

The Role of Visuohaptic Experience in Visually Perceived Depth

Yun-Xian Ho,¹ Sascha Serwe,³ Julia Trommershäuser,³ Laurence T. Maloney,^{1,2} and Michael S. Landy^{1,2}

¹Department of Psychology and ²Center for Neural Science, New York University, New York, New York; and ³University of Giessen, Giessen, Germany

Submitted 9 October 2008; accepted in final form 1 April 2009

Ho Y-X, Serwe S, Trommershäuser J, Maloney LT, Landy MS. The role of visuohaptic experience in visually perceived depth. *J Neurophysiol* 101: 2789–2801, 2009. First published April 2009; doi:10.1152/jn.91129.2008. Berkeley suggested that “touch educates vision,” that is, haptic input may be used to calibrate visual cues to improve visual estimation of properties of the world. Here, we test whether haptic input may be used to “miseducate” vision, causing observers to rely more heavily on misleading visual cues. Human subjects compared the depth of two cylindrical bumps illuminated by light sources located at different positions relative to the surface. As in previous work using judgments of surface roughness, we find that observers judge bumps to have greater depth when the light source is located eccentric to the surface normal (i.e., when shadows are more salient). Following several sessions of visual judgments of depth, subjects then underwent visuohaptic training in which haptic feedback was artificially correlated with the “pseudocue” of shadow size and artificially decorrelated with disparity and texture. Although there were large individual differences, almost all observers demonstrated integration of haptic cues during visuohaptic training. For some observers, subsequent visual judgments of bump depth were unaffected by the training. However, for 5 of 12 observers, training significantly increased the weight given to pseudocues, causing subsequent visual estimates of shape to be less veridical. We conclude that haptic information can be used to reweight visual cues, putting more weight on misleading pseudocues, even when more trustworthy visual cues are available in the scene.

INTRODUCTION

The image of a three-dimensional (3D) surface contains a variety of information that can potentially aid in the estimation of shape and surface relief. This information includes such image features as shading, shadows, specularity (highlights), and occlusions that are highly dependent on factors that are extrinsic to the object such as the pattern of illumination and viewing geometry. The object in Fig. 1 contains surface relief that allows us to identify the object. The image of surface relief depends in large part on the viewing conditions. In Fig. 1 the top part of the figure is in partial shadow, whereas the bottom part is lit more directly, resulting in fewer and lower contrast shadows. As a result, the surface relief in the bottom part of the figure is perceived as flatter than in the rest of the figure. If observers estimate surface properties using measurements of such characteristics of the image—such as mean luminance, contrast, and portion of the image in cast shadow, for example—and do not account accurately for changes in extrinsic factors such as the pattern of illumination, the result will likely be a misperception of surface shape, i.e., a failure of shape constancy.

When visual information leads to failures of shape constancy with changes in illumination, it would be beneficial to turn to a sensory modality that is invariant to changes in illumination and viewpoint such as the haptic system. Illumination and viewing conditions are rarely fixed; thus haptic input may be useful for calibrating the estimation of shape from visual input.

Human visual estimates of 3D surface depth are not perfect. For example, a number of psychophysical studies have reported misestimates of depth for images of shaded 3D surfaces with multiple local minima and maxima consistent with the *bas-relief ambiguity* (for review, see Todd 2004). The *bas-relief ambiguity* describes a class of images resulting from linear (or “affine”) transformations of 3D structure and illumination environments that are indistinguishable when viewed monocularly (Belhumeur et al. 1999). When three or more motion frames or binocular cues to depth are available, the *bas-relief ambiguity* can theoretically be resolved (Longuet-Higgins 1981; Mayhew and Longuet-Higgins 1982). However, when the ambiguity cannot be resolved, observers rely on heuristics that can result in perceived depth reversals, i.e., “hills” perceived as “valleys” and vice versa (Langer and Bühlhoff 2000). Even if stereo or motion cues are available, systematic failures of depth constancy for judgments of metric depth are observed in a variety of tasks (for review, see Todd and Norman 2003).

Judgments of surface roughness depend on illumination conditions, even in the presence of binocular cues that should be sufficient to carry out the task (Ho et al. 2006). In that study, observers compared the perceived roughness of computer-rendered 3D textured surfaces illuminated by a distant punctate light source that varied in its position with respect to the surface. Physical roughness was defined as the variance of the heights of the facets that comprised the simulated surface. Observers consistently judged a surface to be rougher when it was illuminated from a more oblique angle, resulting in more and deeper shadows, even though the stimuli were viewed binocularly so that disparity cues were available. Performance did not improve when additional cues to the illuminant position were provided by adding other objects to the scene (resulting in additional specularity and shadow cues to the location of the illuminant).

Ho et al. (2006) found that observers’ roughness judgments were correlated with changes in *illuminant-variant* measures such as the proportion of the image in cast shadow. These measures varied systematically with changes in both physical roughness and illuminant position. We refer to image information that changes both with the parameter being estimated (in this case, surface roughness) and also with extraneous changes of the scene (i.e., illumination direction) as a “pseudocue.” The discovery that observers use this cast-shadow pseudocue to

Address for reprint requests and other correspondence: M. Landy, Department of Psychology, New York University, 6 Washington Place, Room 961, New York, NY 10003 (E-mail: landy@nyu.edu).



FIG. 1. A 3-dimensional surface showing effects of changes in illumination and viewing geometry on estimates of relief magnitude. This image illustrates how image features (e.g., shadow and highlights) vary with changes in illumination. The bottom portion of the figure is lit more directly than the rest. As a consequence, this part of the image has fewer, lower-contrast shadows and the perceived relief appears to be flatter.

estimate surface roughness is consistent with the idea that human observers are particularly sensitive to dark regions in a noisy two-dimensional (2D) texture (Chubb et al. 2004) and may rely on a “dark-means-deep” heuristic to extract structural information about a surface when disambiguating cues are either unavailable or ignored (Christou and Koenderink 1997; Langer and Bühlhoff 2000).

Although humans may misestimate surface properties because they make use of pseudocues that are affected by changes in viewing conditions, the usage of visual cues might be ameliorated by experience with illuminant- and viewpoint-invariant haptic cues. As Berkeley (1709) once postulated, “touch educates vision.” Some studies suggest that the haptic system trumps the visual system in judgments of smaller-scale surface properties like roughness (e.g., Klatzky et al. 1991, 1993; Lederman and Klatzky 1997; Lederman et al. 1996; however, see Lederman and Abbott 1981). This suggests that the visual system may be relatively unreliable for estimation of such surface properties, so that observers give greater weight to the more reliable source—in this case the haptic system—to optimize overall reliability (Ernst and Banks 2002; Landy et al. 1995).

How might haptic feedback affect observers’ visual judgments of surface shape? It has been suggested that haptic feedback can affect the visual perception of scenes that contain a strong prior via “experience-dependent adaptation” (Adams et al. 2004; Atkins et al. 2001). In the study by Atkins et al. (2001), one visual cue to depth, either 2D texture or motion, was correlated with haptic cues to depth, whereas the other cue was uncorrelated with haptic cues. After training with these stimuli, observers relied more on the cue that was correlated with the haptic cues. Similar results have been found across modalities for the perception of slant (Ernst et al. 2000) and also within modality using judgments of depth paired with other visual cues (Jacobs and Fine 1999). These studies strongly suggest that cue weighting for visual cues to depth is flexible and can be modified by experience with other cues.

There are several ways in which haptic feedback can affect cue combination for subsequent visual judgments. When a haptic signal is in conflict with visual cues to depth and the difference between haptic and visual estimates is stable, the visual system may be *recalibrated* to generate depth judgments that are more consistent with the depth indicated by the haptic cues (Atkins et al. 2003). When two or more noisy visual cues are present and haptic feedback is more correlated with one of these visual cues, a *reweighting* of visual cues may occur, increasing the weight given to the cue that was paired with the haptic feedback (Atkins et al. 2001; Ernst et al. 2000). Here, we investigate the effect of visuohaptic training on the usage of visual cues and pseudocues for scenes with varying illumination.

In the present study, observers compared the depth of cylindrical bumps experienced visually and/or haptically. First, we determine whether varying illuminant position results in visual misperception of bump depth analogous to the lack of shape constancy seen in visual estimation of surface roughness. Second, we determine whether the visual system reweights visual cues and pseudocues to bump depth as a result of visuohaptic training under conditions in which the haptic information is *not* consistent with illuminant-invariant cues but, rather, is artificially correlated with the illuminant-variant visual pseudocues. We refer to this haptic information as haptic *feedback* even though no explicit feedback is provided. The observer might treat the haptic information as “ground truth” information to the task objective, i.e., a standard by which to compare visual information that may be less reliable (Atkins et al. 2001; Ernst et al. 2000). However, the extent to which haptic information is used depends on a number of factors such as the strength of the modality-encoding bias based on the task and the visibility of the hands during exploration (Heller 1970; Lederman et al. 1996). We predict that if haptic feedback to bump depth varies directly with changes in illumination conditions, observers will make greater use of pseudocues to bump depth, resulting in a greater departure from shape constancy across varying illumination.

METHODS

Stimuli

COORDINATE SYSTEMS. We used a Cartesian coordinate system (x , y , z) to define our 3D surfaces (Fig. 2). The origin was in the center of the plane on which the stimulus was presented (the *stimulus plane*). The z -axis was normal to the stimulus plane, the x -axis was horizontal, and the y -axis was vertical in the stimulus plane. We described the position of the observer and the illuminant as vectors from the origin using a spherical coordinate system (ψ , φ , r). Azimuth ψ was defined as the angle between the projection of the vector onto the x - y plane and the negative x -axis, and elevation ψ was the angle between the vector and its projection onto the x - y plane. The punctate illuminant was located at position (ψ_p, φ_p, r_p) and the observer’s viewpoint position was (ψ_v, φ_v, r_v) . The viewpoint was fixed at $(\text{---}, 90^\circ, 45 \text{ cm})$ so that the stimulus plane was viewed frontoparallel and illuminant position was varied within the y - z plane.

SURFACE PATCH. A 110×110 -mm flat, gaze-normal surface was rendered 450 mm away from the observer. An upright cylindrical bump with a fixed height and width of 40 mm (in the x - y plane) was embedded in the surface. The bump was a portion of a vertical circular cylinder produced by placing a 40-mm-tall circular cylinder behind

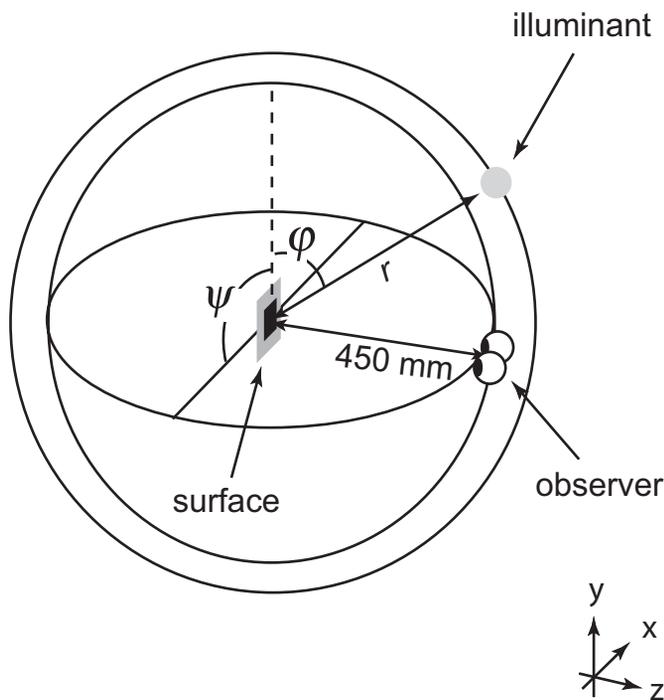


FIG. 2. Coordinate systems. A Cartesian coordinate system was used to define the stimuli. The origin of the coordinate system was at the center of the surface patch containing the bump. The z -axis lay along the line of sight from the observer to the origin; the x -axis was horizontal and tangent to the surface patch; and the y -axis was vertical and tangent to the surface patch. To specify punctate illuminant positions, we used spherical coordinates. The illuminant had an elevation φ (relative to the negative x - y plane) and azimuth ψ (relative to the negative x -axis), and was located at a distance r from the surface patch. In this study, φ was varied, whereas ψ and r were fixed at 90° (directly overhead) and 600 mm, respectively. The observer was at $(x, y, z) = (0, 0, 450 \text{ mm})$ or, in spherical coordinates, $(\psi, \varphi, r) = (90^\circ, 90^\circ, 450 \text{ mm})$.

the flat background surface with a 40-mm-wide portion jutting out of a 40-mm-wide aperture in the background. The shape and maximum depth of the bump were controlled by varying the radius of the circular cylinder. To maintain the 40×40 -mm dimensions of the bump while varying its depth, the distance of the cylinder's center to the surface increased with increasing radius values. Thus a cylinder with a larger radius resulted in a bump with less depth and curvature. The bump subtended a 5.1° visual angle. The bump depth was one of 13 values linearly spaced between 4 and 16 mm indexed as bump-depth level $d = \{1, 2, \dots, 13\}$. A condition in which the surface was flat (i.e., 0 mm) was also included and indexed as $d = 0$. Random fractal noise textures were mapped to the stimulus surface for each bump-depth level d to minimize the possibility that observers used specific patterns in a given texture as cues to bump depth. These surfaces were then rendered in a scene with the illumination parameters described in the next section.

ILLUMINATION ENVIRONMENT. Each surface patch was prerendered under a diffuse illuminant and a punctate illuminant using the Radiance software package (Larson and Shakespeare 1996; Ward 1994; <http://radsite.lbl.gov/radiance/HOME.html>). Surfaces were rendered with Lambertian reflectance and indirect light was allowed to bounce twice in the scene. We asked observers to compare bump depths across two illuminant elevations (φ_p), 45 and 30° . We refer to the illumination condition $\varphi_p = 45^\circ$ as the *test* condition and $\varphi_p = 30^\circ$ as the *match* condition (this is described in more detail in *Procedure*). Only illuminant elevation φ_p was varied; azimuth ψ_p was fixed at 90° and r_p was fixed at 600 mm. Each scene was prerendered twice from slightly different viewpoints (± 30 mm, corresponding to an interpupillary distance of 60 mm) for each tested illumination condition

corresponding to the approximate positions of the observer's eyes (interpupillary distances ranged from 56 to 64 mm for observers in this study). The scenes were viewed binocularly in the experiment. Stereo pairs of typical scenes are shown in Fig. 3 and a representative set of stimuli is shown in Fig. 4.

VISUOHAPTIC TRAINING STIMULI. Haptic bumps were rendered as portions of circular cylinders jutting out of a flat background in an identical manner as the visual stimuli. However, the depths of the bumps presented during visuohaptic training were artificially correlated with the proportion of cast shadow, one of the pseudocues identified in previous work (Ho et al. 2006, 2007). We defined the proportion of cast shadow as the ratio of the area of the cast shadow in the stimulus divided by the total background area. We determined the haptic bump-depth values by first quantifying the amount of cast shadow in each stimulus for every combination of illumination and bump depth. The proportion of cast shadow increases systematically with decreasing illuminant elevation and increasing bump depth (Fig. 5A).

The filled and open circles in Fig. 5B indicate the proportion of cast shadow for the test and match stimuli, respectively. The curve represents the best least-squares second-order polynomial fit to the mean values of proportion of cast shadow for each bump-depth level. To correlate the cast-shadow pseudocue and the depth indicated by the haptic cue, we used this curve as a look-up table for haptic depth as a function of the proportion of cast shadow in the visual stimulus. For example, a match stimulus with a bump depth of 6 mm has 2.6% of the image in cast shadow and would be paired with a haptic bump depth of 7.6 mm in the visuohaptic training trials (indicated by arrows in Fig. 5B). Table 1 shows the resulting haptic bump depth values assigned to each bump depth level for each illumination condition.

Apparatus

Visual stimuli were displayed on an Iiyama Vision Master Pro 514 22-in. CRT monitor suspended from above. Observers viewed the stimuli in a fully reflective mirror reflecting the stereoscopic display using CrystalEyes 3 (Stereographics) liquid-crystal shutter glasses that were synchronized with the monitor's refresh rate at 120 Hz. The monitor's resolution was $1,280 \times 960$ and visual stimuli were generated using an Nvidia GeForce FX 500 graphics card. A look-up table was used to correct monitor nonlinearities. Measurements for the look-up table were taken using a Laser 2000 UK photometer (Ringstead, UK). The maximum luminance achievable on the monitor was 64 cd/m^2 . A head and chin rest was used to limit head movement. Stimuli were presented directly in front of the observer.

A PHANToM 3D Touch interface (SensAble Technologies, Woburn, MA) was used to generate haptic stimuli using force feedback. Use of this apparatus allowed us to manipulate visual and haptic input independently. The device tracks the 3D position of the right index fingertip secured in a thimble attached to the PHANToM arm and generates force fields that simulate haptic properties such as weight, hardness, and friction of virtual haptic stimuli. Our parameter settings simulated a surface that felt like it was made of stiff, smooth, rubber material (stiffness constant $k = 2.5 \text{ N/m}$, friction coefficient 0.6). The hand was not visible to the observer, but the fingertip was represented visually by a small cursor (3-mm-diameter sphere). The apparatus was calibrated to superimpose visual and haptic stimuli in the workspace. Software used for stimulus presentation was programmed in C using the GHOST SDK v.4.0 (SensAble Technologies) tools.

All data were collected at the Department of Psychology, Justus-Liebig University, Giessen, Germany.

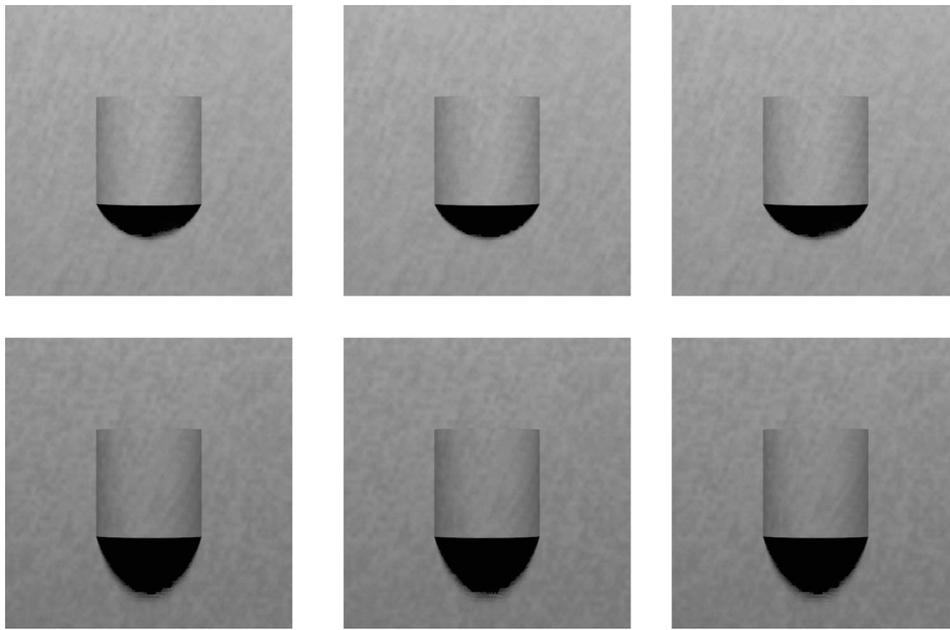


FIG. 3. Examples of stereograms. Two example stimuli with the same physical bump depth (12.0 mm) rendered under the test illuminant position $\varphi = 45^\circ$ (top row) and match, $\varphi = 30^\circ$ (bottom row), presented for crossed (left pair) and uncrossed (right pair) viewing.

Procedure

Observers participated in the following sequence of conditions over the course of 4 days (with at most 2 days between successive experiment days):

- Day 1. Visual-only practice, Visual pretest (two sessions)
- Day 2. Visual pretest (two sessions), Haptic-only
- Day 3. Visuohaptic training, Visual posttest
- Day 4. Visuohaptic training, Visual posttest

In all conditions, a two-interval forced-choice (2-IFC) task was used in which an observer was presented with a test and a match stimulus displayed sequentially and was asked to choose which bump was perceived to be “bigger” (*Welcher Hubbel war größer?*) or “protrudes out of the wall more.” Recall that the test stimulus was a scene illuminated by a light source with elevation $\varphi_p = 45^\circ$ and the match stimulus was a scene illuminated by a light source with elevation $\varphi_p = 30^\circ$. On each trial, the test stimulus could be in either the first interval or the second interval. The rendered depth of the match stimulus was a function of the observer’s previous responses as controlled by a staircase. We used the test–match distinction only in describing how the sequence of trials presented to the observer was affected by his/her judgments in the staircase procedure and we also

used this distinction in analyzing the data. Observers were unaware of the distinction.

A session consisted of 10 interleaved staircases, two for each of the five test bump levels $d_t = 3, 5, 7, 9,$ and 11 . A “1-up, 2-down” and a “2-up, 1-down” staircase were run for each test bump level and observers performed in 20 trials of each staircase type in each session. Match bump-depth levels are indexed by $d_m = \{0, 1, \dots, 13\}$. In all conditions, a session consisted of 200 trials divided into two blocks per session with the order of trials randomized across observers (40 trials per test bump level \times 5 test bump levels).

Before participating, all observers were first tested for stereoscopic vision using two images displayed to the right and left eyes that contained a black square with crossed disparity relative to a larger white square surface. The observer was asked to describe the scene. Only observers who reported that the black square patch appeared in front of the white square participated in the experiment. The observer performed one practice trial of the haptic-only task and the visual-only task under the supervision of the experimenter as an introduction to the experimental environment and task.

VISUAL PRETEST SESSION. In the visual pretest session, the observer viewed two surfaces, a match and a test, sequentially, and indicated which appeared to be larger (2-IFC). On each trial, an “initialization display” was first presented containing four dots arranged in a dia-

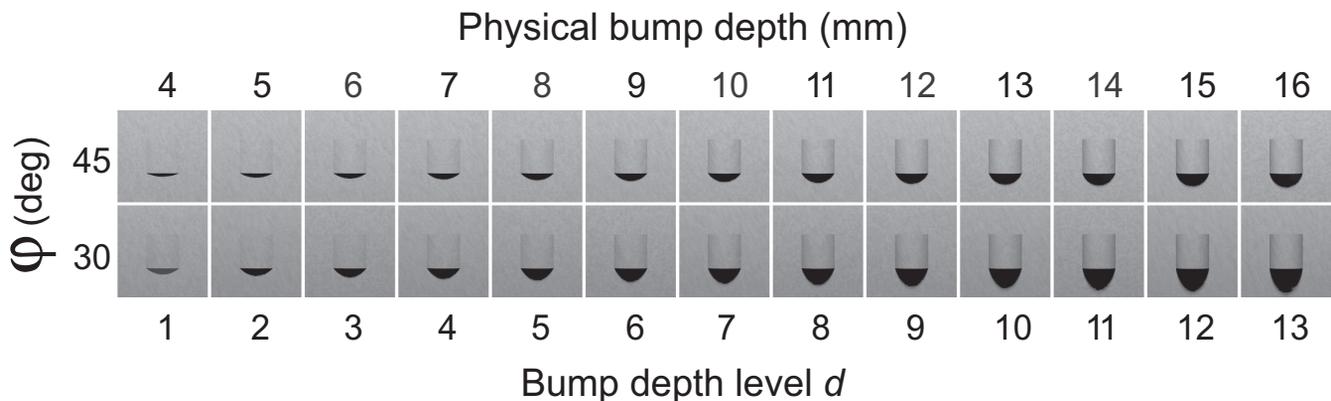


FIG. 4. An example of a stimulus set. Three random fractal noise patterns were mapped to each surface for each bump depth level d . No 2 stimuli contained the same noise pattern. Shown here is one set of scenes rendered under the 2 illuminant positions used in this study (flat surface, $d = 0$, not shown).

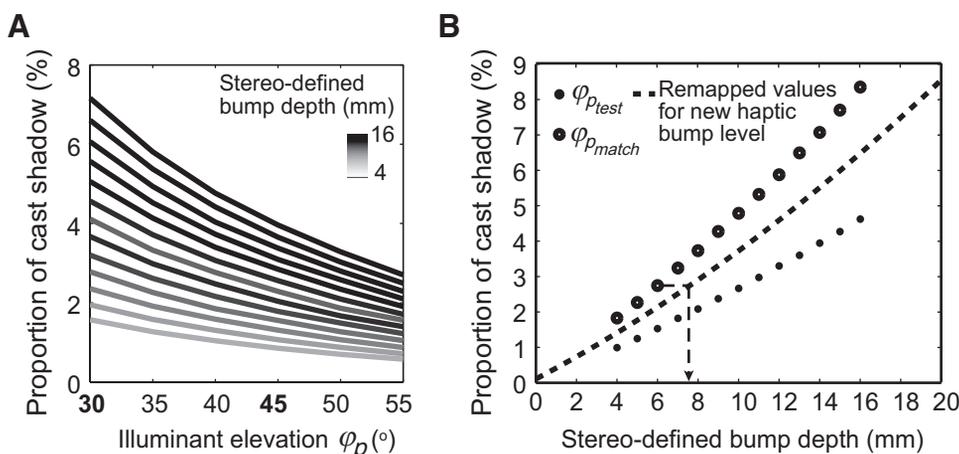


FIG. 5. Pseudocues and remapped haptic bump depths for visuohaptic training. *A*: the proportion of the image that is cast shadow calculated from the stimulus set used in this study plotted as a function of illuminant elevation. Test and match illuminant positions used in this study are indicated in bold. Contours represent different bump depths (indicated by gray level). The proportion of cast shadow increases with increasing bump depth and decreasing illuminant elevation. *B*: the proportion of shadows as a function of visually rendered bump depth, i.e., depth signaled by disparity (filled circles: test; open circles: match). Dotted line: 2nd-order polynomial fit to average proportion of cast shadow for each bump depth. In the visuohaptic training trials, the haptic bump depth was determined by the proportion of cast shadow in the visual stimulus and the correspondence given by the dotted curve (dashed arrow).

mond configuration around a central fixation point for 0.5 s. The dot configuration was designed to allow observers to maintain vergence at the distance at which the stimuli were presented. Next, a test or match stimulus was presented for 1 s, followed by an interstimulus interval (ISI) display, identical to the initialization display, presented for 0.5 s. Finally, the match or test stimulus (respectively) was presented for 1 s. Observers responded by a mouse click. Observers participated in five visual pretest sessions in which staircases were continued across blocks within a session. The first visual pretest session was treated as practice to become familiar with the task and experimental setup and the data from this session were not included in the main data analysis.

HAPTIC-ONLY SESSION. Next, observers ran in a session in which only haptic stimuli were presented. The primary purpose of the haptic-only session was to allow observers to become familiar with the PHANToM apparatus. In this session, each trial was initiated by the observer pressing the dot to the right or left of central fixation in an initialization display visually identical to that of the visual pretest session. The haptic test or match stimulus consisted of a large visible gray wall with a bump centrally embedded that could be explored haptically, but which was not visible to the observer (a square black patch with the height and width of the bump was presented in place of the bump). This bump was displayed for 3.5 s, during which the observer was instructed to move his/her finger across the surface laterally at a comfortable pace. The cursor representing the finger was not visible during the period the finger was over the stimulus. This was done to eliminate visual cues to bump depth from the movement of the cursor. Next, an ISI display was shown identical to the initialization display. The observer pressed either the right or left dot in the ISI display to trigger the next display. Another haptic match or test stimulus (respectively) was then shown for 3.5 s. This was

followed by a response display that contained two buttons; the observer was instructed to press the left button if the first bump was perceived to be bigger and the right button otherwise. Observers ran one haptic-only session.

VISUOHAPTIC TRAINING SESSION. In the visuohaptic training session, observers were presented with both a visual and haptic stimulus presented simultaneously for 3.5 s. The trial sequence was identical to the haptic-only condition except that the stimulus could be both felt and seen. Again, the cursor representing the finger was not visible during the period the finger was over the stimulus and the observer was instructed to move his/her finger across the surface laterally at a comfortable pace. The haptic stimuli were chosen based on the correspondence described earlier (**VISUOHAPTIC TRAINING STIMULI**). A test and a match stimulus were presented sequentially in each trial with the order randomized. Observers ran one visuohaptic training session on each of Days 3 and 4.

VISUAL POSTTEST SESSION. Each visuohaptic training session was followed by a visual posttest session identical to the visual pretest. We refer to this session as a posttest to differentiate it from the sessions that preceded any visuohaptic training (i.e., the pretest). Observers participated in one visual posttest session on each of Days 3 and 4.

Observers

All observers were students recruited from the University of Giesen and paid hourly for their participation. Twelve observers participated in the study. All observers had normal or corrected-to-normal vision and were unaware of the hypothesis under test. Their ages ranged from 20 to 33 yr.

TABLE 1. Remapped haptic bump depths

Bump depth level, <i>d</i>	1	2	3	4	5	6	7	8	9	10	11	12	13	
Visual bump depth, mm	4	5	6	7	8	9	10	11	12	13	14	15	16	
	φ													
	Haptic Bump Depth, mm													
Test	45	—	—	4.3	—	5.8	—	7.4	—	9.0	—	10.5	—	
Match	30	5.2	6.4	7.6	8.8	10.0	11.3	12.4	13.6	14.8	16.0	17.2	18.4	19.6

The physical values for each of the remapped haptically rendered bumps used in the visuohaptic training sessions are shown here for each test and match illumination condition φ and bump level *d*.

RESULTS

Bump-depth constancy

To evaluate the effects of visuohaptic training, we analyzed results from the four visual pre- and two visual posttest sessions. We estimated the point of subjective equality (PSE) for each of the five test bump levels for each of the six sessions by fitting the data with a Weibull distribution and estimating the point at which there was a 50% probability of choosing the match stimulus as the “bigger bump.” Psychometric functions for one observer (AS) are shown in Fig. 6. The black open circles indicate the PSEs obtained for the visual pretest sessions (Sessions 1–4) and the filled black circles indicate the PSEs for the visual posttest sessions (Sessions 5–6). Psychometric functions acquired for the haptic-only session for the same observer are also shown here for comparison (*top row*, light gray).

How does performance compare with that predicted for an observer unaffected by changes in illumination (a “bump-depth-constant” observer)? Estimated PSEs are plotted in Fig. 7 for one observer (NK) and 95% confidence intervals for the PSEs were computed using a parametric bootstrap method; each observer’s data set was resampled 2,000 times and the 2.5th and 97.5th percentiles were calculated from the distribution of estimated PSEs (Efron and Tibshirani 1993). There are systematic differences between test and match bump levels perceived as having equal depth (the PSEs), i.e., a failure of bump-depth constancy. In particular, most PSEs fell below the identity line (or line of constancy, indicated by the dashed line in Fig. 7). We summarize this pattern using a bump-discrimination model that models observers’ trial-by-trial decisions (details of the model are described in the Supplemental data).¹ Using this model and a maximum-likelihood criterion, we estimated three parameters: 1) a bump-depth transfer parameter \hat{c} , 2) an SD of normally distributed noise $\hat{\sigma}$ caused by variability in the observer’s judgments, and 3) a noise-scaling parameter $\hat{\gamma}$ that accounts for any Weber-like stimulus-dependent noise. The bump-depth transfer parameter is simply the slope of the linear fit to the data (or *contour of indifference*) shown in Fig. 7.

This bump-discrimination model was analogous to the modeling of roughness judgments in previous work (Ho et al. 2006, 2007). It might be argued that although the bump-discrimination model provided a simple way of describing observer’s behavior in the task, it did not accurately characterize the data. To address this, we assessed our model’s goodness of fit by regressing the 360 estimated PSEs obtained from our data for all observers to the PSEs predicted by the bump-discrimination model. Figure 8 shows the measured PSEs plotted as a function of the predicted PSEs for all observers. The R^2 value for this regression fit is 0.95 and the data do not seem to deviate from the identity line in any systematic way, suggesting that the model describes the data well. Although we accept the possibility that a nonlinear model with more parameters may fit the data even better, the use of this linear model makes it possible to examine differences in observers’ performance by comparing one parameter, the slope \hat{c} .

If bump-depth constancy holds, the value of \hat{c} should be close to 1. The parameters of this model were obtained for each of 2,000 bootstrapped samples to derive confidence intervals around the parameter estimates. Table 2 shows all estimated slope parameters for each observer ($\hat{\sigma}$ and $\hat{\gamma}$ are not shown or discussed here because they are important for fitting the model to the data, but not directly relevant to the goals of this study). A z test was performed to determine whether each of the slope parameters was significantly different from 1, i.e., whether the observer showed a failure of bump-depth constancy for a given session. Values of \hat{c} that were found to be significantly different from 1 at the Bonferroni-corrected α level of 0.05 for 12 tests ($P < 0.004$) are indicated in boldface.

Failures of depth constancy

The slope parameter estimates \hat{c} are shown in Fig. 9 for each of the pre- and posttest sessions along with 95% confidence intervals obtained by a bootstrapping method (Efron and Tibshirani 1993). We noticed that observers showed no systematic upward or downward trend in the visual pretest sessions. Thus we calculated the mean of the four pretest slope parameters for each observer and compared it to 1 to determine whether observers showed a significant failure of bump-depth constancy before training. Six of 12 observers exhibited a significant failure of constancy in the pretest sessions ($P < 0.004$, corresponding to an overall type I error rate of 0.05, Bonferroni-corrected for 12 tests). Five of these six observers (AJ, CG, JR, NK, and SC) perceived bumps illuminated by the more oblique light source to have increased depth. However, one observer (RZ) showed the opposite trend: bumps illuminated under the *less* oblique light source were perceived to have increased depth.

In the visuohaptic training sessions, the haptic depth was artificially correlated with a pseudocue (the proportion of cast shadow) and correspondingly decorrelated with illuminant-invariant cues (binocular disparity and texture). In Fig. 9, the gray circles indicate the slope parameter estimates from these training sessions, plotted just to the left of the visual posttest that they preceded. The horizontal dotted line indicates the slope estimate that would result if a subject ignored the visual stimulus entirely during these sessions and merely compared haptic stimuli (inferred from the look-up table for haptic bump depth in Table 1). The results for 11 of the 12 subjects suggest that observers integrated the two modalities, resulting in a compromise between visual pretest slopes and the predicted haptic-only slope. Only observer AG’s data indicate that visual input was almost fully ignored during the training sessions.

We predicted that visuohaptic training would increase the weight observers would give to this pseudocue, resulting in greater failures of bump-depth constancy in the visual posttest sessions. Since the match stimuli have the more oblique illuminant, we predicted that after visuohaptic training, these stimuli would be perceived to have increased depth. As a result, even after haptic information was removed in the visual posttest sessions, observers should have required test stimuli to have greater rendered depth to appear equivalent to the match stimuli, resulting in a decrease in slope parameter estimates \hat{c} . We compared the averages of the four pretest to the two posttest slope parameters for each observer using a one-tailed z test. Five of 12 observers had significantly shallower slopes after visuohap-

¹ The online version of this article contains supplemental data.

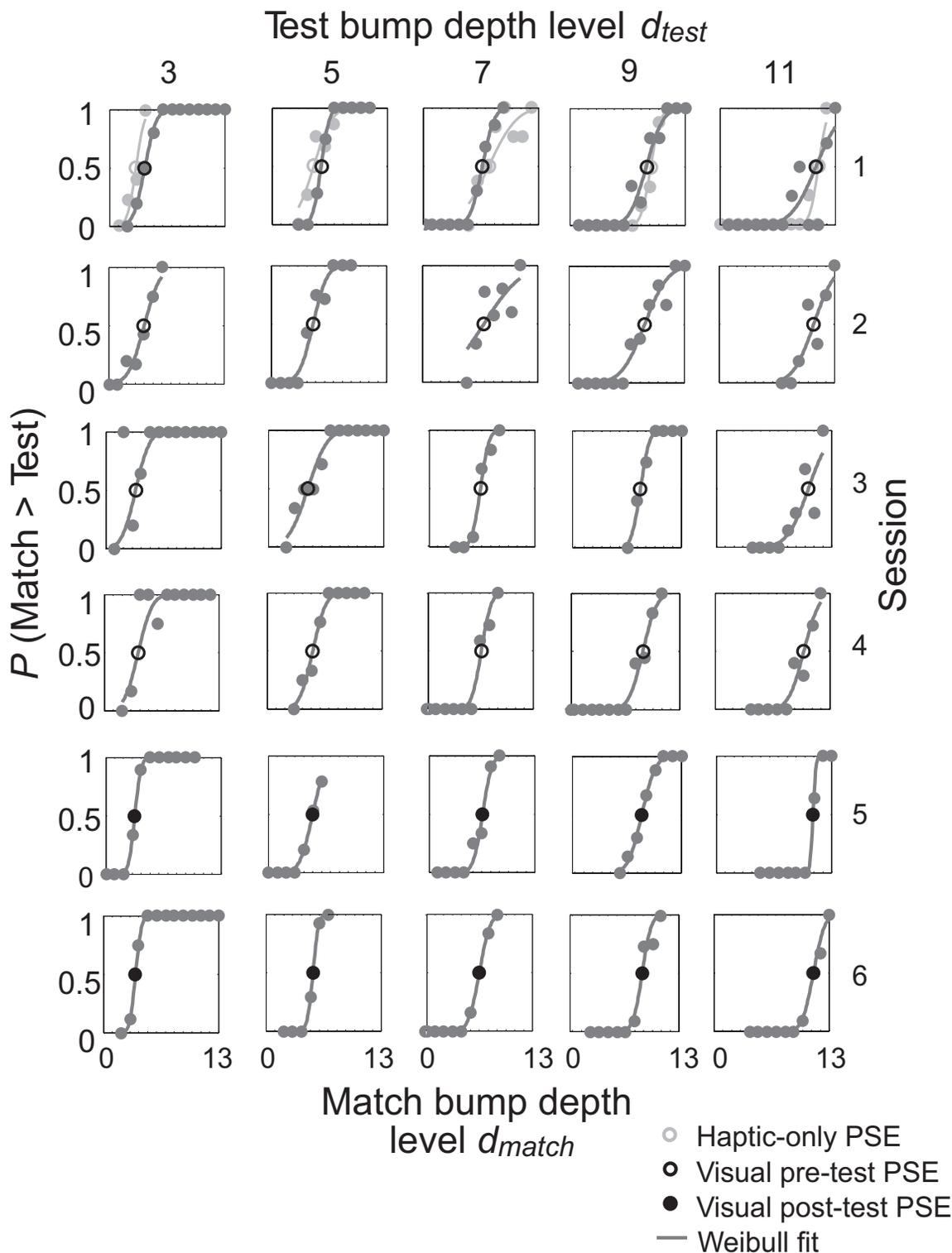


FIG. 6. Results: psychometric functions. We estimated psychometric functions for each of the 5 test bump levels for each pre- and posttest session by fitting a Weibull function to the data and estimated the points of subjective equality (PSEs), i.e., the match bump depth perceived as equivalent to the test depth, for each psychometric function. Shown here are the psychometric functions for one observer (AS). Pretest PSEs are shown as open circles, whereas posttest PSEs are shown as filled circles. Each row corresponds to a session and each column corresponds to a test bump level. As a comparison, psychometric functions obtained for the haptic-only condition are also shown here in light gray (*1st row*). Note that the slopes of the haptic-only functions are comparable to those of the visual pre- and posttest conditions so that haptic discrimination of relief is comparable to visual sensitivity under our conditions.

tic training ($P < 0.004$, corresponding to an overall type I error rate of 0.05, Bonferroni-corrected for 12 tests), indicating that training did result in greater failures of bump-depth constancy in

the predicted direction. The average pretest slope parameter for all observers was 0.92 (ranging from 0.64 to 1.09); the posttest average was 0.79 (ranging from 0.23 to 1.02).

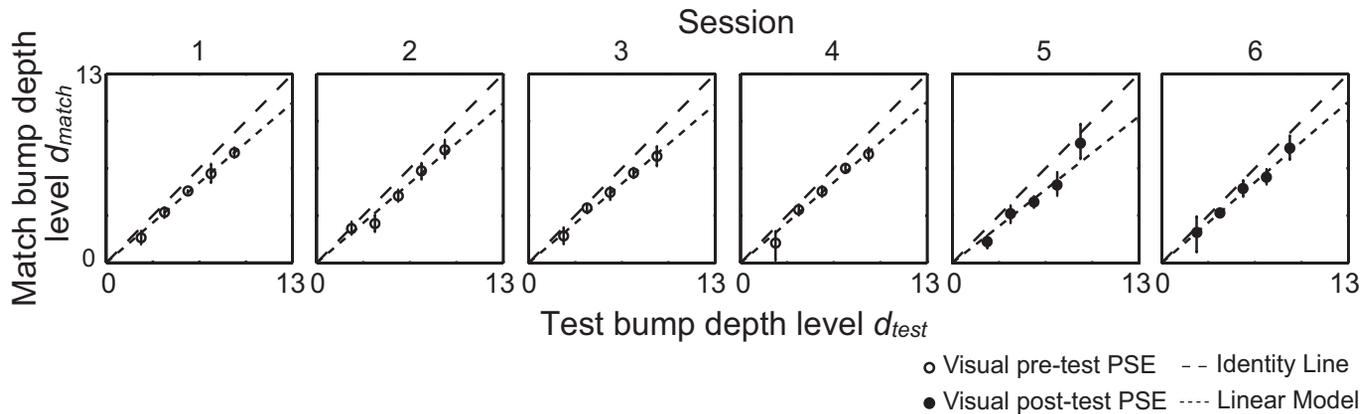


FIG. 7. Results: PSEs and fits of the bump-discrimination model. PSEs are plotted for each of the visual pre- (open circles) and posttest sessions (filled circles) for one observer (NK). Error bars indicate 95% confidence intervals estimated using a bootstrap method (Efron and Tibshirani 1993). The dashed line indicates the line of constancy for which changes in illumination do not change perceived depth. Dotted lines are derived from fits of the bump-discrimination model. For this observer, there are systematic deviations from constancy; all PSEs fall below the identity line.

A pseudocue model

Although the bump-discrimination model fits the data well, it does not provide any insight about the contribution of pseudocues to observers' judgments. If failures of bump-depth constancy were primarily due to the contribution of pseudocues, then we would predict that observers who showed more pronounced failures of bump-depth constancy in the posttest session weighted pseudocues more heavily than the illuminant-invariant visual cues to depth (such as binocular disparity), whereas those who showed nearly bump-depth-constant performance in the posttest session weighted illuminant-invariant cues to bump depth more heavily. In this section, we reanalyze

the data with a model that emphasizes the weighting of pseudocues and illuminant-invariant visual cues to bump depth. We estimate the relative weights in the pre- and postsessions to determine whether training increased the weight given to the particular pseudocue we manipulated: the proportion of shadow in the image.

The pseudocue model suggests that departures from bump-depth constancy are explained by the values of the pseudocues in the stimuli. The details of the pseudocue model are described in detail in the Supplemental data. In brief, the model assumes depth is estimated as a linear combination of the illuminant-invariant visual cues that correctly signal depth in the display (e.g., binocular disparity) and the pseudocue (the proportion of cast shadow in each stimulus). For each PSE in the data, by definition, the observer's depth estimates for the test and match stimuli are equal. Thus any difference in test and match rendered depths in the visual display must be offset by a difference of opposite sign in the pseudocues of the test and match stimuli. That is, we fit a regression equation $\Delta \hat{d}_d = \hat{a}_s \Delta d_s$ to the PSE data, where $\Delta \hat{d}_d$ is the predicted difference in illuminant-invariant visual cues for the PSE (i.e., the distance of the PSE

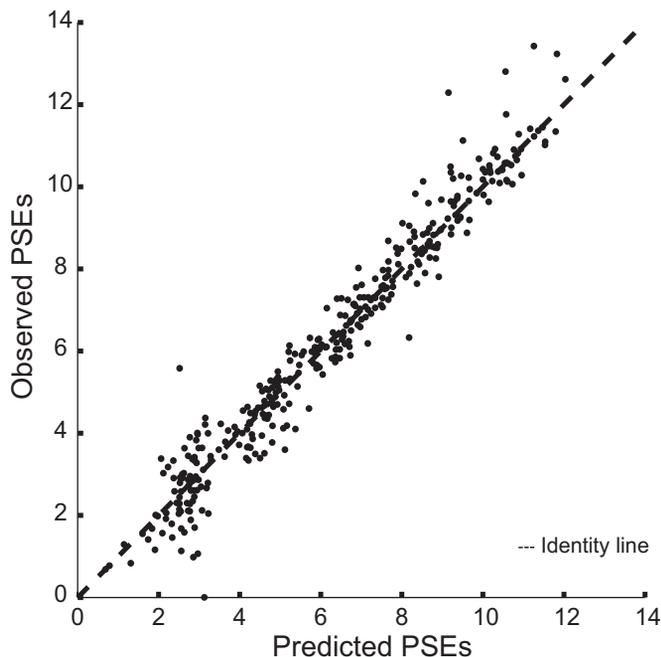


FIG. 8. Bump-discrimination model goodness of fit. Observed PSEs for all observers were regressed to the PSEs predicted by the bump-discrimination model. The dashed line indicates the identity line. Notice that the points fall closely around the identity line and do not seem to exhibit any obvious systematic bias thus suggesting that the bump-discrimination model does a fairly good job of modeling the data. The R^2 value for this regression is 0.95 and is significantly different from 0 ($P < 0.001$).

TABLE 2. Bump-discrimination model slope parameter \hat{c} estimates

Observer	Pretest \hat{c}			Posttest \hat{c}		
AG	0.92	1.03	0.84	0.99	0.23	0.26
AJ	0.83	0.84	0.78	0.79	0.86	0.84
AS	0.99	0.99	0.94	0.98	0.98	1.02
CG	0.91	0.87	0.88	0.93	0.74	0.71
JC	1.02	0.98	0.99	0.91	0.91	0.96
JD	1.04	0.97	0.90	0.93	0.96	0.85
JR	0.86	0.94	0.93	0.95	0.91	0.92
LD	0.96	1.08	1.02	0.96	0.64	0.58
NAK	0.94	0.97	1.00	0.96	0.80	0.92
NK	0.85	0.84	0.84	0.85	0.77	0.85
RZ	1.09	1.07	1.05	1.05	0.95	0.88
SC	0.73	0.64	0.73	0.66	0.82	0.70

The bump-discrimination model pre- and posttest slope parameters \hat{c} for each of the six sessions are shown here for each observer. A bootstrap method was used to obtain 95% confidence intervals. Each observer's performance was resampled 2,000 times and the 2.5th and 97.5th percentiles of the resulting distribution of \hat{c} values were obtained (Efron and Tibshirani 1993). Values of \hat{c} indicated in boldface are significantly different from 1.00, testing at an overall type I error rate of 0.05, Bonferroni-corrected for 12 tests.

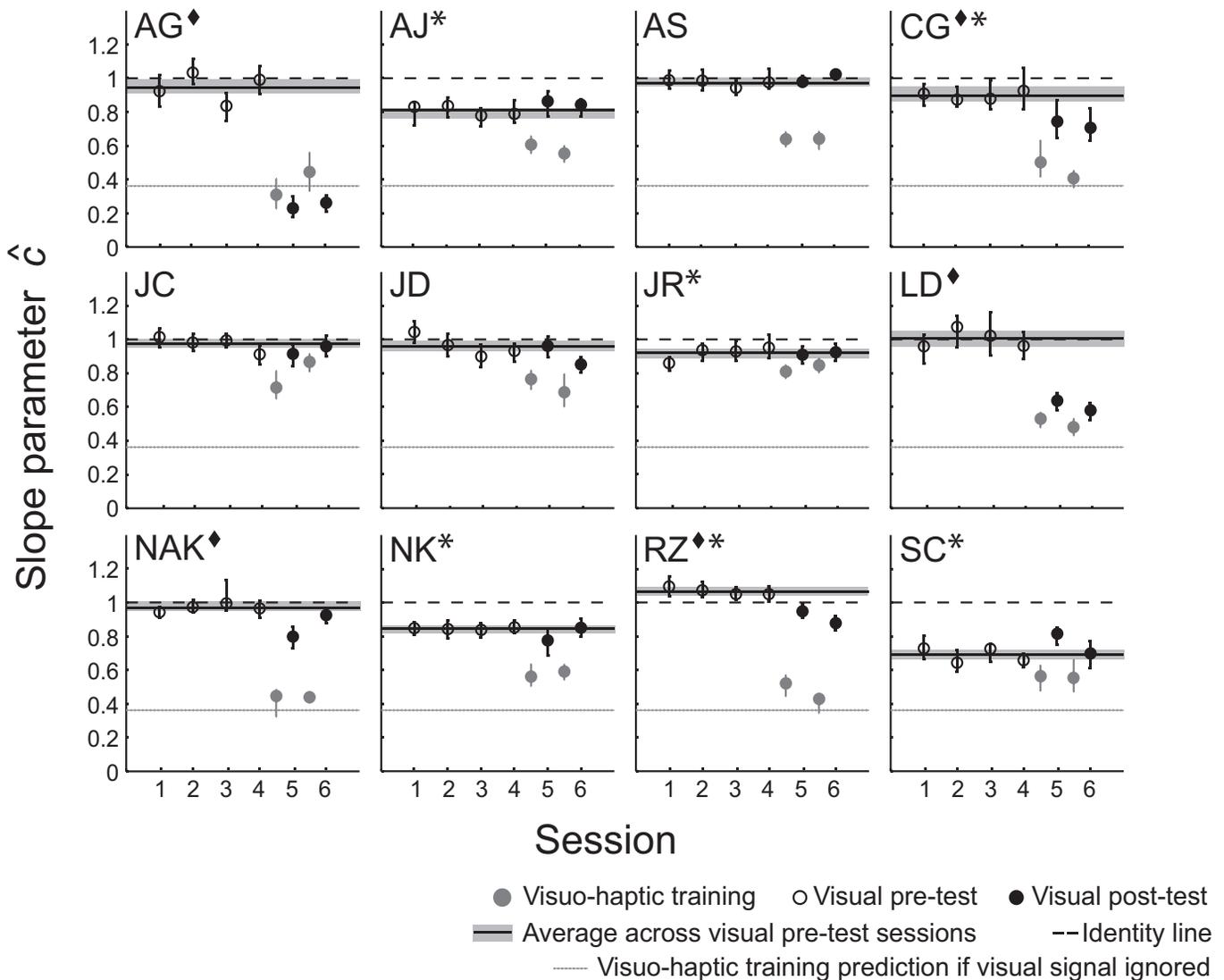


FIG. 9. Results: pre- and posttest comparison. The slope parameters \hat{c} obtained from fits of the bump-discrimination model (e.g., dotted lines in Fig. 7) are plotted for the pre- (open circles) and posttest sessions (filled circles) for each observer. The gray circles indicate slope parameters from the training sessions and are plotted just to the left of the posttest session they preceded. Error bars represent 95% confidence intervals obtained by a bootstrapping method (Efron and Tibshirani 1993). Dashed line: slope value indicating bump-depth constancy. Solid line and gray region: average slope of the 4 pretest sessions and its 95% confidence interval. Dotted line: slope value predicted for an observer who ignores the visual stimulus during the visuohaptic training sessions. The $*$'s indicate pretest slopes that, as a group, were found to be significantly different from 1. The \diamond 's indicate posttest slopes significantly smaller than pretest slopes. (All significance tests performed using a P value of 0.004, corresponding to an overall type I error rate of 0.05, Bonferroni-corrected for 12 tests.)

away from the line of constancy), Δd_s is the difference in the pseudocue for the test and match stimuli at the PSE, and \hat{a}_s is an estimated proportionality constant that takes into account the relative weights the observer gives to the illuminant-invariant visual cues versus the pseudocue, but is also affected by how the observer scales the pseudocue to estimate depth.

Observed values of Δd_d of the PSEs are shown in Fig. 10 as a function of values predicted by the model $\Delta \hat{d}_d$. The proportion of cast shadow in the scene explained a significant amount of the variance in 14 cases of 24, 6 of 12 in the pretest condition, and 8 of 12 in the posttest condition ($P < 0.004$, with an overall type I error rate of 0.05, Bonferroni-corrected for 12 tests; Table 3). In some cases, the model accounted for a relatively low percentage of the variance, which is not surprising, given that in these cases observers exhibited little to no failure of bump-depth constancy.

We evaluated the effect of pseudocues in visual bump-depth judgments by comparing the pre- and posttest values of \hat{a}_s in our model. Although the value of \hat{a}_s is not meaningful on its own, an increase in the value of the posttest \hat{a}_s compared with the pretest \hat{a}_s would suggest that the relative weight of the pseudocue used in comparing bump depths was increased after training. Table 3 indicates that the posttest \hat{a}_s was greater than the pretest \hat{a}_s for 9 of 12 observers. Larger positive differences between pre- and posttest values of \hat{a}_s correspond mostly to observers who exhibited a larger failure of bump-depth constancy in the posttest sessions (the one exception is observer SC). This suggests that these observers' increased the relative weight of pseudocues to illuminant-invariant cues in the posttest session following visuohaptic training. Small differences in pre- and posttest \hat{a}_s values correspond to observers who showed little to no significant change in pre- and posttest

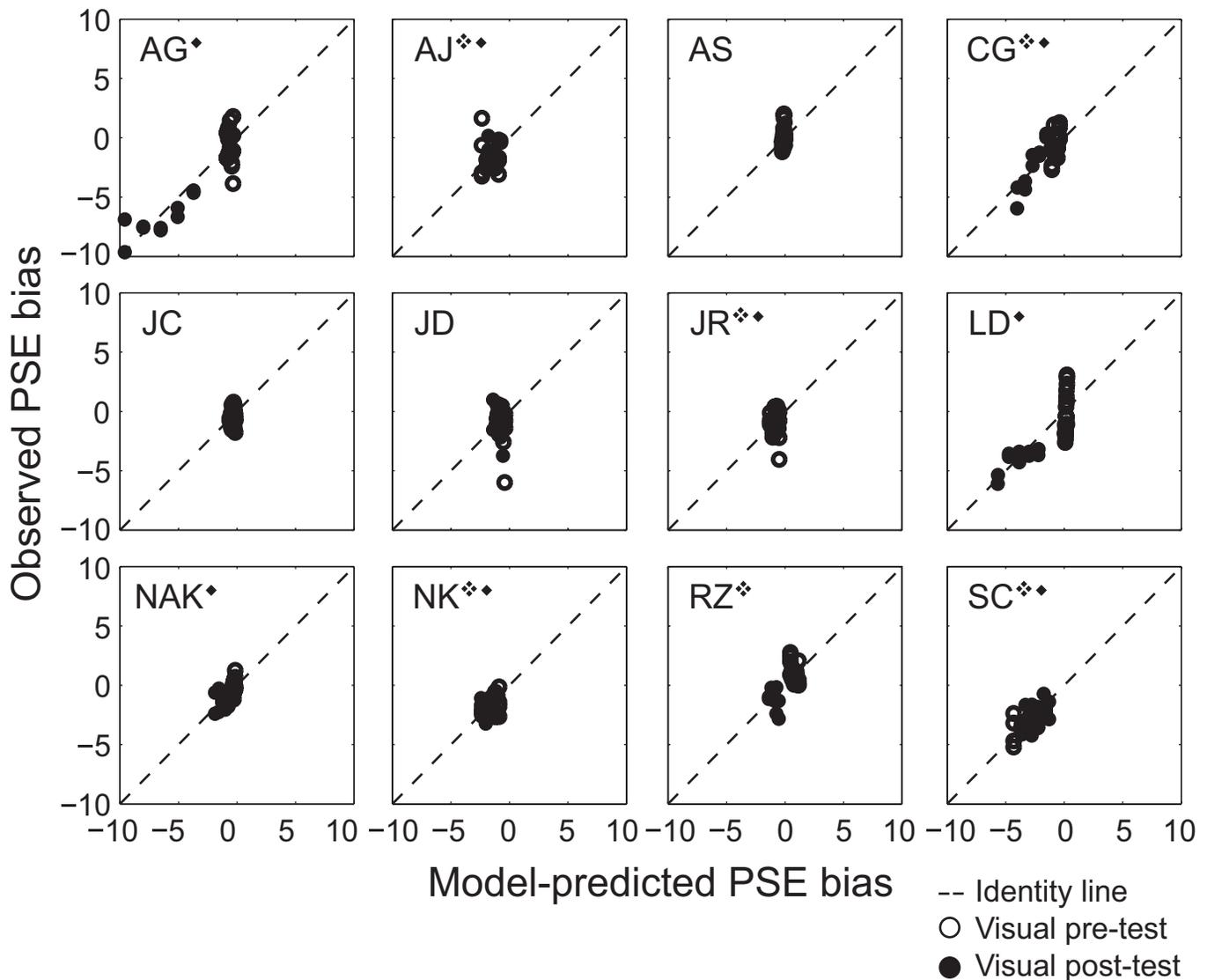


FIG. 10. Predictions of the pseudocue model. Estimates of failure of bump-depth constancy Δd_e are plotted against predicted values $\Delta \hat{d}_e = \hat{a}_e \Delta d_e$ for each observer. The pseudocue model was fit separately to the 20 pretest PSEs (open circles) and 10 posttest PSEs (filled circles). Negative values correspond to PSEs that fell below the line of bump-depth constancy and points clustered tightly around 0 correspond to near bump-depth-constant performance. The \diamond 's and \blacklozenge 's indicate pre- and posttest fits, respectively, with variance-accounted-for values found to be significantly different from 0 ($P < 0.004$, corresponding to an overall type I error rate of 0.05, Bonferroni-corrected for 12 tests). The model does a fairly good job of explaining the pattern of errors exhibited by observers who demonstrated failures of bump-depth constancy.

performance. For these observers, the relative weights of pseudocues and illuminant-invariant cues changed minimally (if at all) between pre- and posttest sessions.

DISCUSSION

In previous studies, we found that most observers perceived rough surfaces viewed under more oblique illuminant positions to be rougher (Ho et al. 2006, 2007), as if the individual bumps in the 3D textures we used were perceived as having increased depth. In the current study, we found that the depth of a single bump was often misperceived in a similar way and that multisensory experience can serve to increase the error in perceived depth. In particular, visuohaptic training can alter perception of depth by shifting the relative weight the visual system applies to available depth cues and pseudocues. More than half of the observers in this study exhibited a failure of bump-depth constancy either in the pre- or posttest sessions.

In the INTRODUCTION, we pointed out that for judgments of roughness, haptic estimates can be more reliable than visual estimates and this might lead observers to calibrate visual estimates to be more consistent with haptic estimates (Atkins et al. 2003; Ghahramani et al. 1997; van Beers et al. 2002). Here, observers made haptic and visual comparisons of depth and we compared the relative reliabilities of visual and haptic estimates. The mean just-noticeable differences (the average slopes of the psychometric functions) from the single-modality sessions indicate that, for 11 of the 12 observers, the reliability of haptic estimates was not significantly different from the reliability of visual estimates ($P > 0.004$, with an overall type I error rate of 0.05, Bonferroni-corrected for 12 tests). For one observer (NK) haptic estimates were significantly *less* reliable than visual estimates ($P < 0.004$). This predicts that NK should not have recalibrated visual cues to match haptic estimates and, indeed, this was the case. However, differences in the relative reliabil-

TABLE 3. Pseudocue model results

Observer	Pretest (20 PSEs)		Posttest (10 PSEs)		Posttest \hat{a}_s - Pretest \hat{a}_s
	\hat{a}_s	VAF, %	\hat{a}_s	VAF, %	
AG	0.22	16.2	2.43	96.9	2.21
AJ	0.60	65.2	0.46	65.2	-0.14
AS	0.07	4.1	-0.03	1.9	-0.10
CG	0.27	39.3	1.01	85.7	0.74
JC	0.11	14.7	0.18	44.8	0.07
JD	0.25	18.6	0.35	33.0	0.10
JR	0.32	39.5	0.31	74.7	-0.01
LD	-0.06	0.9	1.43	96.1	1.49
NAK	0.11	23.4	0.47	74.2	0.36
NK	0.58	86.9	0.61	80.2	0.03
RZ	-0.29	42.7	0.35	49.2	0.64
SC	1.09	95.6	0.84	82.6	2.21

Percentage of variance-accounted-for (VAF) for the pseudocue model for each observer in the pre- and posttest sessions. The estimated coefficient for the proportion of cast shadow, \hat{a}_s , is shown here as well as the VAF. VAF values indicated in boldface were significantly different from 0 at the Bonferroni-corrected α level of 0.05 for 12 tests ($P < 0.004$). Increases in \hat{a}_s from pre- to posttest sessions are an indication of an increase in the pseudocue weight as a result of visuohaptic training.

ity of haptic and visual estimates do not explain individual differences in the training effects for the other observers (Fig. 8).

Judgments of depth for smoothly curved 3D surfaces like the type explored here (e.g., cylinders, ellipsoids) have been found to depend on shading disparities (Bülthoff and Mallot 1988), the presence or absence of cast shadows (Liu and Todd 2004), and viewing distance (Johnston 1991; Johnston et al. 1993, 1994) among others. Varying any of these parameters can produce gross misperceptions of metric depth. However, the difference between these aforementioned studies and the current study is that judgments here do not directly depend on accurate estimates of metric depth. Rather, observers need only to judge relative depth between two stimuli. Even with binocular disparity cues available, many observers gave substantial weight to pseudocues, leading to misjudgments of relative bump depth. This is surprising, given that haptic and visual stimuli were presented within arm's reach (450 mm) where disparities are large and reliable and thus should have been given high weight (Cutting and Vishton 1995; Hillis et al. 2004; Johnston et al. 1994; Landy et al. 1995), resulting in constancy.

Effect of visuohaptic training

Our results demonstrate that visuohaptic training can reinforce pseudocues and lead observers to show a continued, if not more pronounced, failure of bump-depth constancy, but there were large individual differences. In one case (AG), we found that a bump-depth-constant observer can be mistrained so as to exhibit large failures of bump-depth constancy. However, not all bump-depth- or near-bump-depth-constant observers' judgments were affected and the results of mistraining varied substantially. Results for observers who showed initial failure of bump-depth constancy were also variable; some observers showed an even more pronounced failure after training, whereas others maintained the same performance. This is not surprising, given that individuals often differ significantly in the weights applied to different cues (e.g., Oruç et al. 2003).

One possible explanation of the variable effects of training may be the absence of visual guidance during the training session. Recall that the cursor representing hand position was not visible to observers in the area where the stimulus was presented. This was done to avoid conflict between haptic and visual cues caused by presenting the visual cursor simultaneously, in conflict with the visual display during visuohaptic training. Haptic information presented without visual representation of hand position has been found to affect observers' judgments in subsequent visual tasks (Atkins et al. 2001, 2003). However, it has been suggested that superior performance in roughness judgments when both visual and haptic inputs are available is due to the visual information about hand position (Heller 1982). It is therefore possible that for the most effective training to occur, it is necessary to provide visual feedback about hand position.

Another reason visuohaptic training might not have been maximally effective was the fairly large discrepancy between the haptic and visual stimuli presented during visuohaptic training. Although most observers did not notice the difference between visual and haptic input during training, two observers, AG and NK, reported that sometimes the signals did not seem to match in training sessions and this could have led observers to ignore one modality. AG was the one subject who relied almost entirely on haptic input during the training sessions. Moreover, the posttest sessions indicate that, for AG, training led to a complete recalibration so that estimates from visual cues in posttest sessions matched the prediction of haptic-only estimates from the previous training session (Fig. 9, dotted line).

Although the evidence for experience-dependent adaptation was mixed, other studies have suggested that haptic cues can act as a standard to which visual cues are compared (Atkins et al. 2001, 2003; Ernst et al. 2000). In these studies, recalibration or reweighting occurs when visual cues were put in conflict with one another experimentally and one cue was substantially less reliable. In the present study, no direct cue conflict was introduced in the visual stimuli. Instead, the haptic feedback used here during training was presented in conflict with illuminant-invariant cues and artificially correlated with illuminant-variant cues. By doing this, we predicted that the imposed correlation would reinforce shadow size as a cue to visual depth and the visual system would continue to rely on this cue, even in environments where it is not a valid cue to depth.

Recent studies by Backus and colleagues suggest that visual cues can be trained using other cues through *cue recruitment*, a form of associative learning (Backus and Haijiang 2006; Haijiang et al. 2006). They show that the visual system can be trained to use a new arbitrary visual cue such as position in the display to interpret bistable stimuli (e.g., a rotating Necker cube) when this cue is paired with a second, reliable depth cue (e.g., occlusion) via classical Pavlovian conditioning. Ernst (2007) found that observers could be trained to integrate two arbitrary cues (luminance of a visual display and stiffness of a haptic stimulus). In a training session, the two cues were correlated. After training, discrimination was improved when the stimuli to be discriminated were consistent with the correlation learned during training. Thus that study looked at the predictions of cue integration for discriminability, but did not examine how bright or stiff the stimuli were perceived to be.

We found that many observers did not display bump-depth constancy in pretest sessions. This suggests that these observers placed more weight on other cues in the image—that is, pseudocues—to estimate bump depth. It is possible that, over the course of visual development, pseudocues are recruited through visuohaptic experience with 3D textured surfaces under fixed illumination and viewing conditions. We conjecture that overgeneralization of the usage of pseudocues to contexts where they are inappropriate could have accounted for the failures of constancy we found. If illumination in the environment never varied, then shadows and other such images cues would be valid cues to bump depth. Indeed, an alternative interpretation of our results is that training led observers to interpret the shadows as resulting from a light-source position that was less variable (between test and match), implying that the pseudocue was indeed, by inference, more reliable.

Role of pseudocues in 3D surface texture perception

Findings from this study are consistent with our previous studies of the perception of roughness (Ho et al. 2006, 2007), suggesting that the visual system overgeneralizes in using pseudocues where they are not valid cues to depth or roughness, resulting in systematic deviations from constancy. Perceived bump depth and surface roughness both increase under more oblique illumination. The results are consistent with one possible pseudocue, the proportion of cast shadow, but other image statistics could likely explain the behavioral data as well. For example, it has been shown that the human visual system uses the skew of the luminance distribution in making judgments of lightness and gloss (Fleming et al. 2003; Motoyoshi et al. 2007; Nishida and Shinya 1998). Other statistics of the luminance distribution may also serve as pseudocues to depth.

In our study, almost all observers remarked that they noticed the change in illumination between the two scenes presented in each trial. However, when asked what strategy they used to perform the task, most observers mentioned image cues like the change in overall luminance. Even though they may lead to errors, such heuristics provide an easy way to make judgments compared with estimating the illumination and viewing conditions followed by an estimation of shape based on that (i.e., inverse optics).

For some of our observers, reinforcement of pseudocues by haptic feedback led to a more pronounced failure of bump-depth constancy in subsequent visual judgments. Piaget (1952) argued that over the course of early development, motor interactions help give meaning to visual images. Evidence from the developmental literature suggests that a mechanism to integrate information from multiple modalities exists at birth (Sann and Streri 2007; Spelke 1987). Thus visual cues and pseudocues may be learned via cue recruitment during early visuohaptic experience treating the haptic input as a fixed standard.

Several theories of cue combination include the assumption that the relative weights assigned to cues can vary (Bülthoff and Mallot 1988; Landy et al. 1995; Poom and Boerjsson 1999; Porrill et al. 1999; Yuille and Bülthoff 1996). In cross-modal judgments of size, more weight is placed on the more reliable modality (Ernst and Banks 2002) and haptic weight is reduced when visual and haptic locations do not coincide

(Gepshtein et al. 2005). When multiple visual cues are present, haptic feedback can be used to reweight the visual cues (Atkins et al. 2001; Ernst et al. 2000). Thus haptic feedback might be used as a standard against which noisy visual input is measured.

For some subjects, visuohaptic training increased the weight of pseudocues relative to illuminant-invariant cues. If subjects determine cue weights based on relative cue reliability, this effect could be due to visuohaptic training, resulting in an increase in the observer's estimate of the reliability of the shadow cue, a decrease in estimated reliability of illuminant-invariant cues, or both. Our experimental design cannot distinguish between these possibilities. This result is particularly relevant in the perception of 3D surface texture where visual input varies with illumination and viewing conditions, but haptic input does not. This is also consistent with Wallach's theory (Wallach 1985) that in every perceptual domain there exists one primary source that is innate and immutable and other cues are acquired later via correlation with this primary source.

We find that the human visual system may use image pseudocues to judge surface properties, although there are large individual differences. In some cases, the visual system is flexible and can reweight cues and pseudocues in response to multisensory experience. This mechanism might be used to maintain visual constancy. However, this weight learning can occur even when haptic cues reinforce pseudocues, resulting in perception that is further from veridical.

ACKNOWLEDGMENTS

We thank S. Schmidt, T. Schönwetter, and N. Wahl for help with data collection.

GRANTS

This research was supported by National Eye Institute Grants EY-16165 and EY-08266 and the Deutsche Forschungsgemeinschaft (DFG, Emmy-Noether-Programme, Grants TR 528/1-2, 3).

REFERENCES

- Adams WJ, Graf EW, Ernst M. Experience can change the “light-from-above” prior. *Nat Neurosci* 7: 1057–1058, 2004.
- Atkins JE, Fiser J, Jacobs RA. Experience-dependent visual cue integration based on consistencies between visual and haptic percepts. *Vision Res* 41: 449–461, 2001.
- Atkins JE, Jacobs RA, Knill DC. Experience-dependent visual cue recalibration based on discrepancies between visual and haptic percepts. *Vision Res* 43: 2603–2613, 2003.
- Backus BT, Haijiang Q. Competition between newly recruited and pre-existing visual cues during the construction of visual appearance. *Vision Res* 47: 919–924, 2006.
- Belhumeur PN, Kriegman DJ, Yuille AL. The bas-relief ambiguity. *Int J Comput Vis* 35: 33–44, 1999.
- Berkeley G. An essay towards a new theory of vision (G. Berkeley, 1709). In: *Works on Vision*, edited by Turbayne CM. Indianapolis, IN: Bobbs-Merrill Publishing, 1954, p. 277–289.
- Bülthoff HH, Mallot HA. Integration of depth modules: stereo and shading. *J Opt Soc Am A* 5: 1749–1758, 1988.
- Christou CG, Koenderink JJ. Light source dependence in shape from shading. *Vision Res* 37: 1441–1449, 1997.
- Chubb C, Landy MS, Econopouly J. A visual mechanism tuned to black. *Vision Res* 44: 3223–3232, 2004.
- Cutting JE, Vishton PM. Perceiving layout and knowing distances: the integration, relative potency, and contextual use of different information about depth. In: *Perception of Space and Motion*, edited by Epstein W, Rogers SJ. New York: Academic Press, 1995, p. 69–117.

- Efron B, Tibshirani RJ.** *An Introduction to the Bootstrap*. London: Chapman & Hall, 1993.
- Ernst MO.** Learning to integrate arbitrary signals from vision and touch. *J Vis* 7: 1–14, 2007.
- Ernst MO, Banks MS.** Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415: 429–433, 2002.
- Ernst MO, Banks MS, Bühlhoff HH.** Touch can change visual slant perception. *Nat Neurosci* 3: 69–73, 2000.
- Fleming RW, Dror RO, Adelson EH.** Real-world illumination and the perception of surface reflectance properties. *J Vis* 3: 347–368, 2003.
- Gepshtein S, Burge J, Ernst MO, Banks MS.** The combination of vision and touch depends on spatial proximity. *J Vis* 5: 1013–1023, 2005.
- Ghahramani Z, Wolpert DM, Jordan MI.** Computational models for sensorimotor integration. In: *Self-Organization, Computational Maps and Motor Control*, edited by Morasso PG, Sanguineti V. Amsterdam: North-Holland, 1997, p. 117–147.
- Haijiang Q, Saunders JA, Stone RW, Backus BT.** Demonstration of cue recruitment: change in visual appearance by means of Pavlovian conditioning. *Proc Natl Acad Sci USA* 103: 483–488, 2006.
- Heller MA.** Visual and tactual texture perception: intersensory cooperation. *Percept Psychophys* 31: 339–344, 1982.
- Hillis JM, Watt SJ, Landy MS, Banks MS.** Slant from texture and disparity cues: optimal cue combination. *J Vis* 4: 967–992, 2004.
- Ho Y-X, Landy MS, Maloney LT.** How direction of illumination affects visually perceived surface roughness. *J Vis* 6: 634–648, 2006.
- Ho Y-X, Maloney LT, Landy MS.** The effect of viewpoint on perceived visual roughness. *J Vis* 7: 1–16, 2007.
- Jacobs RA, Fine I.** Experience-dependent integration of texture and motion cues to depth. *Vision Res* 39: 4062–4075, 1999.
- Johnston EB.** Systematic distortions of shape from stereopsis. *Vision Res* 50: 1351–1360, 1991.
- Johnston EB, Cumming BG, Landy MS.** Integration of stereopsis and motion shape cues. *Vision Res* 34: 2259–2275, 1994.
- Johnston EB, Cumming BG, Parker AJ.** The integration of depth modules: stereopsis and texture. *Vision Res* 33: 813–826, 1993.
- Klatzky RL, Lederman SJ, Matula DE.** Imagined haptic exploration in judgments of object properties. *J Exp Psychol Learn Mem Cogn* 17: 314–322, 1991.
- Klatzky RL, Lederman SJ, Matula DE.** Haptic exploration in the presence of vision. *J Exp Psychol Hum Percept Perform* 19: 726–743, 1993.
- Landy MS, Maloney LT, Johnston EB, Young M.** Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Res* 35: 389–412, 1995.
- Langer MS, Bühlhoff HH.** Depth discrimination from shading under diffuse lighting. *Perception* 29: 649–660, 2000.
- Larson GW, Shakespeare R.** *Rendering with Radiance: The Art and Science of Lighting and Visualization*. San Francisco, CA: Morgan Kaufmann, 1996.
- Lederman SJ, Abbott SG.** Texture perception: studies of intersensory organization using a discrepancy paradigm, and visual versus tactual psychophysics. *J Exp Psychol Hum Percept Perform* 7: 902–915, 1981.
- Lederman SJ, Klatzky RL.** Relative availability of surface and object properties during early haptic processing. *J Exp Psychol Hum Percept Perform* 23: 1680–1707, 1997.
- Lederman SJ, Summers C, Klatzky RL.** Cognitive salience of haptic object properties: role of modality-encoding bias. *Perception* 25: 983–998, 1996.
- Liu B, Todd JT.** Perceptual biases in the interpretation of 3D shape from shading. *Vision Res* 44: 2135–2145, 2004.
- Longuet-Higgins HC.** A computer algorithm for reconstructing a scene from two projections. *Nature* 293: 133–135, 1981.
- Mayhew JE, Longuet-Higgins HC.** A computational model of binocular depth perception. *Nature* 297: 376–378, 1982.
- Motoyoshi I, Nishida S, Sharan L, Adelson EH.** Image statistics and the perception of surface qualities. *Nature* 447: 206–209, 2007.
- Nishida S, Shinya M.** Use of image-based information in judgments of surface-reflectance properties. *J Opt Soc Am A Opt Image Sci Vis* 15: 2951–2965, 1998.
- Oruç I, Maloney LT, Landy MS.** Weighted linear cue combination with possibly correlated error. *Vis Res* 43: 2451–2468, 2003.
- Piaget J.** *The Origins of Intelligence in Children*. Madison, CT: International Univ. Press, 1952.
- Poom J, Boerjesson E.** Perceptual depth synthesis in the visual system as revealed by selective adaptation. *J Exp Psychol Hum Percept Perform* 25: 504–517, 1999.
- Porrill J, Frisby JP, Adams WJ, Buckley D.** Robust and optimal use of information in stereo vision. *Nature* 397: 63–66, 1999.
- Sann C, Streri A.** Perception of object shape and texture in human newborns: evidence from cross-modal transfer tasks. *Dev Sci* 10: 399–410, 2007.
- Spelke ES.** The development of intermodal perception. In: *Handbook of Infant Perception*, edited by Cohen LB, Salapatek P. New York: Academic Press, 1987, p. 233–260.
- Todd JT.** The visual perception of 3D shape. *Trends Cogn Sci* 8: 115–121, 2004.
- Todd JT, Norman JF.** The visual perception of 3-D shape from multiple cues: are observers capable of perceiving metric structure? *Percept Psychophys* 65: 31–47, 2003.
- van Beers RJ, Wolpert DM, Haggard P.** When feeling is more important than seeing in sensorimotor adaptation. *Curr Biol* 12: 834–837, 2002.
- Wallach H.** Learned stimulation in space and motion perception. *Am Psychol* 40: 399–404, 1985.
- Ward G.** The RADIANCE lighting simulation and rendering system. *Comput Graph* 28: 459–472, 1994.
- Yuille AL, Bühlhoff HH.** Bayesian decision theory and psychophysics. In: *Perception as Bayesian Inference*, edited by Knill DC, Richards W. New York: Cambridge Univ. Press, 1996, p. 123–161.