ECOLOGICAL PSYCHOLOGY, *17*(2), 55–74
Copyright © 2005, Lawrence Erlbaum Associates, Inc.

Calibration of Distance and Size Does Not Calibrate Shape Information: Comparison of Dynamic Monocular and Static and Dynamic Binocular Vision

Geoffrey P. Bingham
Department of Psychology
Indiana University

This study investigated the coupling of distance and size perception as well as the coupling of distance and shape perception. Each was tested in 2 ways using a targeted reaching task that simultaneously yielded measures of distance, size, and shape perception. First, feed-forward reaches were tested without feedback. Errors in size did not covary with errors in distance, but errors in shape did. Second, reaches were tested with visual feedback. Estimated distance and size became more accurate, but shape did not. The evidence indicated that distance and size perception and distance and shape perception are not coupled. These results were replicated 3 times as we also compared performance using dynamic monocular, static binocular, and dynamic binocular vision. Performance was better with binocular than monocular vision both without and with feedback. The presence of a size gradient did not improve monocular distance perception, yielding additional evidence that distance and size perception are not coupled.

The use of vision to control reaching is quite complex. A number of spatial properties of a target object must be perceived to control a reach to grasp the object. The properties include the object's distance, size, and shape. Distance is needed to control reaching. Size and shape are needed to control grasping. It has often been assumed that these properties are coupled in space perception. For instance, the classic size–distance invariance hypothesis assumes that the perception of distance and

Correspondence should be addressed to Geoffrey P. Bingham, Department of Psychology, Indiana University, 1101 East Tenth Street, Bloomington, IN 47405–7007. E-mail: gbingham@indiana.edu

size are coupled and that perception of one of these properties determines the perception of the other. Similarly, shape perception has often been treated as if perceived shape reduced to a set of perceived positions on the surface of an object. In this case, the perceptions of distance and of shape are assumed to be coupled. The coupling of these properties means that if errors occur in the perception of one property, then corresponding errors should occur in perception of the other property. Furthermore, if these respective properties are coupled in perception, then the control of reaching and of grasping would also be coupled, so that errors in one should covary with errors in the other. However, it remains an open question whether distance, size, and shape perception are coupled.

To address this question, we used a reaching paradigm, which yielded simultaneous measures of perceived distance, size, and shape. The strategy was to see if errors in distance perception yielded corresponding errors in size or shape perception. We investigated this in two conditions. In the first, participants performed open-loop reaches. Errors in distance would indeed be expected in this condition. The question was whether these errors would yield corresponding errors in size or shape. In the second condition, participants received visual feedback about their reaches. Here we expected errors to be reduced and the question was whether errors would diminish in the same way for all three properties.

Bingham, Zaal, Robin, and Shull (2000) and Bingham, Bradley, Bailey, and Vinner (2001) investigated open-loop reaches. An unexpected result in both studies was that binocular vision did not yield superior performance as compared to monocular vision. That is, when participants were not allowed to calibrate their reaches, performance using binocular and monocular vision was equally poor. In contrast, Tresilian, Mon-Williams, and Kelly (1999) found that use of binocular vergence with a size gradient yielded relatively accurate reaching (see also Mon-Williams & Tresilian, 1999b; Tresilian & Mon-Williams, 2000). Size gradient refers to the progressive decrease in image size generated when an object of a given size is viewed from progressively greater distances. A size gradient occurs in texture gradients because surface texture elements of constant size appear at a range of distances. In this case, of course, the size gradient is available simultaneously. A size gradient also occurs when a single object is viewed at different distances at different times. In this case, the size gradient is available only over successive times. A size gradient by itself does not provide information about definite distance. It only provides information about relative distance. However, a size gradient might interact with other information about definite distance to yield superior distance perception. Hypothetically, given a size gradient, information about object distance should yield perception of object size. In turn, information about object size should yield perception of distance when coupled with a size gradient (i.e., image size). Given these relations, a size gradient could effectively allow distance information to be stored as an object size estimate and to interact over time and occasions to yield improved distance perception. Essentially, this is a form of the classic size–distance invariance hypothesis (e.g., Boring & Holway, 1940; Hochberg, 1970). If the

size gradient was the reason for the superior results in Tresilian et al. (1999), then with a size gradient, similar improvements should occur in reaching with monocular and binocular vision and performance in the two conditions should remain comparable. To provide a strong test of the effect of a size gradient, we used only a single object size viewed at different distances. This yielded a simple, clear size gradient so that any errors in distance perception should certainly yield corresponding errors in size perception if size and distance perception are coupled.

A second possible reason for the superior results of Tresilian et al. (1999) is that they tested a range of distances, whereas in both studies by Bingham and colleagues only a single distance was tested. The relation between vergence and perceived distance is mediated by an adjusted reference level for vergence (Brenner & Van Damme, 1998; Mon-Williams, Tresilian, & Hasking, in press; Owens & Liebowitz, 1976; von Hofsten, 1976, 1979). The reference vergence level is a function of both luminance level and recently experienced vergence distances (Brenner & Van Damme, 1998; Mon-Williams, Tresilian, & Hasking, in press; Owens & Liebowitz, 1976). With repeated experience of only a single distance, vergence level would become poorly defined as information about distance. We investigated this possibility by testing a range of distances spanning reach space. If this was the reason for the superior results in Tresilian et al. (1999), then we would expect to obtain substantial improvements in the binocular condition but not in the monocular condition.

The perception of shape (and size) is important in the context of reaching and grasping because grasping most often entails contact of the fingers with the back of an object and the location of the back surface can be apprehended only on the basis of shape information obtained from the visible, front portions of the object. Many studies of shape perception have found severe distortions in shape perception (see Todd, Tittle, & Norman, 1995, for a review). Bingham, Bradley, et al. (2001) and Bingham et al. (2000) used reaching measures to evaluate shape perception and they also found distortions in shape perception. We now investigate the extent to which such distortions in shape covary with errors in the perception of distance and size under conditions of open-loop reaching. Of special interest, however, is the possibility that feedback might allow the calibration of shape perception as well as distance and size perception. We investigated the use of visual feedback in the context of both monocular and binocular vision. Bingham, Bradley, et al. (2001) studied reaches performed with visual guidance. They compared monocular and binocular information and found that binocular vision yielded the most accurate performance. Participants used disparity matching to place the visible hand and target in the same depth plane. The implication of this result is that binocular vision should provide the best visual feedback for the calibration of feed-forward reaching. We investigated this by comparing the use of monocular versus binocular information to calibrate feed-forward reaches. The question was whether perceived shapes would also become calibrated. If so, then this would support the possibility that the perception of position and shape are coupled. On the other hand, if feedback calibrates distance and fails to calibrate

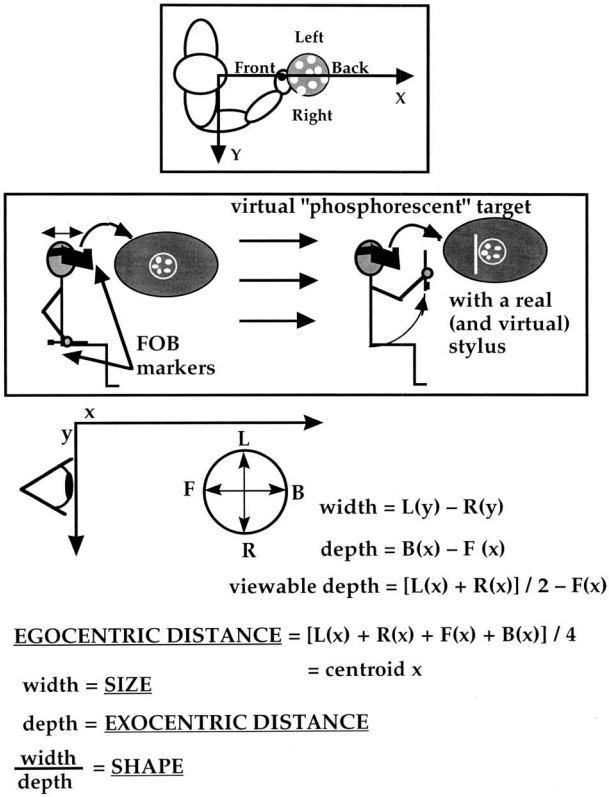

FIGURE 1 Illustration of the reaching task performed in Experiments 1 and 2. The top panel shows the task as seen from above. Participants reach along the X axis to place a stylus held vertically in the hand tangent to the equator of a target sphere at its front, back, left, or right sides. The second panel shows the participant wearing the head-mounted display and viewing a virtual target sphere while holding the stylus in his or her lap and moving his or her head from side to side. Then the participant reaches to touch the sphere with the stylus. A virtual stylus is seen in some conditions after the reach is measured. The third panel shows how the various dependent measures are computed from each block of reaches to the four locations on a target object.

shape, then this would suggest that the perception of position and shape are not coupled and shape does not reduce to the perception of positions.

The target object was tested at five different distances within reach. As shown in Figure 1, participants reached to touch the front, back, and sides of a virtual target sphere with a hand held stylus. We used the centroid of the four positions as a measure of perceived object distance.[1] We used the difference between reaches to

[1]We used the centroid rather than just the front of the object because a grasp is aimed to span an object. The "opposition axis" that extends between the fingers grasping an object often passes through the object centroid (e.g., Iberall, Bingham, & Arbib, 1986).

the left and right sides (width) as a measure of perceived object size. We used the difference between reaches to the front and back (depth) as a measure of perceived object depth. Together, these last two measures yielded a measure of perceived shape, namely the aspect ratio of width to depth. Finally, because use of reaches to the back of the targets entailed the assumption that this position was specified by the visible shape of the front of the target, we used a second measure of depth to investigate this assumption. Viewable depth was computed using the average distance of the reaches to the left and right sides instead of the back.

METHODS

Participants

Twenty-two adults (14 men and 8 women) age 19 to 30 years participated in the experiment. Nine participated in the monocular condition (6 men and 3 women). Six participated in the static binocular condition (4 men, 2 women). Seven participated in the dynamic binocular condition (4 men and 3 women). Participants were paid $5 per hour. All participants had normal or corrected to normal eyesight (using contacts) and normal motor abilities. All were right-handed.

Apparatus

The Virtual Environment Lab consisted of an SGI Octane graphics computer, a Flock of Birds (FOB) motion measurement system with two markers (for head and hand), and a Virtual Research V6 stereo head-mounted display (HMD). Displays in the HMD portrayed a virtual target sphere and hand-held 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 hand held 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 18.5 cm in length and 1 cm in diameter. The 7 cm diameter virtual target sphere was dark with green phosphorescent-like dots and appeared against a dark background so that only the green dots could be seen. The stylus and marker was 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 was 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 .82 kg. The sampling rate of the FOB was 120 Hz. As described in Bingham, Bradley, et al. (2001), we measured the focal distance to

the virtual image, the image distortion, the phase lag and the spatial calibration. The virtual image was at 1 m distance from the eyes. The phase lag was 80 ms. The spatial calibration yielded a resolution of about 2 mm. See Bingham et al. (2001) for additional information about the virtual environment.

Procedure

Participants sat in a wooden chair. In the static binocular condition, the participant rested his or her head on a carved wooden chin rest that sat on top of an aluminum rod. The rod was positioned between the participant's legs and extended from an adjustable clamp on the chair. The rod did not interfere with reaching. The height of the chin rest was adjusted for each participant. Free head movements (no chin rest) were allowed in the dynamic monocular and dynamic binocular conditions. The experimenter first measured the participant's interpupillary distance using a ruler and entered the value into the software. The participant then placed the HMD on his or her head and adjusted the lenses in front of his or her eyes. The participant was allowed a few minutes to move his or her head and hand and to explore and acclimate to the virtual environment. Following this, the maximum reach distance and eyeheight were measured by having the participant hold the stylus out as far as possible in front of his or her face while sitting in the chair and wearing the HMD. The software used the measured values to position the 7 cm virtual sphere at eyeheight and at distances equal to .50, .60, .70, .80, and .90 of the maximum reach.

The task was explained to the participants. Participants were instructed to reach to place the stylus at one of four locations relative to the surface of the target sphere, as shown in Figure 1a. Holding the stylus vertical, they reached to place the midpoint of the stylus tangent to the surface of the sphere at its horizontal equator either to the front, right, left, or back.[2] Only the virtual target sphere could be seen, not the virtual stylus, except at the very end of trials in the feedback conditions, at which point the virtual stylus would be made visible as explained next.

Between trials, the participant sat holding the stylus in his or her lap. At the beginning of each trial, the target appeared at a given distance and the computer announced to the participant the location to be touched on the target (e.g., front, back, left, or right). In the static binocular condition, the participant then simply reached at preferred rates. In the dynamic vision conditions only, the par-

[2]We instructed participants to contact the target with the midpoint of the stylus. In the feedback conditions when participants could see both the stylus and the target, we wanted them to be able to see both the top and bottom of the stylus when it was positioned behind the target. We recorded positions of both the top and bottom of the stylus and computed the mean X and Y coordinates to yield the location of the midpoint. We computed the absolute difference of the top and bottom X coordinates in the monocular no feedback condition to evaluate how much the stylus varied from a vertical orientation. The mean orientation error was .08 rads (SD = .06 rads). If participants misgauged the midpoint of the stylus by ±3 cm, this would have incurred a mean measurement error of only 2.4 mm (SD = 1.8 mm).

ticipant first moved his or her head and torso 10 cm side to side 2 to 3 times at preferred rates while counterrotating the head to keep the target centered in the display and looking at the targeted locus on the surface.[3] Following this, the participant reached at preferred rates. Once the participant had reached the target, he or she said, "okay," and the 3D coordinates of the stylus were recorded. In the no feedback conditions, the participant then placed the stylus back in his or her lap and the next trial was begun. In the feedback conditions, the virtual stylus would become visible (seen together with the target sphere) at the same time that the 3D coordinates of the stylus were recorded. When the stylus was made visible, if its position was incorrect, the participant was allowed to move the stylus to the correct position on the target. Once the participant had done this (which took about 5 sec), he or she placed the stylus back in his or her lap and the next trial was begun (with the stylus invisible once again). A block of trials consisted of reaches to each of the 20 locations (that is, four locations on targets at each of five distances) in a completely random order. Five blocks of trials were performed in each viewing condition. Each participant wore a patch over the left eye during monocular viewing. In the monocular condition, no feedback and feedback conditions were tested on subsequent days, in that order. Participants in both binocular conditions were tested first without feedback and then with feedback in separate sessions on a single day with a 10 to 15 min break between sessions, during which participants removed the HMD and went for a walk around the department.

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 Figure 1b, we used Cartesian coordinates such that depth varied along the X axis and the Y axis lay in a frontoparallel plane. We computed the distance as the X centroid of the four reaches. This distance was reported both as a proportion of maximum reach distance and 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 actual target size (which was computed as the sum of target diameter plus stylus diameter). Shape was computed as the aspect ratio of width to depth. Finally, viewable depth was computed as two times the difference in X between the front and the mean of the left and right.

[3]This also minimized any potential effect of the phase lag in the system.

Design

The independent variables included one between-subjects variable, viewing (dynamic monocular, static binocular, dynamic binocular), and two within-subjects variables, feedback (no feedback and feedback), and block (1–5). The dependent variables were distance, width, depth, shape, and viewable depth.

RESULTS AND DISCUSSION

Results for reach distances are reported first, and then for widths and depths.

Distance

We computed overall mean reach distances for each viewing and feedback condition. These are shown for the dynamic monocular, static binocular, and dynamic binocular conditions (without and with feedback), respectively, in Figures 2, 3, and 4 together with standard error bars representing between-subject variability. In the monocular no feedback condition, participants tended to overreach all targets except the most distant. The presence of a size gradient did not prevent inaccurate perception of distance. Without feedback, the level of performance in both binocular conditions was better than the monocular condition in respect to slopes and intercepts, although only the dynamic binocular condition was better in terms of the overall r^2. Performance in the binocular conditions was relatively accurate. This difference in performance between the monocular and the binocular conditions suggested that a reference vergence and not a size gradient was the important factor yielding good performance in Tresilian et al. (1999). Furthermore, the results from the virtual environment in this experiment were comparable to the results from an actual environment in Tresilian et al. and in particular, the slope and intercept for static binocular vision without feedback were identical to those in Tresilian et al.

With feedback, performance levels were comparable in all respects between static and dynamic binocular conditions. Both binocular conditions were better than the monocular condition, as predicted. This was consistent with the results of Bingham, Bradley, et al. (2001) showing that disparity matching would provide the best feedback information for calibration.

First, we compared performance in the dynamic monocular and static binocular conditions. We performed a multiple regression analysis on the data in the no feedback condition, regressing target distance, viewing condition (monocular and binocular coded as ± 1), block, and vectors representing two-way and three-way interactions on target distance. The result was significant, $p < .001$, $F(7, 392) = 110.8$, $r^2 = .64$. Using a procedure described by Pedhazur (1982), we removed all nonsignificant factors and retested the analysis. The result was significant, $p <$

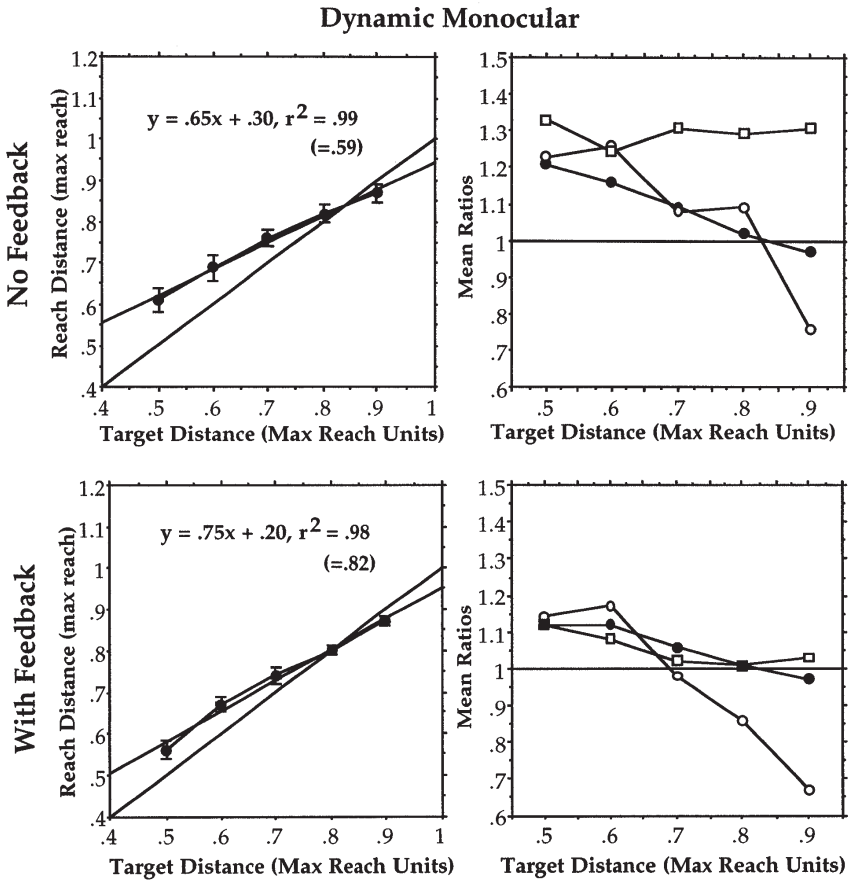

FIGURE 2 Dynamic Monocular: The two panels on the left show mean reach distances graphed as a function of actual target distances. Standard error bars represent between-subject variability. The r^2 accompanying the simple regression equations are for the fits to the means. The values in parentheses are for fits to the trial data and provide some measure of the precision. The two right panels show mean width, depth, and distance, each computed in proportion to actual target values and graphed as a function of target distance. The accurate value of the ratio judged/actual in each case, as shown in this graph, is 1. Width: open squares. Depth: open circles. Distance: filled circles. Top panels: without feedback; Bottom panels: with feedback.

.001, $F(3, 396) = 231.9$, $r^2 = .64$, and the significant factors were target distance ($p < .001$, partial $F = 668.8$), viewing ($p < .01$, partial $F = 19.5$), and the target distance by viewing interaction ($p < .001$, partial $F = 12.0$). We performed separate analyses on the data for each viewing condition with results shown in Figures 2 and 3. The slope and r^2 were .65 and .59, respectively, for monocular and .86 and .66, respectively, for binocular. As shown by the r^2, the precision was only slight better

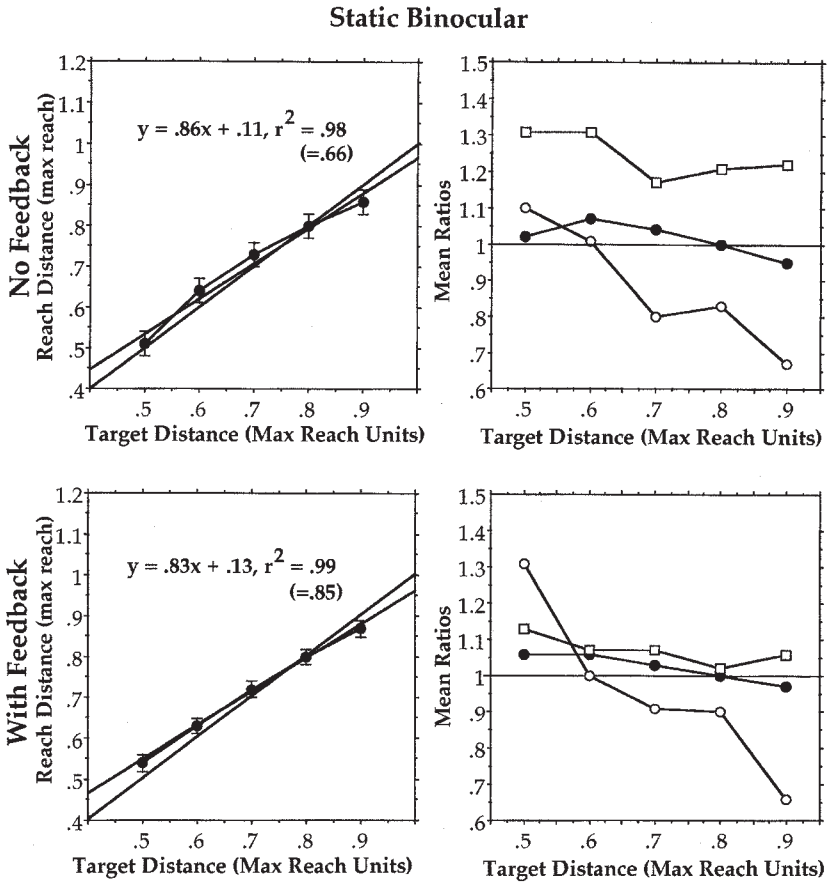

FIGURE 3 Static binocular: The two panels on the left show mean reach distances graphed as a function of actual target distances. Standard error bars represent between-subject variability. The r^2 accompanying the simple regression equations are for the fits to the means. The values in parentheses are for fits to the trial data and provide some measure of the precision. The two right panels show mean width, depth, and distance, each computed in proportion to actual target values and graphed as a function of target distance. The accurate value of the ratio judged/actual in each case, as shown in this graph, is 1. Width: open squares. Depth: open circles. Distance: filled circles. Top panels: without feedback. Bottom panels: with feedback.

with binocular vision. However, as indicated by the slopes, the accuracy was significantly better with static binocular parallax than with dynamic monocular parallax.

Performance improved with feedback. Nevertheless, the final level of performance was still better in the binocular condition than in the monocular. A multiple regression was performed on the combined data of both feedback conditions. Feedback (coded as ± 1) was added as a factor together with its various interactions.

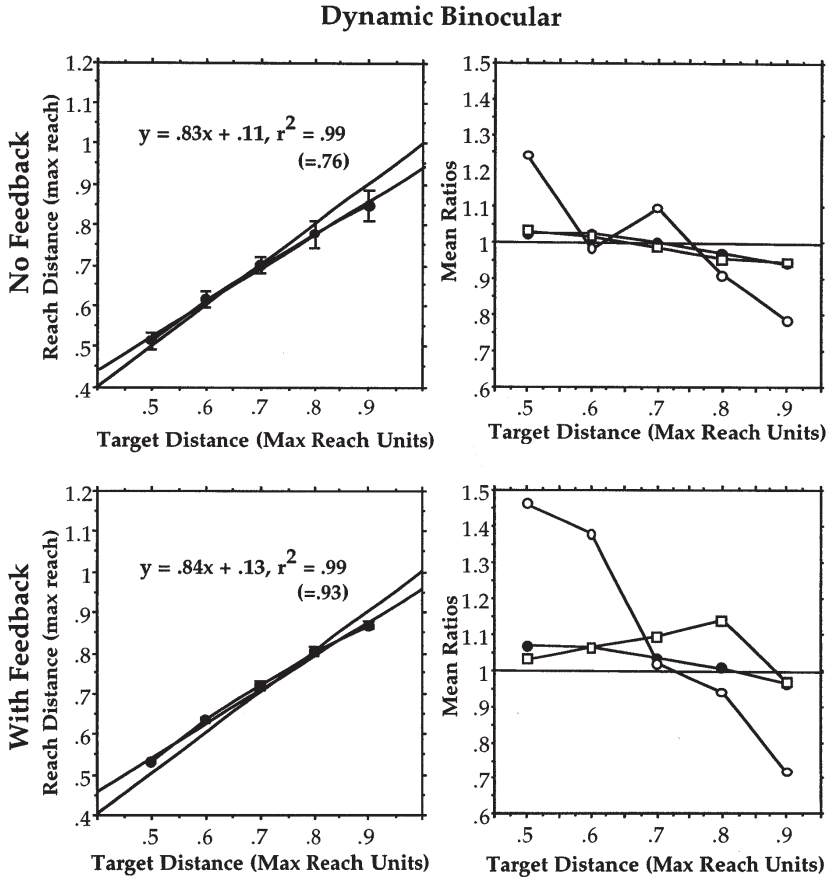

FIGURE 4 Dynamic binocular: The two panels on the left show mean reach distances graphed as a function of actual target distances. Standard error bars represent between-subject variability. The r^2 accompanying the simple regression equations are for the fits to the means. The values in parentheses are for fits to the trial data and provide some measure of the precision. The two right panels show mean width, depth, and distance, each computed in proportion to actual target values and graphed as a function of target distance. The accurate value of the ratio judged/actual in each case, as shown in this graph, is 1. Width: open squares. Depth: open circles. Distance: filled circles. Top panels: without feedback. Bottom panels: with feedback.

The result was significant both before, $p < .001$, $F(15, 784) = 140.7$, $r^2 = .73$, and after nonsignificant factors were removed, $p < .001$, $F(7, 792) = 655.2$, $r^2 = .73$, and the significant factors were target distance ($p < .001$, partial $F = 2045.7$), viewing ($p < .001$, partial $F = 29.8$), target distance by viewing ($p < .001$, partial $F = 18.8$), target distance by feedback ($p < .05$, partial $F = 5.2$), viewing by feedback ($p < .01$, partial $F = 7.4$), block by feedback ($p < .05$, partial $F = 7.4$), and target distance by block by feedback ($p < .05$, partial $F = 6.6$). Block was a significant fac-

tor in the feedback condition as performance improved over blocks. With feedback, the slope and r^2 in the binocular condition were .83 and .85, respectively, and in the monocular condition, they were .75 and .82, respectively.

Next, we compared performance in the static binocular and dynamic binocular conditions. We performed a multiple regression on reach distances with target distance, viewing (static vs. dynamic, coded as ± 1), feedback (coded as ± 1), block, 6 two-way, 4 three-way, and 1 four-way interaction vectors as independent variables. The result was significant ($p < .001, F(15, 634) = 154.3, r^2 = .78$) with all factors and with nonsignificant factors removed, $p < .001, F(3, 646) = 770.9, r^2 = .78$. The significant factors were target distance ($p < .001$, partial $F = 2043.6$), target distance by block ($p < .01$, partial $F = 7.2$), and target distance by block by feedback ($p < .005$, partial $F = 10.6$). Notably, neither viewing nor any of the interactions with viewing were significant. We performed separate analyses in each feedback condition and found that block appeared in a significant interaction with target distance only in the no feedback condition. Target distance was the only significant factor with feedback. Performance was otherwise the same in both viewing and both feedback conditions.

Finally, we computed the absolute reaching errors proportional to target distances. The overall mean proportional absolute errors were as follows: for dynamic monocular vision without feedback, 13% (5.2 cm); for static binocular vision without feedback, 11% (4.4 cm); for dynamic binocular vision without feedback, 8% (3.2 cm); for dynamic monocular vision with feedback, 8% (3.2 cm); for static binocular vision with feedback, 7% (2.8 cm); and for dynamic binocular vision with feedback, 5% (2.0 cm).[4] The error found in Bingham, Bradley, et al. (2001) for both monocular and binocular vision without feedback was 16%, which is comparable to what was found here for monocular vision.

In summary, without feedback, performance was better when participants used binocular vision. This indicated that the experience of targets at multiple distances was important to enable adaptation of the vergence reference level for effective binocular vergence. The difference in performance level between the monocular and binocular conditions and the fact that the performance level with monocular vision was comparable to that found without a size gradient suggested that a size gradient did not greatly contribute to the improved performance level with binocular vision. With feedback, performance was better than without feedback and performance was better when participants used binocular vision. This met our expectation that disparity matching would yield superior feedback and thus, more accurate calibration.

Width

Width and depth were analyzed as proportions of the target values as shown in Figures 2, 3, and 4, where means were plotted together for comparison with mean reach distance as a proportion of target distance. Width is a measure of size perception.

[4]The overall average target distance was 40 cm (= .70 × 57 cm). The values in centimeters were derived by multiplying the given percentages by 40 cm.

Without feedback, size was overestimated by about 30% (2.4 cm) in the monocular condition and by 24% (1.9 cm) in the static binocular condition; it was accurate in the dynamic binocular condition.[5] The estimates of size did not vary with target distance in any condition. If size estimates were governed by distance estimates using image sizes (and thus, the size gradient), then the width ratios plotted in Figures 2 and 3 should have been coincident with the distance ratios. Clearly, they were not. Widths were overestimated, but they did not vary with target distance. This implies that the perception of distance and size were not coupled. Nevertheless, widths became more accurate with feedback just as did reach distances. We performed a multiple regression on width ratios comparing performance in the dynamic monocular and static binocular conditions without and with feedback. If participants were always accurate, then this analysis would account for no variance and would not be statistically significant because we kept object size constant. Nevertheless, the analysis was significant both with all factors included, $p < .001$, $F(15, 784) = 4.6$, $r^2 = .08$, and with nonsignificant factors removed, $p < .001$, $F(2, 797) = 32.4$, $r^2 = .08$, and the significant factors were feedback ($p < .001$, partial $F = 50.0$), and target distance by block ($p < .001$, partial $F = 14.7$). Notably, there was no main effect of target distance in this analysis nor was there a significant target distance by feedback or target distance by viewing interaction. In contrast to the other two viewing conditions, widths in the dynamic binocular condition were accurate on average, both without and with feedback. We performed a multiple regression on widths in this condition with target distance, block, feedback, and the various interaction vectors as factors and the result failed to reach significance ($p > .05$).

To analyze precision, we computed standard deviations of widths for each viewing and feedback condition and for each target distance and participant. The mean standard deviation for all three viewing conditions in both feedback conditions was 24% (1.9 cm).

So the accuracy of size perception was greater on average using dynamic binocular vision. Feedback improved accuracy and did so equally for dynamic monocular and static binocular vision. Size perception did not interact with distance perception. In all viewing and feedback conditions, distance perception was accurate on average for targets at about 80% of maximum reach. Size perception remained as inaccurate at this distance as at other distances. The results indicate that size and distance are relatively independent and imply that a size gradient does not play a strong role in distance perception.

Depth

The plots of mean depth in Figures 2, 3, and 4 exhibit a very different pattern than do the plots of mean width. Whereas widths did not vary with target distance, depths

[5]The actual measured target size was 8.0 cm (7.0 cm plus 1 cm for twice the radius of the stylus when placed tangent to either side of the target). The position of the central axis of the stylus was recorded. The values in centimeters were derived by multiplying the given percentages by 8.0 cm.

did, in all viewing conditions, both without and with feedback. Once again, we performed a multiple regression with the expectation that if participants were accurate, the analysis should be nonsignificant. A multiple regression performed on depths in the dynamic monocular and static binocular conditions was significant with all factors included, $p < .001$, $F(15, 784) = 5.7$, $r^2 = .10$, and with nonsignificant factors removed, $p < .001$, $F(2, 797) = 34.7$, $r^2 = .08$, and the significant factors were target distance ($p < .001$, partial $F = 63.6$) and feedback by viewing by block ($p < .05$, partial $F = 6.0$). The slope as a function of target distance was –1.59. This did not vary as a function of either viewing or feedback conditions.

We performed a multiple regression on depth ratios in the dynamic binocular condition and the result was significant with all factors included, $p < .001$, $F(7, 292) = 12.0$, $r^2 = .21$, and with nonsignificant factors removed, $p < .001$, $F(5, 294) = 15.3$, $r^2 = .21$, and the significant factors were target distance ($p < .001$, partial $F = 44.6$), feedback ($p < .01$, partial $F = 6.8$), block ($p < .001$, partial $F = 14.4$), target distance by block ($p < .001$, partial $F = 15.8$), and target distance by feedback ($p < .05$, partial $F = 5.3$). The slope as a function of target distance increased from –1.00 without feedback to –1.93 with feedback. Feedback actually increased the inaccuracy of the judged depths.

To analyze precision, we computed standard deviations for depths just as we had for widths. The precision of depths was consistently less than that of widths. The mean for monocular viewing with no feedback was 50% (4.0 cm); with feedback it was 39% (3.1 cm). The mean standard deviations for depths in static and dynamic binocular conditions did not vary with feedback condition. They were 46% (3.7 cm) and 42% (3.4 cm), respectively.

Depths were both overestimated and underestimated, but even when overestimated, they often were not overestimated as much as widths were. The result was that shapes were generally compressed in depth relative to width. The width to depth ratios were greater than 1. Width did not vary with target distance and depth did. The result was that compression increased with distance. This is the same as found, for instance, in Norman, Todd, Perotti, and Tittle (1996) and Johnston (1991). Furthermore, these relations were not improved by feedback. Thus, the positional feedback failed to calibrate shape perception. The implication is that shape and position (or distance) are independent.

Viewable Depth

Did our measure of depth really reflect shape perception? The next analysis showed that it did. We had measured reaches to the back of the spherical targets under the assumption that they would reflect shape perception based on optical information projected from the visible front of the objects. To check this assumption, we recomputed depths using the mean X coordinate of reaches to the (visible) left and right of each target, subtracting this from the front and doubling it to yield viewable depth. The viewable depth means are shown in Figure 5 plotted together with the depths from

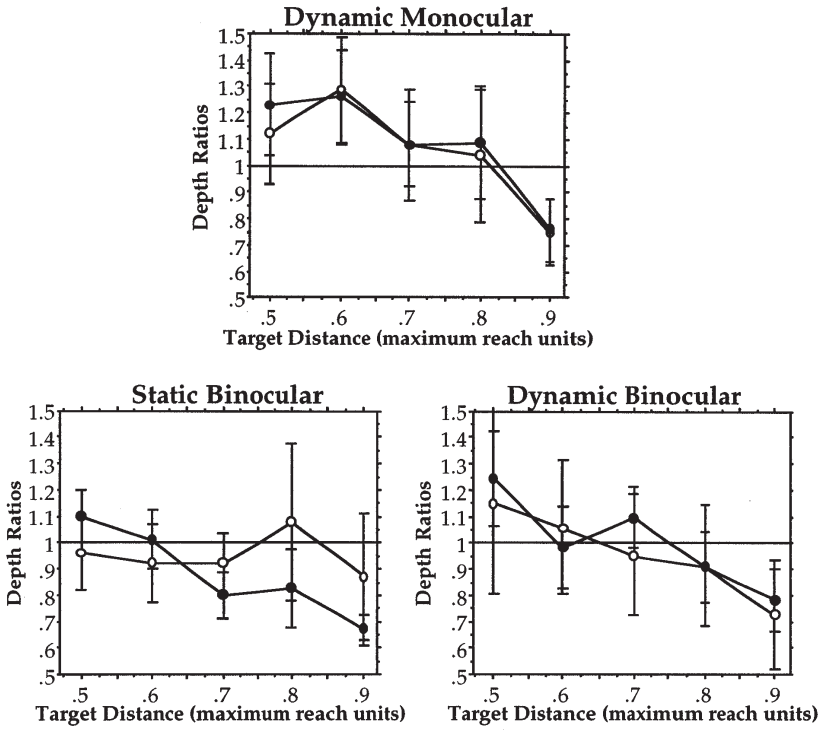

FIGURE 5 Depth and viewable depth: Mean depth and viewable depth values with standard error bars representing between-subject variability. Only data from the no feedback conditions are shown. Depth: filled circles. Viewable depth: open circles.

Figures 2, 3, and 4 (only for the no feedback conditions). We performed a separate analysis in each viewing condition to test whether the two measures were the same as suggested by Figure 5. In each case, we first performed a multiple regression regressing target distance, feedback, block, and interaction vectors on the ratio of the two measures: viewable depth and depth. If the measures are the same, then the ratio should always be 1. The multiple regression failed to reach significance in all three viewing conditions: dynamic monocular, $p > .6$, $F(7, 409) = 0.7$, $r^2 = .01$, mean ratio=.96; static binocular, $p > .5$, $F(7, 335) = 0.8$, $r^2 = .02$, mean ratio=.72; dynamic binocular, $p > .06$, $F(7, 287) = 1.9$, $r^2 = .05$, mean ratio =.84. The ratios were consistent over the various conditions, indicating that the two measures were the same. Second, we regressed depth on viewable depth with feedback, block, and interaction vectors as factors. In the dynamic monocular condition, the result was significant with all factors, $p < .001$, $F(7, 442) = 72.0$, $r^2 = .53$, and with nonsignificant factors removed, $p < .001$, $F(1, 448) = 509.0$, $r^2 = .53$, and the only significant factor was depth. The relation was viewable depth = 1.0 × depth – .10. In the static binocular

condition, the result was significant with all factors, $p < .001$, $F(7, 342) = 41.8$, $r^2 = .46$, and with nonsignificant factors removed, $p < .001$, $F(1, 348) = 286.0$, $r^2 = .45$, and the only significant factor was depth. The relation was viewable depth = 1.2 × depth – .33. In the dynamic binocular condition, the result was significant with all factors, $p < .001$, $F(7, 292) = 18.5$, $r^2 = .31$, and with nonsignificant factors removed, $p < .001$, $F(3, 296) = 41.8$, $r^2 = .30$, and the significant factors were depth ($p < .001$, partial $F = 121.8$), depth by feedback ($p < .05$, partial $F = 4.1$), and feedback by block ($p < .05$, partial $F = 5.8$). The relation in the no feedback condition was viewable depth = 1.08 × depth – .13, and in the feedback condition, it was viewable depth = 0.80 × depth – .13. Given this difference, we tested viewable depth in the dynamic binocular condition, regressing target distance, feedback, block, and interaction vectors on it, but the only significant factor was target distance: viewable depth = –1.35 × target distance + 1.8. Thus, feedback in this condition only affected the placement of the stylus at the back of the targets relative to the front (making the distortion worse) and did not affect placement to the sides. Either way, the depths and thus the shape did not improve in accuracy as did distance and size.

Notably, however, the precision for viewable depth was considerably less with static than with dynamic binocular vision. We performed a mixed design analysis of variance on standard deviations with viewing condition (static vs. dynamic) as a between-subjects factor and measure (depth vs. viewable depth), feedback (without and with), and target distance as within-subject factors. Both measure, $p < .001$, $F(1, 11) = 24.2$, and the measure by viewing interaction, $p < .05$, $F(1, 11) = 4.8$, were significant. The static and dynamic means for depth were 46% (3.7 cm) and 42% (3.4 cm), respectively, and for viewable depth, they were 76% (6.1 cm) and 54% (4.3 cm). Determining the location of the sides of the objects was exceptionally difficult using static binocular vision.

In summary, perceived depth varied with target distance. Perceived depths were generally less than widths, so shapes were compressed in depth and increasingly so at greater distances. Depth estimates were not improved by feedback. These results were confirmed using a second measure of depth that involved only visible positions on the targets, that is, the difference between reaches to the front and sides. This yielded essentially the same results as did the difference between reaches to the front and back. Like accuracy, precision was unaffected by feedback. Together these results indicate that shape perception is not calibrated when position perception is calibrated, and therefore they suggest that shape perception does not reduce to position perception.

GENERAL DISCUSSION

We investigated whether distance, size, and shape perception covary and, in particular, whether a size gradient improves distance perception, and whether shape perception is improved by positional feedback, as is distance perception.

Testing Vision Without Feedback

First, we had found in two previous studies (Bingham, Bradley, et al., 2001; Bingham et al., 2000) that feed-forward reaches performed with binocular vision were not more accurate than reaches performed with monocular vision. This was surprising, especially given the results of Tresilian et al. (1999), Mon-Williams and Dijkerman (1999), Mon-Williams and Tresilian (1999b), Mon-Williams et al. (in press), and Bingham and Pagano (1998), which all had shown that binocular vision yielded accurate feed-forward reaching. We tested two alternative hypotheses. One hypothesis was that the difference in performance was attributable to a size gradient present in all the studies yielding good performance. The size gradient was present because a target object of a single size was tested at a range of distances. The expectation was that the size gradient should allow equal improvements with monocular and binocular vision, if the size gradient is in fact used to better determine distances.

A second hypothesis was that binocular performance had been poor because only a single target distance had been visible and tested. Previous studies of binocular vergence have shown that vergence is evaluated in terms of an adaptable reference vergence distance (Brenner & Van Damme, 1998; Mon-Williams & Tresilian, 1999a, 1999b; Mon-Williams et al., in press; Owens & Liebowitz, 1976; von Hofsten, 1976, 1979). Repeated observation of an object at a single distance with no other object visible would attract the reference vergence to that distance, rendering vergence as rather poorly specified information about distance. If this were the case, then the testing of multiple distances should yield improvements with binocular, but not monocular vision.

The results supported the second hypothesis but not the first one. First, performance in the monocular no feedback condition was not substantially better than found in Bingham, Bradley, et al. (2001). Second, performance using binocular vision without feedback was better than in Bingham, Bradley, et al. (2001) and in this study, it was better than when participants used monocular vision. Performance with static binocular vision was equal to that found by Tresilian et al. (1999), who tested static binocular vision in an actual environment. Third, errors in distance did not yield corresponding errors in size. When reach distances were compared to actual target distances, the slope was significantly different from 1 and the intercept was significantly different from 0. Given these errors in distance perception, the size–distance invariance hypothesis requires that object size should be judged to be different at the different distances. It was not. Therefore, participants were not using the size gradient to apprehend distance. This result makes sense. Consider that if we had manipulated object size, there would have been no information other than distance information to enable participants to distinguish among such objects. So, using a size gradient to determine distance would be a very bad strategy in this case.

Shape Perception and Calibration of Positions

We investigated whether visual feedback about position could be used to calibrate shape perception as well as distance and size perception. We found that size (i.e.,

width) perception did not vary in accuracy as a function of target distance, but object depth was both overestimated and underestimated in a way that varied inversely with target distance. Unlike distance and size perception, perception of object depth was not made accurate by calibration; it even got worse with binocular feedback. Depths were nearly always estimated to be less than widths, with the result that shapes were perceived to be compressed in depth and increasingly so at greater distances. This remained so with feedback.

Using reaches to an occluded position as a measure of perceived object depth entailed the assumption that object shape (and therefore, depth) is specified by the visible portions of the object. (The position of the back of the object must be specified by the combination of object distance, size, and shape.) To control for this assumption and to check the use of position feedback with visible positions, we used a second measure of perceived object depth, viewable depth, which was computed using reaches to the visible sides of the object rather than to the occluded back. The results were essentially the same both with and without feedback. Perceived shape was compressed and remained unchanged by feedback even though perceived distances and sizes became more accurate.

The results show that distance, size, and shape perception are relatively independent. Distance and size are calibrated by feedback. Shape is not. Distance and shape errors covary with actual target distance. Size errors do not. Additional evidence for the independence of position and shape perception has been presented by Crowell, Todd and Bingham (2000) and Bingham, Crowell, and Todd (2001). Furthermore, Norman and Todd (1996) showed that an observers' ability to discriminate higher-order shape properties (i.e., differences in surface orientation) is more precise than the ability to discriminate differences in surface positions. This result is consistent with those found in this work.

The performance level found in this study when participants used binocular vision and feedback was good for distance and size perception, certainly good enough to support common acts of reaching and grasping. The puzzling result at this point is perceived shape. The consistent finding has been that shape perception is inaccurate and imprecise. We have shown that the perception of position at the back of an object is a function of perceived shape using information projected from the visible front of the object. This is the information that must be used to target the fingers in a typical grasp. The level of inaccuracy and imprecision found in these studies is not consistent with results from studies investigating the accuracy and variability of grasping (Paulignan & Jeannerod, 1996; Zaal & Bootsma, 1993). On the other hand, grasping has typically been studied by requiring participants to grasp target objects side to side (or to grasp flat disks front to back, in which case the position of the back is visible). A soda can, on the other hand, is typically grasped by placing the thumb at the front and the fingers to the occluded back. Perhaps other sources of information not yet studied by us may allow more accurate apprehension of object shapes, or perhaps more typical styles of grasping simply are not as accurate.

ACKNOWLEDGMENT

This work was supported by Grant R01 EY11741-01A1 from the National Eye Institute.

REFERENCES

Bingham, G. P., Bradley, A., Bailey, M., & Vinner, R. (2001). Accommodation, occlusion and disparity matching are used to guide reaching: A comparison of actual versus virtual environments. *Journal of Experimental Psychology: Human Perception and Performance, 27, 1314–1344.*

Bingham, G. P., Crowell, J. A., & Todd, J. T. (2001, May 6). *Distortions of distance and shape do not reflect a single continuous transformation on reach space*. Paper presented at the annual meeting of the Vision Sciences Society, Sarasota, FL.

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., 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.*

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

Crowell, J. A., Todd, J. T., & Bingham, G. P. (2000, November 19). *Distinct visuomotor transformations for visually guided reaching.* Paper presented at the annual meeting of the Psychonomic Society, New Orleans, LA.

Ferris, S. H. (1972). Motion parallax and absolute distance. *Journal of Experimental Psychology, 95, 258–263.*

Iberall, T., Bingham, G. P., & Arbib, M. A. (1986). Opposition space as a structuring concept for the analysis of skilled hand movements. *Experimental Brain Research Series 15.* Heidelberg, XXXXX: Springer-Verlag.

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

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). An ordinal role for accommodation in distance perception? *Ergonomics, 43,* 391–404.

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

Mon-Williams, M., Tresilian, J. R., & Hasking, P. (in press). Reduced cue distance perception: A role for vergence and memory. *Perception.*

Norman, J. F., & Todd, J. T. (1996). The discriminability of local surface structure. *Perception, 25,* 381–398.

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

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

Paulignan, Y., & Jeannerod, M. (1996). Prehension movements. In A. M. Wing, P. Haggard, & J. R. Flanagan (Eds.), *Hand and Brain* (pp. 265–282). San Diego: Academic.

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

Poom, L., & Börjesson, E. (1999). Perceptual depth synthesis in the visual system as revealed by selective adaptation. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 504–517.

Richards, W. (1985). Structure from stereo and motion. *Journal of the Optical Society of America A, 2,* 343–349.

Smith, L. B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: The task dynamics of the A-not-B error. *Psychological Review, 106,* 235–260.

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.

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

von Hofsten, C. (1976). The role of convergence in visual space perception. *Vision Research, 16,* 193–198.

von Hofsten, C. (1979). Recalibration of the convergence system. *Perception, 8,* 37–42.

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

Zaal, F. T., & Bootsma, R. J. (1993). Accuracy demands in natural prehension. *Human Movement Science, 12,* 339–345.