

# Visual–Haptic Adaptation Is Determined by Relative Reliability

Johannes Burge,<sup>1</sup> Ahna R. Girshick,<sup>4,5</sup> and Martin S. Banks<sup>1,2,3</sup>

<sup>1</sup>Vision Science Program, <sup>2</sup>Department of Psychology and <sup>3</sup>Wills Neuroscience Institute, University of California, Berkeley, Berkeley, California 94720, and <sup>4</sup>Department of Psychology and <sup>5</sup>Center for Neural Science, New York University, New York, New York 10003

Accurate calibration of sensory estimators is critical for maintaining accurate estimates of the environment. Classically, it was assumed that sensory calibration occurs by one sense changing to become consistent with vision; this is visual dominance. Recently, it has been proposed that changes in estimators occur according to their relative reliabilities; this is the reliability-based model. We show that if cue combination occurs according to relative reliability, then reliability-based calibration assures minimum-variance sensory estimates over time. Recent studies are qualitatively consistent with the reliability-based model, but none have shown that the predictions are quantitatively accurate. We conducted an experiment in which the model could be assessed quantitatively. Subjects indicated whether visual, haptic, and visual–haptic planar surfaces appeared slanted positively or negatively from frontoparallel. In preadaptation, we determined the visual and haptic slants of perceived frontoparallel, and measured visual and haptic reliabilities. We varied visual reliability by adjusting the size of the viewable stimulus. Haptic reliability was fixed. During adaptation, subjects were exposed to visual–haptic surfaces with a discrepancy between the visual and haptic slants. After adaptation, we remeasured the visual and haptic slants of perceived frontoparallel. When vision was more reliable, haptics adapted to match vision. When vision was less reliable, vision adapted to match haptics. Most importantly, the ratio of visual and haptic adaptation was quantitatively predicted by relative reliability. The amount of adaptation of one sensory estimator relative to another depends strongly on the relative reliabilities of the two estimators.

## Introduction

Sensory systems must maintain “internal consistency” to ensure that estimates of an environmental property obtained from different sensors agree with one another. Sensory systems must also maintain “external accuracy” to ensure that estimates are veridical and that motor behavior achieves desired goals. Maintaining internal consistency has often been incorrectly modeled as equivalent to maintaining external accuracy. We ask how human sensory systems maintain internal consistency among sensor estimates.

Consider two sensors, visual ( $V$ ) and haptic ( $H$ ), estimating environmental property  $S$ . We represent this as follows:  $\hat{S}_V = f_V(S)$  and  $\hat{S}_H = f_H(S)$ , where  $f_V$  and  $f_H$  are the estimators and  $\hat{S}_V$  and  $\hat{S}_H$  are their estimates. The sensory system is internally consistent when  $E(\hat{S}_V) = E(\hat{S}_H)$ , where  $E$  is the expectation of each estimate.

Several investigators claim that vision determines the combined estimate. In such “visual capture,” what is seen determines what is felt or heard (Rock and Victor, 1964; Hay et al., 1965; Warren and Rossano, 1991). Others claim that the combined estimate is a statistically optimal, minimum-variance, weighted

average. Given two Gaussian-distributed estimates  $\hat{S}_V$  and  $\hat{S}_H$ , the combined estimate is as follows:

$$\hat{S}_{VH} = w_V \hat{S}_V + w_H \hat{S}_H \quad (1)$$

$$w_V = \frac{r_V}{r_V + r_H}; w_H = \frac{r_H}{r_V + r_H} \quad (2)$$

$$r_i = 1/\sigma_i^2, \quad (3)$$

where  $w_V$  and  $w_H$  are the combination weights and  $r_V$  and  $r_H$  are the reliabilities of the estimates, where reliability is defined as the inverse of the measurement variance,  $\sigma^2$  (Yuille and Bülthoff, 1996; Jacobs, 1999). Combining information this way provides a better empirical account of sensory combination than does visual capture (Ernst and Banks, 2002; Gepshtein and Banks, 2003; Alais and Burr, 2004).

If estimators in the optimal combination model (Eqs. 1–3) are not calibrated, problems arise. Consider presenting a surface with a slant of  $0^\circ$  to the eye and hand. Suppose that vision and touch are equally reliable, but that distorting spectacles have biased vision by  $20^\circ$  (Adams et al., 2001). When the stimulus is seen and felt, the bias causes an inaccurate combined estimate of  $10^\circ$  (Eq. 1), undercutting a potential benefit of cue combination: increased accuracy. Thus, the combination model that is optimal for unbiased estimators generally yields suboptimal estimates when the estimators are biased. Internal inconsistency introduces another problem: If the stimulus is seen but not felt, its perceived slant will be  $20^\circ$ . If it is felt but not seen, it will be  $0^\circ$ . An achievement of perception is undermined: perception of an environ-

Received Dec. 27, 2009; revised March 5, 2010; accepted April 27, 2010.

This work was supported by American Optometric Foundation's William C. Ezell Fellowship to J.B. and Research Grants NIH R01-EY12851 and NSF BCS-0617701 to M.S.B. We thank Marc Ernst for helpful discussion.

Correspondence should be addressed to Johannes Burge at the above address. E-mail: jburge@mail.cps.utexas.edu.

DOI:10.1523/JNEUROSCI.6427-09.2010

Copyright © 2010 the authors 0270-6474/10/307714-08\$15.00/0

mental property as constant despite changes in the proximal stimuli used to estimate the property.

These problems show why cue calibration is essential for cue combination. The classic studies on cue calibration examined how vision and another sense respond to miscalibration. The data often suggest that nonvisual estimators, like touch, calibrate to match vision in spatial tasks, a behavior called “visual dominance” (Rock and Victor, 1964; Hay and Pick, 1966; Knudsen and Brainard, 1991; Knudsen, 2002). In contrast, some recent studies show that visual recalibration occurs (Lewald, 2002; van Beers et al., 2002; Atkins et al., 2003). Perhaps a reliability-based model of cue calibration (Ghahramani et al., 1997) accounts for human behavior better than visual dominance does, just as a reliability-based model of cue combination is superior to visual capture. This model is given by the following:

$$\Delta f_V(S) = \eta p_V (\hat{S}_H - \hat{S}_V)$$

and

$$\Delta f_H(S) = \eta p_H (\hat{S}_V - \hat{S}_H) \quad (4)$$

$$p_V = w_H; p_H = w_V, \quad (5)$$

where  $\Delta f_V$  and  $\Delta f_H$  are additive changes in the visual and haptic estimators,  $p_V$  and  $p_H$  are the proportions of visual and haptic calibration, and  $\eta$  is a small positive constant determining adaptation rate. Note that the proportion of visual adaptation equals the haptic weight, not the visual weight; therefore, if the haptic weight is high, vision should adapt more and vice versa. Assuming that estimators do not drift, this model maximizes the reliability of the combined estimate and minimizes the change in bias after discrepancy is introduced.

## Materials and Methods

**Overview.** A strong test of the reliability-based model of sensory calibration should have three features. (1) The experimental design should include measurements of the reliability of each estimator so that the quantitative predictions of the reliability-based model can be computed. (2) The relative reliabilities of the estimators should be manipulated across a reasonable range so that different reliability-based outcomes can be generated. (3) A purely perceptual response should be used so that one can distinguish sensory from motor recalibration.

Our experimental design incorporates all three features. With respect to the first issue, we measured the reliabilities of the visual and haptic estimators separately before adaptation began. With respect to the second issue, we manipulated the relative reliabilities of vision and haptics and examined adaptation for different relative reliabilities. In one condition, vision was more reliable, so the reliability-based model predicted more haptic adaptation. In another condition, haptics was more reliable, so more visual adaptation was predicted. With respect to the third issue, subjects compared their slant percepts to an internal standard: they indicated whether visual and haptic surfaces were slanted right-side back (positive slant) or left-side back (negative) from frontoparallel. Thus, changes in either the visual or haptic estimator could be detected and measured. Three previous studies contained some of these design features (Ghahramani et al., 1997; Lewald, 2002; van Beers et al., 2002), but to our knowledge, none contained all three. We expand on this point in the Discussion.

The experiment consisted of two distinct but closely related studies: a cue-combination study and a cue-calibration study. The combination study measured estimator reliabilities. Those reliabilities were used to construct the conditions in the calibration study. The calibration study examined whether visual dominance or the reliability-based model provides a better description of human cue calibration. Visual dominance predicts that all adaptation will be haptic for all relative reliabilities. In other words, it predicts that the haptic estimator will change during

adaptation and that the visual estimator will not, even when the reliability of the haptic estimator is greater than that of the visual estimator. The reliability-based model predicts that the less reliable estimator will adapt more, and by an amount predicted by the reliability ratio (Eqs. 4, 5).

**Subjects.** Ten subjects, 22–33 years of age, participated. All were naive to the experimental purpose, were right handed, and had normal stereopsis and normal or corrected-to-normal visual acuity.

**Stimuli.** The stimulus was a  $110 \times 110 \text{ mm}^2$  plane that was slanted about the vertical axis (tilt =  $0^\circ$ ). Slant ranged from  $-10$  to  $10^\circ$  with an average of  $0^\circ$ . To estimate slant, subjects viewed the plane binocularly and/or touched it with the index finger and thumb.

The haptic stimulus was generated by two PHANTOM (SensAble Technologies) force-feedback devices, one each for the index finger and thumb. Finger and thumb movements had all six degrees of freedom in the  $20 \text{ cm}^3$  workspace. The three-dimensional (3D) positions of the tips of the finger and thumb were monitored, and appropriate forces were applied when they reached the positions of the simulated haptic object. The PHANTOMS provided a compelling simulation of the orientation, size, and stiffness of the stimulus plane.

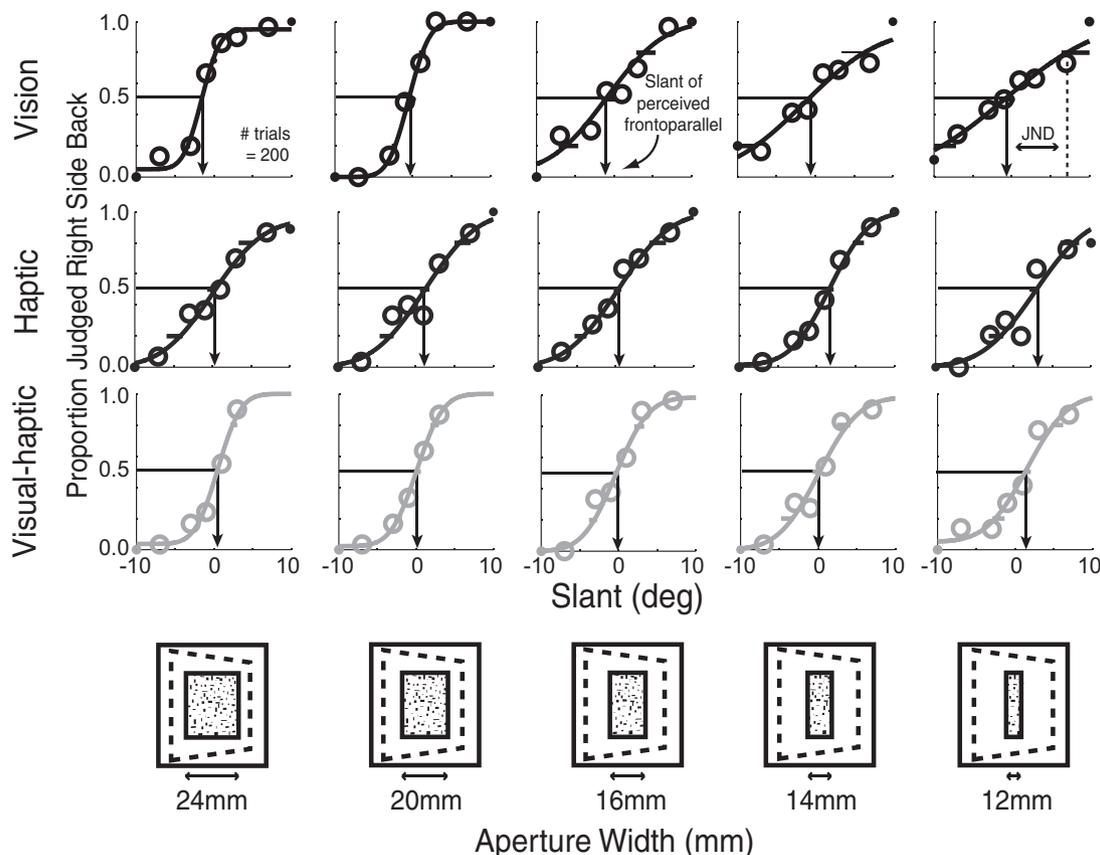
The visual stimulus was a random-dot stereogram simulating a slanted plane. The dots subtended 8 arcmin at the 49 cm viewing distance. Dot density was  $\sim 3 \text{ dots/deg}^2$ . New dots were displayed with each presentation. Finger and thumb positions were tracked by the PHANTOMS and indicated by small 3D markers that were visible between trials and invisible during trials. A rectangular aperture was positioned  $3 \pm 0.5 \text{ cm}$  in front of the slanted plane's center. The aperture was 92 mm high and ranged from 12 to 60 mm in width. With a narrower aperture, fewer dots were visible, and visual reliability decreased (verified by measuring slant-discrimination thresholds).

The apparatus was calibrated to align the visual and haptic stimuli spatially. Trials were initiated by the subject placing the finger and thumb within a visible sphere, 4 cm in diameter, that was  $5 \pm 1 \text{ cm}$  in front of the center of the slanted plane. In the vision-alone trials, stimulus presentation began 400 ms after the fingers reached the starting location. In haptic-alone trials, stimulus presentation began when the thumb and finger contacted the plane, which for most subjects was  $\sim 400 \text{ ms}$  after their digits entered the starting sphere. Visual–haptic trials were initiated in the same way as haptic-alone trials. In all cases, stimulus duration was 1 s.

**Procedure.** Perceived slant was measured using a one-interval, two-alternative forced-choice procedure. In each interval, a plane was presented (visual-alone, haptic-alone, or visual–haptic). At the end of the presentation, the subject indicated whether he/she perceived the plane to be slanted positively or negatively from frontoparallel. No feedback was provided. Slant was varied from trial to trial according to the method of constant stimuli. Using a maximum-likelihood criterion (Wichmann and Hill, 2001), we found the cumulative Gaussian function that best fit the psychometric data. Those data were the proportion of trials in which the subject indicated that the plane was positively slanted as a function of its physical slant. The mean of the best-fitting function was our estimate of the slant that looked frontoparallel and the SD of the function was our estimate of the just-noticeable difference (JND).

In each block of the cue-combination study, visual-alone, haptic-alone, and visual–haptic trials were randomly interleaved. Two hundred trials were presented in each of the three conditions at each of five aperture widths. Each block consisted of 300 trials (100 per condition). Two blocks were run at each aperture width for a total of 10 blocks.

In the cue-calibration study, two aperture widths were chosen for each subject based on their results in the cue-combination study. We used the two widths that made the visual–haptic reliability ratio ( $r_V:r_H$ ) equal to 3:1 and 1:3. By using these ratios, we could readily determine whether changes during adaptation were reliability based or not. Each experimental session lasted  $\sim 60 \text{ min}$  and consisted of three phases: preadaptation, adaptation, and postadaptation. Each subject went through four such sessions, one for each pairing of two aperture widths (reliability ratio = 3:1 and 1:3) and two visual–haptic discrepancies during adaptation ( $-9$  and  $+9^\circ$ ), where discrepancy is defined as  $S_V - S_H$ . Sessions were always separated by at least 24 h to minimize possible carryover from previous sessions.



**Figure 1.** Cue-combination data from one representative subject. Visual-alone, haptic-alone, and visual–haptic psychometric data are shown as a function of aperture width. The proportion of trials in which the slanted plane was perceived as right-side back from frontoparallel is plotted against the slant of the plane for visual-alone (top row), haptic-alone (middle row), and visual–haptic (bottom row) trials. The icons at the bottom depict how the variation of the width of the aperture through which the slanted planes were seen. Arrows indicate the slant that is perceived as frontoparallel. The JND is one SD from perceived frontoparallel; shallower slopes indicate larger JNDs.

In the preadaptation phase, the visual-alone, haptic-alone, and visual–haptic stimuli were each presented 100 times for a total of 300 trials. We used this phase to determine the visual-alone, haptic-alone, and visual–haptic surface slants that were perceived as frontoparallel before adaptation began.

Only visual–haptic stimuli were presented in the adaptation phase. The discrepancy between the visually and haptically specified slants was increased by  $1.5^\circ$  every 96 trials until it reached a maximum of  $\pm 9.0^\circ$ . By gradually increasing the discrepancy, we made it less apparent to the subjects and achieved more complete adaptation (Kagerer et al., 1997). One hundred twenty trials were run at the largest discrepancy, so there were 600 trials in the adaptation phase. The slants of the visual–haptic stimuli (i.e., the average of the visually and haptically specified slants) ranged from  $-10$  to  $10^\circ$ , but the average of those slants across the adaptation phase was  $0^\circ$ . The adaptation phase lasted  $\sim 30$ – $45$  min.

The postadaptation phase was identical to the preadaptation phase except that the visual–haptic trials retained the largest discrepancy presented in the adaptation phase ( $-9.0$  or  $9.0^\circ$ ). By presenting the visual–haptic discrepant stimuli, we assured that adaptation was maintained and therefore that subjects did not regress to their preadaptation calibration. We assessed visual adaptation by comparing preadaptation and postadaptation responses to the visual-alone stimuli and haptic adaptation by comparing preadaptation and postadaptation responses to the haptic-alone stimuli.

The subjects were questioned at the end of the four experimental sessions. None noticed the visual–haptic discrepancy during the adaptation and postadaptation phases.

## Results

### Recalibration experiment

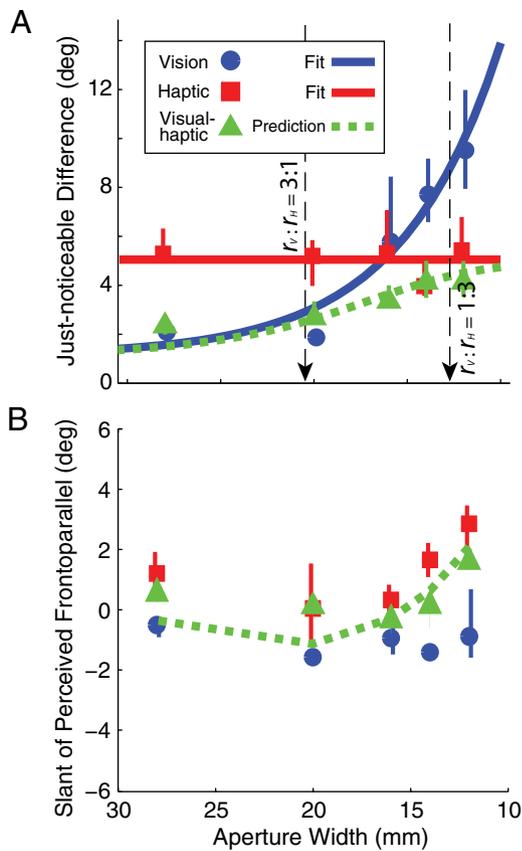
Figure 1 shows psychometric data for one representative subject for the visual-alone (upper row), haptic-alone (middle row), and

visual–haptic (lower row) stimuli in the cue-combination study. The proportion of trials in which the subject indicated the stimulus was slanted positively is plotted as a function of stimulus slant. The slopes of the visual-alone functions were lower for narrow apertures than for wide ones. Thus, the visual slant estimates were, as we hoped, less reliable with narrow apertures. The slopes of haptic-alone data were unaffected by the width of the aperture.

Figure 2 provides summary data for the same subject. Figure 2A plots JNDs as a function of aperture width for visual-alone, haptic-alone, and visual–haptic stimuli. The visual–haptic JNDs were well predicted by the optimal cue-combination model (Eqs. 1–3). Figure 2B plots the slants that were perceived as frontoparallel as a function of aperture width. The subject had small biases of opposite signs in the visual and haptic estimates of slant. As predicted by the optimal combination model, visual–haptic estimates were similar to the reliability-based averages of the visual and haptic estimates (Eqs. 1–3).

Figure 3 plots the data from all the subjects. Observed JNDs (Fig. 3A) and visual–haptic slants perceived as frontoparallel (Fig. 3B) are plotted as a function of the values predicted by the optimal combination model. The cue-combination model, which weights each cue according to its reliability, provided an excellent account of the data (but see Rosas et al., 2005).

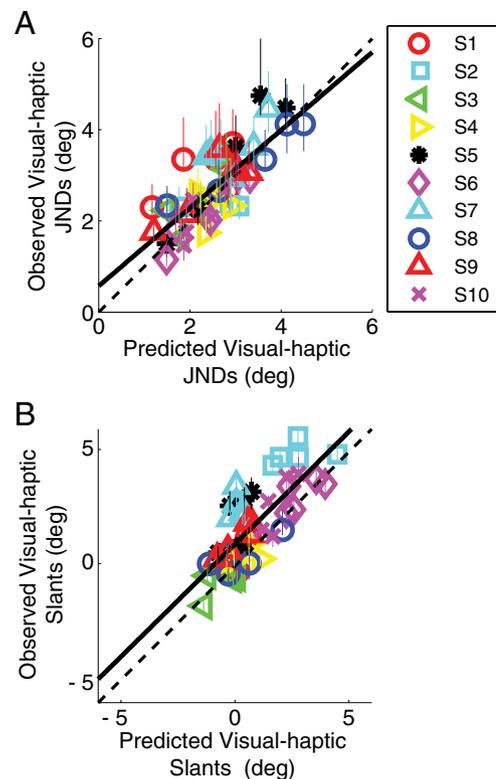
From the cue-combination data, we found the aperture widths that would produce reliability ratios of 3:1 and 1:3 for each subject. These ratios allowed us to most readily distinguish the



**Figure 2.** Cue-combination results for one representative subject. **A**, Visual-alone slant discrimination thresholds (blue circles) decreased systematically as aperture width increased. As expected, haptic-alone discrimination thresholds (red squares) remained constant with aperture changes. The blue line is an exponential fit to the visual thresholds. The red line is a horizontal line fit to haptic JNDs. The black arrows mark aperture widths for which the ratios of visual and haptic reliabilities are 3:1 and 1:3 (for this subject: 20.5 and 12.6 mm, respectively). The visual–haptic thresholds (green triangles) and the reliability-based prediction for combined thresholds are also shown (green dashed line). The error bars are 95% confidence intervals. **B**, Visual-alone, haptic-alone, and visual–haptic slants of perceived frontoparallel. The green line is the zero-free-parameter reliability-based prediction of the visual–haptic slants of perceived frontoparallel based on the unimodal thresholds and unimodal slants of perceived frontoparallel. Across subjects, the visual bias was not significantly different from zero (*t* test,  $t_{(9)} = -1.43, p = 0.19$ ). The haptic bias was significantly different from zero (*t* test,  $t_{(9)} = 3.1372, p < 0.012$ , mean = 2.1°, CI = 0.58–3.59), although the change in bias with aperture size was not significant (ANOVA,  $F_{(1,9)} = 1.66, p = 0.18$ ). The haptic bias is presumably due to slight differences between the position tracked by the force-feedback device and the actual surface of the finger.

predictions of the visual-dominance and reliability-based models. If visual dominance occurs, the haptic estimator should exhibit a significant change and the visual estimator should not change at all, regardless of the reliability ratio. If reliability-based recalibration occurs, vision should change by one-third as much as haptics when the reliability ratio is 3:1 ( $r_V:r_H$ ) and by three times as much as the haptics when the ratio is 1:3.

Figure 4 shows the cue-calibration data. Figure 4A plots the change in the visual estimator between the preadaptation and postadaptation phases on the abscissa and the change in the haptic estimator on the ordinate. The visual-dominance prediction is the vertical dotted line (no change in vision). The reliability-based predictions are the magenta and purple dashed lines for the 3:1 and 1:3 reliability ratios, respectively. The magenta circles are the data when the reliability ratio was 3:1, and the purple squares are the data when the ratio was 1:3. The data are much more



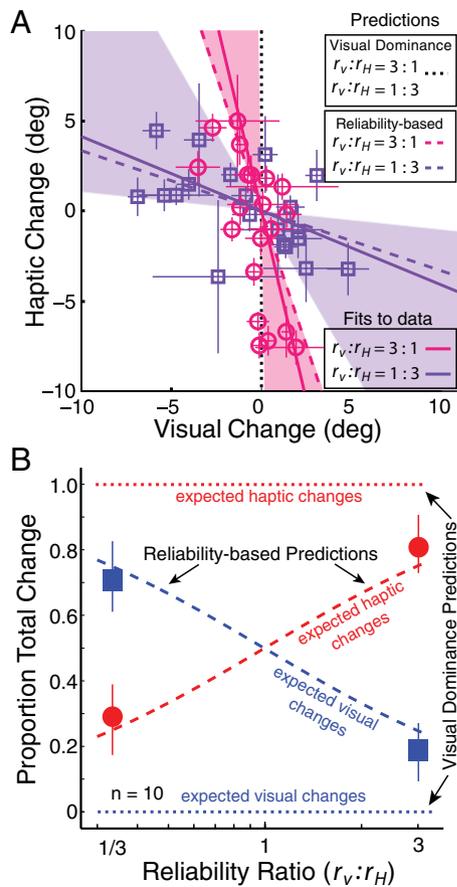
**Figure 3.** Observed versus predicted data for the cue-combination study for all subjects. **A**, Observed JNDs plotted against predicted JNDs ( $R^2 = 0.603$ ). The solid black line is a linear regression to the data ( $y = 0.85x + 0.57$ ). **B**, Slants that were perceived as frontoparallel plotted against the predictions. Subjects performed in a manner consistent with the predictions ( $R^2 = 0.605$ ). The solid black line is a linear regression to the data ( $y = 0.99x + 0.95$ ). The positive shift of the predicted slants of perceived frontoparallel reflects the slight haptic bias.

consistent with the reliability-based model than with visual dominance. Figure 4B shows the predictions and observations in another format. The proportion of change between preadaptation and postadaptation is plotted as a function of the reliability ratio ( $r_V:r_H$ ). Again the reliability-based model provides a much more accurate account of the data than visual dominance.

**Modeling of recalibration**

An observed discrepancy between two sensory estimates of the same environmental property could be due to random measurement error to which the system should not recalibrate, or due to systematic error to which the system should recalibrate. The problem of determining whether two sensory estimators are internally consistent is equivalent to the statistical problem of determining whether two random variables have the same means (Burge et al., 2008). How should the nervous system adapt to minimize systematic error, thereby maintaining internal consistency over time? That is, what is the optimal strategy? In the absence of external feedback, one cannot determine which of two estimators is inaccurate. We therefore investigated the strategy that minimized the change in bias and variance of the combined estimate.

We simulated recalibrating systems under two conditions. In the first simulation, we assumed that individual estimators do not randomly drift. In the second simulation, we considered the effects of drifting estimators. In both simulations, the reliability ratio of the two estimators ( $r_V:r_H$ ) was set to 1:3 and a discrepancy of fixed size was introduced between the visual and haptic estimators at time 0. We simulated the effects of the introduced



**Figure 4.** Adaptation predictions and data. **A**, Adaptation of each estimator in the conditions with high (magenta circles) and low visual reliability (purple squares). Each symbol represents the adaptation observed for one subject in one condition. The adaptation of each estimator was quantified as the difference between the preadaptation and postadaptation slants of perceived frontoparallel ( $PSE_{V_{post}} - PSE_{V_{pre}}$ ,  $PSE_{H_{post}} - PSE_{H_{pre}}$ ). Error bars on the data points represent  $\pm 1$  SD of this difference. The visual-dominance model predicts that all the data would lie on the vertical dotted line. The reliability-based model (Eq. 5) predicts that the data from the two reliability conditions will fall on the dashed purple (3 times more visual than haptic adaptation) and dashed magenta (1/3 times more visual than haptic adaptation) lines. The solid lines represent the best fits to the data that passed through zero. The shaded areas around the solid lines represent 95% confidence intervals for the best-fit lines from 1000 bootstrapped datasets. **B**, Group-average predictions and results. The average proportion of visual and haptic adaptation is plotted in the two reliability conditions: high and low visual reliability ( $r_V:r_H = 3:1$  and  $r_V:r_H = 1:3$ , respectively). The horizontal dotted blue and red lines represent the visual dominance predictions for visual and haptic changes, respectively, as a function of reliability ratio. All of the change is predicted to be haptic, regardless of reliability ratio. The diagonal dashed blue and red lines show the predictions of the reliability-based model for visual and haptic changes as a function of reliability ratio. The blue squares and red circles represent the visual and haptic data. Error bars represent 95% confidence intervals computed from 1000 bootstrapped datasets.

discrepancy under all possible permutations of visual combination weights,  $w_V$ , and proportions of visual change,  $p_V$ . Under the present assumptions of the model, and in the absence of additional information about the probability of an external disturbance, there is no principled prediction for the adaptation rate. We therefore set the adaptation rate  $\eta$  to an arbitrary but plausible value: 0.05. Other adaptation rates yielded the same qualitative results. We then tracked the changes to the two estimators and the combined estimate as the estimators adapted over time. We repeated this procedure 1000 times each for a range of combination weights and calibration proportions.

The simulation results for the optimal calibration rule and a suboptimal rule are in Figure 5. The optimal calibration rule

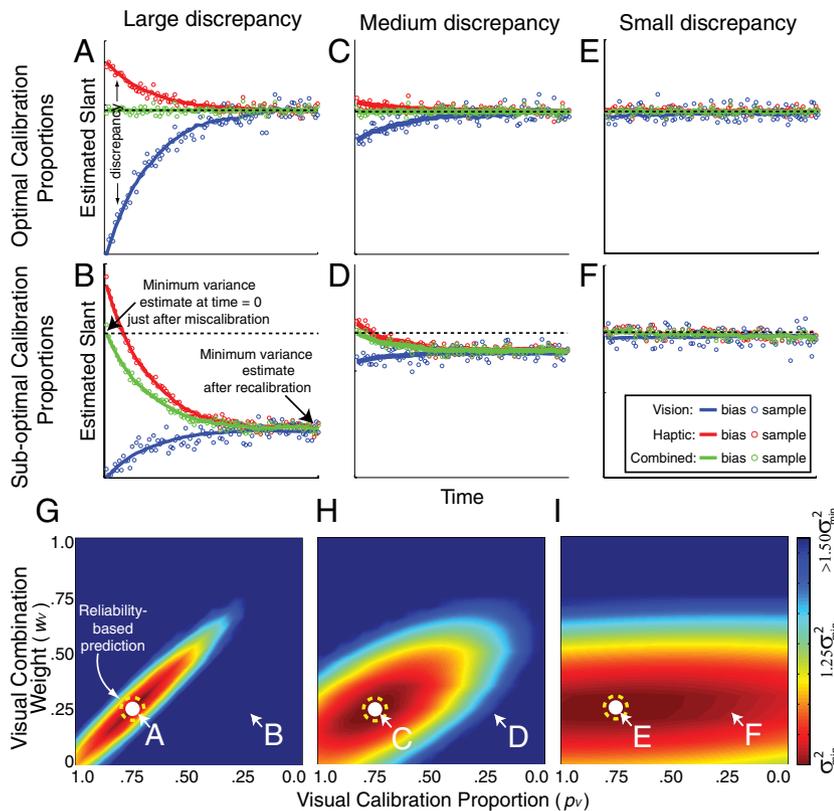
minimizes the variance and change in bias of the combined estimates as the visual and haptic estimators change to restore internal consistency. Given that the reliability ratio ( $r_V:r_H$ ) was set to 1:3, the optimal combination weights ( $w_V$  and  $w_H$ ; Eq. 2) are 0.25 and 0.75. The upper row shows the results when the calibration rule and combination rule were matched (“optimal calibration rule,” i.e.,  $p_V = 1 - w_V$ ), the visual weight ( $w_V$ ) was 0.25, and the proportion of visual change ( $p_V$ ) was 0.75. The middle row shows the results for a suboptimal rule in which the combination and calibration rules were not matched. When the calibration and combination rules match, the combined estimate remains constant over time. When the rules do not match, the combined estimate’s bias changes as the system adapts over time. The amount of change of the combined estimate depends strongly on the size of the initial discrepancy.

The lower row of Figure 5 summarizes the results for all possible combination weights and calibration proportions. When the calibration and combination rules match (optimal rule; i.e.,  $p_V = 1 - w_V$ ), the variance of the combined estimator over time is minimized. When the rules do not match (suboptimal rule), the variance is higher. When the discrepancy introduced between the visual and haptic estimators is large, the combined estimator variance is critically dependent on matching the rules. When the discrepancy is small, rule matching is not nearly as consequential.

The simulation shows that matching the calibration and combination rules minimizes the variance of the combined estimates as the visual and haptic estimators change to restore internal consistency. Previous work has shown that the combination weights in sensory cue-combination are essentially optimal (Ernst and Banks, 2002; Alais and Burr, 2004). Thus, our simulation shows that reliability-based cue calibration is needed to minimize variance over time (assuming that the estimators do not drift). Why is it advantageous to minimize the variance of the combined estimate over time? To maintain external accuracy, the adapting system has to compute error signals: the difference between the system’s current estimates and external feedback. Clearly, the variance of the error signal depends on the variance of the current estimate. It is easier to detect systematic changes in low-variance signals than in high-variance signals. Thus, by minimizing the variance of the combined estimate, even during adaptation, the system becomes better able to detect the need to recalibrate when external feedback becomes available (Burge et al., 2008). We conclude that the reliability-based model of cue calibration is optimal in that it yields the lowest variance combined estimate and makes the system most able to detect systematic changes between sensory estimators.

The analysis thus far has assumed that estimators are generally stable and that significant recalibration occurs only in the face of external disturbances. It is perhaps more likely that individual estimators drift randomly from trial to trial (Baddeley et al., 2003), so we now examine the consequences of such drift.

In a second simulation, we added random drift,  $d$ , to each estimator after the visual and haptic changes ( $\Delta f_V$  and  $\Delta f_H$ ) had been calculated. After each slant estimate on trial  $t$ , the visual estimator was changed by  $\Delta f_{V_t} \pm d_{V_t}$  and the haptic estimator by  $\Delta f_{H_t} \pm d_{H_t}$ . We varied the total amount of drift ( $d_V + d_H$ ) and the ratio of the drift magnitudes ( $d_V:d_H$ ) and observed the effects by repeating the simulation in Figure 5 with one additional wrinkle: The adaptation rate was set according to the Kalman-filter update rule, the optimal adaptation rule for dynamic linear systems with Gaussian-distributed noise (Kalman, 1960; Maybeck,



**Figure 5.** Simulation results for estimators with the reliability ratio ( $r_V:r_H$ ) set to 1:3 and an adaptation rate of 0.05. Different adaptation rates did not change the qualitative effects. For different combination weights and calibration proportions, we calculated the variance of the combined estimate over time for 1000 simulated trials following an initial discrepancy. The left, center, and right columns represent the results when the visual–haptic discrepancy introduced at time 0 was large ( $5\sigma_V^2$ ), medium ( $\sigma_V^2$ ), and small ( $\sigma_V^2/5$ ), respectively. The estimators did not drift randomly. The panels in the upper two rows plot the visual, haptic, and combined estimates over time (blue, red, and green, respectively). The upper row shows adaptation when the visual calibration proportion ( $p_V$ ) was appropriate for reliability-based adaptation ( $p_V = 1 - w_V$ ). The middle row shows adaptation when the visual calibration proportion was inappropriate for reliability-based adaptation ( $p_V = w_V$ ). The bottom row summarizes adaptation for a wide variety of parameters. The abscissa is the visual calibration proportion ( $p_V$ ). The ordinate is the visual combination weight ( $w_V$ ). Color represents the variance of the combined estimate over time, dark red representing the smallest variance. The white circles indicate the position of minimum variance in each figure. The white-lettered labels **A–F** indicate the parts of those surfaces that are shown in the panels of the upper two rows. The dashed yellow circles indicate the optimal calibration rule given that  $r_V:r_H$  equaled 1:3: the visual calibration proportion and visual combination weight that should yield minimum variance according to the reliability-based model.

1979). The results are plotted in Figure 6. We observed two effects of interest:

- (1) When the drift was small, the minimum variance was predicted by the estimator reliability ratio ( $r_V:r_H$ ). When the drift was large, it was predicted by the drift ratio ( $d_V:d_H$ ). Specifically, the visual and haptic calibration proportions were predicted by the following:

$$p_V = \frac{d_V}{d_V + d_H}; p_H = \frac{d_H}{d_V + d_H}. \tag{6}$$

- (2) The estimator reliability ratio was a better predictor of the minimum-variance point when the introduced discrepancy was large. The drift ratio was a better predictor when the introduced discrepancy was small.

In summary, the reliability-based model accurately predicts the adaptation proportions that yield the minimum-variance combined estimate when two conditions are present: (1) the amount of random drift is small relative to the measurement variance (Eq. 3), and (2) the discrepancy is large relative to the drift. If these two condi-

tions do not hold, the adaptation proportions yielding the minimum-variance estimate are better predicted by the relative drifts defined in Equation 6. This result makes sense. If the drift is large relative to the measurement variances (i.e., the inverse reliabilities), the measurement variance will have very little effect. If, on the other hand, the drift is small relative to the measurement variances, the measurement variance will have more effect, so the reliability-based model provides a more accurate account. These simulation results are consistent with many years of research on Kalman filter behavior.

Our main empirical result is that the reliability-based model provides an excellent account of visual–haptic adaptation (Figs. 2–4). Our simulations show that the reliability-based prediction is observed when the drift of the estimators is small relative to the measurement noise associated with the estimators. From these two findings, we infer that the estimator drift in the human experiment must have been small relative to the measurement noise, under the assumption that the model is correct. We also found that the amount of visual adaptation was slightly less than predicted by the reliability-based model and that haptic adaptation was slightly more than predicted (these trends were not statistically significant). From the simulation results, we see that this behavior is consistent with a visual estimator that is more stable (less susceptible to drift) than the haptic estimator.

## Discussion

### Previous work and reliability-based calibration

Several investigators have proposed that the relative reliability of sensory estimators determines the amount by which the estimators change during adaptation (e.g., Witten and Knudsen, 2005). Three sets of investigators have tested this proposition (Ghahramani et al., 1997; Lewald, 2002; van Beers et al., 2002). Ghahramani et al. (1997) examined visual–auditory adaptation in a localization task. They measured the relative reliabilities of the visual and auditory estimators. The visual reliability was much higher than the auditory, so visual dominance and the reliability-based model made essentially the same predictions. They observed that essentially all of the measurable adaptation occurred in the auditory estimator, which is consistent with both the visual-dominance and reliability-based models. Lewald (2002) also looked at visual–auditory adaptation. He found statistically significant evidence for visual adaptation, which is inconsistent with visual dominance. He did not, however, measure relative reliability, so one cannot tell whether his data were quantitatively consistent with the reliability-based model. van Beers et al. (2002) measured visual–proprioceptive adaptation. They also observed significant visual adaptation,

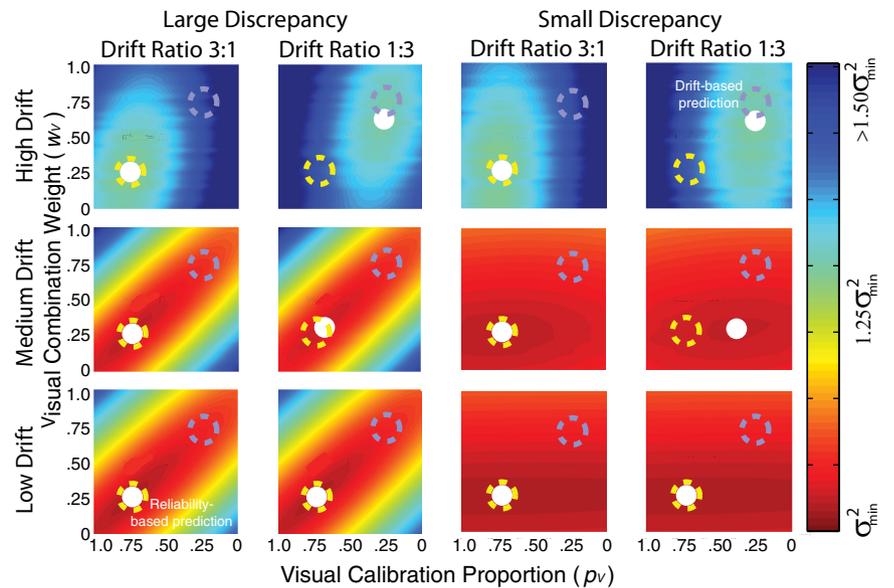
which is again inconsistent with visual dominance. They also tested different conditions that were likely to change relative reliabilities of the two estimators and observed sensible changes in the amounts of visual and proprioceptive adaptation. They did not, however, measure estimator reliabilities, so again one cannot determine whether the data were quantitatively consistent with the reliability-based model.

Ghahramani et al. (1997), Lewald (2002), and van Beers et al. (2002) all assessed adaptation with a pointing task. When motor responses are used, one generally cannot distinguish sensory from motor recalibration. Consider, for example, the prism-adaptation study (von Helmholtz, 1867). A prism is placed in front of the eye and the accuracy of pointing to a visual target is measured. When the prism is first introduced, pointing is displaced in the direction of prism displacement. After a few trials, accuracy is restored, which shows that adaptation has occurred. One cannot easily determine whether the adaptation reflects a change in the visual estimate of direction, in the motor commands to point, or in the motor system's response to a given command [but see Lewald (2002), Zwiers et al. (2003), and Hernandez et al. (2008) for indirect methods that resolve this issue]. In the Ghahramani et al. (1997), Lewald (2002), and van Beers et al. (2002) studies, some combination of sensory and motor adaptation could have produced the adaptation effects they observed. Thus, by using a pointing response, they were unable to determine how much of the adaptation they observed was sensory.

In the work reported here, we assessed whether the reliability-based model of cue calibration accounts quantitatively for human behavior. We measured the relative reliability of the visual and haptic estimators, manipulated their relative reliabilities, and chose relative reliabilities that allowed the predictions of visual dominance and reliability-based models to differ sufficiently. We also used a purely perceptual task, so we could be confident that the observed adaptation was all sensory. Our experiment was thus able to demonstrate convincingly that the visual-dominance model does not account for cue calibration and, more importantly, that the reliability-based model provides a quantitatively accurate account of sensory cue calibration.

### External accuracy

Perceptual and motor systems must maintain calibration for effective interaction with the environment. The literature on recalibration has therefore focused primarily on how to achieve and maintain the external accuracy of sensory and sensorimotor systems. However, external accuracy cannot be maintained without external feedback, and because external feedback is not always available, it is important to understand how the system performs without it. We focused here on calibration in the absence of feedback: i.e., on the maintenance of internal consistency.



**Figure 6.** Simulation results showing the variance of the combined estimate when the estimates undergo a random walk. The abscissa in each panel is the visual calibration proportion ( $p_v$ ) and the ordinate is the visual combination weight ( $w_v$ ). Color represents the variance of the combined estimate over time, dark red representing the smallest variance. The visual and haptic estimates were characterized by random walks with drifts of  $d_v$  and  $d_h$ . The recalibration rate was set with a Kalman filter. The estimator reliability ratio ( $r_v/r_h$ ) was 1:3, as in the simulation in Figure 5. The top, middle, and bottom rows show the results when the drift magnitude,  $d_v + d_h$ , was large ( $\sigma_v^2 + \sigma_h^2$ ), medium [ $(\sigma_v^2 + \sigma_h^2) \times 10^{-2}$ ], and small [ $(\sigma_v^2 + \sigma_h^2) \times 10^{-4}$ ], respectively. The two columns on the left show the results when the discrepancy introduced between the visual and haptic estimates was large ( $5\sigma_v^2$ ), and the two columns on the right show the results when the introduced discrepancy was small ( $1/5\sigma_v^2$ ). The first and third columns display the results when the drift ratio ( $d_v/d_h$ ) was 3:1 and the second and fourth columns the results when the ratio was 1:3. The dashed yellow circles mark the predictions if minimum variance were determined by the reliability-based model (i.e., by the ratio of estimator reliabilities  $r_v/r_h$ ); they are always at (0.75, 0.25) because the reliability ratio was always 1:3. The dashed blue circles represent the predictions if the minimum variance were determined only by the drift ratio  $d_v/d_h$ . The white circles indicate the calibration proportions and combination weights that yielded the combined estimate with the lowest variance.

To illustrate the importance of both internal consistency and external accuracy, it is useful to examine a situation in which reasonable models of each make different predictions. This occurs when a very reliable estimator becomes externally inaccurate. Adams et al. (2001) created such a situation by placing a horizontal magnifier in front of one eye for every waking hour for 1 week. Due to the magnifier, frontoparallel surfaces created binocular disparities at the retinas that are normally associated with slanted surfaces. The magnifier did not alter any other visual or nonvisual estimator. The predicted behaviors for internally consistent and externally accurate models are quite different. Reliability-based recalibration, which maintains internal consistency, predicts that the disparity cue should not adapt because disparity was more reliable than other depth cues. In contrast, a model of external accuracy predicts that disparity should adapt because it was the only cue made inaccurate by the magnifier. Adams et al. (2001) observed the second outcome: all of the change occurred in the disparity estimator. Thus, calibration did not follow the reliability-based scheme of Equations 4 and 5.

The Adams et al. (2001) result does not conflict with our analysis or results. During the week in which the subjects adapted, they received feedback from everyday interactions with the environment. The feedback was sufficient for the system to figure out that the disparity estimator had become miscalibrated and therefore to adapt that estimator to achieve external accuracy. What would have happened if subjects had been unable to interact with the environment (e.g., they sat in one place without the ability to make movements to visible objects) and disparity

and texture were the only available slant estimators? Without feedback about the accuracy of their percepts, the subjects could hope only to perceive surface slants as constant and unchanging, regardless of which slant estimators were used. Before internal consistency was achieved, disparity would yield a different estimate than texture, and both would be different from the combined estimate. If the texture estimator adapted to match disparity, the subject would have a stable, minimum-variance slant estimate, but it would be externally inaccurate.

When more than two estimators are available (e.g., disparity, texture, motion parallax), it would be rational to adjust the outlier (disparity). While adjusting the outlier may appear rational, it does not guarantee external accuracy. There still must be a method for incorporating external feedback to achieve external accuracy.

In summary, reliability-based adaptation yields the minimum-variance combined estimate (assuming random estimator drift is small or nonexistent from trial to trial), and internally consistent estimators ensure rigid percepts of the world. However, strict adherence to reliability-based recalibration can lead to internally consistent, but externally inaccurate, estimators. More research is needed on how sensory cue calibration, as distinct from sensorimotor calibration, should incorporate external feedback to guide the relative adaptation of different sensory estimators.

### Neural mechanisms of sensory cue calibration

The neural mechanisms underlying sensory cue calibration are not well understood. It has perhaps been worked out the best in the barn owl (Knudsen and Knudsen, 1989a,b; Knudsen and Brainard, 1991). In these animals, vision and audition are both used to direct flight toward prey. In analogy to the Helmholtz prism experiments, displacing prisms were attached to the owls' heads and their flight paths were measured. Paths were initially deviated in the visually specified direction and end-point accuracy decreased. After prolonged adaptation, flight paths and end-point accuracy became similar to baseline behavior, which means that the owls had recalibrated their sensorimotor behavior (Knudsen and Knudsen, 1989a,b). In the owl brain, two adjacent spatial maps seem to be involved in estimating location: the inferior colliculus, fed primarily by auditory signals, and the optic tectum, fed primarily by visual signals. The connections between these areas were examined after adaptation and compared to the connections in normal animals. The topographic map in the inferior colliculus reorganized after adaptation to match the map in the optic tectum (Knudsen and Brainard, 1991). If this reorganization underlies the maintenance of internal consistency, it appears that adaptation was implemented by changes in the auditory estimator alone. In humans, audition is generally a less reliable indicator of location than vision is (Alais and Burr, 2004). If vision is more reliable in barn owls too, the results of Knudsen and Knudsen (1989a,b) are consistent with both visual dominance and reliability-based adaptation. To distinguish the two models, one needs to decrease the relative reliability of the visual estimator, which could be done by degrading the visual input. In this case, the reliability-based model would predict shifts in the optic tectum rather than the inferior colliculus.

### References

- Adams WJ, Banks MS, van Ee R (2001) Adaptation to three-dimensional distortions in human vision. *Nat Neurosci* 4:1063–1064.
- Alais D, Burr D (2004) The ventriloquist effect results from near-optimal cross-modal integration. *Curr Biol* 14:257–262.
- Atkins JE, Jacobs RA, Knill DC (2003) Experience-dependent visual cue recalibration based on discrepancies between visual and haptic percepts. *Vision Res* 43:2603–2613.
- Baddeley RJ, Ingram HA, Miall RC (2003) System identification applied to a visuomotor task: near-optimal human performance in a noisy changing task. *J Neurosci* 23:3066–3075.
- Burge J, Ernst MO, Banks MS (2008) The statistical determinants of adaptation rate in human reaching. *J Vis* 8:20.1–19.
- Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415:429–433.
- Gepshtein S, Banks MS (2003) Viewing geometry determines how vision and haptics combine in size perception. *Curr Biol* 13:483–488.
- Ghahramani Z, Wolpert DM, Jordan MI (1997) Computational models of sensorimotor integration. In: *Self-organization, computational maps and motor control* (Morasso PG, Sanguineti V, eds). Amsterdam: Elsevier.
- Hay JC, Pick HL Jr (1966) Visual and proprioceptive adaptation to optical displacement of the visual stimulus. *J Exp Psychol* 71:150–158.
- Hay JC, Pick HL, Ikeda K (1965) Visual capture produced by prism spectacles. *Psychon Sci* 2:215–216.
- Hernandez TD, Levitan CA, Banks MS, Schor CM (2008) How does saccadic adaptation affect visual perception? *J Vis* 8:3.1–16.
- Jacobs RA (1999) Optimal integration of texture and motion cues to depth. *Vision Res* 39:3621–3629.
- Kagerer FA, Contreras-Vidal JL, Stelmach GE (1997) Adaptation to gradual as compared with sudden visuo-motor distortions. *Exp Brain Res* 115:557–561.
- Kalman RE (1960) A new approach to linear filtering and prediction problems. *J Basic Eng* 82:35–45.
- Knudsen EI (2002) Instructed learning in the auditory localization pathway of the barn owl. *Nature* 417:322–328.
- Knudsen EI, Brainard MS (1991) Visual instruction of the neural map of auditory space in the developing optic tectum. *Science* 253:85–87.
- Knudsen EI, Knudsen PF (1989a) Visuomotor adaptation to displacing prisms by adult and baby barn owls. *J Neurosci* 9:3297–3305.
- Knudsen EI, Knudsen PF (1989b) Vision calibrates sound localization in developing barn owls. *J Neurosci* 9:3306–3313.
- Lewald J (2002) Rapid adaptation to auditory-visual spatial disparity. *Learn Mem* 9:268–278.
- Maybeck P (1979) *Stochastic models, estimation, and control, Vol 1*. New York: Academic.
- Rock I, Victor J (1964) Vision and touch: an experimentally created conflict between the two senses. *Science* 143:594–596.
- Rosas P, Wagemans J, Ernst MO, Wichmann FA (2005) Texture and haptic induces in slant discrimination: reliability-based cue weighting without statistically optimal cue combination. *J Opt Soc Am A* 22:801–809.
- van Beers RJ, Wolpert DM, Haggard P (2002) When feeling is more important than seeing in sensorimotor adaptation. *Curr Biol* 12:834–837.
- von Helmholtz H (1867) *Handbuch der Physiologischen Optik*. Leipzig: Leopold Voss.
- Warren DH, Rossano MJ (1991) Intermodality relations: vision and touch. In: *The psychology of touch* (Heller MA, Schiff W, eds). London: Erlbaum.
- Wichmann FA, Hill NJ (2001) The psychometric function: I. Fitting, sampling and goodness-of-fit. *Percept Psychophys* 63:1293–1313.
- Witten IB, Knudsen EI (2005) Why seeing is believing: merging auditory and visual worlds. *Neuron* 48:489–496.
- Yuille AL, Bülthoff HH (1996) Bayesian decision theory and psychophysics. In: *Bayesian approaches to perception* (Knill D, Richards W, eds). Cambridge, UK: Cambridge UP.
- Zwiers MP, Van Opstal AJ, Paige GD (2003) Plasticity in human sound localization induced by compressed spatial vision. *Nat Neurosci* 6:175–181.