virtual environments. The virtual environment entails accommodation to a virtual image at a focal distance beyond reach. It also requires the decoupling of accommodation from vergence because the eyes converge at the target distance. Accordingly, we predicted relative overestimation of egocentric distance in the virtual environment. The reason was that previous studies on accommodation and vergence show that vergence should be pulled in the direction of the virtual image, that is, beyond the target distance with resulting overestimation of egocentric distance. Relative overestimation is exactly what we found. Egocentric distances of reaches were 15% of the target distance farther (≈6 cm) in the virtual than in the actual environment. This occurred in both monocular and binocular viewing conditions. In addition, the rate of outward drift was tripled with use of monocular vision in the virtual environment, to yield a total drift of 10% (≈4 cm). The increased rate of drift did not occur when participants used binocular vision.

Next, to control for potential effects of memory and delay, we used the virtual environment to preserve continuous perception of the target during the reaches while eliminating relative distance and feedback information. Participants viewed a virtual target but not a virtual stylus (or hand). However, the results were the same as the previous results. Therefore, the distortions and instability were not a product of memory use, but simply reflected uncalibrated perception.

Next, we directly tested the effect of the perturbation to accommodation and vergence in the virtual environment. Participants performed reaches to a visible target with an invisible stylus while wearing −2D glasses in the HMD. The glasses reduced the focal distance to the target distance, and therefore the prediction was that reaches should overestimate egocentric distance by significantly less than they had without the glasses. The results confirmed the prediction. Overreaching was reduced by half, from ≈14% of the target distance to only ≈7%. This simultaneously confirmed our analysis of the perturbing effect of the virtual environment and demonstrated the importance of accommodation and vergence in guiding reaches.

Finally, we found significant individual differences in perceived distance. These might be attributed to individual differences found in the rest posture of vergence under low illumination (e.g., Owens & Liebowitz, 1980; see also Peli, 1999). Rest vergence varies in individuals between 50 cm and infinity with a mean of about 100 cm. Especially for monocular vision in low illumination, vergence is biased toward the rest posture. This would have occurred in both the actual and virtual environments, and therefore cannot account for the overshoot found in the virtual as compared with the actual environment.

Use of Relative Distance Information

We investigated reaching when participants were allowed vision of both target and stylus while they reached. We tested this in an actual environment, and we found that egocentric distances were overreached by 6% when participants used monocular vision, but reaches were accurate when participants used binocular vision. In the binocular condition, reaches were also highly precise and stable, whereas in the monocular condition, they were more variable. In the monocular condition, shape was expanded in depth by over 40% and 3-D size was overestimated by ≈50%. In the binocular condition, shape and size were accurate and precise.

We found performance to be the same in the virtual environment with two exceptions. First, there was a stronger tendency in the monocular condition to overestimate depth, egocentric distance, and 3-D size. Second, in the binocular condition, reaches to the back of the target were inaccurate (although precise). Participants placed the stylus somewhat beyond the center of the target instead of at the back. The result was that both egocentric distance and 3-D size were somewhat underestimated, whereas shape was strongly compressed in depth.

Next, we investigated whether the superior binocular performance could be attributed to disparity matching. We isolated disparity matching by eliminating other information about contact. We created displays in which the virtual stylus could no longer be occluded by the target sphere. The result was that performance in the binocular condition remained very good, whereas[5] in the monocular condition, egocentric distances were strongly overestimated (by ≈25%), and outward drift was accelerated, yielding ≈12% (≈5 cm) drift. Drift was also found in width, depth, shape, and 3-D size. Clearly, the use of disparity matching was responsible for the high performance level exhibited for binocular vision.

On Method and the Use of Virtual Environments

Tasks performed in virtual environments are remarkably similar to tasks commonly performed in actual environments. One can move around and observe an unmoving rigid object from different perspectives and different distances. Objects have a virtual presence. They appear to occupy (more or less) specific locations and to have (more or less) specific 3-D shapes and sizes. What is remarkable about all this is that the objects in question do not exist. Rather, when viewing a virtual object, one is viewing images drawn on displays by a computer. One is viewing computer graphics.

Computer graphics are widely used to study vision and visual information and have been for nearly 40 years. Green and Braun-

---

[5] Although 8 of the participants in the binocular condition immediately used disparity matching to achieve and maintain accurate reaching, a 9th participant was strongly affected by the lack of occlusion information and initially overreached the target by more than 40%. Nevertheless, this participant gradually adjusted over blocks until she was performing accurate reaches by the end of the experimental session. A simple regression of block number on egocentric distances for this participant yielded a slope of 4% and an R^2 of .90, and her final egocentric distances were ≈1. This case is interesting because it shows that a reduction in the relative amount of mismatch in disparity can act over the time scale of a set of reaches to yield gradual improvement in accuracy.

stein published their first studies in 1961 and 1962, respectively, but displays consisting of drawings or projected film images were used long before this. Researchers using such displays were well aware that the resulting conditions are unrepresentative because participants must look at a special surface on which the displayed images appear. This surface is, of course, flat and at a distance that is typically different from that of the objects and surfaces portrayed in the images. Accordingly, investigators have worried about cues to flatness, or alternatively, they have struggled to eliminate information about the presence of a display surface, by creating displays populated by bright dots or patches, in a dark field and viewed in the dark. Head movement with respect to the display surface has often been restricted to prevent the generation of optic flow information about the display surface itself. A collimating lens has sometimes been used to place the surface at effectively infinite distance. This has limited study to the perception of objects at large distances from the observer, in which parallel perspective can be used to model optical transformations. Alternatively, perspective transformations in optics have been studied under highly unrepresentative conditions. Along with the presence of the display surface, the coupling between voluntary self-motion and resulting optical transformations is severed by standard computer graphics displays, and measures of the perception of surface layout (distances, sizes, and shapes) have been limited to varieties of explicit judgment (i.e., magnitude estimation or matching). Action measures have been excluded.

The power of computer graphics is that it allows control and manipulation of optical information. Increases in computer speed and power have made possible the coupling of displays to voluntary head motion and the generation of resulting optical transformations in real time. Finally, with the development of motion measurement systems and miniaturization of displays, optics can be rejoined with the actions that both generate optical transformations and are coordinated and controlled using them. Virtual environments have put the animal and optics back together. Somatosensation (muscle sense, skin sense, and vestibular system) and vision are rejoined as are perception and action.

However, participants are still viewing computer graphics displays. As in studies using simple computer graphics, investigators might control for the presence of the display by placing both the displays and the virtual object surfaces at large viewing distance, that is, using parallel perspective, but this would be to jettison all the real advantages offered by this technology. Instead, the need is to confront a methodological issue that has been lying behind nearly all extant studies of perception. Nearly all methods for the study of perception entail perturbations of perception. Experiments manipulate perception by perturbing the available information. Typically, they strive to isolate and manipulate a single type of hypothesized information. For instance, motion parallax might be isolated from binocular information and from static texture gradients, lighting and shading gradients, and other pictorial information. Such conditions entail multiple perturbations required to remove information that normally covaries with the information under study.[6] Furthermore, the experimental circumstances required to control information entail accessory perturbations, and such accessory perturbations are nearly always confounded with the perturbations of immediate interest. The presence of a display surface is an excellent example.

To evaluate the effect of a perturbation, it is necessary for investigators to measure performance under unperturbed and otherwise representative conditions and to compare performance to that obtained with the perturbation. To evaluate the effect of multiple perturbations, one can create a continuum of conditions ranging from unperturbed, representative conditions, to conditions that isolate information. This is the strategy that we have adopted (see also Bingham & Pagano, 1998; Bingham et al., 2000; Pagano & Bingham, 1998; Wickelgren et al., 2000).[7]

In the current study, we used a virtual environment to isolate and manipulate disparity matching, optic flow generated by self-motion, and occlusion. In a virtual environment, observers view displays, and therefore they must be accommodated to the focal distance of the display, not to the distance of the virtual objects. Much of the information about a display surface is controlled in an HMD. The display is fixed to the head of the observer so that there is no relative motion between head and display (except that due to eye movements; e.g., see Bingham, 1993a). The (miniature) display is viewed through a lens that places it at a larger distance beyond reach. (In some HMDs, the focal distance is set at infinity by such lenses.) There is no visible surface texture. Observers have no real awareness that they are viewing a surface. But they must accommodate it nevertheless.

Therefore, we tested the effect of this perturbation. We found that it did not have a significant effect on shape perception (at least according to our measures). It did have an effect on the perception of egocentric distance. The effect was predictable. It amplified inaccuracy and instability normally present in the perception of actual environments. Distances were overestimated, and outward drift in egocentric distances was accelerated. Thus, the effect was to heighten the already-present need for calibration.

Visual Information for Calibration

Bingham et al. (2000) studied the calibration of reaches in an actual environment with haptic feedback from contact with targets. They found that such feedback stabilized the egocentric distance of reaches and allowed reaches to become accurate (see also Bingham & Pagano, 1998; Pagano & Bingham, 1998; Wickelgren et al., 2000). However, haptic feedback did not correct shape distortions.

In the current experiments, we did not investigate calibration itself. Instead, we investigated what visual information might be

---

[6] Removal of information is often achieved by holding the relevant variable constant. For instance, the control of static texture gradients in the context of a study on motion perspective gradients typically entails the presence of a static gradient equal to zero. This specifies a slant of zero, that is, an upright surface. The result is contradictory information, which is truly a perturbation. Likewise, when we eliminated occlusion, we effectively created information specifying that the target remained beyond the stylus.

[7] In formulating the conditions that we have studied thus far, we have made an effort to be representative of conditions studied in the literature so that our results may be placed in context. Thus, we have studied perception of patch-light objects floating in empty space at various reachable distances from the observer. These conditions are therefore not entirely representative of the conditions under which visually guided reaching most commonly takes place (i.e., for instance, with the objects sitting on a support surface).

used to calibrate reaches under conditions of either monocular or binocular viewing. Clearly, it is relative distance information that must be used for visual calibration. Such information can be used in continuous visual guidance to bring the hand into contact with a target, to achieve a relative distance of zero. This is important because zero distance is the only relative distance value that is definite. Once the hand is positioned at the target distance, then either proprioceptive information about the arm position or the experience of the difference between the result of feedforward positioning of the hand and the result of the continuous guidance might be used to calibrate reaching.

The potential usefulness of relative distance information depends on the ability to achieve contact (i.e., zero distance) accurately. We found that without head movement, monocular vision was less useful in this regard than binocular. Normally, a person does not make head movements during a reach. This limits monocular feedback information to occlusion and relative size. Use of relative size requires that the actual relative sizes of the hand and a target be known in advance. Occlusion does not require such knowledge in advance. We eliminated occlusion in Experiment 4 and found that monocular perception of shape and distance was completely destabilized. Reaching was highly inaccurate and drifted rapidly outward over successive reaches. When occlusion information was available to participants in Experiment 1, reaches were more stable and accurate. However, the information was really effective only for reaches to the front of the target. Reaches to the sides and back continued to overestimate the distance to the targeted locations.

Overestimation might have been expected for reaches to the back simply because occlusion information would be ineffective. We checked whether performance was the same when participants reached to the sides of the target. We had measured perceived object depth by computing the difference between reaches to the front and back. We computed another measure using the mean of reaches to the left and right sides and computing twice the difference between this and reaches to the front. We computed means for each participant and performed a two-tailed, paired t test comparing the two ways of measuring depth in the virtual environment (i.e., using the back or using the sides). The result failed to reach significance, $p > .8$, and the mean difference was only 1% (2 mm). A similar result was obtained in the actual environment. Therefore, poor performance using monocular vision was not merely due to an inability to use occlusion information to guide reaches to the back.

Binocular vision introduces the possibility that disparity matching could be used while viewing the stylus and target with stabilized head. This information was isolated in Experiment 4 and the result (with a single exception) was that reaches were just as accurate as they had been in Experiment 1, in which occlusion information was also present. Unfortunately, shape was still distorted in the configuration of the reaches because participants reached not to the back but to the center of the target when matching the stylus to the disparity of the visible contour of the target.

Finally, we tested the use of monocular parallax matching in a single participant. Use of this information is rather impactical because making head movements while reaching is difficult to coordinate and time-consuming. Nevertheless, we found that it is possible to use this information, and furthermore, the result was accurate in all respects. Reaches in this condition did not exhibit shape distortions (although this participant did produce the characteristic shape compression in the binocular condition). Why this might be deserves further investigation.

The bottom line is that disparity matching appears to be the most effective information about relative distance that can be used to guide the hand continuously to a target, so that the position of the hand at the target can be used to calibrate other information about definite distance, size, and shape. Otherwise, reaching performance was not significantly different with use of binocular as compared with monocular vision; both required calibration.

References

Aglioti, S., DeSouza, J., & Goodale, M. (1995). Size-contrast illusions deceive the eye but not the hand. *Current Biology, 5,* 679–685.

Baird, J. C., & Biersdorf, W. R. (1967). Quantitative functions for size and distance judgments. *Perception & Psychophysics, 2,* 161–166.

Barfield, W., & Furness, T. A. (Eds.). (1995). *Virtual environments and advanced interface design.* Oxford, England: Oxford University Press.

Bingham, G. P. (1993a). Optic flow from eye movement with head immobilized: "Ocular occlusion" beyond the nose. *Vision Research, 33,* 777–789.

Bingham, G. P. (1993b). Perceiving the size of trees: Form as information about scale. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 1–23.

Bingham, G. P., & Pagano, C. C. (1998). The necessity of a perception–action approach to definite distance perception: Monocular distance perception to guide reaching. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 145–168.

Bingham, G. P., & Romack, J. L. (1999). The rate of adaptation to displacement prisms remains constant despite acquisition of rapid calibration. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1331–1346.

Bingham, G. P., & Stassen, M. G. (1994). Monocular distance information in optic flow from head movement. *Ecological Psychology, 6,* 219–238.

Bingham, G. P., Zaal, F., Robin, D., & Shull, J. A. (2000). Distortions in definite distance and shape perception as measured by reaching without and with haptic feedback. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1436–1460.

Braunstein, M. L. (1962). Depth perception in rotating dot patterns: Effects of numerosity and perspective. *Journal of Experimental Psychology, 64,* 415–420.

Brenner, E., & Van Damme, W. J. M. (1998). Judging distance from ocular convergence. *Vision Research, 38,* 493–498.

Bridgeman, B., Kirch, M., & Sperling, A. (1981). Segregation of cognitive and motor aspects of visual function using induced motion. *Perception & Psychophysics, 29,* 336–342.

Bridgeman, B., Peery, S., & Anand, S. (1997). Interaction of cognitive and motor-oriented maps of visual space. *Perception & Psychophysics, 59,* 456–469.

Elliott, D., & Madalena, J. (1987). The influence of pre-movement information on manual aiming. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 39*(A), 541–559.

Eriksson, E. S. (1974). Motion parallax during locomotion. *Perception & Psychophysics, 16,* 197–200.

Faugeras, O. (1993). *Three-dimensional computer vision: A geometric viewpoint.* Cambridge, MA: MIT Press.

Fermuller, C., Cheong, L., & Aloimonos, Y. (1997). Visual space distortion. *Biological Cybernetics, 77,* 323–337.

Fisher, S. K., & Ciuffreda, K. J. (1988). Accommodation and apparent distance. *Perception, 17,* 609–621.

Foley, J. M., & Held, R. (1972). Visually directed pointing as a function of target distance, direction and available cues. *Perception & Psychophysics, 12,* 263–268.

Franz, V., Gegenfurtner, K., Bülthoff, H., & Fahle, M. (2000). Grasping visual illusions: No evidence for a dissociation between perception and action. *Psychological Science, 11,* 20–25.

Gentillucci, M., & Negrotti, A. (1994). Dissociation between perception and visuomotor transformation during reproduction of remembered distances. *Journal of Neurophysiology, 72,* 2026–2030.

Gilinsky, A. (1951). Perceived size and distance in visual space. *Psychological Review, 58,* 460–482.

Gogel, W. C. (1977). The metric of visual space. In W. Epstein (Ed.), *Stability and constancy in visual perception* (pp. 129–181). New York: Wiley.

Gogel, W. C., & Tietz, J. D. (1979). A comparison of oculo-motor and motion parallax cues of egocentric distance. *Vision Research, 19,* 1161–1179.

Goodale, M. A., Jakobson, L. S., & Servos, P. (1996). The visual pathways mediating perception and prehension. In A. M. Wing, P. Haggard, & J. R. Flanagan (Eds.), *Hand and brain: The neurophysiology and psychology of hand movements* (pp. 15–31). San Diego, CA: Academic Press.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience, 15,* 20–25.

Graham, J. K., Bradshaw, M. F., & Davis, A. M. (1997a). Does temporal delay affect pointing accuracy in open-loop pointing in adults and children? *Perception, 26,* 1331–1332.

Graham, J. K., Bradshaw, M. F., & Davis, A. M. (1997b, October). *How does temporal delay affect open-loop pointing accuracy? A comparison of adults and children aged six to ten years.* Paper presented at the Conference on Vision for Reach and Grasp, University of Minnesota, Minneapolis, MN.

Graham, J. K., Bradshaw, M. F., & Davis, A. M. (1998). The effect of pre-movement delays on pointing accuracy in middle childhood. *Perception, 27,* 1379–1387.

Green, B. F. (1961). Figure coherence in the kinetic depth effect. *Journal of Experimental Psychology, 62,* 272–282.

Gregory, R. L. (1970). *The intelligent eye.* New York: McGraw-Hill.

Haffenden, A., & Goodale, M. (1998). The effect of pictorial illusion on prehension and perception. *Journal of Cognitive Neuroscience, 10,* 122–136.

Held, R., & Durlach, N. (1991). Telepresence, time delay and adaptation. In S. R. Ellis, M. K. Kaiser, & A. J. Grunwald (Eds.), *Pictorial communication in virtual and real environments* (pp. 232–246). London: Taylor & Francis.

Held, R., Efstathiou, A., & Greene, M. (1966). Adaptation to displaced and delayed visual feedback from the hand. *Journal of Experimental Psychology, 72,* 887–891.

Hochberg, J. E. (1978). *Perception* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis.* New York: Oxford University Press.

Jeannerod, M. (1988). *The neural and behavioral organization of goal-directed movements.* Oxford, England: Oxford University Press.

Johansson, G. (1973). Monocular movement parallax and near space perception. *Perception, 2,* 135–146.

Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research, 31,* 1351–1360.

Kocian, D. F., & Task, H. L. (1995). Visually coupled systems hardware and the human interface. In W. Barfield & T. A. Furness (Eds.), *Virtual environments and advanced interface design* (pp. 175–257). Oxford, England: Oxford University Press.

Leibowitz, H. W. (1966). Role of changes in accommodation and convergence in the perception of size. *Journal of the Optical Society of America, 8,* 1120–1123.

Leibowitz, H. W. (1972). Oculomotor adjustments and constancy. *Perception & Psychophysics, 12,* 497–500.

Liang, J., Shaw, C., & Green, M. (1991, November). *On temporal–spatial realism in the virtual reality environment.* Paper presented at the Fourth Association for Computing Machines Symposium on User Interface Software and Technology, Hilton Head, SC.

Loomis, J. M., DaSilva, J. A., Fujita, N., & Fukusima, S. S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 906–921.

Marotta, J., DeSouza, J., Haffenden, A., & Goodale, M. (1998). Does a monocularly presented size-contrast illusion influence grip aperture? *Neuropsychologia, 36,* 491–497.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action.* Oxford, England: Oxford University Press.

Mon-Williams, M., & Dijkerman, H. C. (1999). The use of vergence information in the programming of prehension. *Experimental Brain Research, 128,* 578–582.

Mon-Williams, M., & Tresilian, J. R. (1999a). *Ordinal depth information from accommodation.* Unpublished manuscript.

Mon-Williams, M., & Tresilian, J. R. (1999b). Some recent studies on the extraretinal contribution to distance perception. *Perception, 28,* 167–181.

Nakayama, K., & Loomis, J. M. (1974). Optical velocity patterns, velocity sensitive neurons, and space perception: A hypothesis. *Perception, 3,* 63–80.

Norman, J. F., & Todd, J. T. (1993). The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. *Perception & Psychophysics, 3,* 279–291.

Norman, J. F., Todd, J. T., Perotti, V. I., & Tittle, J. S. (1996). The visual perception of 3-D length. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 173–186.

Ono, H., & Steinbach, M. J. (1990). Monocular stereopsis with and without head movement. *Perception & Psychophysics, 48,* 179–187.

Owens, D. A., & Liebowitz, H. W. (1980). Accommodation, convergence, and distance perception in low illumination. *American Journal of Optometry & Physiological Optics, 57,* 540–550.

Pagano, C. C., & Bingham, G. P. (1998). Comparing measures of monocular distance perception: Verbal and reaching errors are not correlated. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1037–1051.

Parker, A. J., Cumming, B. G., Johnston, E. B., & Hurlbert, A. C. (1995). Multiple cues for three-dimensional shape. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 351–364). Cambridge, MA: MIT Press.

Pavani, F., Boscagli, I., Benvenuti, F., Rabufetti, M., & Farne, A. (1999). Are perception and action affected differently by the Titchener circles illusion? *Experimental Brain Research, 127,* 95–101.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research* (2nd ed.). Fort Worth, TX: Harcourt Brace.

Peli, E. (1999). Optometric and perceptual issues with head-mounted displays. In P. Mouroulis (Ed.), *Visual instrumentation: Optical design and engineering principles* (pp. 205–276). New York: McGraw-Hill.

Perotti, V. J., Todd, J. T., & Norman, J. F. (1996). The visual perception of rigid motion from constant flow fields. *Perception & Psychophysics, 58,* 666–679.

Phillips, F., & Todd, J. T. (1996). Perception of local three-dimensional shape. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 930–944.

Post, R., & Welch, R. (1996). Is there dissociation of perceptual and motor responses to figural illusions? *Perception, 25,* 569–581.

Rieser, J. J., Pick, H. L., Ashmead, D. H., & Garing, A. E. (1995).

Calibration of human locomotion and models of perceptual–motor organization. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 480–497.

Rogers, B. J., & Collett, T. S. (1989). The appearance of surfaces specified by motion parallax and binocular disparity. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 41*(A), 697–717.

Rogers, B., & Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception, 8,* 125–134.

Rogers, S., & Rogers, B. J. (1992). Visual and nonvisual information disambiguate surfaces specified by motion parallax. *Perception & Psychophysics, 52,* 446–452.

Sedgwick, H. A. (1986). Space perception. In K. R. Boff, L. Kauffman, & J. P. Thomas (Eds.), *Handbook of perception and human performance V: Sensory processes and perception* (Vol. 1, pp. 21.1–21.57). New York: Wiley.

Smith, W. M., & Bowen, K. F. (1980). The effects of delayed and displaced visual feedback on motor control. *Journal of Motor Behavior, 12,* 91–101.

Steinbach, M. J., & Ono, H. (1991). Motion parallax judgments of depth as a function of the direction and type of head motion. *Canadian Journal of Psychology, 45,* 92–98.

Swenson, H. A. (1932). The relative influence of accommodation and convergence in the judgment of distance. *Journal of General Psychology, 7,* 360–380.

Tharp, G., & Liu, A. (1992). Timing considerations of head mounted display performance. In SPIE (Ed.), *Human vision, visual processing and digital display III* (Vol. 1666, pp. 10–13). San Jose, CA: SPIE.

Tittle, J. S., & Braunstein, M. L. (1993). Recovery of 3-D shape from binocular disparity and structure from motion. *Perception & Psychophysics, 54,* 157–169.

Tittle, J. S., & Perotti, V. J. (1997). The perception of shape and curvedness from binocular stereopsis and structure from motion. *Perception & Psychophysics, 59,* 1167–1179.

Tittle, J. S., Todd, J. T., Perotti, V. J., & Norman, J. F. (1995). Systematic distortion of perceived three-dimensional structure from motion and binocular stereopsis. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 663–678.

Todd, J. T., & Bressan, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics, 48,* 419–430.

Todd, J. T., Tittle, J. S., & Norman, J. F. (1995). Distortions of three-dimensional space in the perceptual analysis of motion and stereo. *Perception, 24,* 75–86.

Toye, R. C. (1986). The effect of viewing position on the perceived layout of space. *Perception & Psychophysics, 40,* 85–92.

Tresilian, J. R., Mon-Williams, M., & Kelly, B. M. (1999). Increasing confidence in vergence as a cue to distance. *Proceedings of the Royal Society of London, Series B, 266,* 39–44.

Ullman, S. (1979). *The interpretation of visual motion.* Cambridge, MA: MIT Press.

Vindras, P., & Viviani, P. (1998). Frames of reference and control parameters in visuomanual pointing. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 569–591.

Vishton, P. M., Rea, J. G., Nunez, L. N., & Cutting, J. E. (1999). Comparing effects of the horizontal–vertical illusion on grip scaling and judgment: Relative versus absolute, not perception versus action. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1659–1672.

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

Wann, J. P., Rushton, S., & Mon-Williams, M. (1995). Natural problems for stereoscopic depth perception in virtual environments. *Vision Research, 35,* 2731–2736.

Welch, R. B. (1978). *Perceptual modification: Adapting to altered sensory environments.* New York: Academic Press.

Wickelgren, E. A., McConnell, D., & Bingham, G. P. (2000). Reaching measures of monocular distance perception: Forward versus side-to-side head movements and haptic feedback. *Perception & Psychophysics, 62,* 1051–1059.

Appendix

Measurement of HMD Focal Distance

We measured the optical properties of the multicomponent lenses in the V6 HMD. First, the focal distance and magnification were measured in a manner that included the eye in a representative configuration, as shown in Figure A1. This approach guaranteed relatively small measurement error. A black 1-cm square grid was printed on white paper with 11 line elements spaced 1 mm apart in both the *x* and *y* directions. The grid was placed behind the lens at the distance of the display in the HMD, that is, 1.3 cm from the back surface of the lens system. The eye was positioned at about 1 cm from the front of the lens. A small, 3-mm cube-shaped beam splitter was placed between the eye and the lens so that the observer could see the virtual image of the grid superimposed on a large poster board surface located to the side of the observer. The observer instructed experimenters, who placed this surface at the distance of the virtual image. The distance from the surface to the beam splitter and from the beam splitter to the eye and lens were all measured, respectively. Locations of the images of line crossings were marked on the poster board surface. These values were used to determine the focal distance and magnification of the lens and to evaluate potential lens distortion in the central part of the 60° field of view of the HMD display. The virtual image was found to lie at 98 cm from the eye. The average magnification of the lens was 32.75.

Measurement of HMD Image Distortion

The grid used above was back lit to project a real image through the lens and an artificial pupil onto a sheet of white poster board placed in front of the lens, as shown in Figure A2. The image locations of all line crossings were marked on the poster board. The distances between these locations were measured and analyzed to evaluate the lens distortion over the full field. The lens exhibited a pincushion distortion that left the central portions of the image relatively distortion free. The distortion was proportional to the percentage distance from the center to the edge of the field with the proportionality constant of .10 (e.g., 2% at 20% and 5% at 50%). Only about 70% of the lens field is used to view the HMD displays, so the maximum distortion is about 7% at the edge of the displays. We might have inverted and eliminated this distortion as part of the computation of the display images, but such computation would add to the loop delay time. We avoided this because the distortion was small (that is, less than 3% within the central 20° field). Effectively, it was as if our participants were wearing weak glasses and experiencing the associated distortion.

Measurement of Virtual Environment Phase Delay

We built an apparatus to measure the phase delay, that is, the delay between FOB measurement of head–hand motion and computer trans-

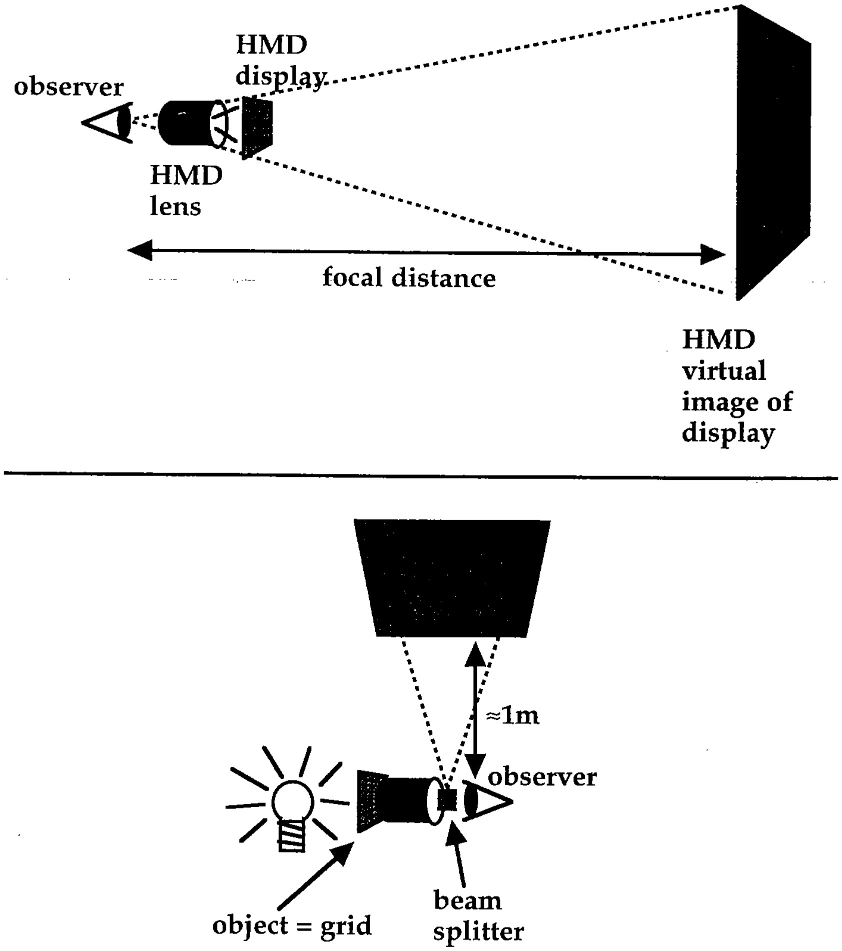


Figure A1. Top: An observer (on the left) looks through a lens to view a display and sees the virtual image of the display at a distance determined by the focal distance of the lens. Bottom: The apparatus used to measure the focal distance of our lens. An observer (on the right) views a display consisting of a black grid printed on white paper that is lit from behind. The grid is the size of the actual display in the head-mounted display (HMD) and at the same distance from the back of the lens as is the display in the HMD. The observer views the display through the lens (which is located about 2 cm from the eye) and a tiny beam splitter (which lies between the eye and the lens). A large, lighted, white-textured cardboard surface is placed at eye height to the right of the observer so that, by virtue of the beam splitter, the observer simultaneously sees the display grid and the cardboard surface. The observer instructs the experimenters, who adjust the distance of the cardboard surface so that the observer is able to focus both the grid and the cardboard and so that the grid appears to lie on the cardboard surface. See the text for remaining details.

formation of the display. A FOB marker was placed on a 20 cm long (2.5-cm diameter) wooden dowel that was mounted standing vertically near the outer edge of a turntable. The marker was mounted on the dowel to be well above the turntable to avoid potential interference. A fin was fixed to the edge of the platter next to the FOB marker. As shown in the top panel of Figure A3, the turntable was placed so that when the FOB marker crossed a plane in the measurement volume, the fin simultaneously interrupted a light-activated switch causing a clock to start. Independently, the change in coordinates, measured by the FOB as the marker passed the plane, also caused a region of the display to switch from black to white after an image with a specified number of polygons had been drawn elsewhere in the display. A photodiode was placed on the display and detected when the display turned white. This caused the clock to stop. The marker and fin on the turntable were placed by hand so that a mere hand tremor caused the clock and screen to flicker on or off. Then the turntable was set in rotation to measure a delay with each cycle. The marker passed the plane in the volume, and simultaneously, the fin interrupted the switch that started the clock. The

(*Appendix continues*)

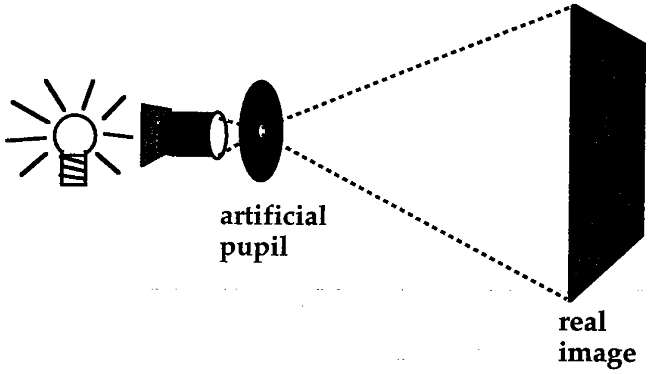


Figure A2. The apparatus used to measure distortion in the image of a display that would be observed through the lens. The same back-lit grid on paper was used as in the measurements of focal distance. A real image of this display was formed on a large paper projection surface by placing an artificial pupil between the lens and the image surface to yield a focused image. See the text for remaining details.

delay occurred between this event and the resulting change in the display that caused the clock to stop (through the photodiode). Fifty cycles were measured for each number of polygons drawn. The lower panel of Figure A3 shows mean measured delay (with standard deviation bars) as the number of polygons drawn in the image was increased. Our research required images containing about 5,000 polygons, yielding a delay of about 50 ms on this function. We found a 30-ms increase when we plugged in the HMD and Octane Channel Option Board (Silicon Graphics Incorporated, Mountain View, CA) yielding a final phase lag of 80 ms. This, in turn, should be small enough to avoid significant perturbation of visually guided reaching under conditions of preferred movement speeds (Barfield & Furness, 1995; Held & Durlach, 1991; Held et al., 1966; Kocian & Task, 1995; Liang et al., 1991; Smith & Bowen, 1980; Tharp & Liu, 1992).

Spatial Calibration of the FOB Motion Measurement System

We built a Plexiglas calibration frame and used it to spatially calibrate the FOB. A rectangular array of 6 × 4 locations spaced 10 cm apart, to cover a 50-cm × 30-cm area, was positioned horizontally at each of four different heights, spaced 10 cm apart, starting 10 cm below the FOB emitter. The FOB marker was placed at each of the 22 locations (we could not reach two locations on the Plexiglas to drill holes) in the array, at each of the four heights, and measured. FOB-measured coordinates were regressed on the physically measured coordinates of the locations on the Plexiglas. The x, y, and z FOB values were entered into each of three multiple regressions to predict physical x, y, or z measured with a T square, a carpenter's level, and a meter stick. The R^2 values in all cases were better than .999. The square root of the mean of the squared residuals provides a measure of the error of the estimate. The largest value was obtained for the bottom-most height, farthest from the FOB emitter where the greatest inaccuracy would be expected. This value was 1.5 mm. The FOB was accurate to within this value and free of measurable distortion (which we tested by regressing the physical coordinates on the residuals to find no relation or regular pattern).

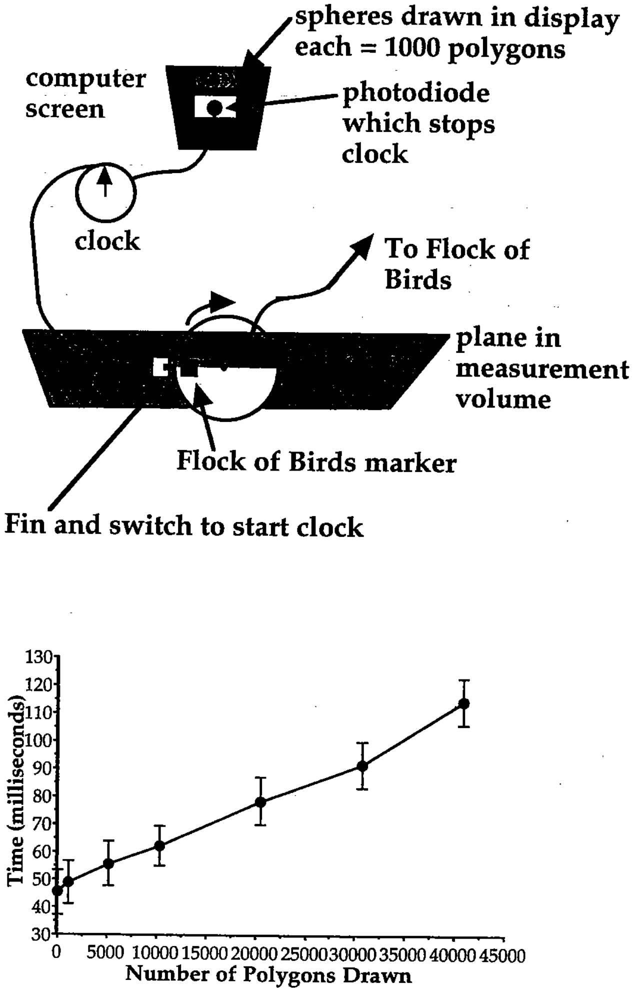


Figure A3. Top: The apparatus used to measure phase delay in the virtual reality system. See text for details; the Flock of Birds measurement system is manufactured by Ascension Technology Corporation, Burlington, VT. Bottom: Mean phase delays (with standard deviation bars) plotted as a function of the number of polygons drawn in the display.

Received March 30, 2000
Revision received August 31, 2000
Accepted January 22, 2001 ■