

# Humans Use Predictive Kinematic Models to Calibrate Visual Cues to Three-Dimensional Surface Slant

Peter Scarfe and Andrew Glennerster

School of Psychology and Clinical Language Sciences, University of Reading, Reading, Berkshire, United Kingdom RG6 6AL

When the sensory consequences of an action are systematically altered our brain can recalibrate the mappings between sensory cues and properties of our environment. This recalibration can be driven by both cue conflicts and altered sensory statistics, but neither mechanism offers a way for cues to be calibrated so they provide accurate information about the world, as sensory cues carry no information as to their own accuracy. Here, we explored whether sensory predictions based on internal physical models could be used to accurately calibrate visual cues to 3D surface slant. Human observers played a 3D kinematic game in which they adjusted the slant of a surface so that a moving ball would bounce off the surface and through a target hoop. In one group, the ball's bounce was manipulated so that the surface behaved as if it had a different slant to that signaled by visual cues. With experience of this altered bounce, observers recalibrated their perception of slant so that it was more consistent with the assumed laws of kinematics and physical behavior of the surface. In another group, making the ball spin in a way that could physically explain its altered bounce eliminated this pattern of recalibration. Importantly, both groups adjusted their behavior in the kinematic game in the same way, experienced the same set of slants, and were not presented with low-level cue conflicts that could drive the recalibration. We conclude that observers use predictive kinematic models to accurately calibrate visual cues to 3D properties of world.

**Key words:** 3D slant; calibration; psychophysics; stereopsis

## Introduction

To effectively control our behavior, our sensorimotor systems need to maintain external accuracy with respect to the world. Previous research has shown that perceptual attributes can be recalibrated after extended experience of altered sensory feedback, such as when wearing laterally displacing or magnifying prisms (von Helmholtz, 1925; McLaughlin and Webster, 1967; Adams et al., 2001). There are, however, many possible causes of the recalibration. In addition to changing the sensory consequences of a person's actions, prisms introduce low-level cue conflicts and alter the overall statistics of incoming sensory data. Both of these are known to cause recalibration, but in both instances this can either increase or decrease a cue's external accuracy (Ernst et al., 2000; Atkins et al., 2001; Adams et al., 2004; Burge et al., 2010; Seydell et al., 2010; van Beers et al., 2011). This is possible because neither the perceptual estimates provided by a cue, nor the reliability associated with those estimates,

can signal whether a cue is well calibrated (Smeets et al., 2006; Ernst and Di Luca, 2011; Scarfe and Hibbard, 2011).

Despite these challenges to accurate calibration, human observers exhibit expert knowledge of the physics governing the environment, such as mass, gravity, and object kinematics (Battaglia et al., 2013; Smith and Vul, 2013; Smith et al., 2013). Knowledge describing these dynamics could provide the error signals required to maintain accurate calibration. Our experiment was designed to test this idea. To do this we manipulated the kinematics of a 3D game in which observers altered the slant of a surface so that a ball would bounce off the surface and through a target hoop. By altering the bounce angle of the ball we were able to make the surface behave as if it had a different slant to that signaled by visual cues. This allowed us to determine whether observers' expectations of the ball's bounce, based on an understanding of object kinematics, could drive recalibration of perceived 3D surface slant.

We also examined slant recalibration when the ball's altered bounce was coupled with the ball spinning in such a way that it could "explain away" its altered trajectory (Battaglia et al., 2010; Shams and Beierholm, 2010; Clark, 2013). The role of internal physical models and sensory prediction has been most closely studied in the domain of visuomotor control where it has been shown that error signals based on sensory prediction can be used to update information about the current body state (Wolpert et al., 1998; Cressman and Henriques, 2011; Henriques and Cressman, 2012). More recently, research has suggested that internal models of physical laws may be a much more general part of our perceptual experience than previously thought (Hamrick et al., 2011; Battaglia et al., 2013; Smith and Vul, 2013; Smith et al.,

Received March 11, 2014; revised June 17, 2014; accepted June 23, 2014.

Author contributions: P.S. and A.G. designed research; P.S. performed research; P.S. analyzed data; P.S. and A.G. wrote the paper.

This work was supported by the Wellcome Trust (086526/A/08) and EPSRC (EP/K011766/1). We thank Loes van Dam, Paul Hibbard, Marc Ernst, and Mario Kleiner for helpful discussion.

The authors declare no competing financial interests.

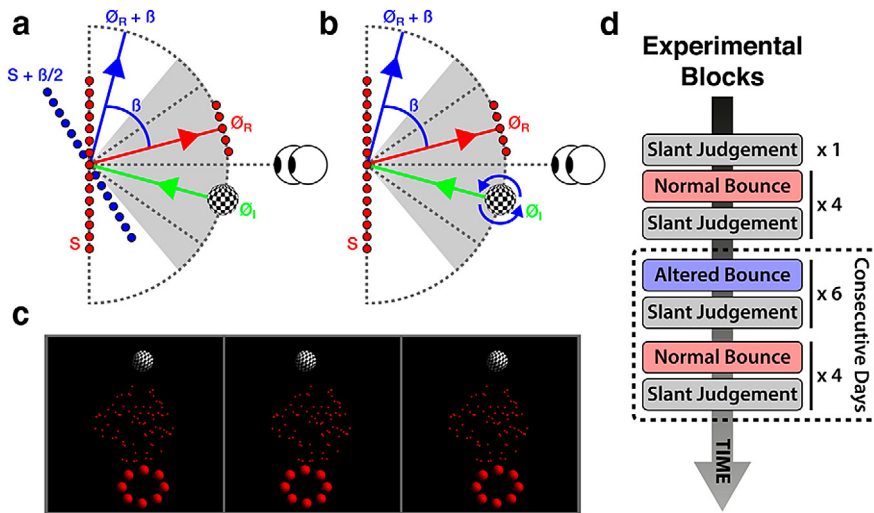
This article is freely available online through the *JNeurosci* Author Open Choice option.

Correspondence should be addressed to Dr. Peter Scarfe, School of Psychology and Clinical Language Sciences, University of Reading, Early Gate, Whiteknights Road, Reading, Berkshire, UK RG6 6AL. E-mail: p.scarfe@reading.ac.uk.

DOI:10.1523/JNEUROSCI.1000-14.2014

Copyright © 2014 Scarfe and Glennerster

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.



**Figure 1.** Schematic illustration of the kinematic ball task for (a) the nonspinning ball and (b) the spinning ball. During the experiment on each trial the position of the ball took a random position within a  $\pm 50$  degree range on an imaginary circle 12 cm in diameter, centered on the center of the slanted surface. This range is represented by the shaded region in a and b. Similarly, on each trial the hoop took a different random position within a  $\pm 35$  degrees range on the same imaginary circle. This range is represented by the dashed diagonal lines in a and b. c, Stereogram of the ball game stimuli (left and middle images for divergent fusion, middle and right image for crossed fusion). d, Sequence of experimental blocks during the experiment.

2013). The experiments reported here provide an unambiguous test of the extent to which sensory prediction, based on internal physical models, can be used to calibrate sensory cues so that they provide accurate information about the world.

**Materials and Methods**

**Apparatus.** Stimuli were displayed on a spatially calibrated, gamma corrected, CRT (1600 × 1024 pixels, 85 Hz refresh rate) viewed at a distance of 50 cm in a completely dark room. At this distance the monitor spanned ~51 × 34 degrees of visual angle. Eye height was matched to the vertical center of the CRT and head position maintained directly in front of the center of the monitor with a chin rest. Stereoscopic presentation was achieved using Stereographics Crystal Eyes CE-3 shutter goggles. All stimuli were tailored to the interocular distance of each observer and were rendered in OpenGL using MATLAB and the Psychtoolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). Stimuli were rendered in red to exploit the fastest phosphor of the monitor. No crosstalk between the two eyes view was visible.

**Participants.** There were 12 participants in total (9 male, 3 female). Six participants were assigned to each condition (four males and two females in condition one, five males and one female in condition two). All participants were naive to the purpose of the experiment, except for P.S. Sample size was determined on the basis of the consistency of the effects across participants in a pilot study. The University of Reading Research Ethics Committee approved the study.

**Kinematic ball game.** Observers performed two different tasks: a 3D kinematic ball game and a frontoparallel slant judgment task. In the kinematic game they had to adjust the slant of a disparity-defined surface on-line with a mouse so that a moving checkerboard-textured ball (2 cm diameter) bounced off the surface and through a target hoop (Fig. 1). Rendered distance to the surface matched that of the screen. The surface was elliptical with a fixed width of 10 cm, but with a variable height selected from a uniform distribution of 7.5–12.5 cm (when frontoparallel, this equated to 8.58 and 14.25 degrees, respectively). Varying the height degraded the use of vertical angular subtense as a cue to slant. The dot density of the surface was 0.6 dots per cm<sup>2</sup>. Each dot had a diameter of 3.4 pixels (~0.11 degrees) and was positioned with subpixel precision. The slant of the surface could be adjusted by  $\pm 50$  degrees around its horizontal axis using lateral mouse movements.

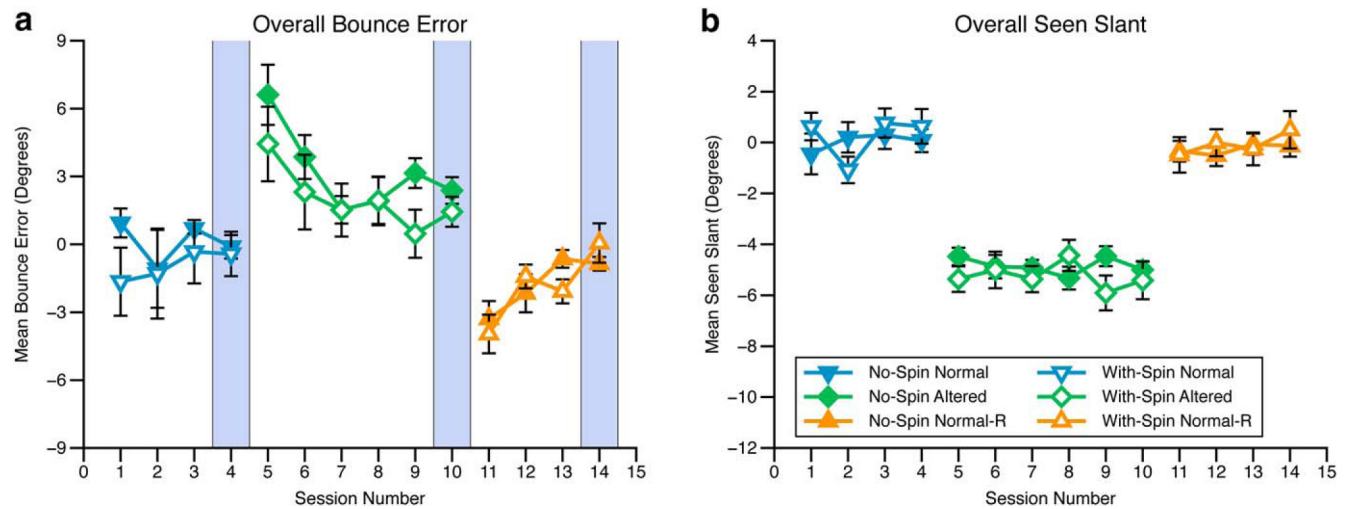
There were 80 trials in each block. On each trial, the starting position of the ball took a random value within  $\pm 50$  degrees of straight ahead, on an

imaginary circle with a radius of 12 cm centered on the center of the slanted surface (Fig. 1a,b, shaded region). The position of the hoop took a random value between  $\pm 35$  degrees of straight ahead on the same imaginary circle (Fig. 1a,b, diagonal dashed lines). The ball and hoop took random positions within different angular ranges to ensure that when the ball’s bounce bias was introduced (see below) observers could always get the ball directly through the hoop, i.e., they were never given an impossible task. The orientation of the ball around its center (in all three dimensions) was randomized on each trial. The target hoop consisted of eight balls (1 cm in diameter) equally spaced around a circle 4 cm in diameter and was always rendered tangential to the imaginary radius on which it was placed. The hoop and checkerboard-textured ball were illuminated with a single point light source off to the upper right. The random-dot surface was unaffected by this light source.

At the beginning of each trial, the checkerboard ball would pause for 1 s in its starting position before launching directly toward the center of the surface at a constant speed of 6 cm s<sup>-1</sup>. Initially, for the first four blocks, the ball bounced off the surface vertically with a mirror reflection, such that  $\phi_r = \phi_i$  (Fig. 1). This is, to a first approximation, how nonspinning balls and planar surfaces behave in real life. For the next six blocks of trials, a bias was added to the ball’s bounce such that  $\phi_r = \phi_i + \beta$  (exaggerated for the purpose of illustration in Fig. 1). The value of  $\beta$  was 15 degrees (sign consistent for a single observer, but counterbalanced across observers). During this period, for any slant  $S_\beta$ , the surface now behaved as if it had a slant of  $S_\beta + \beta/2$ . For the final four blocks of trials, the bounce bias was removed. Observers reported that, throughout the experiment, they remained entirely unaware of the introduction and removal of  $\beta$  despite having to adapt their behavior in the game to continue getting the ball through the hoop (for further discussion see Results, Equivalence of the learning signal across groups).

After bouncing off the surface, the ball continued moving at the same speed as before the bounce until its center intersected the imaginary circle upon which it started. The game then paused for 1 s to give observers feedback about the magnitude of their error. If the center of the ball passed within the hoop’s radius the hoop turned a lighter shade of red, if it passed within 1.25–1.5 times its radius it stayed the same shade of red, and if it passed outside this range it turned a darker shade of red. An error of zero was achieved if the ball passed directly through the hoop. It was predicted that, if high-level knowledge can drive sensory recalibration in addition to adapting behavior, observers might also recalibrate their perception of slant with introduction and removal of  $\beta$ . This would happen if observers assumed the laws of kinematics governing the task were invariant over time and instead used errors in the bounce task to infer that their encoding of slant from disparity had slipped out of calibration.

**Spinning ball.** In a separate condition, for a separate set of observers, everything was identical to that described above except that when the scene appeared the checkerboard ball began spinning around its horizontal axis at a speed of 300 degrees/s. The ball’s spin was in the same direction as the bias added to its bounce. The aim here was that observers would have to adapt their behavior in the game in the same way, because  $\beta$  was the same, and would therefore experience the same set of surface slant; however, the ball’s spin might offer an alternative physical explanation for its altered bounce and thus remove any need for recalibration. Discounting of this sort need not be a binary decision, but could instead reflect the degree of certainty the observer has that the alternative explanation is valid. This type of process has been termed “explaining away” in the context of Bayesian decision theory and can be modeled within a probabilistic framework (Battaglia et al., 2011).



**Figure 2.** *a*, Mean bounce error across sessions during the ball bounce task. *b*, Mean seen slant during each session of the ball bounce task (note: Normal-R = Normal bounce regained). Error bars show SEM. The blue vertical stripes highlight the final block of each phase of the experiment. These are included on the data plots to ease comparison across this figure and Figure 3.

**Frontoparallel slant judgment task.** To track any recalibration of slant that occurred, all observers completed a frontoparallel slant judgment task interleaved between game blocks (Fig. 1*d*). In this task, observers judged whether a random dot stereogram of a slanted surface (identical to that used in the kinematic ball game) was sloped toward or away from them (top near–bottom far, and top far–bottom near, respectively). During this task, the hoop and ball were not present, but the surface was otherwise identical to that used in the kinematic ball game, including the variation in its vertical height on each trial. At the start of each trial, a fixation point 13 pixels ( $\sim 0.42$  degrees) in diameter was presented for 1 s centrally in the plane of the screen. The 3D slanted surface was then presented, also for 1 s. This was replaced by the fixation point, which remained on the screen until observers indicated their response by pressing one of two keyboard buttons. Keyboard responses were used to avoid any carryover effects from using the mouse for the ball-bounce task.

The slant of the surface was varied using the method of constant stimuli (seven slant values, each presented in a randomized order 33 times). The number of slant values and number of repetitions had provided well fit psychometric functions in a pilot study. The slant range and slant center was manually adjusted for each observer as needed throughout the experiment to ensure that their point of subjective equality (PSE) remained within the range and that the function encompassed high and low performance levels to obtain a good psychometric function fit (Wichmann and Hill, 2001*a, b*; Prins and Kingdom, 2009; Kingdom and Prins, 2010). This was essential as different observers adapted their perception of slant by different magnitudes and these differences could not be predicted a priori (Fig. 4, described in more detail later in the text). We note also, that we chose a task that observers could understand readily and unambiguously; indeed, subjectively, all found the task exceptionally easy to complete.

**Procedure.** There were 14 sessions (Fig. 1*d*) each lasting around half an hour. Each session consisted of two separate tasks: first, the 3D kinematic game and second, the frontoparallel, slant judgment task. The exception to this was the first session, which had an additional block of slant judgment trials before the first block of ball task trials. To maximize the chances of measuring any recalibration, the last 10 sessions had to be completed over consecutive days. Furthermore, given that sleep can enhance the effects of perceptual learning with stereoscopic stimuli (van Ee, 2001), observers were encouraged to spread these last 10 sessions over a full working week of 5 d. Therefore, typically, each observer completed the last 10 blocks, 2 blocks per day for 5 consecutive days.

## Results

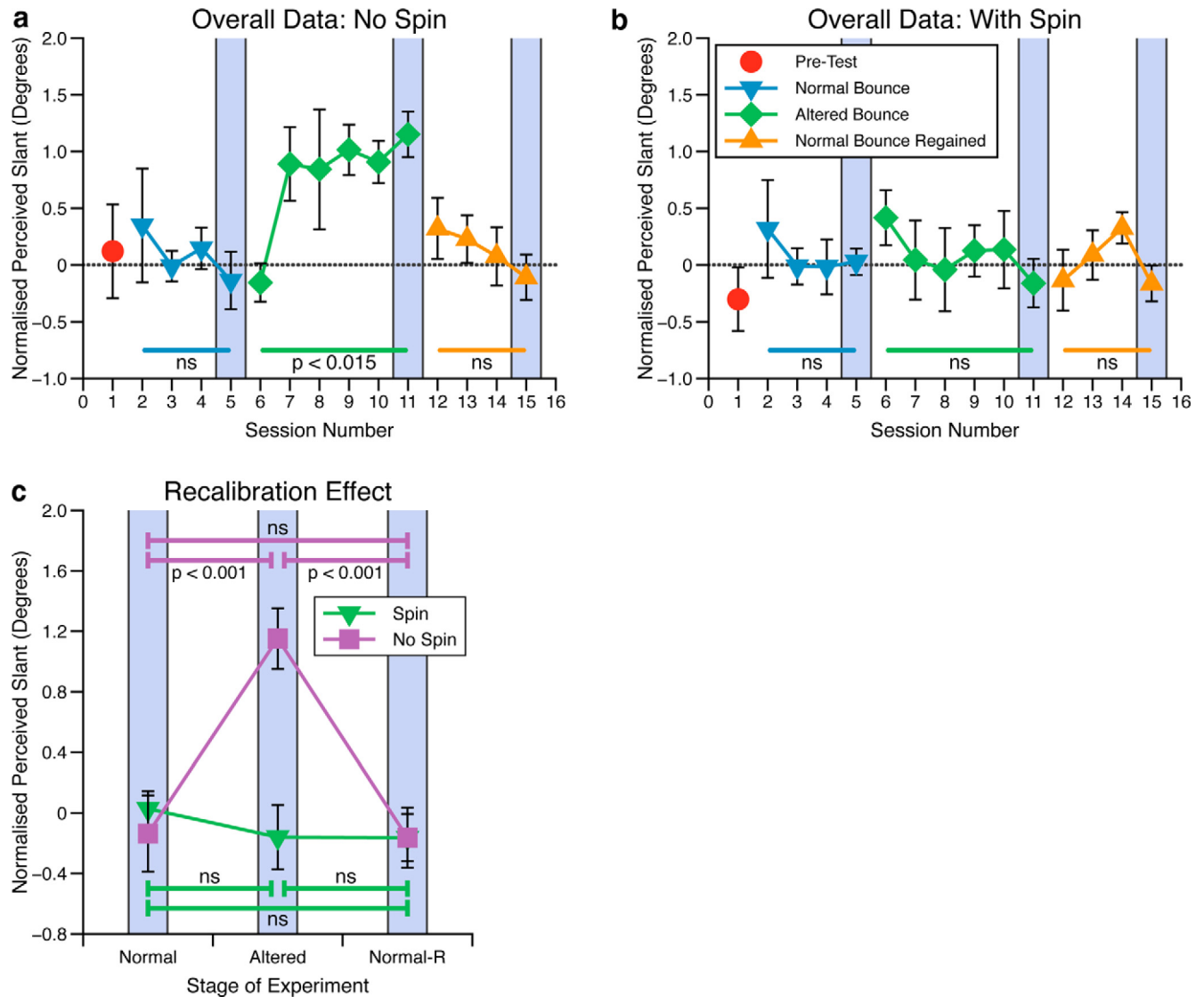
### Kinematic ball game

Observers' average error across blocks of the ball game was calculated, as well as the average slant presented throughout the task

(Fig. 2). Because half of the observers experienced a positive bounce bias and half a negative bounce bias, in all graphs, for those observers where  $\beta$  was negative we have multiplied their data by  $-1$ . As can be seen, observers in the two conditions (non-spinning and spinning ball) learned to change their behavior in response to the altered bounce in the same way and at the same rate over the full course of the experiment, including with the introduction and removal of  $\beta$  (Fig. 2*a*). This was confirmed by a within-subjects ANOVA, with experimental block as a between-subjects factor, which showed that there was a significant main effect of experimental block ( $F_{(13,130)} = 16.50, p < 0.001$ ), but no significant effect of group ( $F_{(1,10)} = 0.72, p = 0.42$ ), and no interaction ( $F_{(13,130)} = 1.0, p = 0.46$ ). As a result of this, observers in each group experienced the same slants during the ball game (Fig. 2*b*; this is also true if medians are used instead of means). We can therefore be confident that any differences in recalibration observed across the two groups cannot be caused by differences in the slants seen during the ball game or differences in the way observers in each group learned to respond to the altered bounce. For example, there is no evidence in our data that the ball spinning disrupted how observers adapted to its altered bounce. Indeed, the magnitude of  $\beta$  was specifically selected on the basis of piloting so that this would not occur.

### Recalibration of perceived slant

Cumulative Gaussian functions were fitted to the slant judgment data by a maximum likelihood procedure in MATLAB using the Palamedes software package (Prins and Kingdom, 2009; Kingdom and Prins, 2010; Prins, 2012). The PSE and 95% confidence intervals were estimated via bootstrapping. The PSE represents the slant that observers would perceive as frontoparallel. Figure 3 shows the mean PSEs for observers in each group across the 15 sessions. Here, to remove any small constant biases in perceived frontoparallel, each observer's data have been normalized to the mean slant seen as frontoparallel across blocks 3–5 (mean shift of 0.46 degrees across observers). Additionally, as for the ball game data, where  $\beta$  was negative the observer's data have been multiplied by  $-1$  to make the bounce biases' polarity comparable across observers.



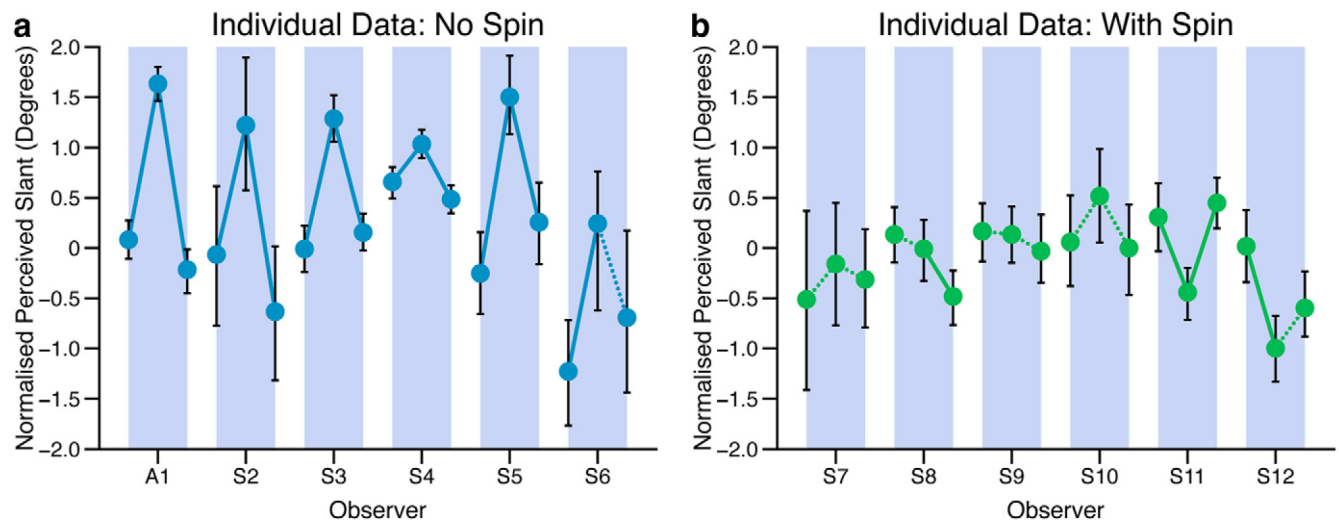
**Figure 3.** Slant perceived as frontoparallel, normalized across observers, as measured with the frontoparallel slant judgment task interleaved between ball bounce blocks. This is shown for (a) the nonspinning ball group and (b) the spinning ball group. c, Overall recalibration effect averaged across observers in each group. This represents the slant perceived as frontoparallel at the end of each phase of the experiment, i.e., sessions 5, 11, and 15, as highlighted in a and b. Error bars all show SEM.

**Effects over time**

As can be seen in Figure 3a, when the bounce bias was introduced (blocks 6–11) observers in the nonspinning ball group recalibrated their perception of slant so that a perceptually frontoparallel surface became more like the surface that behaved like a frontoparallel surface in the ball game. Upon removal of the bounce bias observers were able to more rapidly switch back to the previous sensory mapping for perceived slant, as would be predicted by previous research (Welch et al., 1993; van Dam et al., 2013). Consistent with this, within-subjects ANOVA showed that for the nonspinning ball condition there was no significant recalibration across sessions in the normal bounce phase ( $F_{(3,15)} = 0.43, p = 0.73, ns$ ), there was significant recalibration during the altered bounce phase ( $F_{(5,25)} = 3.54, p < 0.015$ ), and no significant recalibration over the phase where the normal bounce was regained ( $F_{(3,15)} = 0.84, p = 0.49, ns$ ). In contrast, with the spinning ball there was no significant recalibration in any phase of the experiment (normal bounce phase:  $F_{(3,15)} = 0.33, p = 0.81, ns$ ; altered bounce phase:  $F_{(5,25)} = 2.03, p = 0.11, ns$ ; and when the normal bounce was regained:  $F_{(3,15)} = 3.11, p = 0.06, ns$ ).

**Perceived slant after full experience**

We would expect the maximum learning effect to be evident at the end of each phase of the experiment where the bounce was altered, i.e., after blocks 11 and 15. Figure 3c shows the slant perceived as frontoparallel in these cases compared with that in block five, which was the last block of the normal bounce phase, before introduction of the altered bounce. As can be seen, with a nonspinning ball, experience of a biased bounce resulted in significant recalibration of perceived slant; however, this recalibration was eliminated when the introduction of the bounce bias was coupled with the ball spinning. Between-subjects ANOVA showed there to be a significant main effect of condition (spinning vs nonspinning ball;  $F_{(1,30)} = 6.39, p < 0.017$ ) and experimental phase (normal bounce, altered bounce, and normal bounce regained,  $F_{(2,30)} = 6.26, p < 0.005$ ), as well as a significant condition by phase interaction ( $F_{(2,30)} = 8.41, p < 0.001$ ). This interaction arose because perceived frontoparallel was the same after each phase of the experiment with the spinning ball, but not with the nonspinning ball (Fig. 3c).



**Figure 4.** Recalibration effect for each observer in the (a) nonspinning ball and (b) spinning ball groups. This is in the same format as Figure 3c, and shows the slant perceived as frontoparallel at the end of each phase of the experiment, i.e., sessions 5, 11, and 15, as highlighted in Figure 3a and b. Error bars show bootstrapped 95% confidence intervals from the psychometric function fit. Solid lines show differences between PSEs at the  $p < 0.05$  level, dashed lines show no significant difference. Participant labeled A1 is P.S. Blue bands delineate the data from each observer.

### Consistency of effects over observers

We were interested in how consistent the recalibration effect was across observers. In Figure 4 we plot perceived slant at the end of each phase of the experiment, for each observer, in each condition. Here error bars show bootstrapped 95% confidence intervals around the perceived slant PSEs. As can be seen visually, the predicted recalibration effect was highly consistent across all observers in the no-spin condition. In contrast, the recalibration effect was absent in all observers in the spinning ball condition. To test this statistically, on a per observer basis, we used a bootstrapped likelihood ratio (LR) test to compare the PSEs generated by each observer at the end of each stage of the experiment (Kingdom and Prins, 2010; i.e., normal bounce phase compared with altered bounce phase, and altered bounce phase compared with normal bounce regained phase).

For each observer, we first determined the likelihood of the data from both conditions assuming a single psychometric function, with a single PSE (1PSE model), versus two separate psychometric functions, with different PSEs and, potentially, different slopes (2PSE model). The 1PSE model assumes that there is no effect of adaptation and that any difference between the data in the two conditions is simply due to sampling noise. In contrast, the 2PSE model assumes that the difference between PSEs is due to a significant recalibration effect and not just sampling noise. The ratio of these two likelihoods (1PSE/2PSE) gives us the LR, which is a measure of the relative goodness of fit of the 1PSE and 2PSE models. If there is indeed no difference between conditions (i.e., no recalibration effect), a generative 1PSE model needs to be able to generate an LR as small as that derived from the experimental data (Kingdom and Prins, 2010).

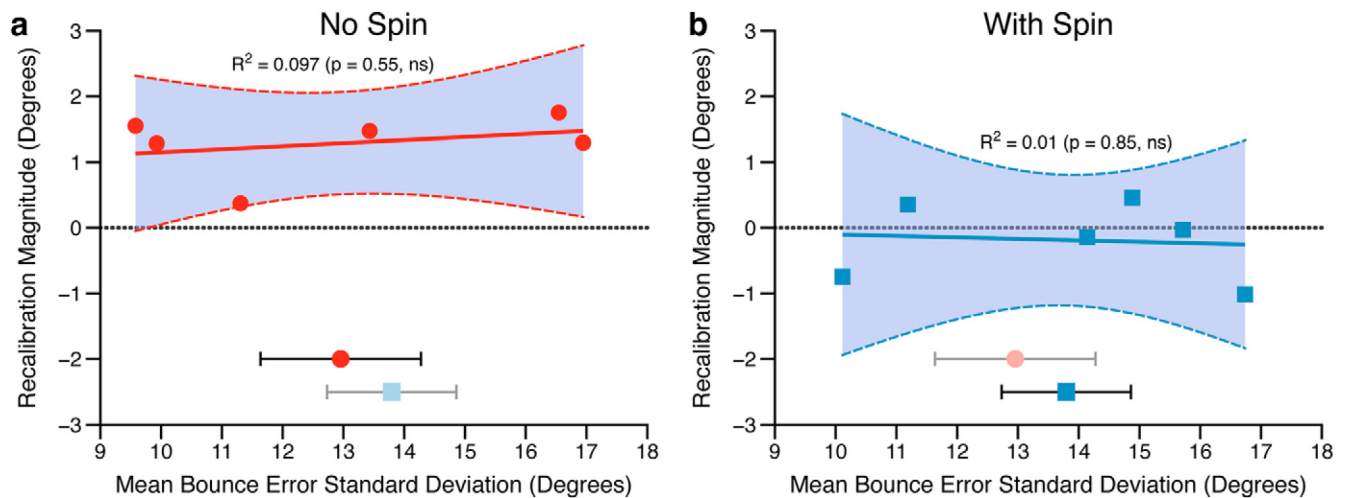
To assess this we used the Palamedes toolbox (Prins and Kingdom, 2009) to simulate each observer in an experiment with the same stimulus intensities as in the real experimental conditions, 1000 times, each time generating responses in accordance with the 1PSE model and recalculated the LR. From these simulations we determined the probability with which an observer whose PSEs do not differ (1PSE model) could generate an LR as small as that generated by a real observer in the experiment. This determines the solid and dashed lines in Figure 4. Solid lines indicate that 5% or less of the 1000 simulations of the 1PSE model pro-

duced an LR less than or equal to that of the experimental data (i.e., a difference at the  $p < 0.05$  level). Conversely, dashed lines indicate no significant difference ( $p > 0.05$ ). From this analysis it is clear that the group effects we report are also highly robust and consistent across individual observers in the experiment.

### Equivalence of the learning signal across groups

The experiment was designed, as far as possible, to equate the learning signal across the two experimental conditions. One possible difference that it is important to exclude is that the spinning ball simply caused observers to be more uncertain as to its trajectory and that this provided a weaker learning signal for recalibration. It has been shown previously that participants in both groups adapted in the same way and at the same rate to the altered bounce (Fig. 2). This suggests that the learning signal was equally effective in both conditions. However, the reliability of the learning signal can also be examined directly by looking at the variability of participants' behavior in each group. Specifically, if it really were participant uncertainty about the ball's trajectory that determined the extent of slant recalibration, participants in the spinning ball group should show greater variability in the ball task during the altered bounce phase, and, regardless of group, those participants who were more variable should show less recalibration of slant.

As a measure of variability, the mean SD of bounce errors was calculated across participants in each group during the altered bounce phase of the experiment, i.e., across the whole period in which they experienced the learning signal that triggered the initial recalibration. As can be seen in Figure 5, there was on average no difference in variability across groups (between-subjects  $t$  test,  $t = 0.5$ ,  $p = 0.63$ , ns), and no relationship between how variable an observer was during the ball task and the magnitude of their recalibration of slant (magnitude of recalibration being calculated as the difference between the PSE in block 15 and PSE in block 11, just as in Fig. 4). This was true for both the spinning ball ( $R^2 = 0.01$ ,  $p = 0.85$ , ns) and nonspinning ball groups ( $R^2 = 0.097$ ,  $p = 0.55$ , ns). We can therefore conclude that the magnitude of slant recalibration was not determined by differences in the variability of the learning signal across groups. Note also that



**Figure 5.** The extent of individual observers' recalibration of slant as a function of their variability in the 3D kinematic game. The magnitude of recalibration is calculated as perceived slant after block 11 minus perceived slant after block 5 (just as in Fig. 4). This is plotted against the average SD of bounce errors in the ball game during the altered bounce phase of the experiment. This is shown for (a) the nonspinning ball group and (b) the spinning ball group. The solid line in both plots shows a least-squares linear fit to the data, with 95% confidence intervals of the fit shown with dashed lines (statistics of linear regression inset). The dotted horizontal lines show zero recalibration. In the lower part of both a and b the group mean bounce error SD during the altered bounce phase of the experiment is shown for both groups, with less opacity for the condition shown on each graph. Symbols for averages match those of the individual data. Error bars show SEM.

our conclusions remain unchanged if an alternative measure of variability, such as the SD in the last altered bounce block, is used.

Despite this analysis, some might argue that the mere fact that the spinning ball trials are noticeably “different” in some unspecified way might disrupt the learning signal required for slant recalibration, while leaving no measurable effect on observers' behavior. While this is possible, without specifying a putative mechanism for this disruption this would be hard, if not impossible, to objectively test. Furthermore, exactly the same criticism applies to similar experiments that explore explaining away (Knill and Kersten, 1991; Battaglia et al., 2010). This is due to the fact that for explaining away to occur there always has to be a difference between conditions, i.e., a signal that does the “explaining” in one condition that is absent in another condition. One can therefore always counter that an unspecified and unknown property of that signal, other than that identified, could account for the explaining away. We have therefore taken the pragmatic approach of equating the reliability of the learning signal across groups and demonstrating that there is no statistical difference between experimental conditions and no relationship between uncertainty in the learning signal and the magnitude of slant recalibration.

It is also worth noting that in pilots for this experiment, if the bounce bias was too large participants immediately noticed it and attempted to consciously “correct” for it. This resulted in highly erratic behavior and no recalibration of slant. It was therefore essential that observers did not consciously detect the altered bounce. With the magnitude of  $\beta$  used, this was the case for all of the observers. What observers did notice was that in some blocks they missed the hoop more often, but they universally attributed this to having a “bad day” or “bad block of the game.” None had any idea that this drop in performance was caused by a biased bounce. Of course, failing to notice an altered bounce consistent with a 7.5 degree difference in slant is very different from distinguishing two slants separated by 7.5 degrees, as in a slant threshold task.

## Discussion

Our results show that human observers recalibrate their perception of 3D surface slant when they experience altered visual feed-

back consistent with a surface having a different slant to that signaled by disparity cues. Observers therefore appear to have assumed that the physical laws governing the kinematic task were invariant over time and used prediction errors caused by the altered bounce to recalibrate their perception of slant to bring it back into alignment with the physical behavior of the surface. Furthermore, when an alternative for the altered feedback presented itself (the ball spinning), this pattern of recalibration was eliminated, suggesting that observers used the ball's spin to explain away (Battaglia et al., 2010) its altered bounce. Importantly, the recalibration we observed (or lack thereof, in the case of the spinning ball) could not be accounted for in terms of response biases, low-level cue conflicts, uncertainty in the learning signal, or alterations to the statistics of incoming sensory data. It was especially important to exclude these in our experiment as they have been shown in the past to cause recalibration (Ernst et al., 2000; Atkins et al., 2001; Burge et al., 2010), and this control has not been possible in previous research using traditional techniques such as prism adaptation (Ogle, 1950; Pick et al., 1969; Adams et al., 2001).

The use of error signals based on sensory prediction has long been studied in the domain of visuomotor control (Blakemore et al., 1998; Wolpert et al., 1998; Blakemore et al., 2001). Within this domain there are similarities between explaining away (Battaglia et al., 2010), as described here, and a mechanism that has been termed “credit assignment” (Berniker and Kording, 2008). Credit assignment acts to assign motor errors across different effectors or actions for the purposes of recalibration (Franklin and Wolpert, 2011; Wolpert and Landy, 2012). However, the importance of sensory prediction in calibrating cues, including those to 3D object properties, has not been fully appreciated until now. Our results suggest that the ability to compare incoming sensory signals to expected values, in this case using predictive kinematic models (Smith and Vul, 2013; Smith et al., 2013), is a critical component in keeping visual cues such as slant from disparity well calibrated. This is understandable given that cues themselves carry no information about their accuracy (Berkeley,

1709; Smeets et al., 2006; Burge et al., 2010; Ernst and Di Luca, 2011; Scarfe and Hibbard, 2011; Zaidel et al., 2011).

We see two additional areas where sensory prediction may be important: first, in determining how novel cues to object properties are learned and, second, in determining when existing cues to object properties should be combined. Previous research has shown that novel cues to object properties can be learned through a process of paired association with existing “trusted” cues (Haijiang et al., 2006; Ernst, 2007). However, this type of learning cannot guarantee that the learned cue is providing accurate information about the world. One way in which this could be achieved is by learning cues that provide behavioral predictability, even in the absence of correlations with existing trusted cues. Once learned, past behavior could also indicate whether cues are likely to have a common cause and hence should be combined. Previous research has shown that cues that are spatially (Gepshtein et al., 2005) and temporally (Parise et al., 2012) correlated are more likely to be combined. However, spurious correlations could also arise with discrepant sensory stimuli that should remain segregated. Using behavioral predictability as a cue to a common cause would avoid this problem.

It is worth noting that the recalibration we observed was not sufficient to completely account for the behavior of the ball (it was ~16% of “full” recalibration). This is consistent with previous studies demonstrating explaining away (Battaglia et al., 2010, 2011). The extent of explaining away is likely to be affected by numerous variables that influence the observer’s beliefs about the causal structure of the stimuli (Körding et al., 2007; Battaglia et al., 2011). As discussed in the Materials and Methods section, explaining away need not necessarily be an all-or-none phenomenon. Our data indicate a significant difference between the spinning and nonspinning ball cases (Figs. 3c, 4), demonstrated by testing the null hypothesis that there is no recalibration. However, given an appropriately rich dataset, a full Bayesian model could test a much wider range of outcomes that include a variable degree of explaining away.

There is debate about the extent to which the realism of stimuli is helpful in distinguishing between hypotheses in experiments (De Gelder and Bertelson, 2003; Felsen and Dan, 2005; Rust and Movshon, 2005; Battaglia et al., 2013; Scarfe and Hibbard, 2013). However, in our experiment the fact that observers used the ball’s spin to explain away the need for recalibration, even with a simplified simulation of kinematics, suggests that the information provided to observers was sufficient for them to infer the causal structure of the stimuli and use this to determine the extent of sensory recalibration.

It is also important to note that observers had far greater experience of real-world physics between experimental sessions than they did of the altered physics within experimental sessions, so the magnitude of recalibration was unlikely to fully account for the ball’s altered bounce. This is consistent with previous research on perceptual learning and sensory adaptation of 3D object properties (Adams et al., 2004; Ernst, 2007). As such, it suggests that the learning we observed must have been context specific; otherwise observers’ experience of real-world physics between sessions would have overridden any learning driven by altered physics within sessions. This type of specificity has been demonstrated in previous studies on perceptual learning where an adaptation state can be yoked to the context in which it was learned (Welch et al., 1993; Martin et al., 1996; Osu et al., 2004; Kerrigan and Adams, 2013; van Dam et al., 2013).

We also note that it is an open question as to whether the recalibration observed altered the mapping between disparity

and perceived slant, or the mapping between perceived slant and the observer’s response. Previous research suggests that perceptual learning and adaptation may be driven by both cue-specific and cue-invariant mechanisms (Adams et al., 2001; Ivanchenko and Jacobs, 2007). However, the fact that observers remained consciously unaware of the altered bounce and that the recalibration was measured in a different task, with a different mode of response, provides evidence in favor of the hypothesis that recalibration altered the mapping between disparity and perceived slant. It is also clear that, in addition to perceptual recalibration, some response recalibration also occurred, as observers in the spinning ball group were able to improve their performance in the altered bounce phase of the experiment in exactly the same way as the no-spin group, despite the fact that they exhibited no recalibration of perceived slant.

Overall, our results, together with those on visuomotor adaptation, highlight the fact that cues are simply arbitrary patterns of sensory data that allow us to make accurate predictions about their hidden world causes (Berkeley, 1709; Gibson, 1950). This predictability is embodied in our understanding of physical laws (Smith and Vul, 2013; Smith et al., 2013) and our own bodily mechanics (Wolpert et al., 1998). Throughout our discussion we have drawn the distinction between sensory cues and error signals produced by a discrepancy between predicted and observed sensory feedback. While this distinction has been profitable (Atkins et al., 2001, 2003; Cressman and Henriques, 2011; Henriques and Cressman, 2012), ultimately, a close iterative relationship between the two must exist, as the only information available from which to derive error signals is that provided by the senses (von Helmholtz, 1925; Gregory, 1980; Knill and Richards, 1996; Rao and Ballard, 1999; Clark, 2013). The spatial and temporal invariance of the physics governing the world is one way to constrain possible solutions to this open-ended self-calibration problem.

## References

- Adams WJ, Banks MS, van Ee R (2001) Adaptation to three-dimensional distortions in human vision. *Nat Neurosci* 4:1063–1064. [CrossRef Medline](#)
- Adams WJ, Graf EW, Ernst MO (2004) Experience can change the ‘light-from-above’ prior. *Nat Neurosci* 7:1057–1058. [CrossRef Medline](#)
- Atkins JE, Fiser J, Jacobs RA (2001) Experience-dependent visual cue integration based on consistencies between visual and haptic percepts. *Vision Res* 41:449–461. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Battaglia PW, Di Luca M, Ernst MO, Schrater PR, Machulla T, Kersten D (2010) Within- and cross-modal distance information disambiguate visual size-change perception. *PLoS Comput Biol* 6:e1000697. [CrossRef Medline](#)
- Battaglia PW, Kersten D, Schrater PR (2011) How haptic size sensations improve distance perception. *PLoS Comput Biol* 7:e1002080. [CrossRef Medline](#)
- Battaglia PW, Hamrick JB, Tenenbaum JB (2013) Simulation as an engine of physical scene understanding. *Proc Natl Acad Sci U S A* 110:18327–18332. [CrossRef Medline](#)
- Berkeley G (1709) *An essay towards a new theory of vision*. Cirencester, UK: The Echo Library.
- Berniker M, Körding K (2008) Estimating the sources of motor errors for adaptation and generalization. *Nat Neurosci* 11:1454–1461. [CrossRef Medline](#)
- Blakemore SJ, Wolpert DM, Frith CD (1998) Central cancellation of self-produced tickle sensation. *Nat Neurosci* 1:635–640. [CrossRef Medline](#)
- Blakemore SJ, Frith CD, Wolpert DM (2001) The cerebellum is involved in predicting the sensory consequences of action. *Neuroreport* 12:1879–1884. [CrossRef Medline](#)

- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436. [CrossRef Medline](#)
- Burge J, Girshick AR, Banks MS (2010) Visual-haptic adaptation is determined by relative reliability. *J Neurosci* 30:7714–7721. [CrossRef Medline](#)
- Clark A (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci* 36:181–204. [CrossRef Medline](#)
- Cressman EK, Henriques DY (2011) Motor adaptation and proprioceptive recalibration. *Prog Brain Res* 191:91–99. [CrossRef Medline](#)
- De Gelder B, Bertelson P (2003) Multisensory integration, perception and ecological validity. *Trends Cogn Sci* 7:460–467. [CrossRef Medline](#)
- Ernst MO (2007) Learning to integrate arbitrary signals from vision and touch. *J Vis* 7(5):1–14. [CrossRef Medline](#)
- Ernst MO, Di Luca M (2011) Multisensory perception: from integration to remapping. In: *Sensory cue integration* (Trommershauser J, Körding KP, Landy MS, eds), pp 224–250. New York: Oxford UP.
- Ernst MO, Banks MS, Bühlhoff HH (2000) Touch can change visual slant perception. *Nat Neurosci* 3:69–73. [CrossRef Medline](#)
- Felsen G, Dan Y (2005) A natural approach to studying vision. *Nat Neurosci* 8:1643–1646. [CrossRef Medline](#)
- Franklin DW, Wolpert DM (2011) Computational mechanisms of sensorimotor control. *Neuron* 72:425–442. [CrossRef Medline](#)
- Gepshtein S, Burge J, Ernst MO, Banks MS (2005) The combination of vision and touch depends on spatial proximity. *J Vis* 5(11):1013–1023. [CrossRef Medline](#)
- Gibson JJ (1950) *The perception of the visual world*. Boston: Houghton Mifflin.
- Gregory RL (1980) Perceptions as hypotheses. *Philos Trans R Soc Lond B Biol Sci* 290:181–197. [CrossRef Medline](#)
- Haijiang Q, Saunders JA, Stone RW, Backus BT (2006) Demonstration of cue recruitment: change in visual appearance by means of Pavlovian conditioning. *Proc Natl Acad Sci U S A* 103:483–488. [CrossRef Medline](#)
- Hamrick J, Battaglia PW, Tenenbaum JB (2011) Internal physics models guide probabilistic judgments about object dynamics. In: *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*.
- Henriques DY, Cressman EK (2012) Visuomotor adaptation and proprioceptive recalibration. *J Mot Behav* 44:435–444. [CrossRef Medline](#)
- Ivanchenko V, Jacobs RA (2007) Visual learning by cue-dependent and cue-invariant mechanisms. *Vision Res* 47:145–156. [CrossRef Medline](#)
- Kerrigan IS, Adams WJ (2013) Learning different light prior distributions for different contexts. *Cognition* 127:99–104. [CrossRef Medline](#)
- Kingdom FAA, Prins N (2010) *Psychophysics: a practical introduction*. London: Academic.
- Kleiner M, Brainard D, Pelli D (2007) What's new in Psychtoolbox-3? *Perception* 36:Abstr. [CrossRef](#)
- Knill DC, Kersten D (1991) Apparent surface curvature affects lightness perception. *Nature* 351:228–230. [CrossRef Medline](#)
- Knill DC, Richards W (1996) *Perception as Bayesian inference*. Cambridge: Cambridge UP.
- Körding KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L (2007) Causal inference in multisensory perception. *PLoS One* 2:e943. [CrossRef Medline](#)
- Martin TA, Keating JG, Goodkin HP, Bastian AJ, Thach WT (1996) Throwing while looking through prisms. 2. Specificity and storage of multiple gaze-throw calibrations. *Brain* 119:1199–1211. [CrossRef Medline](#)
- McLaughlin SC, Webster RG (1967) Changes in the straight-ahead eye position during adaptation to wedge prisms. *Percept Psychophys* 2:37–44. [CrossRef](#)
- Ogle KN (1950) *Researches in binocular vision*. Philadelphia: W. B. Saunders Company.
- Osu R, Hirai S, Yoshioka T, Kawato M (2004) Random presentation enables subjects to adapt to two opposing forces on the hand. *Nat Neurosci* 7:111–112. [CrossRef Medline](#)
- Parise CV, Spence C, Ernst MO (2012) When correlation implies causation in multisensory integration. *Curr Biol* 22:46–49. [CrossRef Medline](#)
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442. [CrossRef Medline](#)
- Pick HL, Warren DH, Hay JC (1969) Sensory conflict in judgments of spatial direction. *Percept Psychophys* 6:203–205. [CrossRef](#)
- Prins N (2012) The psychometric function: the lapse rate revisited. *J Vis* 12(6):pii:25. [CrossRef Medline](#)
- Prins N, Kingdom FAA (2009) Palamedes: Matlab routines for analyzing psychophysical data. <http://www.palamedestoolbox.org>.
- Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 2:79–87. [CrossRef Medline](#)
- Rust NC, Movshon JA (2005) In praise of artifice. *Nat Neurosci* 8:1647–1650. [CrossRef Medline](#)
- Scarfe P, Hibbard PB (2011) Statistically optimal integration of biased sensory estimates. *J Vis* 11(7):pii:12. [CrossRef Medline](#)
- Scarfe P, Hibbard PB (2013) Reverse correlation reveals how observers sample visual information when estimating three-dimensional shape. *Vision Res* 86:115–127. [CrossRef Medline](#)
- Seydell A, Knill DC, Trommershäuser J (2010) Adapting internal statistical models for interpreting visual cues to depth. *J Vis* 10(4):1.1–27. [CrossRef Medline](#)
- Shams L, Beierholm UR (2010) Causal inference in perception. *Trends Cogn Sci* 14:425–432. [CrossRef Medline](#)
- Smeets JB, van den Dobbelaars JJ, de Grave DD, van Beers RJ, Brenner E (2006) Sensory integration does not lead to sensory calibration. *Proc Natl Acad Sci U S A* 103:18781–18786. [CrossRef Medline](#)
- Smith KA, Vul E (2013) Sources of uncertainty in intuitive physics. *Top Cogn Sci* 5:185–199. [CrossRef Medline](#)
- Smith KA, Dechter E, Tenenbaum JB, Vul E (2013) Physical predictions over time. In: *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, pp 1–6.
- van Beers RJ, van Mierlo CM, Smeets JB, Brenner E (2011) Reweighting visual cues by touch. *J Vis* 11(10):pii:20. [CrossRef Medline](#)
- van Dam LC, Hawellek DJ, Ernst MO (2013) Switching between visuomotor mappings: learning absolute mappings or relative shifts. *J Vis* 13(2):pii:26. [CrossRef Medline](#)
- van Ee R (2001) Perceptual learning without feedback and the stability of stereoscopic slant estimation. *Perception* 30:95–114. [CrossRef Medline](#)
- von Helmholtz H (1925) *Helmholtz's treatise on physiological optics*. Bristol, UK: Thoemmes.
- Welch RB, Bridgeman B, Anand S, Browman KE (1993) Alternating prism exposure causes dual adaptation and generalization to a novel displacement. *Percept Psychophys* 54:195–204. [CrossRef Medline](#)
- Wichmann FA, Hill NJ (2001a) The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept Psychophys* 63:1293–1313. [CrossRef Medline](#)
- Wichmann FA, Hill NJ (2001b) The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Percept Psychophys* 63:1314–1329. [CrossRef Medline](#)
- Wolpert DM, Landy MS (2012) Motor control is decision-making. *Curr Opin Neurobiol* 22:996–1003. [CrossRef Medline](#)
- Wolpert DM, Miall RC, Kawato M (1998) Internal models in the cerebellum. *Trends Cogn Sci* 2:338–347. [CrossRef Medline](#)
- Zaidel A, Turner AH, Angelaki DE (2011) Multisensory calibration is independent of cue reliability. *J Neurosci* 31:13949–13962. [CrossRef Medline](#)