

Reaching for visual cues to depth: The brain combines depth cues differently for motor control and perception

David C. Knill

Center for Visual Science and
Department of Brain and Cognitive Sciences,
University of Rochester, Rochester, NY, USA



Vision provides a number of cues about the three-dimensional (3D) layout of objects in a scene that could be used for planning and controlling goal-directed behaviors such as pointing, grasping, and placing objects. An emerging consensus from the perceptual work is that the visual brain is a near-optimal Bayesian estimator of object properties, for example, by integrating cues in a way that accounts for differences in their reliability. We measured how the visuomotor system integrates binocular and monocular cues to 3D surface orientation to guide the placement of objects on a slanted surface. Subjects showed qualitatively similar results to those found in perceptual studies—they gave more weight to binocular cues at low slants and more weight to monocular cues like texture at high slants. We compared subjects' performance in the visuomotor task with their performance on matched perceptual tasks that required an observer to estimate the same 3D surface properties needed to control the motor behavior. The relative influence of binocular and monocular cues changed in qualitatively the same way across stimulus conditions in the two types of task; however, subjects gave significantly more weight to binocular cues for controlling hand movements than for making explicit perceptual judgments in these tasks. Thus, the brain changes how it integrates visual cues based not only on the information content of stimuli, but also on the task for which the information is used.

Keywords: visuomotor control, cue integration, perception and action, reaching and grasping, stereopsis, texture, binocular vision

Introduction

Consider the most mundane motor tasks: reaching to pick up an object, touching a button on the television, or placing an object down on a slanted surface. All of these tasks require the brain to integrate diverse visual cues about the three-dimensional (3D) geometry of objects to generate appropriate motor commands. Recent perceptual work has demonstrated that human observers make judgments about object size, shape, and orientation by integrating visual cues in close to a statistically optimal way (Hillis, Watt, Landy, & Banks, 2004; Jacobs, 1999; Knill & Saunders, 2003; Saunders & Knill, 2001). They rely more heavily on whatever cues are most reliable in a given stimulus. For example, under some conditions, monocular cues for 3D surface orientation (e.g., surface texture and the outline shape of a figure) are more reliable than binocular cues; subjects correspondingly give more weight to those cues when making surface orientation judgments (Hillis et al., 2004; Knill & Saunders, 2003). Similar results have also been reported for integrating information from different sensory modalities (e.g., touch and sight) (Alais & Burr, 2004; Battaglia, Jacobs, & Aslin, 2003; Ernst & Banks, 2002; van Beers, Sittig, & van der Gon, 1999). These kinds of results suggest that the structure of the information in the visual stimulus is the principle factor determining how the brain integrates sensory cues.

Research on cue integration has generally relied on subjective reports of an observer's perceptual experience; thus, it does not directly address how the brain combines visual depth cues to control motor behavior. Early studies comparing motor performance when viewing objects with one or two eyes suggested that the brain relies primarily on binocular information (e.g., retinal disparities) to control goal-directed hand movements in depth (Marotta, Behrmann, & Goodale, 1997; Servos, Goodale, & Jakobson, 1992). A number of more recent studies, however, have shown that the brain can accurately control some aspects of hand movements when only one eye is open. While pre-shaping hand grip is significantly affected by closing one eye, the kinematics of hand transport (how the hand moves from one point to another) is relatively unaffected, at least in some stimulus conditions (Watt & Bradshaw, 2000, 2003). Similarly, the kinematics of subjects' movements to place an object on a slanted surface were qualitatively similar when only texture cues are available and when only binocular cues (e.g., disparity) are available (Knill & Kersten, 2003). Just as importantly, subjects' accuracy in orienting the object prior to placement was approximately as accurate in these two conditions. Little improvement was seen when stimuli contained both reliable texture cues and binocular cues. This suggests that good texture cues to slant can be effective in driving aspects of motor behavior (e.g., hand orientation) that depend on slant information.

While these studies question the idea that binocular cues are predominant for visuomotor control, they do not in fact provide direct evidence bearing on the question of how binocular and monocular cues contribute to motor control when they are both available in a stimulus, as is usually the case. Knill and Kersten (2003) showed that accuracy in the object placement task they studied was significantly impacted by motor noise, making it a poor probe into how visual cues are integrated when estimating slant for motor control. Several studies have shown that subjects give significant weights to monocular cues for making some types of perceptual judgments about an object (e.g., orientation and curvature) (Buckley & Frisby, 1993; Frisby & Buckley, 1992; Hillis et al., 2004; Johnston, Cumming, & Landy, 1994; Knill & Saunders, 2003; Tittle, Norman, Perotti, & Phillips, 1998); however, whether these results can be generalized to performance on motor tasks is an open question. The current experiments apply the cue perturbation paradigm commonly used in perceptual studies to quantify how the brain integrates cues to depth for controlling a simple goal-directed hand movement. In the standard perceptual paradigm, an experimenter would present stimuli that contain small conflicts between cues and then correlate subjects' judgments of depth (or curvature, slant, etc.) with the particular values suggested by the individual cues. Our analysis correlated measurements of hand movements with the information provided by each cue in cue-conflict stimuli to infer how subjects weighted binocular and monocular cues for controlling their movements. We then compared these "visuomotor" weights to the weights that the brain gives to the cues for making analogous perceptual judgments. At question is whether the way that humans integrate cues depends only on how reliable the cues are or whether it also depends on the behavioral function for which the cues are used; that is, does it depend on the output of the system as well as the input to the system.

Because numerous perceptual studies have shown that human observers give a significant weight to monocular cues for making judgments of a surface's orientation in 3D space (Hillis et al., 2004; Knill & Saunders, 2003; Saunders & Knill, 2001), we studied a motor task that depends on visual estimates of surface orientation: placing an object onto a slanted, planar surface. Figure 1 illustrates the task and the experimental apparatus. Subjects viewed a textured figure in a stereoscopic virtual display (a circular disk in Experiment 1) and were asked to place a cylinder flush onto the surface (a robot arm aligned a real surface with the virtual surface so subjects actually were placing a cylinder onto a real surface). An Optotrak system recorded the positions of infrared markers placed on the cylinder, so we could compute the position and 3D pose of the cylinder in real-time during a trial.

Binocular cues were provided by the vergence angles of subjects' eyes (set by the optical distance of the monitor) and by the pattern of disparities created by viewing the disk through stereoscopic glasses. Monocular cues were provided by the outline shape of the figure (an ellipse for the

circular disk) and the foreshortening of the texture pattern within the disk. To quantify the relative contributions of the cues to controlling the movement, we presented subjects with a subset of stimuli containing small cue conflicts, so binocular and monocular cues suggested slightly different slants (the orientation of the surface away from the frontoparallel). To compute cue weights, we correlated the orientation of the cylinder at different points in time during a movement with the slant suggested by each cue on that trial.

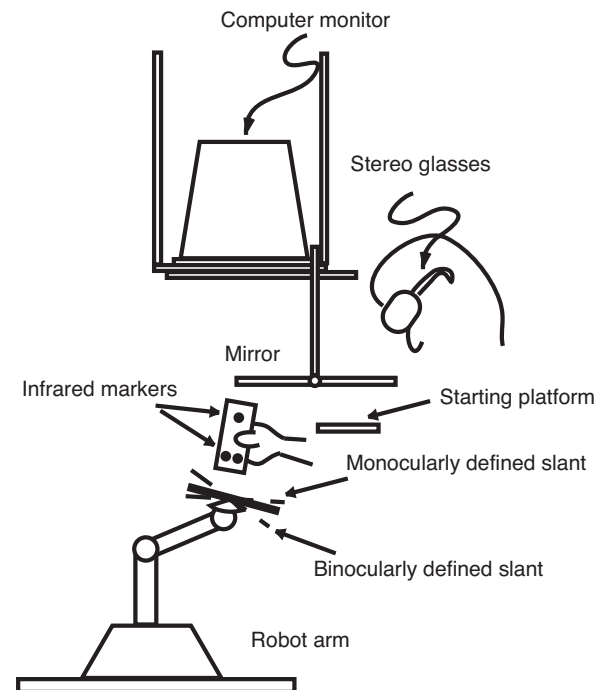


Figure 1. The apparatus used for both experiments reported here. Subjects viewed a stimulus presented on an inverted monitor through a mirror, so a virtual surface appeared under the mirror. On cue-conflict trials, the monocular cues in the stimulus were made to suggest one slant and the disparity cues were made to suggest another slant. Subjects moved a cylinder from a starting platform positioned to the right of the target stimulus to place it flush onto the target. (They had to move the cylinder from right to left to place it on the target surface.) Infrared markers on the cylinder were tracked by an Optotrak system to compute the position and orientation of the cylinder in real-time.

Experiment 1

The first experiment was designed to measure cue weights for stimuli containing presumably strong monocular cues—slanted circular disks filled with randomly tiled textures (see Figure 2). Previous results have shown that the reliabilities of figural cues like texture change markedly with surface slant (Blake, Bulthoff, & Sheinberg, 1993; Knill, 1998) (angle away from the frontoparallel) and that subjects give more weight to texture cues at high slants than low when making perceptual judgments about surface ori-

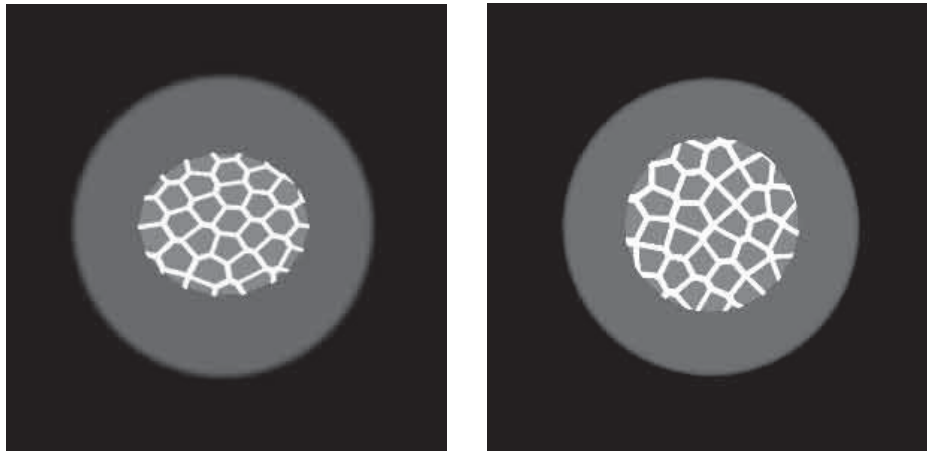


Figure 2. Examples of stimuli used in the experiment. Because they were close to the eyes, the aperture appeared blurred when subjects focused on the target figure.

entation (Knill & Saunders, 2003; Hillis et al., 2004). We therefore measured cue weights for stimuli around two different slants, 20° and 40° . To assess whether visual feedback from the moving cylinder had an impact on the relative contributions of binocular and monocular cues, we ran subjects in both open loop and closed loop conditions. In the closed loop condition, we rendered a virtual cylinder co-aligned with the real cylinder as it moved within the workspace. Because we had subjects view stimuli through circular occluders to eliminate contextual cues provided by the monitor, feedback was still limited to the terminal phase of movements. In the open loop condition, we did not render the cylinder and subjects saw only the target surface.

Methods

Visual stimuli

Visual displays were presented on a computer monitor viewed through a mirror (Figure 1) using Crystal Eyes shutter glasses to present different stereo views to the left and right eyes. Displays had a resolution of 1024×768 pixels and a refresh rate of 118 Hz (59 Hz for each eye's view). Stimuli were drawn in red to take advantage of the comparatively faster red phosphor of the monitor and prevent interocular cross-talk. Figure 3 shows a view of the virtual display from the top and side, with important dimensions and distances indicated on the figure.

Subjects viewed the virtual display through a pair of adjustable, circular occluders positioned in front of both eyes and adjusted so the subject could just see a circle of radius 8 cm (15.6° of visual angle) through each eye separately. This eliminated contextual cues to the orientation of the CRT screen in space. The target surface was rendered in a position centered on the center of the virtual image of the CRT in 3D space. Stimuli consisted of planar, circular disks filled with random Voronoi textures. The disks had a radius of 6 cm, so the horizontal extent of the projected

figures subtended 11.9° from the point of view of a subject (because the figure was always rotated around the horizontal axis). The sizes of the figures along the vertical dimension varied from trial to trial as a function of the slant of the stimulus. For the cue-conflict stimuli, the vertical extent of the figures depended on the slant specified for the mo-

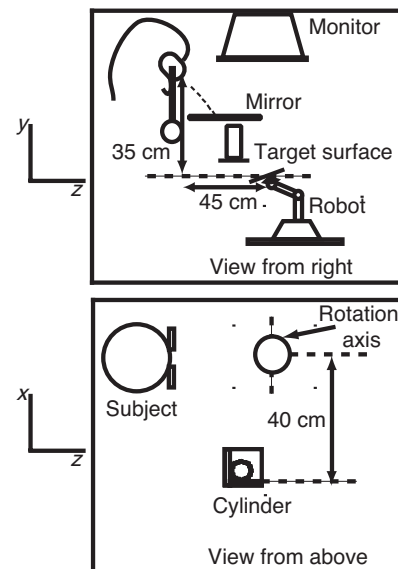


Figure 3. Side and top views of the physical arrangement of the experimental apparatus. The target surface appeared at a distance of 57 cm from the viewer. We defined the slant of the surface as its orientation around the horizontal (x) axis relative to the subjects' line of sight. A slant of 0° was defined as frontoparallel; a slant of 90° would have been an edge-on view of the target surface. Given the viewing geometry, a horizontal (tabletop) surface would have appeared as a slant of 37° . The starting plate, on which the cylinder was placed at the beginning of a trial, was 40 cm to the right of the target surface, 20 cm closer to the subject than the target surface, and 10 cm above the target surface (all measured from the subject's point of view). This placed the cylinder below shoulder level at its starting point.

nocular cue. Textures were created by positioning points in the plane at random using a stochastic diffusion process and then drawing the Voronoi polygons for the resulting random lattice (for a detailed description of the process, see Knill & Saunders, 2003). Twenty different extended Voronoi textures were used in the experiment. On each trial, a sample was taken from a random position within a texture and rotated by a random angle in the plane of the figure. Thus, no texture pattern was repeated exactly in the experiment.

On cue-consistent trials, the textured target shapes were rendered on each trial at the specified slant. Cue conflicts were generated by rendering a distorted copy of the figure and texture at the slant specified for the binocular cue. The figure and texture were distorted so when projected from the binocular slant to a point midway between a subject's two eyes (the cyclopean view), the projected figure and texture suggested the slant specified for the monocular cue on that trial. We determined the figure and texture distortion in two stages. First, we projected the positions of the figure and texture vertices into the virtual image plane of a cyclopean view of a surface with the slant specified for the monocular cue. We then back-projected the projected vertex positions onto a surface with the binocular slant to generate the new, distorted texture vertices.

Spatial calibration of the virtual environment required computing the coordinate transformation from the reference frame of the Optotrak to the reference frame of the computer monitor as well as the location of a subject's eyes relative to the monitor. These parameters were measured at the start of each experimental session using an optical matching procedure. The backing of the half-silvered mirror was temporarily removed, so subjects could see their hand and the monitor simultaneously, and subjects aligned an Optotrak marker to a sequence of visually cued locations. Cues were presented monocularly, and matches were performed in separate sequences for left and right eyes. Thirteen positions on the monitor were cued, and each position was matched twice at different depth planes. The combined responses for both eyes were used to determine a globally optimal combination of 3D reference frame and eye position. After the calibration procedure, a rough test was performed in which subjects moved a marker viewed through the half-silvered mirror and checked that a rendered spot appeared co-aligned with the marker.

The 3D position of the cylinder was tracked in real-time by an Optotrak 3020 system at 120 Hz. Four infrared markers were placed on the cylinder. Using the recorded 3D positions of the markers, we computed the 3D position of the center of mass of the cylinder as well as its orientation in space. The markers were positioned to allow full recovery of the cylinder's 3D pose even when subjects rotated the cylinder to make one or another of the markers invisible to the Optotrak camera.

When subjects were moving the cylinder onto the target, we rendered a cylinder to appear coextensive with the true stimulus. Because of the approximately 1-1/2 video

frame (25 ms) delay between measurement of marker positions on the cylinder (see below) and the appearance of the cylinder in the virtual image, we used linear extrapolation of the position and orientation of the cylinder from previous frames to predict the position and orientation of the cylinder at the time it appeared in the display. Except at the very end of a movement, when accelerations were high, this procedure left no perceptually detectable visual error between the image of the real cylinder (when viewed through a half-silvered mirror) and the virtual image of the cylinder.

Apparatus

Figure 3 shows the geometry of the physical apparatus. Subjects started a trial by placing the cylinder on the starting plate. They tucked the cylinder into a notch at the back corner of the plate, so the starting position was the same on every trial. A PUMA 260 robot arm positioned a circular metal plate (the target surface) to be coextensive with the virtual image of the figure on each trial with a random variation of $\pm 2^\circ$ added to the slant of the target surface. On cue-consistent trials, the random slant perturbations were added to the simulated slant of the stimulus, whereas on cue-conflict trials, they were added to the slant midway between the monocular and binocular slants. On cue-conflict trials, the slants suggested by the binocular and monocular cues differed by 4° ; therefore, the $\pm 2^\circ$ variation was equivalent to positioning the target plate at a slant chosen from a uniform distribution within the interval defined by the monocular and binocular slants. A metal plate attached to the bottom of the cylinder was connected to a 5-volt source. The metal plates on the starting and target surfaces were connected through parallel resistor circuits to ground, so when the circuit between the cylinder and one of the plates was closed, the voltage input to an A-to-D port flipped from 5 to 0 volts. By reading the signal levels at the two ports connected to the starting plate and target plate, respectively, we were able to determine when the cylinder left the starting plate and when it first made contact with the target surface.

Procedure

The beginning of each trial was triggered by the closing of the circuit between the bottom of the cylinder and the starting plate (indicating that the cylinder was on the starting plate). At this point, the robot arm moved the target surface to the chosen orientation and after a period of 1 s, a new target stimulus was displayed. After 750 ms, an audible beep was given to signal the subject to move the cylinder and place it flush onto the target surface. Closing of the circuit between the bottom of the cylinder and the target plate signaled the end of the trial. If the cylinder did not make contact with the target plate within 1-1/2 s of the go signal, two successive beeps were generated to signal an error, and the trial was discarded. The condition for that trial was then randomly swapped with another of the remaining trials. At 1-1/2 s after the go signal, the target stimulus disappeared, signaling to subjects that they could move back

to the starting plate. This process was repeated until the end of a block.

Sixteen slant conditions were used. Eight were cue-consistent conditions in which target stimuli varied from 16° away from the frontoparallel to 44° away from the frontoparallel in 4° increments. Eight were cue-conflict conditions with the following pairs of monocular/binocular slants: $20 / 24$, $20 / 16$, $24 / 20$, $16 / 20$, $40 / 44$, $40 / 36$, $44 / 40$, and $36 / 40$. Subjects ran in four sessions of two blocks each. Each block contained 256 trials (16 trials per condition), giving a total of 128 trials per condition. Two of the sessions were open loop (the cylinder was not rendered) and two were closed loop (the cylinder was rendered). Open loop and closed loop sessions were run in an ABBA order and were counterbalanced across subjects (four subjects ran in the ABBA order and four ran in the BAAB order). Subjects were told that they could take breaks whenever they felt tired by simply holding the cylinder in their lap at the end of a trial (this effectively stopped the progress of the experiment, because new trials were triggered by placement of the cylinder on the starting plate).

Data analysis

The behavioral data for the experiment was provided by the Optotrak recordings of the four markers mounted to the side of the cylinder. These were used to compute the orientation of the cylinder at each time frame of the recording, expressed as its slant (the angle of the main axis of the cylinder out of the frontoparallel plane of the observer) and tilt (the angle of the main axis projected into the frontoparallel plane). As has been shown previously for this task (Knill & Kersten, 2003), the tilt trajectories (tilt as a function of time) did not correlate strongly with the slant of the target surface; thus, our analysis focused on the slant of the cylinder.

Our principle measure of performance on the task was the slant of the cylinder just prior to making contact with the surface. It was therefore critical that we accurately determined the time at which the cylinder first made contact with the target surface. Most important was that we used an estimate of the contact time that was not after the true contact time, because the physical interaction between the cylinder and the target surface would force the slant of the cylinder after contact to the true slant of the surface. Were we to have a late bias in our estimate of the contact time, this effect would bias our cue weight estimates toward being 50/50. The time at which the circuit between the bottom of the cylinder and the target plate first closed provided an initial estimate of the contact time; however, due to the rise time of the voltage signal, this was not perfect. We improved the estimate by using the observation that the acceleration profile of the cylinder showed a spike at contact with the surface (see Figure 4). The time of this spike provided an estimate of the contact time for the cylinder. We searched backward in time from the time marked by the closing of the target-cylinder circuit to find

the first appearance of this spike. We found that it invariably occurred 0-3 Optotrak frames prior to the closing of the circuit. We marked the Optotrak frame just preceding the spike in acceleration as the contact time for the cylinder.

Subjects

Subjects were eight undergraduates at the University of Rochester who were naive to the goals of the experiment. Subjects had normal stereo vision. Data from one subject was uninterpretable (the subject did not change the orientation of the cylinder as a function of target surface slant) and was discarded from the analysis.

Results

To visualize subjects' hand movements in the experiments, we used the movement data from the Optotrak recordings to reconstruct videos of the motion of the cylinder on a sample trial drawn from one subject's data. [Movie 1](#) shows the stimulus as seen from the point of view of the observer during the trial. The apertures are not explicitly drawn in the video, but are apparent from the appearance of the cylinder as it comes within the view through the aperture. [Movie 2](#) was made by from the same data by simulating a camera positioned off to the side of the experimental apparatus. This movie gives a better view of the movement of the cylinder, particularly how the subject rotates it, throughout the trial.

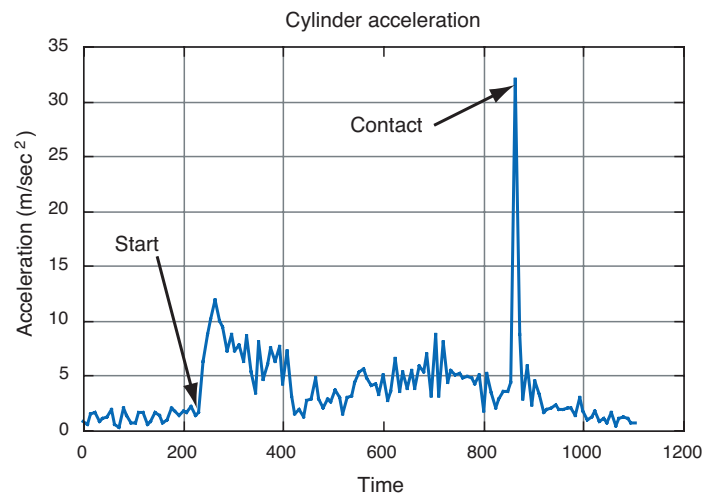
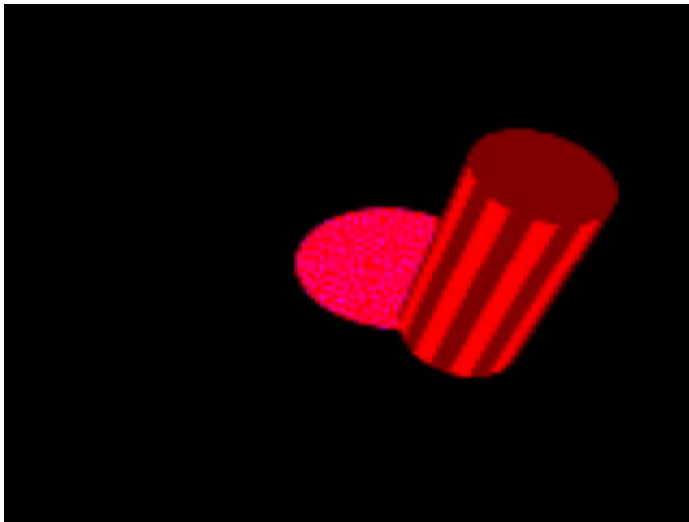
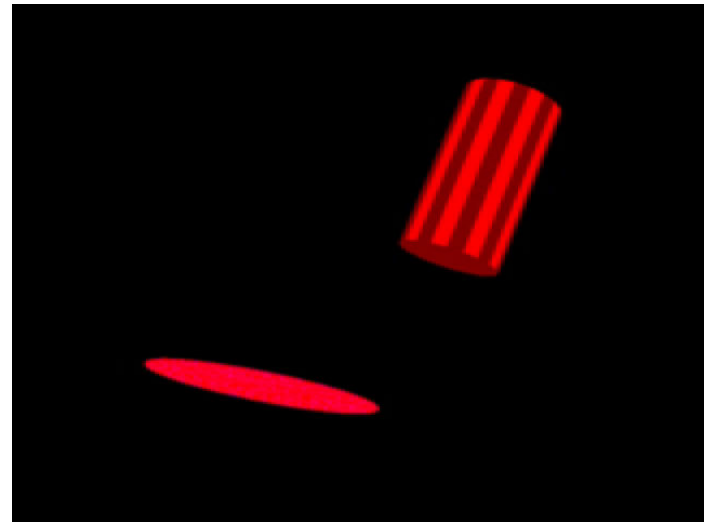


Figure 4. Acceleration profile for the cylinder on one trial. Contact is clearly marked by a sharp spike in acceleration. In this case, the spike appeared two optotrak frames prior to detecting closure of the circuit between the bottom of the cylinder and the target surface. The start of the movement is also clearly apparent in the rise in acceleration just after 200 ms.

Figure 5a plots one subject's average cylinder slant trajectories (the slant of the cylinder as a function of time) for the target stimuli with consistent cues. Each curve shows how the slant of the cylinder changed over time for one of the target surface slants used in the experiment. Figure 5b



Movie 1. The display as seen by a subject during one experimental trial. The movie shows the display exactly as it appeared to the subject, except that it is shown here slowed down by a factor of approximately 4.



Movie 2. The movement of the cylinder on one trial as seen from a viewpoint to the left of the subject.

and 5c and show slant trajectories for the cue-conflict stimuli for the same subject. Each graph shows how changing one of the cues affected the slant trajectories of the cylinder. At 40°, changes in binocular and monocular cues led to approximately equal changes in the subject's movements. At 20°, however, changing the binocular slant led to much larger changes in movements than did changing the monocular slant. This suggests that monocular cues affected the movements more for target surfaces with slants near 40° than for surfaces with slants near 20°. These trajectories were drawn from the closed loop condition. The kinematics of movements in the open loop condition were qualitatively indistinguishable from the closed loop kinematics.

To quantify the relative contributions of the cues to controlling the orientation of the cylinder, we modeled the

slant trajectory of the cylinder on any given trial (the orientation of the main axis of the cylinder out of the frontoparallel plane) as a function of the estimated slant of the target disk plus some independent motor noise. Assuming that cues are combined linearly and that movement trajectories vary linearly within the neighborhood of a given target surface slant, we can write the slant trajectory on a given trial as a weighted sum of the average trajectories that subjects would have generated for target disks at each of the slants suggested by the two cues, $\bar{s}_{\text{cylinder}}(t; \sigma_{\text{bin}})$ and $\bar{s}_{\text{cylinder}}(t; \sigma_{\text{mono}})$, plus some added noise

$$s_{\text{cylinder}}(t) = w_{\text{bin}} \bar{s}_{\text{cylinder}}(t; \sigma_{\text{bin}}) + w_{\text{mono}} \bar{s}_{\text{cylinder}}(t; \sigma_{\text{mono}}) + \Omega(t), \quad (1)$$

where $s_{\text{cylinder}}(t)$ is the observed slant of the cylinder as a

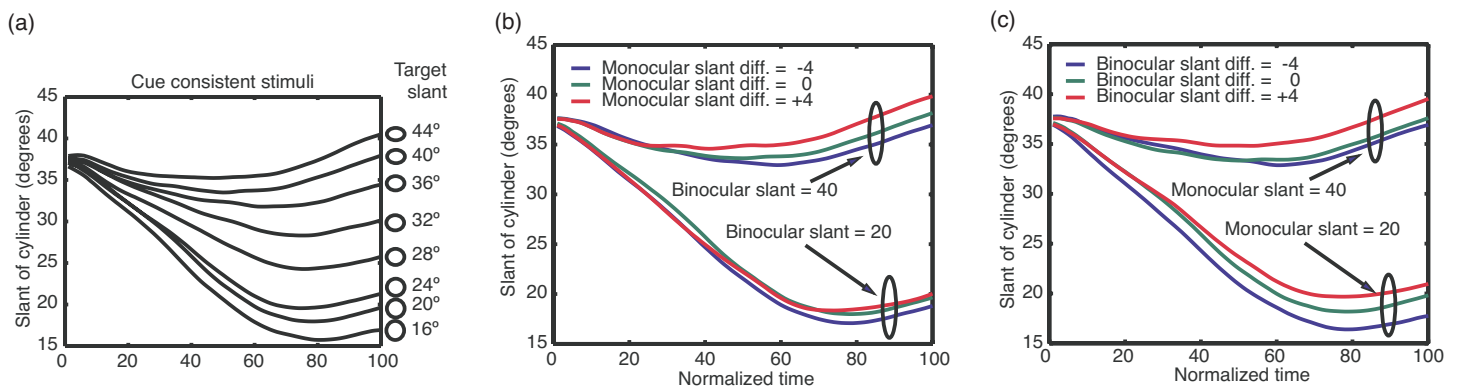


Figure 5. Results from Experiment 1. (a). The average slant of the cylinder (the angle between the main axis of the cylinder and the frontoparallel plane) as a function of time for each of the eight cue-consistent target slant conditions. The ellipses to the right of the figure represent the slant of the target for each average trajectory. Trajectories were averaged by first stretching or compressing each trajectory to a common duration (arbitrarily labeled 100). (b). Average slant trajectories for cue-conflict stimuli in which the monocular cues suggested a fixed slant (either 20° or 40°) and the slant suggested by the binocular cues varied around the monocularly defined slant by $\pm 4^\circ$. (c). The same as (b), but for stimuli in which the binocular cues defined a fixed slant and the monocular cues were made to vary around that slant.

function of time. The noise term, $\Omega(t)$, subsumes time-varying motor noise, trial-to-trial variations in motor strategies, noise due to variability in perceptual estimates of surface slant, and noise in the motion measurements. In this formulation, the perceptual weights, w_{bin} and w_{mono} are constrained to sum to 1.

Fitting the model to the movement data has several difficulties. First, it does not easily accommodate trial-to-trial variability in movement duration, though assuming motor invariance, we can remove this source of variation by normalizing the trajectories to a constant duration (as we did to compute the average trajectories shown in Figure 5). Second, it requires estimating the covariance of the noise process, which is poorly constrained due to the small sample sizes used in the experiment (relative to the dimensionality of the trajectories). Fortunately, we have previously developed methods to deal with this problem and have shown that for this task, the slant of the disk just prior to contact with the target surface (its *contact slant*) contains all of the discriminant information in the trajectory about the slant of the target surface (Knill & Kersten, 2003). We therefore used only the contact slants to fit the relative cue weights that subjects used to estimate target surface slant for generating the movement trajectories. These are given by fitting the linear equation

$$s_{\text{cylinder}}(t_{\text{contact}}) = k(w_{\text{bin}}\sigma_{\text{bin}} + w_{\text{mono}}\sigma_{\text{mono}}) + b \quad (2)$$

to the data, with the sum of cue weights set to 1.

Figure 6 shows the binocular weights computed from subjects' movements for target surface slants at 20° and 40° (the monocular weights are simply 1 minus the binocular weights). The weights reflect the behavior evident in the full slant trajectories: subjects gave more weight to binocular information at low slants than at high slants. A two-way

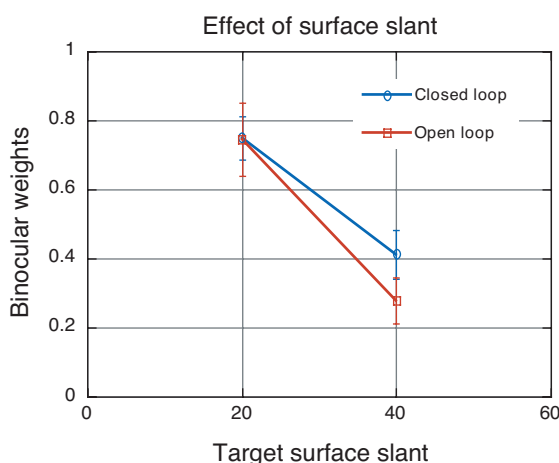


Figure 6. The average weights that subjects gave to binocular cues for orienting the cylinder for target slants of 20° and 40°. The binocular weights are normalized so the binocular and monocular weights sum to 1. A value of .5 indicates equal influences of binocular and monocular cues on the final contact slant of the cylinder.

ANOVA with target slant and feedback condition as factors revealed a significant effect of target slant, $F(1,24) = 30.6$, $p < .0001$, but no significant effect of feedback, $F(1,24) = 0.98$, $p > .33$, and no significant interaction, $F(1,24) = .64$, $p > .41$.

Discussion

Biases that apply to the mapping between perceived slant and contact slant are all absorbed in the constants, k and b , in Equation 2. These include perceptual biases in perceived slant that are independent of visual cues, the impact of other cues (e.g., blur) on estimated slant, and biases in the mapping between perceived slant and subjects movements. Subjects did show movements that regressed toward a slant somewhere in the middle of the full range of surface slants tested in the experiment; however, we cannot distinguish from the data whether these biases are perceptual or motor in origin.

More significant for the interpretation of the cue weights is the fact that cue-specific biases cannot be distinguished from cue weights in a putative cue-integration mechanism. Multiplicative biases in slant estimates from either the binocular or monocular cues change the relative weights fit to the cues. Perceptual experiments consistently show that subjects perceive relative depth from stereopsis at the close distance used here to be magnified (Foley, 1980; Johnston, 1991). This effect would lead to a gain greater than 1 on subjects' estimates of slant from stereopsis, which would in turn be reflected in a greater apparent weight for the binocular cue. Similar data on perceived slant from figural and texture cues are not available; however, work on how humans integrate perceptual disparity and skew symmetry information to estimate surface slant is consistent with the hypothesis that the gain on perceived slant from disparities is greater than the gain on perceived slant from the outline shape of a figure at a viewing distance similar to the one used in the current experiment (Saunders & Knill, 2001). Existing data would therefore suggest that the weights measured here reflect some amount of slant scaling from stereopsis.

The results are qualitatively similar to previous perceptual studies showing an increase in the perceptual weighting of monocular cues at high slants (Hillis et al., 2004; Knill & Saunders, 2003). This is consistent with optimal cue integration, resulting from the fact that relative uncertainty of monocular cues like texture decreases as surface slant increases. The aforementioned depth-from-disparity scaling effects, however, corrupt the expected relationship between measured weights and cue uncertainty. To the extent that the scaling effects are due to overestimates of viewing distance (Johnston, 1991), their impact on measured binocular cue weights is approximately the same at 20° and 40° slant. Thus, it remains the case that the proportional change in binocular cue weights as a function of surface slant is qualitatively consistent with optimal cue integration.

Cue weights were the same in open loop and closed loop conditions; however, this should not be read as evidence that the cues are used similarly in the presence or absence of visual feedback. In our experimental set-up, the cylinder only appeared in view during the last 240-340 ms of each movement, so there was little time, given delays in the sensorimotor loop, for special-purpose visual feedback control processes to affect the outcome of a movement. In the more common situation, in which vision of the hand is continuously available, one might well see the influence of visual feedback mechanisms in subjects' trajectories.

The data from this experiment were too noisy to measure the effects of haptic feedback on cue weights. This could appear in the data in two forms. First, haptic feedback could support adaptive estimation of the viewing distance to the surface, thus normalizing the scaling of disparities discussed above. Second, because the target surface was positioned at a slant midway between the two cues (with random variations), the haptic feedback could have pushed subjects to more of a 1:1 weighting of cues. We consider these effects in more depth in the analysis of results from [Experiment 2](#).

Experiment 2

To directly compare how the brain weights binocular and monocular depth cues for motor control with how it weights them for perception, we measured cue weights for perceptual judgments that had the same informational demands as the visuomotor task and that were performed under the same stimulus conditions. To simplify comparisons across tasks, we reran the visuomotor task using only one viewing condition—the closed loop condition. We chose this over the open loop condition because subjects commented that the open loop version of the experiment seemed less natural. We used two perceptual tasks, a visual matching task, in which subjects aligned a thin, cylindrical probe to appear perpendicular to the target surface ([Figure 7](#)), and a haptic matching task that replicated the motor task in all aspects except that subjects did not place the cylinder onto the target surface; rather, they held the cylinder in position over the target surface and oriented it to the position in which it “felt” perpendicular to the target surface. At this point, they pressed a mouse button and the orientation of the cylinder was recorded. In the haptic matching experiment, we did not render a virtual cylinder as we did in the visuomotor task.

As in [Experiment 1](#), the target surface was shown at a range of slants, but cue conflicts were introduced only around 36° . We chose this slant as a trade-off between two considerations. First, we wanted to use a high enough slant that subjects would give a reasonable weight to the monocular cues presented. Second, we wanted to avoid a “flattening” effect that appeared in the data in [Experiment 1](#); subjects did not seem to vary the orientation of the cylinder much at high slants. This was likely due to the fact that ori-

entering the cylinder to place it on a surface that was slanted away from the subject by much more than a tabletop surface (approximately 37°) felt somewhat awkward. Whatever the cause, it added some uncertainty to our estimates of cue weights around 40° . (The SEs of weight estimates in [Experiment 2](#) were half as large on average as those in [Experiment 1](#) for approximately the same number of trials.)

We varied the reliability of the monocular cues by using two types of target surface: a textured, circular disk (reliable), as in the first experiment, and a textured, randomly shaped figure (unreliable) (see [Figure 7](#)). Because subjects far preferred the closed loop condition in [Experiment 1](#) (it felt much more natural) and no significant effect of feedback was found, we ran subjects in the visuomotor task using the closed loop version of the experiment in which a virtual cylinder was rendered in the workspace.

To avoid across-task learning effects, we used a between-subjects design rather than a within-subjects design. We therefore ran three different groups of subjects in the motor task and the two perceptual matching tasks. Although we would have preferred to use a within-subjects design, we decided against it for several reasons. First, we would not expect learning effects to be symmetric across the order of doing the tasks. The visuomotor task provides haptic feedback in support of learning that is not provided in the perceptual tasks. Thus, one might expect more of a learning effect to appear if the visuomotor task were performed first. Second, performance in the haptic matching task, which is most closely matched to the visuomotor task, could be severely impacted by prior exposure to the visuomotor task. Having run a task in which they placed the cyl-



Figure 7. Stimuli in [Experiment 2](#) included both textured disks slanted in depth (as in [Experiment 1](#)) and textured, randomly shaped figures like the one shown here. In the visual matching task, subjects used the computer mouse to adjust the 3D orientation of a probe (shown here) to appear perpendicular to the textured figure. Both the probe and the figure were presented stereoscopically to subjects. In the haptic matching task, subjects saw just the textured figure.

inder onto a surface prior to running the haptic matching task, subjects might well learn the strategy of mimicking the object placement task when performing the haptic match.

Methods

Stimuli

Stimuli were the same as those used in [Experiment 1](#), with the addition of 20 different randomly shaped, smooth figures to use in the random shape condition. Each figure was generated by randomizing the coefficients of a sum of cosine and sine waves defining the radial distance of the boundary from the center point of the figure. We then computed the second order moments of inertia of each figure and stretched the figures appropriately to make the moments of inertia isotropic (so the best-fitting ellipse to each figure was a circle) and to make the average radius 6 cm. On each trial of the experiment, a figure was randomly chosen from this set and then rotated by a random angle within the plane defined by the monocular slant. These were also filled with random Voronoi textures.

Procedure (object placement task)

Ten slant conditions were used. Six were cue-consistent conditions in which the slant of the target stimuli varied from 20° to 40° in 4-deg increments. Four were cue-conflict conditions with the following pairs of monocular/binocular slants: 36 / 40, 36 / 32, 40 / 36, and 32 / 36. In addition, two different figure conditions were used, corresponding to circular or randomly shaped figures. This gave a total of 20 stimulus conditions.

Subjects ran in two sessions. For four of the subjects, each session contained two blocks of 160 trials (8 trials per condition), giving a total of 32 trials per condition. For the other three subjects, each session contained three blocks of 100 trials (5 trials per condition), giving a total of 30 trials per condition.

Procedure (visual match task)

Stimulus conditions were equivalent to those used in the object placement task. As shown in [Figure 7](#), a probe figure was added to the stimulus, which subjects could rotate using a mouse. The probe was rendered as a 4-mm wide and 6-cm tall cylinder, with balls placed on either end. The balls had a diameter of 8 mm. The bottom of the probe was centered on the center of the target surface. Subjects ran in two sessions of four blocks, each containing 100 trials (5 per stimulus condition) giving a total of 40 trials per condition. On each trial, the target stimulus was presented for 750 ms, after which an audible beep signaled subjects to adjust the probe until it appeared perpendicular to the target surface. Subjects adjusted the probe by moving a mouse over a tabletop surface placed under the mirror. When they were satisfied with a setting, they pressed the mouse button, the stimulus disappeared for 1 s and then a new trial began. If they pressed the mouse but-

ton before the go signal, the trial was discarded. Subjects took an average of 1 to 1-1/2 s to make the adjustment.

Procedure (haptic match task)

Stimulus conditions were equivalent to those used in the object placement experiment. Subjects ran in two sessions of four blocks, each containing 100 trials (5 per stimulus condition) giving a total of 40 trials per condition. The robot arm was removed from behind the mirror and subjects held the cylinder in place behind the mirror. The cylinder was not rendered in the virtual display, so the only information available about the orientation of the cylinder was proprioceptive information from a subject's arm and hand. On each trial, the target stimulus was presented for 750 ms, after which an audible beep signaled subjects to adjust the cylinder until it felt as if it was perpendicular to the target surface. When satisfied, subjects pressed a button on a mouse held in their free hand. If they pressed the mouse button before the go signal, the trial was discarded. Subjects took an average of 1 to 1-1/2 s to make the adjustment.

Subjects

Subjects were 21 undergraduates at the University of Rochester who were naive to the goals of the experiment. Subjects had normal stereo vision. Subjects were split into three groups of seven subjects each. Each group ran in one of the three tasks.

Results

[Figure 8](#) shows the relative binocular cue weights computed for each of the three tasks. A two-way ANOVA showed a main effect of task, $F(2, 36) = 12$, $p < .00001$, and a main effect of figure type, $F(1,36) = 33.85$, $p < .00001$. The interaction was not significant, $F(2,36) = .83$, $p > .44$. Subjects weighted binocular cues on all three tasks more for the random figure than for the circular figure, reflecting the reduced reliability of the figural information provided by the random figures. The more striking result is that for both types of figure, binocular cues contribute much more to subjects' performance in the motor task than they do to their performance in either of the perceptual tasks [post hoc ANOVA, visuomotor vs. visual match, $F(2,24) = 21.42$, $p < .0001$; visuomotor vs. haptic match, $F(2,24) = 21.03$, $p < .0001$]. The brain effectively gives 2.6 times more weight to binocular cues than to monocular cues for controlling the motor task, averaged across the two types of figures used in the experiment. When making perceptual judgments, however, subjects relied more heavily on monocular cues. When averaged across the two types of figures, monocular and binocular cues influenced perceptual judgments almost equally.

The binocular weights for the circle figure in the visuomotor task appear substantially larger than the results from [Experiment 1](#) for the similar, closed loop, 40°-slant condition (.61 vs. .42). The results of [Experiment 1](#), however, suggest a rapid decrease in the binocular weight with

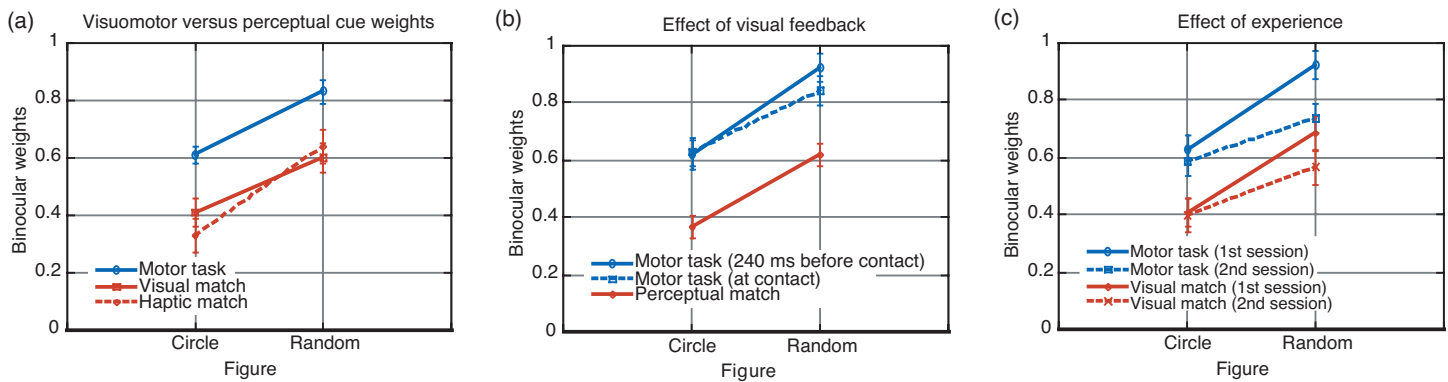


Figure 8. Results from Experiment 2. (a). Plots of the normalized binocular cue weights derived from the object placement data (motor task) and from the two perceptual matching tasks (visual match and haptic match), for both circles and random figures. As expected, subjects gave more weight to binocular cues for the random figures than for the circles. They also gave much more weight to binocular cues for the motor task than for either perceptual task. (b). Plots of cue weights derived from the object placement task using the slant of the cylinder either 240 ms before it made contact with the target surface or just before it made contact. The average cue weights for the two perceptual tasks are shown for comparison. Because the cylinder came into a subject's view between 240 and 340 ms. before the cylinder made contact with the target surface, the first set of cue weights could not reflect any influence of visual feedback (assuming a fast estimate of 80 ms for the visuomotor delay). The data show no significant effect of visual feedback on cue weights, $F(1,16) = 1.1$, $p > .31$. (c). Plots of the cue weights derived from both the cylinder placement task and the visual matching task, computed separately for the first and second sessions of the experiment. Both tasks show equivalent effects of experience on subjects' ability to use the monocular cues for the randomly shaped figures, but the difference between visuomotor and perceptual weights for early and late sessions is unchanged.

increasing slant. A simple linear interpolation of the data from Experiment 1 gives an estimated binocular weight in that experiment at 36° of .48. This estimate, however, is based on the assumption that the weights change linearly. A quadratic change in weights would be more consistent with the results of the two experiments. In experiments reported elsewhere, we have found average cue weights at 35° of .58 in closed loop conditions similar to those used here (Greenwald, Knill, & Saunders, 2004). This is very close to the value of .61 found here.

Discussion

It is possible that subjects performed the visual matching task using a relative orientation judgment (perpendicularity of the surface and probe); however, given that the monocular cues for the orientation of the probe were designed to be weak (only the length of the probe is foreshortened), one would expect such a strategy to have biased subjects to use binocular cues more heavily, the opposite of what was found. Furthermore, the weights derived from the haptic matching task, which requires the same estimate of absolute slant relative to the observer as the motor task, are not significantly different from those for the visual matching task and are much different from the weights derived from the motor task.

The results would appear to imply that the brain uses qualitatively different cue-weighting strategies for motor control than it does for computing perceptual representations; however, we must consider several simpler explanations before accepting this account. The first is that in the

motor control task, subjects had visual feedback from the cylinder in the end stages of the movement, so the difference in cue weighting may have reflected the use of relative disparity information between the cylinder and the target surface to adjust the orientation of the cylinder at the end of the movement. In Experiment 1, the observed differences between the conditions with and without final stage feedback from the cylinder were not significant; however, to reliably discount this explanation, we replaced the contact slant of the cylinder in the linear regression with the slant of the cylinder at a time at which the cylinder would have just come into view (measured backwards from the time of contact with the surface). As shown in Figure 8b, the differences in cue weights measured in this way did not significantly differ from those measured using the contact slant of the cylinder; thus, the increased binocular weighting seen in the motor task cannot be attributed to special-purpose visual feedback mechanisms.

A second concern is that subjects may have adjusted their cue weights over time in the motor task based on haptic feedback from the target surface. Randomization of the slant of the target surface between the slants suggested by the two cues should have minimized the haptic information available for such adaptive learning; however, this information did indicate that each cue was equally correlated with the true slant of the surface. Subjects might have adjusted their cue weights over time to match the haptic feedback. Subjects' cue weights do not accord with this explanation for the differences found here, which would have the binocular weights increase over time from an initial level equal

to the perceptual weights toward 0.5; however, it remains possible that a learning effect could have influenced the results. To measure any such effect, we computed cue weights separately for the first and second days of testing in the object placement and visual matching tasks. Due to high levels of variability in some subjects' haptic judgments in the first session, we were not able to compute a useful measure of the effect of experience on the haptic task. The results are shown in [Figure 8c](#). Subjects show a decrease in the binocular cue weight for the random figures, but the decrease is equivalent for both the visuomotor task and the perceptual matching task. This indicates that passive experience with the random shapes leads to a greater weight being given to figural cues for those figures, but that haptic feedback in the motor task has little to do with the effect. Subjects learned to use the compression of the random figure as a cue to slant over time, perhaps because it was consistently correlated with the binocular cue.

A particular form of learning that could have selectively impacted the weights measured in the visuomotor task is adaptive scaling of the viewing distance used to determine surface slant from disparity information. The measured cue weights include multiplicative biases in the estimated slant from each cue used in the experiment. Subjects are known to mis-scale relative depth from disparity at near viewing distances (Foley, 1980; Johnston, 1991, p. 2679) but might well adapt their scaling based on haptic feedback after just a few trials. This sort of fast learning effect would not appear in a between-sessions comparison. The difficulty with this account for the observed effects is existing data clearly show that perceived depth is "stretched" at the near viewing distances used here, an effect that would lead to an initial overweighting of the binocular cue. Learning based on haptic feedback would therefore be expected to decrease that scaling and lead to a lower weight being measured for the binocular cue when haptic feedback is available (the visuomotor task) than when it is not (the perceptual tasks). Thus, this type of adaptive scaling of disparity information would predict the opposite of what we found; binocular weights for perceptual tasks, when no feedback about viewing distance is available, should be higher than binocular weights for the visuomotor task. One could construct a similar hypothesis for adaptive scaling of slant from figure and texture cues. Although such an account cannot be entirely discounted, previous work suggesting that the gain on slant from stereopsis is greater at this viewing distance than the gain on slant from figural cues (Saunders & Knill, 2001) argues against it.

General discussion

Early studies of pointing and grasping movements using only one or both eyes suggested that binocular cues are critical for efficient movements in 3D space. A number of more recent studies have questioned the generality of this result. Watt and Bradshaw (2000, 2003) have shown, for

example, that monocular cues like motion parallax can by themselves support accurate scaling of hand transport velocities (but not grip aperture). Similarly, both the accuracy and the shapes of movement trajectories in the object placement task used here are similar under binocular and monocular viewing (Knill & Kersten, 2003). That it is possible to guide movements effectively with monocular information is effectively illustrated by the many people who successively navigate their world without binocular vision. Several people with only one eye have even succeeded at high levels of athletics (e.g., a recent Division 1 college basketball player had lost one eye early in life). None of these observations, however, tells us about the relative contribution of binocular and monocular cues to motor control when both are present in a stimulus. We have shown that monocular cues about 3D surface orientation can contribute significantly to motor control even in the presence of binocular cues; however, visuomotor control of object placement relies much more heavily on binocular cues than does the perceptual system in tasks requiring estimates of the same surface property.

A number of authors have suggested that the brain performs different visual computations for perception and motor control (Milner & Goodale, 1995). The most commonly cited behavioral evidence for this hypothesis has come from studies that show an attenuation of illusory visual effects when measured using motor behavior rather than explicit perceptual report (Aglioti, DeSouza, & Goodale, 1995; Brenner & Smeets, 1996; Haffenden & Goodale, 1998). Recent studies, however, have cast doubt on these conclusions on methodological (Franz, 2001; Franz, Fahle, Bulthoff, & Gegenfurtner, 2001) or conceptual grounds (Smeets, Brenner, de Grave, & Cuijpers, 2002). Even if one were to reliably find that a perceptual illusion is attenuated in observed motor behavior, such an effect could be (and usually is) interpreted as reflecting differences in the representations on which perceptual judgments and motor behavior rely (e.g., object-centered vs. viewer-centered) rather than on differences in the intermediate computations used to derive the representations (Smeets et al., 2002). Because our experiments studied the cue-integration process that gives rise to estimates of an object's 3D properties for both perceptual judgments and for motor control, the results reflect differences in the internal computations that lead up to the representations on which both types of behavior are based.

Do our results imply that the brain processes visual depth information independently for visuomotor control and perception as suggested by Milner and Goodale (1995)? Such an account does not explain why one should obtain different cue weights for the two types of functions. The optimal cue-integration strategy should be the same for the visuomotor and perceptual tasks used here, as it depends only on the information content of the stimuli (because both types of task required estimates of viewer-centered surface slant). Thus, it would appear that if the weights we measured for one task were optimal, they would be subop-

timal for the other task. What rational basis would exist for visuomotor control relying more on binocular information than on perceptual judgments?

One possibility is that different cue weights might be optimal for different tasks when one considers the specific demands of different tasks. In Bayesian decision theory, task demands are enforced by specifying cost functions associated with a task and estimating a parameter like slant to minimize the expected cost of performance errors (Maloney, 2002; Yuille & Bulthoff, 1996). In the context of motor control, the cost (or gain) associated with performance is a combination of estimation errors, motor errors, and the costs or gains associated with each possible movement. That subjects adjust their motor strategies based on the costs and gains associated with motor performance has been demonstrated in pointing tasks (Trommershauser, Maloney, & Landy, 2003a, 2003b). It is difficult, however, to construct a scenario in which changing the cost function for the task leads to significant changes in the measured cue weights used to combine cues. An optimal estimator derives its estimate by applying a cost function to the combined information from both cues. If the likelihood functions associated with slant estimates from each cue are approximately Gaussian, applying different cost functions for different tasks amounts to applying a point nonlinearity to the weighted average of slants derived from each cue separately. This has little effect on relative cue weights when a linear model is fit to the result.

Another possibility is that our results reflect more the properties of online visual control than motor planning. Even were motor planning based on the same visual estimates of slant as were perceptual judgments, the weights that we derived from contact slants might have been influenced by visual estimates of surface slant computed during the online control phase of movements (even without visual feedback from the hand). Glover and Dixon have argued that illusions influence motor planning much more than online control of hand movements, suggesting that the visual processes underlying the two stages of motor control may be distinct (S. Glover & P. Dixon, 2001; S. R. Glover & P. Dixon, 2001). Whether or not different visual computations subserve the two control phases, the tight time constraints under which online control must operate could affect the relative contributions of binocular and monocular cues during that part of a movement. If the visual system processes binocular cues more quickly than monocular cues (or at least, the monocular cues used here), we might expect the system to effectively give more weight to binocular cues during online control. In another study, we have found that for the object placement task used here, subjects do appear to process binocular cues to slant more quickly than monocular cues when making online adjustments in their movements.

The mechanisms underlying visuomotor and perceptual differences in cue weighting are necessarily a matter of speculation at this point. Nevertheless, our results suggest that task-specific computational constraints on visual

mechanisms other than those imposed by the available image information influence how the brain integrates different sensory cues about the world for guiding behavior. In the terms used in the [Introduction](#), cue weighting is affected not only by the information in the input to the system, but also by the function for which the information is used—the system's output.

Acknowledgments

Commercial relationships: none.

Corresponding author: David C. Knill.

Email: knill@cvs.rochester.edu.

Address: Center for Visual Science and Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY, USA.

References

- Aglioti, S., DeSouza, J. F., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Current Biology*, 5(6), 679-685. [[PubMed](#)]
- Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, 14(3), 257-262. [[PubMed](#)]
- Battaglia, P. W., Jacobs, R. A., & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America A*, 20(7), 1391-1397. [[PubMed](#)]
- Blake, A., Bulthoff, H. H., & Sheinberg, D. (1993). Shape from texture: Ideal observers and human psychophysics. *Vision Research*, 33(12), 1723-1737. [[PubMed](#)]
- Brenner, E., & Smeets, J. B. (1996). Size illusion influences how we lift but not how we grasp an object. *Experimental Brain Research*, 111(3), 473-476. [[PubMed](#)]
- Buckley, D., & Frisby, J. P. (1993). Interaction of stereo, texture and outline cues in the shape perception of three-dimensional ridges. *Vision Research*, 33(7), 919-933. [[PubMed](#)]
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), 429-433. [[PubMed](#)]
- Foley, J. M. (1980). Binocular distance perception. *Psychological Review*, 87(5), 411-434. [[PubMed](#)]
- Franz, V. H. (2001). Action does not resist visual illusions. *Trends in Cognitive Science*, 5(11), 457-459. [[PubMed](#)]
- Franz, V. H., Fahle, M., Bulthoff, H. H., & Gegenfurtner, K. R. (2001). Effects of visual illusions on grasping. *Journal of Experimental Psychology: Human Perception and Performance*, 27(5), 1124-1144. [[PubMed](#)]

- Frisby, J. P., & Buckley, D. (1992). Experiments on stereo and texture cue combination in human vision using quasi-natural viewing. In G. A. Orban & H. H. Nagel (Eds.), *Artificial and biological visual systems*. Berlin: Springer-Verlag.
- Glover, S., & Dixon, P. (2001). The role of vision in the on-line correction of illusion effects on action. *Canadian Journal of Experimental Psychology*, 55(2), 96-103. [PubMed]
- Glover, S. R., & Dixon, P. (2001). Dynamic illusion effects in a reaching task: Evidence for separate visual representations in the planning and control of reaching. *Journal of Experimental Psychology: Human Perception and Performance*, 27(3), 560-572. [PubMed]
- Haffenden, A. M., & Goodale, M. A. (1998). The effect of pictorial illusion on prehension and perception. *Journal of Cognitive Neuroscience*, 10(1), 122-136. [PubMed]
- Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optional cue combination. *Journal of Vision*, 4(12), 967-992, <http://journalofvision.org/4/12/1/>, doi:10.1167/4.12.1. [PubMed][Article]
- Jacobs, R. A. (1999). Optimal integration of texture and motion cues to depth. *Vision Research*, 39(21), 3621-3629. [PubMed]
- Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research*, 31(7-8), 1351-1360. [PubMed]
- Johnston, E. B., Cumming, B. G., & Landy, M. S. (1994). Integration of stereopsis and motion shape cues. *Vision Research*, 34(17), 2259-2275. [PubMed]
- Knill, D. C. (1998). Discrimination of planar surface slant from texture: Human and ideal observers compared. *Vision Research*, 38(11), 1683-1711. [PubMed]
- Knill, D. C., & Kersten, D. (2003). Visuomotor sensitivity to visual information about surface orientation. *Journal of Neurophysiology*. [PubMed]
- Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research*, 43(24), 2539-2558. [PubMed]
- Maloney, L. T. (2002). Statistical theory and biological vision. In D. Heyer & R. Mausfeld (Eds.), *Perception and the physical world: Psychological and philosophical issues in perception* (pp. 145-189). New York: Wiley.
- Marotta, J. J., Behrmann, M., & Goodale, M. A. (1997). The removal of binocular cues disrupts the calibration of grasping in patients with visual form agnosia. *Experimental Brain Research*, 116(1), 113-121. [PubMed]
- Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford, England: Oxford University Press.
- Saunders, J. A., & Knill, D. C. (2001). Perception of 3D surface orientation from skew symmetry. *Vision Research*, 41(24), 3163-3183. [PubMed]
- Servos, P., Goodale, M. A., & Jakobson, L. S. (1992). The role of binocular vision in prehension: A kinematic analysis. *Vision Research*, 32(8), 1513-1521. [PubMed]
- Smeets, J. B., Brenner, E., de Grave, D. D., & Cuijpers, R. H. (2002). Illusions in action: Consequences of inconsistent processing of spatial attributes. *Experimental Brain Research*, 147(2), 135-144. [PubMed]
- Tittle, J. S., Norman, J. F., Perotti, V. J., & Phillips, F. (1998). The perception of scale-dependent and scale-independent surface structure from binocular disparity, texture, and shading. *Perception*, 27(2), 147-166. [PubMed]
- Trommershauser, J., Maloney, L. T., & Landy, M. S. (2003a). Statistical decision theory and the selection of rapid, goal-directed movements. *Journal of the Optical Society of America A*, 20(7), 1419-1433. [PubMed]
- Trommershauser, J., Maloney, L. T., & Landy, M. S. (2003b). Statistical decision theory and trade-offs in the control of motor response. *Spatial Vision*, 16(3-4), 255-275. [PubMed]
- van Beers, R. J., Sittig, A. C., & van der Gon, J. J. D. (1999). Integration of proprioceptive and visual position-information: An experimentally supported model. *Journal of Neurophysiology*, 81(3), 1355-1364. [PubMed]
- Watt, S. J., & Bradshaw, M. F. (2000). Binocular cues are important in controlling the grasp but not the reach in natural prehension movements. *Neuropsychologia*, 38(11), 1473-1481. [PubMed]
- Watt, S. J., & Bradshaw, M. F. (2003). The visual control of reaching and grasping: Binocular disparity and motion parallax. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2), 404-415. [PubMed]
- Yuille, A., & Bulthoff, H. H. (1996). Bayesian decision theory and psychophysics. In D. C. Knill & W. Richards (Eds.), *Perception as Bayesian inference*. Cambridge, England: Cambridge University Press.