Disambiguation of Necker cube rotation by monocular and binocular depth cues: Relative effectiveness for establishing long-term bias

Sarah J. Harrison *, Benjamin T. Backus, Anshul Jain

Graduate Center for Vision Science, SUNY College of Optometry, 33 West 42nd Street, New York, NY 10036, USA

ABSTRACT

The apparent direction of rotation of perceptually bistable wire-frame (Necker) cubes can be conditioned to depend on retinal location by interleaving their presentation with cubes that are disambiguated by depth cues (Haijiang, Saunders, Stone, & Backus, 2006; Harrison & Backus, 2010a). The long-term nature of the learned bias is demonstrated by resistance to counter-conditioning on a consecutive day. In previous work, either binocular disparity and occlusion, or a combination of monocular depth cues that included occlusion, internal occlusion, haze, and depth-from-shading, were used to control the rotation direction of disambiguated cubes. Here, we test the relative effectiveness of these two sets of depth cues in establishing the retinal location bias. Both cue sets were highly effective in establishing a perceptual bias on Day 1 as measured by the perceived rotation direction of ambiguous cubes. The effect of counter-conditioning on Day 2, on perceptual outcome for ambiguous cubes, was independent of whether the cue set was the same or different as Day 1. This invariance suggests that a common neural population instantiates the bias for rotation direction, regardless of the cue set used. However, in a further experiment where only disambiguated cubes were presented on Day 1, perceptual outcome of ambiguous cubes during Day 2 counter-conditioning showed that the monocular-only cue set was in fact more effective than disparity-plus-occlusion for causing long-term learning of the bias. These results can be reconciled if the conditioning effect of Day 1 ambiguous trials in the first experiment is taken into account (Harrison & Backus, 2010b). We suggest that monocular disambiguation leads to stronger bias either because it more strongly activates a single neural population that is necessary for perceiving rotation, or because ambiguous stimuli engage cortical areas that are also engaged by monocularly disambiguated stimuli but not by disparity-disambiguated stimuli.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Biases in the visual system are invaluable for rapid and successful interpretation of retinal input. Rather than attempt evaluation of near-threshold disparities, for instance, a quick-and-dirty decision about relative distance from the observer may be made on the basis of familiarity with an object (Ittelson, 1951) or likely object shape (Pizlo, Li, & Steinman, 2008). It has recently been shown that the visual system can learn new long-term biases, when presented with new environmental contingencies: Through an associative learning paradigm, the perceived direction of rotation of a perceptually bistable wire-frame (Necker) cube at stimulus onset becomes dependent on its retinal location (Haijiang et al., 2006; Harrison & Backus, 2010a).

Training of the retinal location bias was originally achieved through a conditioning paradigm whereby subjects viewed interleaved presentation of perceptually ambiguous cubes with cubes that were disambiguated by binocular disparity and occlusion depth cues (Haijiang et al., 2006): Disambiguated cubes rotated in opposite directions at two retinal locations, and the perceived rotation direction of ambiguous cubes adopted the same location–rotation contingency. It has since been shown that a combination of monocular depth cues, in the absence of disparity information in the stimulus, is sufficient to disambiguate rotation direction and establish the perceptual bias (Harrison & Backus, 2010a). In both cases, the long-term nature of the bias was demonstrated through resistance to counter-conditioning, in the form of reverse-contingency training, on a consecutive day.

In the experiments that follow, we directly compared learning with the "monocular-only" cue set to that with the original "disparity-plus-occlusion" set, with two aims in mind. First, we aimed to evaluate and understand the relative effectiveness of the two cue sets, in terms of strength of the established bias. Previous studies cannot unequivocally answer this question, as biases were assessed by reverse-training using the same cue set as was used in initial training. Our secondary question was to examine the extent to which the learned biases are in common between the two cue sets or are in fact specific to the cue set used

* Corresponding author. Fax: +1 212 938 5760.
E-mail addresses: sharrison@sunyopt.edu (S.J. Harrison), bbackus@sunyopt.edu (B.T. Backus), ajain@sunyopt.edu (A. Jain).

when the bias was first learned. That is, are we training one and the same bias in both cases, or two different biases?

The first question, of the relative effectiveness of the two cue sets in training long-term bias, could be considered within a framework whereby the visual system learns which biases are appropriate by association with "ecologically-valid" (Brunswick, 1956) or "long-trusted" (Backus & Haijiang, 2007) cues. Presumably, the more "trusted" a cue is, the more effective it should be in our learning paradigm. Under some circumstances, binocular disparity could be more "trusted", and hence a more effective cue than monocular cues such as shadows and occlusion: Whereas light-sources are rarely fixed, and contrast changes can give misleading impressions of occlusion, disparity is a low-level geometrical property of an image that, at the supra-threshold levels we use in the following experiments, provides direct evidence about relative depth within a shape. Although the monocular cues we used also provide sufficient information to disambiguate rotation direction, the visual system may not, in general, place as much weight on this qualitative form of depth information.

Our second question, of whether the two cue sets train two independent biases, pertains to whether the biases established by the two cue sets are instantiated in the same or different neural populations. Although the monocular-only cue set and the disparity-plus-occlusion cue set both depict motion and depth, this does not necessitate that the percept of rotation is supported by activity in identical populations in both cases. The extent to which biases trained using the monocular-only cue set are affected by reverse-training using the disparity-plus-occlusion cue set, and vice versa, will address this issue.

2. General methods

2.1. Hardware and software

Experiments were programmed in Python using the Vizard platform version 3.11 (© WorldViz, Santa Barbara, CA, USA) on a Dell Precision T3400 computer. Stimuli were rear-projected onto a screen, using a Christie Mirage S + 4 K projector.

2.2. Cube stimuli

Stimuli were identical to those used previously (Harrison & Backus, 2010a). Simulated rotating cube stimuli were light against a dark background. All stimuli were viewed through red–green glasses, in order to present disparity information and to control the eye of presentation.

Two distinct sets of depth cues were used, to control the direction of rotation of disambiguated cubes (Fig. 1). The first cue set was binocular disparity and occlusion. Disparity cues were created by the use of red–green anaglyphic images. Cubes were presented with geometrically correct disparity, resulting in a maximum disparity of 1.0° between the nearest and farthest point of the rotating cube. The occlusion cue consisted of a central column around which the cube rotated. The column was a 2-dimensional vertical strip of width 4.0 cm, extending from the top to the bottom of the screen area. Far portions of the cube were occluded as they moved around the back of the column, whereas closer portions of the cube were visible in front of the column. The column was presented stereoscopically at the screen distance, hence provided an additional reference point for relative disparity of different edges of the cube.

The second cue set consisted of monocular depth cues only. Cube frame edges had increased width and breadth, which permitted the creation of a depth-from-shading cue through use of a directional light-source, and an internal occlusion cue. The effect of the light-source on different faces of the cube edges was further manipulated by use of the Vizard 'fog' function, which simulates the effect of haze or 'aerial perspective' whereby contrast decreases with distance. The occlusion cue provided by the central column was once again present, but the column was presented monocularly.

The edges of the cube were solid rectangular parallelepipeds with length of 20.0 cm (outside edges), and width and breadth of 0.3 cm (disparity-plus-occlusion cue set) or 2.0 cm (monocular-only cue set). Each transparent face of the cube contained 25 randomly placed dots, which stabilized appearance on ambiguous trials as a single rigid rotating body. Cubes rotated about a vertical axis at a rate of 45 degrees s^{-1}, from a starting orientation at stimulus onset such that their front and back edges were vertical and co-incident at the center of the image (45° of yaw). The roll and pitch angles, which determine whether the cube appeared to be viewed from above or below at stimulus onset, were either both +25° or both −25°. This parameter was balanced across both disambiguated and ambiguous trials to remove effects of confounding cues (Haijiang et al., 2006). Cubes were viewed at a distance of 1.0 meter; hence cube edges subtended approximately 11.5° of visual angle when perpendicular to the line of sight.

Ambiguous cubes were presented monocularly to observers' right eyes (i.e. only the green image), hence had no disparity information, and also contained no other monocular cues to depth. The dimensions of ambiguous cubes were always identical to those of the disambiguated cubes with which they were interleaved. All disambiguated and ambiguous cubes were presented using orthographic projection so that perspective cues did not indicate front and back of the cube in the ambiguous case.

2.3. Procedure and task

The trial composition was identical to that of Harrison and Backus (2010a): In all experiments, a 2.0 cm × 2.0 cm square-outline fixation marker was presented binocularly, at the screen depth. The fixation marker remained on the screen at all times. Subjects were instructed to achieve fixation of the marker prior to initiating each trial with a key press, and to then maintain fixation rather than look directly at the rotating cube stimulus. On initiation of each trial, a rotating cube appeared centered either 12° above or below the fixation marker. Simultaneously, a comparison dot repeated cycles of horizontal motion through the fixation marker. Dot speed was 15.7 cm s^{-1}, and dot direction was randomized, with equal probability for leftward and rightward motion. The dot was presented at fixation depth on disambiguated trials and monocularly on ambiguous trials.

Subjects' task was to indicate whether the direction of motion of the comparison dot (leftward or rightward) was the same as the direction of motion of the front of the cube or the back of the cube (keypress '2' and '8' respectively). Due to random assignment of dot direction, the measure of interest, perceived direction of rotation, was not correlated with the keypress response. None of the subjects who took part in the following experiments reported being aware that cube rotation was dependent on cube location even though location correlated perfectly with rotation direction on disambiguated trials. The cube and comparison dot remained on the screen for a minimum of 1.5 s and a maximum of 6.0 s; subjects' response terminated the presentation at any time after 1.5 s. The order of presentation of stimuli for each experiment is described below.

2.4. Subjects

Subjects were adult members of the public who were paid for their time. Subjects' vision was normal or corrected-to-normal with non-bifocal lenses. All subjects were required to have a
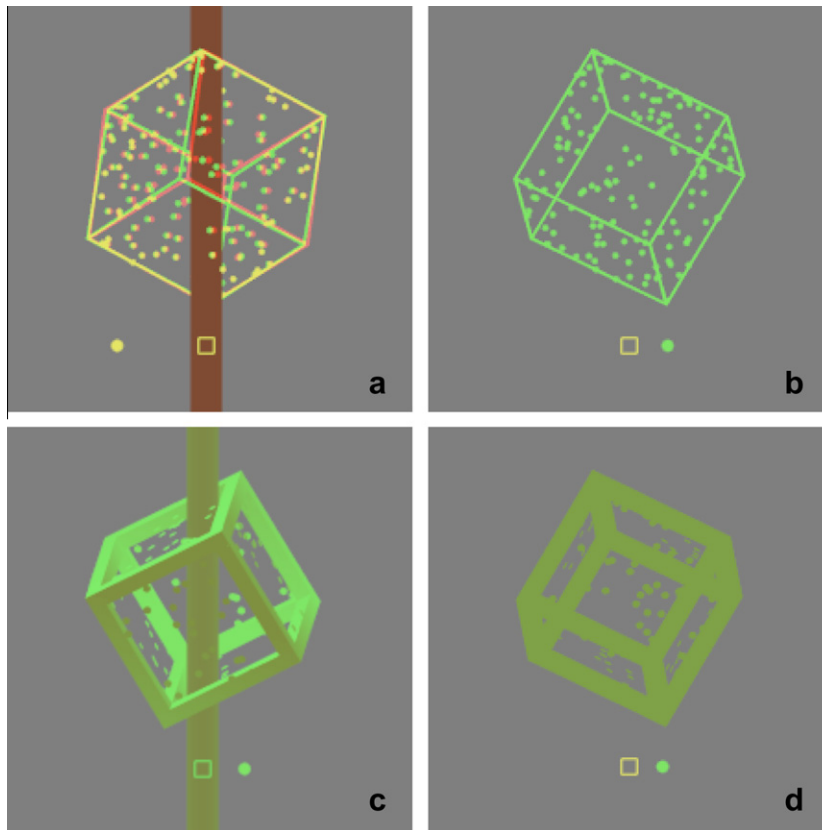
Fig. 1. Screen shots (cropped) showing the various depth cues used. (a) Cube disambiguated by disparity and occlusion cues; (b) corresponding ambiguous cube, presented monocularly; (c) CUbe disambiguated by occlusion, internal occlusion, depth-from-shading, and haze; (d) Corresponding test cube. Screen shots also show the location of the fixation square and path of the comparison dot for a "top" position cube in both experiments.

minimum stereoacuity of 240 s of arc (TNO stereoacuity test) regardless of whether the experimental condition in which they took part used disparity as a depth cue. However, our critical measure of subjects' suitability for the experiment, in terms of both stereoacuity in dynamic displays (Rouse, Tittle, & Braunstein, 1989) and task comprehension, was their performance on Day 1 disambiguated trials.

Eighteen subjects, who met other criteria but did not reach performance levels on disambiguated trials of 95% or over on Day 1, were excluded from the study (Experiment 1, two subjects who viewed the monocular-only cue set, three subjects who viewed the disparity-plus-occlusion cue set; Experiment 2, six subjects who viewed the monocular-only cue set, six subjects who viewed the disparity-plus-occlusion cue set; Additional control experiment, 1 subject excluded). Finally, as our study aimed to compare the strength of retained bias on Day 2 as a result of equivalent perceptual outcomes on Day 1 (rather than comparing the relative ease with which subjects interpret various depth cues), we required subjects to have been equally "trained" on Day 1 to perceive ambiguous trials in the same directions as disambiguated trials; two subjects from Experiment 1 who reported less than 70% of ambiguous cubes as rotating in the same direction as cubes disambiguated by disparity-plus-occlusion were therefore excluded. A total of 72 subjects completed experiments presented in the study.

2.5. Analysis

The percent of cubes seen as rotating in the direction specified by the Day 1 disambiguated cubes, at each of the two locations, was transformed into a z-score, i.e. we used a probit (inverse-cumulative-normal) transformation (Backus, 2009; Dosher, Sperling, & Wurst, 1986). This is a measure of the likelihood of

the observations given normally distributed noise in a decision variable. For the purpose of analysis, saturated values (100% or 0%) were replaced with a z-score of 2.394. This is equivalent to two nonconforming responses within 240 observations, or one response in 120 observations. For each subject, z-scores for the top and bottom locations were summed, giving a "zDiff" measure of the extent to which perceived rotation differed between the two locations. In the case of ambiguous cubes, zDiff is a measure of training-induced bias, which is independent of any global, preexisting bias for rotation direction.

3. Experiment 1

The efficacy of disparity-plus-occlusion and monocular-only cue sets, both for disambiguating rotation direction and for establishing long-term bias for perceived rotation direction, has previously been demonstrated (Harrison & Backus, 2010a). Our aim here was to directly compare the strength of the bias established by equivalent perceptual outcomes achieved under the two cue sets, and the overlap in the neural instantiation of the bias, by use of a cross-over design.

On Day 1, subjects viewed randomly-interleaved ambiguous and disambiguated cubes. On Day 2, subjects once again viewed randomly-interleaved ambiguous and disambiguated cubes, but the location–rotation contingency depicted by disambiguated cubes was reversed. The extent to which the bias learned on Day 1 is still present on Day 2, is demonstrated through resistance of the perceptual outcome in the ambiguous case to this reverse-training.

In previous studies, the long-term nature of the learned bias was evaluated through reverse-training using the same disambiguating cues as had been used during initial training. Thus biases

trained using disparity-plus-occlusion cues were shown to be resistant to reverse-training using the same disparity-plus-occlusion cues, and equivalently for the monocular-only cue set. The strength of the long-term biases learned from the two cue sets cannot be directly compared in this way, as the reverse-training procedure is different in each case; a valid assessment of the relative effectiveness of the two cue sets in training the bias on Day 1 requires that the biases are assessed identically on Day 2. However, this raises the additional question of whether the learned bias might be tied to the disambiguating cues used in initial training: If one cue set reverse-trained a bias initially established using the other cue set, would this be because the second cue set was stronger, or because the new cue set was training an entirely new bias?

To disentangle these possibilities, we tested 16 subjects in a cross-over design: Eight subjects received training using the (binocular) disparity-plus-occlusion cue set and reverse-training using the monocular-only cue set (BM), and eight subjects received training and reverse-training using the reverse order of cue sets (MB). Additionally, our analysis used data from an earlier study (published in Harrison & Backus, 2010a; Experiment 1 "same" room subjects, Experiment 4 all subjects). that used identical stimuli and methods. The data was from two independent groups of eight subjects, who received either disparity-plus-occlusion training and reverse-training (BB), or monocular-only training and reverse-training (MM).

3.1. Methods

On Day 1, subjects were trained under either "top-clockwise (CW), bottom-counterclockwise (CCW)" contingency, or vice versa. All subjects completed one session consisting of 240 disambiguated trials and 240 ambiguous trials. Disambiguated and ambiguous trials were randomly-interleaved, within balanced sets of 8 trials that contained all combinations of ambiguous/disambiguated trial type, top/bottom cube location, and above/below cube viewpoint (as detailed in Section 2). The first set of trials was constrained to present the eight possible trials in a disambiguated-ambiguous sequence, alternating between locations and cube viewpoints. The order of presentation of these first eight trials was counterbalanced across subjects. On Day 2, trials were organized as before, but subjects viewed disambiguated cubes with the opposite location–rotation contingency to that viewed on Day 1.

3.2. Results

3.2.1. Ambiguous trials

Results from the four groups are presented in Fig. 2. We tested for differences in perceptual outcomes of Day 1 ambiguous trials between the two cue set groups using an independent-samples t-test. Note that effective recruitment of the location cue on Day 1 was a prerequisite for subject inclusion. We found no difference between the two cue sets (disparity-plus-occlusion, mean = 4.30, s.e.m. = .14; monocular-only, mean = 4.00, s.e.m. = .19; $t(30)$ = 1.27, p = .21), confirming that on Day 1, there was no difference between groups in the extent to which ambiguous cubes adopted the location–rotation contingency depicted by disambiguated cubes.

All four groups retained a significant bias on Day 2, as shown by comparison of perceptual outcomes of Day 2 ambiguous trials for each group, with the (sign-reversed) perceptual outcomes of Day 1 ambiguous trials for all four groups (p < .01 for all groups, corrected for unequal variance).

Next, we analyzed perceptual outcomes of Day 2 ambiguous trials, using a 2-way ANOVA with orthogonal factors of Day 1 cue set and Day 2 cue set. Neither Day 1 nor Day 2 cue set, nor their interaction, were found to be significant (Day 1 cue set, $F(1, 28)$ = 1.08, p = .31; Day 2 cue set, $F(1, 28)$ = .34, p = .57; Day 1 × Day 2, $F(1, 28)$ = .02, p = .88; BB, mean = 1.61, s.e.m. = .59; BM, mean = 1.30, s.e.m. = .67; MB, mean = .97, s.e.m. = 1.01; MM, mean = .45, s.e.m. = .48).

To further verify the lack of difference between the cue sets, in establishing the long-term bias and in their ability to reverse-train the bias, we calculated the change in ambiguous trial outcomes between Day 1 and Day 2. Although performance had not been significantly different between the two cue set groups on Day 1, we sought to confirm that minor differences between subjects on Day 1 were not hiding a difference in the extent to which the bias was retained on Day 2. For each subject, we calculated the difference between the Day 1 zDiff and Day 2 zDiff. As before, these measures were analyzed using a 2-way ANOVA. Again, neither Day 1 cue set nor Day 2 cue set, nor their interaction, were significant (Day 1 cue set, $F(1, 28)$ = .37, p = .55; Day 2 cue set, $F(1, 28)$ = .22, p = .65; Day 1 × Day 2, $F(1, 28)$ = .28, p = .60; BB, mean = 2.48, s.e.m. = .52; BM, mean = 3.21, s.e.m. = .66; MB, mean = 3.32, s.e.m. = 1.14; MM, mean = 3.27, s.e.m. = .37).

At first glance, the results of Experiment 1 suggest that there is no significant difference in the strength of long-term bias learned during Day 1 training with the disparity-plus-occlusion vs. monocular-only cue sets. However, interpreting the results may not be straightforward in this respect: Subjects viewed a mixture of ambiguous and disambiguated cubes on Day 1. A difference in the effectiveness of the two cue sets in establishing the bias that was assessed on Day 2 could have been masked by learning caused by the perceptual outcomes of Day 1 ambiguous cubes (Harrison & Backus, 2010b; see also, van Dam & Ernst, 2010). This problem is addressed in Experiment 2, where we compare groups of subjects who viewed only disambiguated trials on Day 1, with disambiguation provided either by disparity-plus-occlusion or by the monocular-only cue set.

Regardless of the relative effectiveness of the two cue sets, Experiment 1 conclusively demonstrates that the learned bias is not specific to the cue set used to train it: A bias established with the monocular-only cue set was evident during reverse-training with the disparity-plus-occlusion cue set, and vice versa. This suggests, at the very least, a large overlap in the neural population underlying biases established by the two cue sets in Experiment 1.

3.2.2. Disambiguated trials

In addition to analyzing perceptual outcomes on ambiguous trials, we also routinely checked subjects' comprehension and task compliance by analyzing reported percepts of disambiguated trials. Disambiguated trial outcomes were not significantly different between the two cue sets on Day 1, as expected from our use of subject inclusion criteria in the study. However, we found that Day 2 outcomes for supposedly disambiguated cubes were not always in accord with the disambiguating cues (Fig. 3). A 2-way ANOVA showed that both Day 1 and Day 2 cue sets had a significant effect on the perceptual outcome of disambiguated cubes on Day 2 ($F(1, 28)$ = 6.53, p = .02; $F(1, 28)$ = 23.32, p < .01), as did their interaction ($F(1, 28)$ = 5.46, p = .03): Monocular-only trials were frequently misreported, apparently even more so if the bias was established using the disparity-plus-occlusion cue set on Day 1. In contrast, the percept of the disparity-plus-occlusion trials was reliably reported in the same direction as depicted by the cues.

We surmise that the bias learned on Day 1, both during training with disparity-plus-occlusion and with monocular-only cues, was in some cases strong enough to overcome the rotation direction specified by the reverse-contingency monocular depth cues on Day 2. This finding suggests a difference in the strength of disambiguation between the disparity-plus-occlusion and monocular-only cue sets, which was not revealed by presenting disambiguated
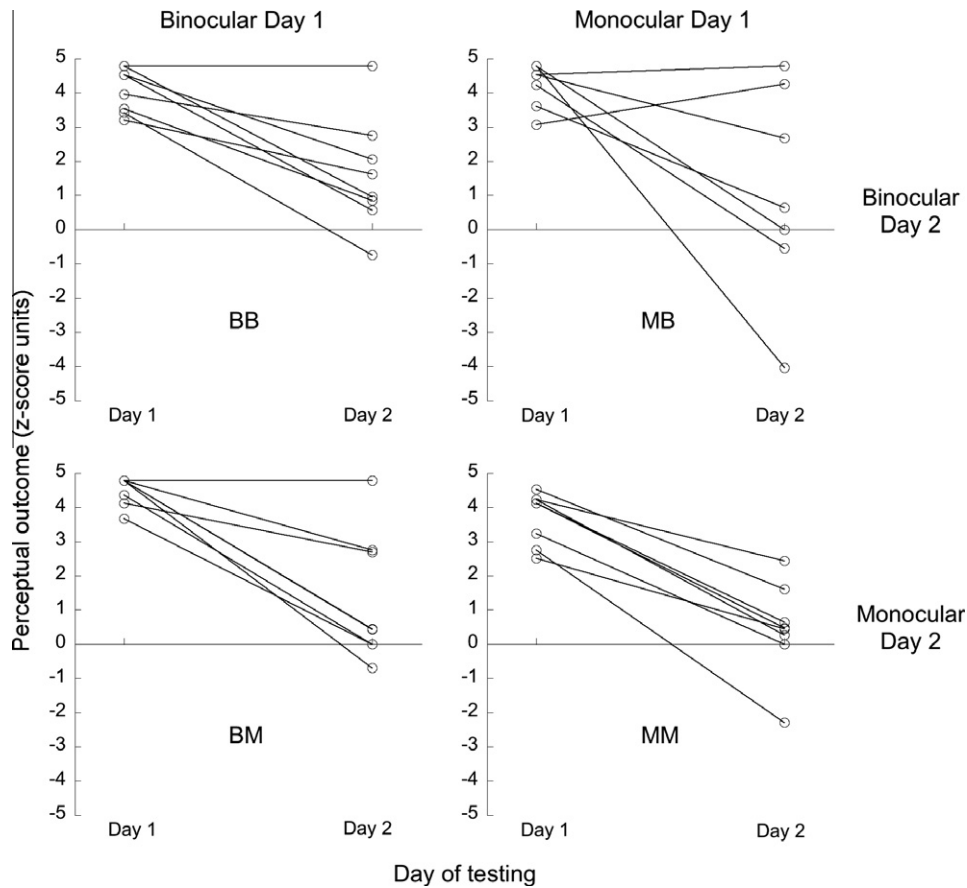
Fig. 2. zDiff scores for ambiguous trials on Days 1 and 2, for four groups of eight subjects representing all combinations of training and reverse-training cue sets. Each connected pair of points are data for an individual subject. Outcomes are assessed relative to Day 1 location–rotation contingency.

cubes in the absence of a bias (as we did on Day 1). However, the misreporting on Day 2 of monocular-only disambiguated trials, when Day 1 conditioning had used the disparity-plus-occlusion cue set rather than the monocular-only cue set, suggests that disparity-plus-occlusion in fact establishes the stronger bias. This is not consistent with our conclusion based on the outcomes of ambiguous trials, which we expected to be a sensitive indicator of bias. We have no explanation for the result regarding the outcomes of disambiguated trials in Experiment 1, but note that it is not substantiated in Experiment 2.

4. Experiment 2

The aim of Experiment 2 was to provide a stronger test of the relative effectiveness of the two cue sets, by comparing the long-term influence of viewing disambiguated trials only. As in Experiment 1, a full cross-over design was completed, with both cue sets used for Day 1 conditioning and Day 2 counter-conditioning. In addition, in Experiment 1, the edge elements of monocular-only cubes had a greater width and breadth than did those of disparity-plus-occlusion cubes, which led to greater total luminance and contrast of the stimuli. These configural properties could potentially cause confounding differences in the effectiveness of the stimuli for establishing bias. Hence, in Experiment 2, the dimensions of the Day 1 cubes were matched between the two cue sets, removing configural confounds between the conditioning stimuli and allowing a direct comparison between the effectiveness of the two cue sets: In this experiment, the only difference between the conditioning stimuli on Day 1 was the cue set used to disambiguate the rotation direction.

4.1. Methods

All aspects of the stimulus configuration were identical to that used in Experiment 1: As before, subjects were trained on Day 1 under either "top-clockwise (CW), bottom-counterclockwise (CCW)" contingency, or vice versa. On Day 1, all subjects completed one session consisting of 480 disambiguated trials, and no ambiguous trials. Sixteen subjects viewed cubes with rotation direction disambiguated by the monocular-only cue set. A further sixteen subjects viewed cubes with rotation direction disambiguated by disparity-plus-occlusion; cubes had edges with increased width and breadth of 2.0 cm to match that of the monocular-only cubes. On Day 2, half of each group received reverse-training using wire-frame disparity-plus-occlusion cubes with the smaller width and breadth of 0.3 cm. The other half of the subjects received reverse-training using the monocular-only cue set.

4.2. Results

4.2.1. Ambiguous trials

The perceptual outcomes of Day 2 ambiguous trials are presented in Fig. 4 a & b. As there were no ambiguous trials on Day 1 in Experiment 2, we verified that all groups retained a significant bias on Day 2 by comparison of perceptual outcomes of Day 2 ambiguous trials for each group, with the (sign-reversed) perceptual outcomes of Experiment 1 Day 1 ambiguous trials for all four groups, ($p < .01$ for all groups, corrected for unequal variance).

A 2-way ANOVA, with orthogonal factors of Day 1 cue set and Day 2 cue set, showed that Day 1 cue set was highly significant ($F(1, 28) = 22.94$, $p < .01$), with monocular-only conditioning
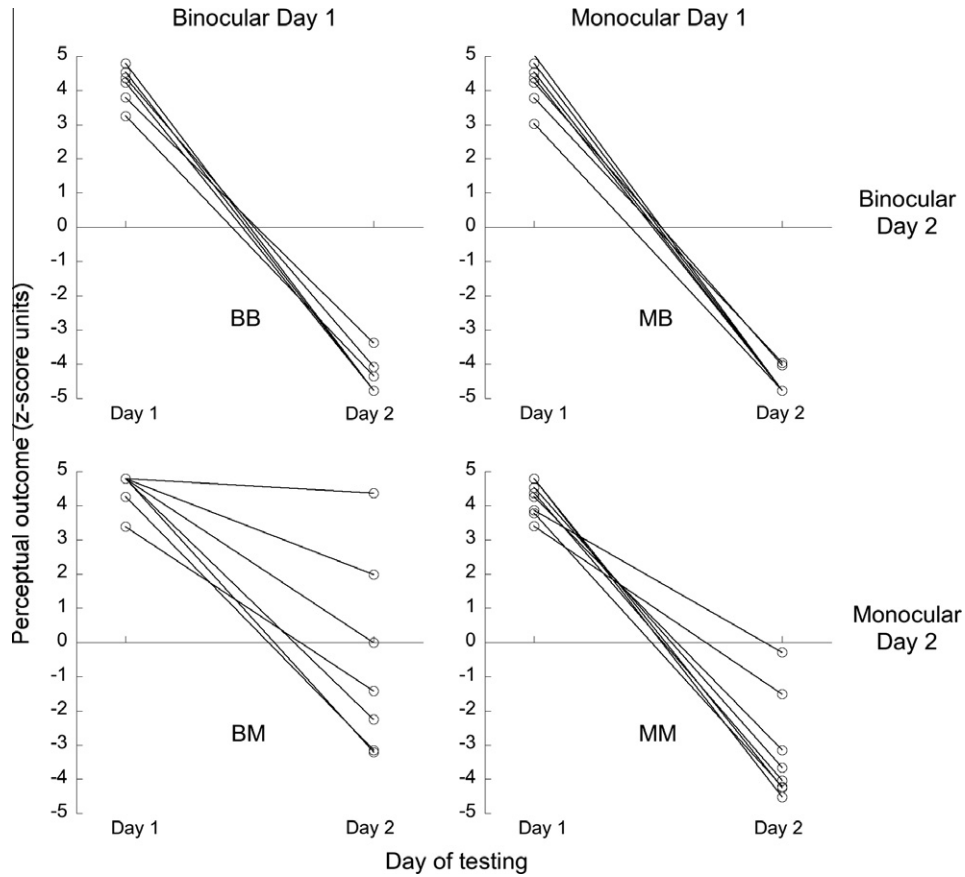
Fig. 3. zDiff scores for disambiguated trials on Days 1 and 2, for four groups of eight subjects representing all combinations of training and reverse-training cue sets. Each connected pair of points are data for an individual subject. Outcomes are assessed relative to Day 1 location–rotation contingency. Disambiguated trials had the opposite contingency on Day 2 to that on Day 1, hence "correct" performance would result in zDiffs of equal magnitude but opposite sign on the 2 days.

leading to far greater retained bias than disparity-plus-occlusion (zDiff = 1.79 vs. −1.89). However, neither Day 2 cue set ($F(1, 28) = .02$, $p = .90$) nor the interaction of Day 1 and Day 2 cue set ($F(1, 28) = .15$, $p = .70$) were even marginally significant. In summary, Day 1 monocular-only conditioning led to greater retained

bias, regardless of the form of Day 2 counter-conditioning (BB, mean = −2.15, s.e.m. = .78; BM, mean = −1.63, s.e.m. = .54; MB, mean = 1.75, s.e.m. = .97; MM, mean = 1.83, s.e.m. = .71).

To control for the fact that cube element size was different on Day 1 and Day 2 for subjects who received disparity-plus-occlusion
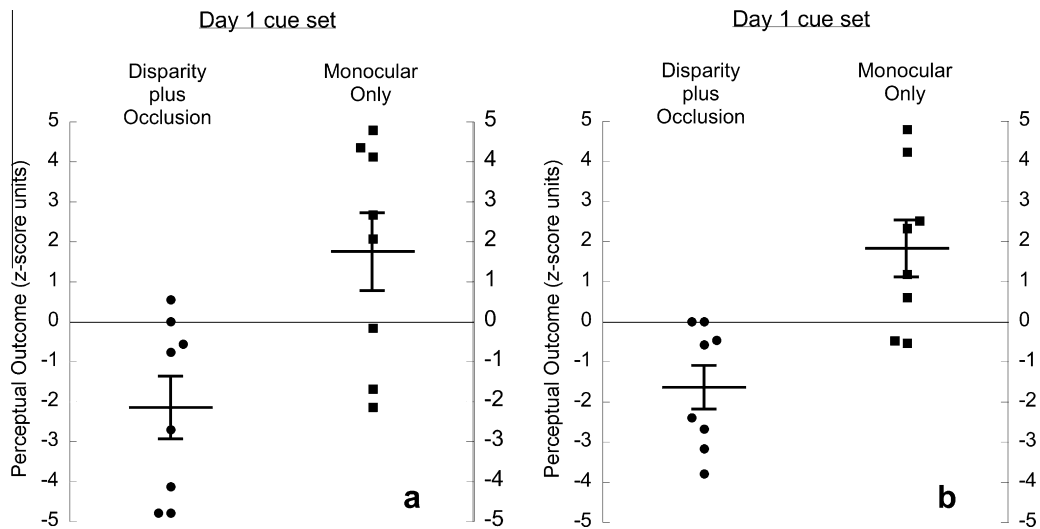

Fig. 4. zDiff scores for ambiguous trials on Day 2. Data points for individual subjects are overlaid with means and standard errors of the mean. Outcomes are assessed relative to Day 1 contingency. (a) Subjects viewed the disparity-plus-occlusion cue set on Day 2; two groups of eight subjects. (b) Subject viewed the monocular-only cue set on Day 2; two groups of eight subjects. It can be seen that the two graphs are highly similar, indicating that the cue set used for counter-conditioning on Day 2 did not affect the retained bias; Day 1 conditioning using the monocular-only cue set established a greater perceptual bias than did the disparity-plus-occlusion cue set.

conditioning and counter-conditioning, we collected additional data from eight subjects who viewed cubes with standard width (0.3 cm) edges on both days, i.e. identical to group BB above, but with no configural difference between the conditioning stimuli on Day 1 and Day 2. We found no significant difference in the bias established using "standard" or "wide" disparity-plus-occlusion stimuli ($t(14) = .30$, $p = .77$), providing evidence that configural differences are not an important influence in the biases we are measuring. Hence, the results of Experiment 2 conclusively show that the monocular-only cue set establishes a greater long-term bias than does the disparity-plus-occlusion cue set. This difference cannot be attributed to configural differences between the stimuli such as cube element size or total stimulus contrast. This finding is all the more notable insofar as the Day 1 disparity-plus-occlusion stimuli in Experiment 2 had *greater* luminance than the Day 1 monocular-only stimuli, as the red and green images had fixed (unequal) luminance across all conditions.

MB subjects in Experiments 1 and 2 viewed identical disambiguating stimuli on Day 1, and received identical counter-conditioning on Day 2. Likewise for BB subjects in Experiment 1 and in the control experiment described above. The difference between the two groups in each case was in the proportion of disambiguated stimuli presented on Day 1. Hence, we were able to test for differences between groups that viewed a 50:50 mix of ambiguous and disambiguated stimuli on Day 1 (Experiment 1), and those that viewed disambiguated cubes only on Day 1 (Experiment 2). We found that subjects who viewed the disparity-plus-occlusion cue set on Day 1 had greater retained bias on Day 2 if they had viewed a mixture of ambiguous and disambiguated trials on Day 1 than if they had viewed disambiguated trials only ($t(14) = 4.32$, $p < .01$,

zDiff = 1.61, s.e.m. = .59; zDiff = −1.86, s.e.m. = .55 respectively). This confirms our previous result (Harrison & Backus, 2010b). In contrast, for the monocular-only cue set on Day 1, there was no difference in long-term bias between subjects who viewed the mixture of trials and those that viewed disambiguated trials only ($t(14) = −.31$, $p = .76$, zDiff = .97, s.e.m. = 1.01; zDiff = 1.45, s.e.m. = 1.16 respectively). We interpret these results as indicating that our monocular-only and ambiguous stimuli are more similar to each other in their ability to establish long-term bias than are our disparity-plus-occlusion and ambiguous stimuli.

4.2.2. Disambiguated trials

Disambiguated trial outcomes were not significantly different between the two cue sets on Day 1, as expected from subject inclusion criteria for the study. However, once again, Day 2 outcomes were not always in accord with the disambiguating cues, with the vast majority of instances being for the monocular-only cue set (Fig. 5). This adds to evidence from Experiment 1 that monocular depth cues are more readily over ruled by the visual system.

A 2-way ANOVA showed that Day 2 cue set had a significant effect on the perceptual outcome of disambiguated cubes on Day 2 ($F(1, 28) = 22.37$, $p < .01$) while Day 1 cue set did not ($F(1, 28) = 2.20$, $p = .15$). The interaction between Day 1 and Day 2 cue set was not significant ($F(1, 28) = 1.29$, $p = .27$). Hence, in this experiment, where Day 2 biases are directly attributable to experience of Day 1 disambiguated trials (as opposed to a combination of disambiguated and ambiguous Day 1 trials, as in Experiment 1), we did *not* find that Day 2 monocular-only disambiguated trials were more likely to be misreported when the Day 1 conditioning cue set was disparity-plus-occlusion. To the contrary, it
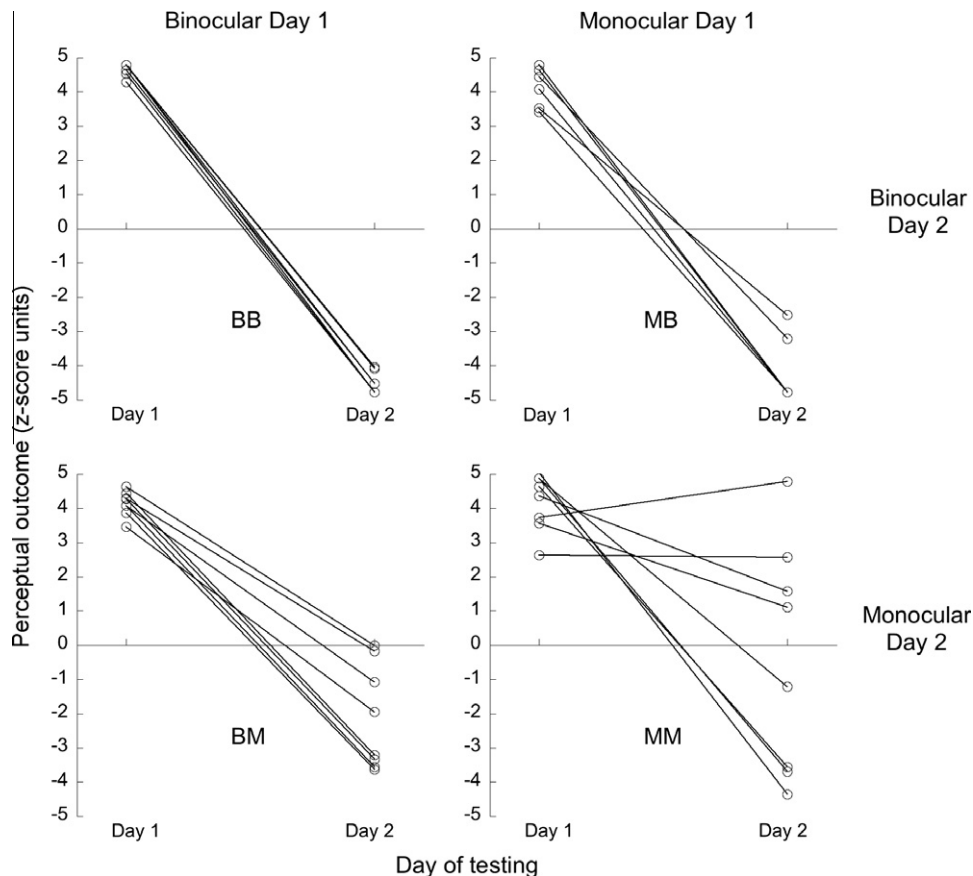

Fig. 5. zDiff scores for disambiguated trials on Days 1 and 2, for four groups of eight subjects representing all combinations of training and reverse-training cue sets. Each connected pair of points are data for an individual subject. Outcomes are assessed relative to Day 1 location–rotation contingency. Disambiguated trials had the opposite contingency on Day 2 to that on Day 1, hence "correct" performance would result in zDiffs of equal magnitude but opposite sign on the 2 days.

is evident in Fig. 5 that, although the comparison didn't reach significance, there was a greater tendency to misreport Day 2 monocular-only disambiguated trials when the Day 1 cue set - was monocular-only.

5. Discussion

Experiment 1 showed that the previously-demonstrated learned bias for rotation-in-depth is resistant to reverse-training regardless of whether the disambiguating cue set is the same or different from the set used when the bias was established. Hence, the bias is not specific to the cues used to initially define it. Our second experiment suggested that training was more effective in establishing the long-term bias when subjects viewed cubes disambiguated by our combination of monocular depth cues than when they viewed combined depth cues of binocular disparity and occlusion.

The implication of the first finding is that the bias established by interleaved ambiguous and disambiguated trials is instantiated in a common neural population, regardless of which cue set is used in training. The equivalence between cue sets cannot be attributed to the common occlusion cue, as we found large differences in Experiment 2. Findings from numerous studies converge on MT+/V5 as the likely site of this neural population: Single-cell recordings in macaque show that MT responses predict the perceived rotation direction of ambiguous structure-from-motion stimuli (Bradley, Chang, & Andersen, 1998; Dodd, Krug, Cumming, & Parker, 2001; Grunewald, Bradley, & Andersen, 2002), through activity in neurons jointly tuned for the direction of motion and binocular disparity of local stimulus components (DeAngelis, Cumming, & Newsome, 1998; DeAngelis & Uka, 2003; Maunsell & Van Essen, 1983a, 1983b). Likewise, fMRI studies in humans have shown that activity in MT+ correlates with perceptual outcome for ambiguous structure-from-motion stimuli such as used here (Brouwer & van Ee, 2007). Further, psychophysical studies have demonstrated that the perceptual outcome of ambiguous Necker cubes shows retinal location specificity (Harrison & Backus, 2010a; Knapen, Brascamp, Adams, & Graf, 2009), pointing to a strongly retinotopically-organized, rather than high-level, cortical area as the main locus of the bias. Hence, it is seems likely that the observed bias for rotation direction is instantiated in the responses of MT+ neurons, and is an adaptation caused by previous activity in the same population. This does not however preclude the involvement of other cortical areas in the perceptual decision process (e.g. Parker & Krug, 2003).

Our second finding addresses our first question, of the relative effectiveness of the two cue sets in establishing the long-term bias; however, it could be interpreted at many levels. The most conservative explanation would simply state that our chosen depth cues happened to differ in their effectiveness, and that another set of cues could cause an arbitrary amount of learning, either greater or lesser in magnitude than the two cue sets used here. This statement is somewhat conservative, and places no value on our observation of differences in the robustness of the two cue sets: In both experiments, we unexpectedly found that depth information provided by our monocular cue set on Day 2 was frequently "over ruled" by the bias learned on Day 1, whereas the disparity-plus-occlusion cue set universally "trumped" the bias.

We consider it highly unlikely that cubes disambiguated by monocular-only cues were more effective because both they and the ambiguous stimuli (in which the bias was observed) were depicted without disparity. We admit the possibility that the percept for cubes disambiguated by monocular-only cues elicited activity in a neural population more strongly overlapping that for ambiguous stimuli than did cubes disambiguated by disparity-plus-occlusion (see discussion below). However it is improbable that a simple monocular/binocular distinction underlies our results for three reasons: (i) in Experiment 1, the two cue sets led to a bias that was invariant to the combination of conditioning and counter-conditioning cue set; (ii) we have previously shown that the bias is not specific to the eye of origin (Harrison & Backus, 2010a), hence would conjecture that it is unlikely to be instantiated in a population of neurons that is differentially sensitive to monocular vs. binocular information; and (iii) the bias for perceived rotation in stimuli such as used here is widely held to be instantiated in MT neurons which, while holding traces of ocular dominance, are not thought to be selective for monocular vs. binocular input (as detailed above).

We conjecture that the monocular cue set in fact caused greater learning *because* it provided less robust information for the task at hand – namely, forming a rapid judgment of motion in depth. This relationship could be interpreted as opposite to that predicted by a "learning from trusted cues" framework (Haijiang et al., 2006; Helmholtz, 1910/1925), but is possibly consistent with our previous finding of stronger learning from ambiguous trials than from disambiguated trials (Harrison & Backus, 2010b; see also, van Dam & Ernst, 2010). In that study, viewing ambiguous trials on Day 1 led to learning that was even greater (in terms of magnitude of retained bias on Day 2) than we found here for viewing the monocular-only cue set trials. Ambiguous cubes surely provide the least information about true rotation possible, and yet they apparently establish the strongest long-term perceptual bias. An analogous relationship between strength of ambiguity and strength of (short-term) priming has been reported by Pastukhov, Ludwig, and Braun (2010) for not only perceptual outcome of Necker cubes, but for binocular rivalry and the kinetic depth effect.

If there is in fact an inverse relationship between how robust the rotation-in-depth signal is and the strength of the long-term bias that the stimulus establishes, how could this be explained? Is it sufficient simply to describe our finding as showing that a bias is learned under conditions where it is most evident to the system that a bias will be needed to disambiguate future occurrences? Our monocular-only cue set clearly disambiguated rotation direction on Day 1. Hence, the perceptual outcomes themselves provided the visual system with no greater "evidence" that a bias was needed than did the disparity-plus-occlusion depth cue. Instead, we propose that inferential processes are the "evidence" that the visual system is using here, and that these processes in some way drive the observed learning. For instance, *imagining* a stimulus can cause learning (Tartaglia, Bamert, Mast, & Herzog, 2009), and can influence the outcome of binocular rivalry (Pearson, Clifford, & Tong, 2008); perhaps other types of internally-generated percepts have similar effects.

While the disparity in the disparity-plus-occlusion cue set can provide direct, bottom-up, sensory evidence to MT+ as to which moving elements of the cube stimulus are near and which are far, the monocular-only cues conceivably require more complex processing to assign depth to the moving elements and to jointly encode motion and depth. The stronger bias after monocular training could be the result of activating MT+ neurons by indirect instead of direct pathways, i.e. (monocular depth cue) activation by indirect pathways could cause greater changes in a neural population that completely overlaps with that activated by (binocular depth cue) direct, bottom-up, pathways.

An alternative possibility is that the more complex processing required to reach a rotation-in-depth percept for the monocular-only stimuli could evoke activity in a greater range of cortical areas, which are also active in resolving the rotation direction of

ambiguous stimuli. Indeed, an MEG study of apparent motion found neural activity that was not only attributable to the establishment and maintenance of the dominant percept, but was in fact related to the ambiguity of the stimulus, with greater ambiguity resulting in greater levels of activity (Kaneoke, Urakawa, Hirai, Kakigi, & Murakami, 2009). In the current instance, where we are concerned to know whether additional neural substrates might be active in supporting the percept of 3D rotation in our stimuli, we note that mental rotation in 2D has been shown to activate not only MT+ but also several regions in the parietal and frontal cortex (Seurinck, de Lange, Achten, & Vingerhoets, 2010), and that areas such as LOC are thought to attribute volume to 2D images (Moore & Engel, 2001). Areas such as these could therefore also be active in our task.[1]

Under this second scenario, whereby the monocular-only cue set elicits activity in additional cortical areas as compared to the binocular-plus-disparity cue set, we would indeed conclude that the two cue sets established equivalent biases in Experiment 1 due to the overriding training influence of the ambiguous trials themselves (Harrison and Backus, 2010b), which constituted half of the trials in both the monocular-only and disparity-plus-occlusion condition. The conjecture that ambiguous cubes and cubes disambiguated by the monocular-only cue set elicit activity in more closely overlapping populations than do ambiguous and disparity-plus-occlusion cubes is further supported by our comparison of Experiments 1 and 2, demonstrating no difference in long-term bias established by monocular-only or by a mixture of ambiguous and monocular-only stimuli on Day 1.

In short, resolving the direction of rotation using monocular (pictorial) cues may require greater inferential resources from brain areas that interpret these cues, as compared to using disparity cues. These inferential processes may be the key to establishing a strong bias for perceptual outcome, as was observed in the ambiguous Necker cube. Additional work will be needed to determine whether, as a general rule, stimuli that are disambiguated by "inference-demanding" cues cause greater learning of perceptual bias than stimuli that are disambiguated by "easy-to-interpret" low-level cues.

6. Conclusion

We conclude that a combination of monocular cues is more effective than binocular disparity-plus-occlusion for establishing a long-term bias for perceived rotation direction at onset of a Necker cube stimulus. Additionally, the learned bias was not specific to the training cue set, hence is likely instantiated by, at a minimum, overlapping neural populations in the two cases. We speculate that inferential processes are critical in the formation of the perceptual bias.

Acknowledgment

---

[1] While arguments have been presented (Siegel & Allan, 1992) against the necessity of the existence, and adaptation, of neurons that are individually jointly tuned to two stimulus properties (here rotation direction and retinal location), contingent effects such as we see here are usually taken to be strong evidence in favor a neural mechanism that jointly encodes both properties (Barlow, 1990; Braddick, Campbell & Atkinson, 1978). Hence, the means by which the cortical areas described above could be involved in instantiating a rotation bias that is contingent on retinal location is unclear.

References

Backus, B. T. (2009). The mixture of Bernoulli experts: A theory to quantify reliance on cues in dichotomous perceptual decisions. *Journal of Vision, 9*(1), 1–19. 6.

Backus, B. T., & Haijiang, Q. (2007). Competition between newly recruited and pre-existing visual cues during the construction of visual appearance. *Vision Research, 47*(7), 919–924.

Barlow, H. (1990). A theory about the functional role and synaptic mechanism of visual aftereffects. In C. Blakemore (Ed.), *Vision: coding and efficiency*. Cambridge, UK: Cambridge University Press.

Bradley, D. C., Chang, G. C., & Andersen, R. A. (1998). Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature, 392*(6677), 714–717.

Braddick, O., Campbell, F., & Atkinson, J. (1978). Channels in vision: Basic aspects. In R. Held, H. W. Leibowitz, & H. L. Teuber (Eds.), *Handbook of sensory physiology, VIII* (pp. 3–38). Berline: Springer-Verlag.

Brouwer, G. J., & van Ee, R. (2007). Visual cortex allows prediction of perceptual states during ambiguous structure-from-motion. *Journal of Neuroscience, 27*(5), 1015–1023.

Brunswick, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley, CA: University of California Press.

DeAngelis, G. C., Cumming, B. G., & Newsome, W. T. (1998). Cortical area MT and the perception of stereoscopic depth. *Nature, 394*(6694), 677–680.

DeAngelis, G. C., & Uka, T. (2003). Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *Journal of Neurophysiology, 89*(2), 1094–1111.

Dodd, J. V., Krug, K., Cumming, B. G., & Parker, A. J. (2001). Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *Journal of Neuroscience, 21*(13), 4809–4821.

Dosher, B. A., Sperling, G., & Wurst, S. A. (1986). Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Research, 26*(6), 973–990.

Grunewald, A., Bradley, D. C., & Andersen, R. A. (2002). Neural correlates of structure-from-motion perception in macaque V1 and MT. *Journal of Neuroscience, 22*(14), 6195–6207.

Haijiang, Q., Saunders, J. A., Stone, R. W., & Backus, B. T. (2006). Demonstration of cue recruitment: Change in visual appearance by means of Pavlovian conditioning. *Proceedings of the National Academy of Sciences of the United States of America, 103*(2), 483–488.

Harrison, S. J., & Backus, B. T. (2010a). Disambiguating Necker cube rotation using a location cue: What types of spatial location signal can the visual system learn? *Journal of Vision, 10*(6), 1–15.

Harrison, S. J., & Backus, B. T. (2010b). Uninformative visual experience establishes long term perceptual bias. *Vision Research, 50*(18).

Helmholtz, H. V. (1910/1925). *Treatise on physiological optics* (Vol. 3). New York: Optical Society of America. 1925.

Ittelson, W. H. (1951). Size as a cue to distance: Static localization. *American Journal of Psychology, 64*(1), 54–67.

Kaneoke, Y., Urakawa, T., Hirai, M., Kakigi, R., & Murakami, I. (2009). Neural basis of stable perception of an ambiguous apparent motion stimulus. *Neuroscience, 159*(1), 150–160.

Knapen, T., Brascamp, J., Adams, W. J., & Graf, E. W. (2009). The spatial scale of perceptual memory in ambiguous figure perception. *Journal of Vision, 9*(13), 1–12.

Maunsell, J. H., & Van Essen, D. C. (1983a). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology, 49*(5), 1127–1147.

Maunsell, J. H., & Van Essen, D. C. (1983b). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology, 49*(5), 1148–1167.

Moore, C., & Engel, S. A. (2001). Neural response to perception of volume in the lateral occipital complex. *Neuron, 29*(1), 277–286.

Parker, A. J., & Krug, K. (2003). Neuronal mechanisms for the perception of ambiguous stimuli. *Current Opinion in Neurobiology, 13*(4), 433–439.

Pastukhov, A., Ludwig, K., & Braun, J. (2010). Ambiguous, high-contrast and unambiguous, low-contrast primes induce a common perceptual memory. *Perception, 39*, 152 (ECVP Abstract Supplement).

Pearson, J., Clifford, C. W., & Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Current Biology, 18*(13), 982–986.

Pizlo, Z., Li, Y., & Steinman, R. M. (2008). Binocular disparity only comes into play when everything else fails; a finding with broader implications than one might suppose. *Spatial Vision, 21*(6), 495–508.

Rouse, M. W., Tittle, J. S., & Braunstein, M. L. (1989). Stereoscopic depth perception by static stereo-deficient observers in dynamic displays with constant and changing disparity. *Optometry and Vision Science, 66*(6), 355–362.

Seurinck, R., de Lange, F. P., Achten, E., & Vingerhoets, G. (2010). Mental rotation meets the motion after effect: The role of hV5/MT+ in visual mental imagery. *Journal of Cognitive Neuroscience*, 1–10.

Siegel, S., & Allan, L.G. (1992). Pairings in learning and perception: Pavlovian conditioning and contingent aftereffects. In: D. Medin (Ed.), The psychology of learning and motivation (vol. 28, pp. 127–160).

Tartaglia, E. M., Bamert, L., Mast, F. W., & Herzog, M. H. (2009). Human perceptual learning by mental imagery. *Current Biology, 19*(24), 2081–2085.

van Dam, L. C. J., & Ernst, M. (2010). Preexposure disrupts learning of location-contingent perceptual biases for ambiguous stimuli. *Journal of Vision, 10*, 15.