



Published in final edited form as:

*Vision Res.* 2013 April 19; 82: 13–21. doi:10.1016/j.visres.2013.02.006.

## Generalization of Cue Recruitment to Non-moving Stimuli: Location and Surface-texture Contingent Biases for 3-D Shape Perception

Anshul Jain<sup>1</sup> and Benjamin T. Backus<sup>1</sup>

<sup>1</sup>Graduate Center for Vision Research, SUNY College of Optometry, 33 W. 42<sup>nd</sup> Street, New York, NY 10036, USA and SUNY Eye Institute, 550 Harrison Street, Suite 340, Syracuse, NY 1320

### Abstract

Long-lasting perceptual biases can be acquired through training in cue recruitment experiments (e.g. Backus, 2011, Haijiang, Saunders, Stone & Backus, 2006). Stimuli in previous studies contained motion, so the learning could be explained as an idiosyncrasy in some specific neuronal population such as the middle temporal (MT) area (Harrison & Backus, 2010a). The current study addresses the generality of cue recruitment by testing whether motion is necessary for learning a cue-contingent perceptual bias. We tested whether location and a novel cue, surface texture, would be recruited as cues to disambiguate perceptually bistable stationary 3-D shapes. In Experiment 1, stereo and luminance cues were used to disambiguate shape according to location in the visual field, and observers' (N=10) percepts on ambiguous test trials became biased in favor of the contingency during training. This bias lasted into the following day. This result together with previous studies that used moving stimuli suggests that location-contingent biases are easily learned by the visual system. In Experiment 2, location was fixed, and instead the new cue to be recruited was a surface texture. Learning did not occur when stimuli were para-foveal, texture was task-irrelevant, and disparity was continuously present in training stimuli (N=10). However, learning did occur when stimuli were central, task was texture-relevant, and disparity was transient (N=8). Thus, we show for the first time that an abstract cue, surface texture, can also be learned without motion.

### Keywords

Cue Recruitment; Perceptual Learning; Perceptual Bias; 3D Shape; Bistable Perception

### 1. Introduction

As an individual's environment changes over time, optimal perception would require that the individual's perceptual system adapt. Learned biases reflect this adaptation: they show what the system believes to be the most likely interpretation of the sense data (Brunswik, 1956, Helmholtz, 1910/1925). The learned biases can be described within the framework of Bayesian inference as a change in prior belief, with examples including changes in the light-

© 2013 Elsevier Ltd. All rights reserved.

Correspondence should be addressed to: Dr. Anshul Jain, Graduate Center for Vision Research, SUNY College of Optometry, 33W 42<sup>nd</sup> Street, New York City, NY 10036, USA, Phone: 212-938-5777, ajain@sunyo.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

from-above prior (Adams, Graf, & Ernst, 2004), the convexity prior (Champion & Adams, 2007), and the stationarity prior (Jain & Backus, 2010). Visual cues constrain perceptual interpretations, and changes in the way the visual system uses visual cues to construct perceptual appearance are an important form of adaptation. Learning a new use for a visual cue, so that it affects some attribute of appearance that it did not affect before, is called cue recruitment (Backus, 2011).

A series of cue-recruitment studies have shown that the rotation direction of a perceptually bistable 3D object can be made contingent on new signals such as its translation direction (Haijiang et al., 2006), location (Backus & Haijiang, 2007, Haijiang et al., 2006, Harrison & Backus, 2010a), or shape (Harrison & Backus, 2012). Location was also recruited as a cue upon which the stationarity prior became contingent (Jain & Backus, 2010). The apparent rotation direction of a cylinder can be made contingent on binocular vertical disparities (Di Luca, Ernst & Backus, 2010). In these studies, all stimuli contained motion. Thus, it could be argued that motion is a critical requirement for this form of learning. Since these stimuli engaged motion sensitive areas, such as MT (Born & Bradley, 2005, review) and MST (Saito, Yukiie, Tanaka, Hikosaka, Fukada & Iwai, 1986, Tanaka & Saito, 1989), it is therefore important to know whether cue recruitment is an idiosyncratic phenomenon within the motion perception system.

First, we examined whether motion signals are necessary for cue recruitment by measuring the strength of learned location-contingent bias using stimuli that did not contain motion (Experiment 1). Second, we tested whether a bias contingent on surface-texture can be acquired to affect appearance of a static 3D shape. The significance of an acquired texture-contingent bias is that, like motion, there could be something special about location that makes location particularly easy to learn (i.e. recruit) as a cue. This study is not the first to look at other cues besides location; other recruited cues include shape (Harrison & Backus, 2012, Sinha & Poggio, 1996), vertical disparity (Di Luca et al., 2010), translation direction (Haijiang et al., 2006), and motion within the display that is not part of the object itself (Backus, Jain & Fuller, 2011), but all of these studies used moving objects to measure the acquired cue-contingent perceptual bias.

## 2. General Methods

### 2.1. Subjects

Thirty-two observers participated in the study, twelve in Experiment 1, twelve in Experiment 2A and 8 in Experiment 2B. Data from four observers, two each in Experiment 1 and Experiment 2A were discarded because they could not perform the task reliably (i.e. their answers on Training trials did not agree with the visual cues that were intended to control appearance on those trials). All observers were naïve to the purpose of the experiments. The experiments were conducted in compliance with the standards set by the IRB at the Graduate Center for Vision Research, SUNY College of Optometry. Observers were paid for their participation. All observers had normal or corrected-to-normal vision and a stereo acuity better than 4 min of arc (TNO stereo-acuity test).

### 2.2. Apparatus

The experiments were implemented using the Python-based virtual reality software toolkit Vizard™ 3.11 (WorldViz LLC, Santa Barbara, CA, USA) on a Dell Precision T3400 computer running the Windows XP operating system. Stimuli were rear projected onto a 1.8 m × 2.4 m screen using a Christie Mirage S+4k projector. The display refresh rate was fixed at 120 Hz and the screen resolution was set at 1024 × 768 pixels. Observers were seated at a distance of 1.5 m from the screen and wore red-green anaglyphs to view the stimuli.

### 3. Experiment 1 – Location Contingent Bias

#### 3.1. Stimuli

The stimuli consisted of a dihedral right angle constructed by joining two squares along one of their edges to mimic the outline of an “open book”. The edges of the squares were struts (rectangular parallelepipeds). Each face of the ‘book’ contained 20 randomly placed dots to stabilize the percept of a rigid 3-D object. Each edge of the square was 15 cm in length before projection and the stimuli were viewed from a distance of 1.5 m thus subtending a visual angle of 5.7 degrees when perpendicular to the line of sight. The stimuli could be perceived in one of two configurations, an open book facing towards an observer or an open book facing away from the observer (Figure 1). On any given trial, stimuli were presented vertically centered 5.7 degrees above or below a central fixation square.

#### 3.2. Procedure

The experiment consisted of two types of trials: Training trials and Test trials. On training trials, the observer’s percept was controlled using three cues that specified depth relations: binocular disparity, proximity-luminance covariance (Doshier, Sperling & Wurst, 1986), and occlusion (occlusion bar as well as internal occlusions). Importantly, we put stimulus configuration (as determined by depth cues) into correlation with location. Thus, on training trials, observers were presented with the “facing away” configuration above fixation and the “facing towards” configuration below fixation, or vice versa (counter-balanced across observers). Observers pressed a key to initiate a trial. After this key-press, the stimulus was displayed for 1.15 s. After the first 0.5 s, one of the two faces within the dihedral angle stimulus (randomly chosen) was highlighted (edge thickness was increased by a factor of 1.5) for 0.25 s and then the stimulus returned to its previous state for the rest of the trial. The observer’s task was to report whether the highlighted face appeared closer or farther away than the other face, which uniquely determined the perceived configuration. Because the face to be highlighted was chosen randomly, observers’ responses were uncorrelated with both the perceived configuration and the stimulus location. The task was chosen to discourage observers from using cognitive strategies to make their response rather than rely on their percept. Observers were instructed to report the trial as a “missed trial” by pressing a third key if they were unsure how to respond, because either they failed to notice the probe or the stimulus or both, and were told that by being attentive they could minimize their fraction of missed trials. Observers did not receive any feedback. Figure 2 shows typical trial sequences for two training trials and a test trial. The fixation-cross disappeared after the observer responded and it appeared again after 1 s indicating that the observer could initiate the next trial.

Perceptually bistable stimuli such as the one used in this experiment are known to switch perceptual states spontaneously (Attneave, 1971, Blake & Logothetis, 2002). Further, transients like the probe used in this experiment have shown to cause a perceptual switch (Kanai, Moradi, Shimojo & Verstraten, 2005). To minimize this effect, which would have reduced the apparent magnitude of learning, observers were instructed to respond based on the percept at stimulus onset in case their percept switched during the trial or use the “missed trial” key if they were not sure. In post-experiment interviews observers were asked about this issue explicitly. They universally reported that there were very few instances of switching and that they felt they were able to follow the instructions to respond according to their percept at stimulus onset. Thus we do not believe the learning effects were actually any larger than revealed by the data.

Each subject collected data over two sessions on two consecutive days. Each session began with 80 training trials followed by a pseudorandom mixture of 200 training and 200 test

trials presented in five blocks of 80 trials each. For the training trials on Day 2, the correlation between configuration, as specified by the disambiguating cues, and location, was reversed as compared to Day 1 for each observer. This gave us a second measure of the strength of learning that occurred on Day 1, in terms of resistance to retraining on Day 2 (Haijiang et al., 2006).

### 3.3. Results

For statistical analyses, observers' proportions (i.e. proportion of trials seen as "open-away" or "as trained on Day 1") were transformed to a z-score (Backus, 2009, Doshier et al., 1986). Results were converted back to proportions for plotting. The proportion of each percept on ambiguous test trials was computed based on the expected response as predicted by the location contingency during training. Saturated proportions (100% and 0%) were assigned a z-score of  $\pm 2.326$ , corresponding to a consistency of 99% or 1%.

Figures 3A and 3B show mean proportions for observers on training and test trials as a function of number of training trials on Day 1 and Day 2, respectively. Individuals' whole-session proportions are plotted as Z-scores in figure 3C. Binocular disparity, occlusion and proximity-luminance covariance were effective at controlling observers' percepts on training trials and observers reported 98% of training trials on average to be seen as specified by the depth cues ( $t(9) = 18.5$ ,  $p \ll 0.001$ , mean = 98.1 %, s.e.m. = 0.7 %). Critically, observers' percepts on ambiguous test trials were biased in a manner consistent with the location-configuration contingency on Day 1 ( $t(9) = 4.75$ ,  $p < 0.01$ , mean accuracy = 81 %, s.e.m = 6 %), showing that the visual system learned location-dependent biases to perceive 3D shapes in the absence of motion signals. This bias was acquired fairly quickly and did not build up further over the duration of the session, as shown by the fact that observers' proportion-seen-as-trained did not vary between blocks of trials for either training (repeated measures ANOVA:  $F(5,45) = 2.01$ ,  $p = 0.10$ ) or test trials ( $F(4,36) = 0.55$ ,  $p = 0.70$ ). It is quite common for effects to be visible immediately and grow little during the session in a cue recruitment experiments. A possible explanation is a strong trial-to-trial priming effect that decreases over time as long term contingent learning increases (Jain, Fuller & Backus, 2010).

A second measure of the strength of the bias learned on Day 1 was the resistance to re-learning of the reverse location-configuration contingency on the following day. On Day 2's training trials, the location-configuration contingency was opposite to that on Day 1. Again, we were able to control observers' percepts on the training trials using the disambiguating cues. Observers reported 97% of training trials to have the configuration specified by the cues ( $t(9) = 15.2$ ,  $p \ll 0.001$ , mean accuracy = 96.7 %, s.e.m. = 1.7 %). Critically, observers did not perceive ambiguous test trials as predicted by the location-configuration contingency on Day 2 ( $t(9) = 0.89$ ,  $p = 0.4$ , mean accuracy = 60.84 %, s.e.m. = 7.26 %). In fact, observers retained a bias to perceive the ambiguous trials as predicted by the training on Day 1. One of the 10 observers did effectively unlearn the Day 1 bias on Day 2, but this observer was an exception. The mean proportion seen, without the outlier, was significantly better than chance in favor of the bias learned on Day 1 ( $t(8) = 3.8$ ,  $p < 0.01$ , mean accuracy = 67.42 %, s.e.m. = 3.4 %). Observers' proportions on the training trials (repeated measures ANOVA:  $F(5,45) = 0.42$ ,  $p = 0.83$ ) as well as their learned bias ( $F(4,36) = 0.70$ ,  $p = 0.60$ ) did not vary significantly between blocks of trials on Day 2. To summarize, the results from Experiment 1 show that the human visual system can learn location-contingent biases in the absence of motion cues, and that similar to the location-contingent bias for stimuli with motion cues, the learned bias is strong and lasts at least until the following day as indicated by the resistance to learn the reverse contingency presented on Day 2. We also measured the proportion of trials where subjects responded that the highlighted surface was closer. Only

one out of 10 observers showed a significant overall bias ( $t(1) = 16.2, p = 0.04$ ) to perceive highlighted surface as closer on ambiguous test trials over two sessions.

## 4. Experiment 2 – Feature Contingent Bias

### 4.1. Stimuli

Stimuli consisted of a Necker cube that had two non-adjacent (opposite) faces textured with horizontal and vertical square wave gratings, respectively, of unequal thickness (Figure 4). The cube was presented such that the two textured faces were approximately frontal with one being closer to the observer than the other. In Experiment 2A, the cube edges were 25 cm in length thus subtended a visual angle of 9.4 degrees when perpendicular to the line of sight (Figure 4A). Each edge of the cube in Experiment 2A was a rectangular parallelepiped of thickness 10 mm (0.38 deg). A fixation square 0.76 degrees in size was presented at the center of the screen to aid fixation and the cube was presented 9.4 degrees above fixation. The thick and thin lines used in the surface textures were 5 mm (0.19 deg) and 2.5 mm (0.09 deg) wide, respectively. In Experiment 2B, the frame of the object was a rectangular parallelepiped with 8 short edges of 25 cm each in length and 4 long edges of 37.5 cm. The two textured surfaces were near and far faces, connected by long edges (Figure 4B). Each edge of the cube in Experiment 2B was a parallelepiped of thickness 15 mm (0.57 deg). The stimulus was presented at center of the screen. The thick and thin lines used in the surface textures were 6 mm (0.22 deg) and 3 mm (0.11 deg) wide, respectively.

In both experiments, the cuboid was orthographically rendered in two orientations with a yaw, pitch and roll of (25, 12.5, 0) and (-25, -12.5, 0) degrees, respectively. For each of these orientations, the cube could be perceived in one of two configurations: as seen from above or as seen from below. Similar to Experiment 1, Experiment 2 consisted of two kinds of trials – disambiguated training trials and ambiguous test trials. On training trials, the cube was disambiguated using two cues: disparity and proximity-luminance covariance. On training trials, one of the textures--always the same one for a given observer during a given session-- was made to appear closer than the other using the disambiguating cues. In Experiment 2A, both disambiguating cues persisted for the entire duration of the training trials, while in Experiment 2B the disparity signal was turned off after 120 ms by turning off the stimulus to the left eye. This was done to engage perceptual maintenance mechanisms and we predicted that this would increase the learning of the new cue for 3D shape (texture orientation and thickness). Cue recruitment in previous experiments was stronger when disambiguating cues were weaker (Backus et al., 2011, Harrison & Backus, 2010b, Harrison, Backus & Jain, 2011).

### 4.2. Procedure

Experiment 2 was conducted during a single session for each observer. Similar to Experiment 1, the session began with 80 training trials followed by a pseudorandom mixture of 200 training and 200 test trials presented in five blocks of 80 trials each. The first block consisted of only training trials followed by a 1:1 mixture of test and training trials for the rest of the blocks. In Experiment 2A, the observer initiated a trial by pressing a key. The cube was presented for 0.5 s, after which one of the two textured surfaces was highlighted for 0.25 s. The cube stayed on for another 0.25 s before disappearing. Observers were instructed to respond after the cube disappeared and their task was to report whether the highlighted face was closer or farther away from them as compared to the other face. Again, this task ensured that observers' responses were uncorrelated with both the perceived configuration and the stimulus orientation, thus reducing the chance that observers would "figure out" the experimental contingency and respond according to this rule instead of by following the instructions to use their percept. The fixation-cross disappeared after the



observer responded and appeared again after 1 s to indicate that the observer could initiate the next trial. In Experiment 2B, the stimulus was presented for 1s on each trial and the inter-trial interval lasted for 1s after the observer's response. Further, instead of highlighting one face by thickening its lines, we introduced a diagonal break in one of the textures and asked observers to report whether the face with the streak in the texture was closer or farther away from them as compared to the other face. This task also dissociated responses from perceived configuration and additionally forced observers to attend to the texture. We hypothesized that attending to the texture itself may cause the visual system to learn to use texture features as a cue to 3D shape. Earlier work has shown that in a similar setup location-contingent learning of rotation direction was stronger when observers attended to the location of the cube (Backus & Fuller, 2010). Observers did not receive any feedback.

### 4.3. Results

Figure 5A shows mean percent perceived as trained for the observers on training and test trials as a function of the number of elapsed training trials for Experiment 2A. These data are re-plotted separately for each observer, for the whole session, as Z-scores in Figure 5B. The disambiguating cues were effective at determining observers' percept on training trials ( $t(9) = 21.05$ ,  $p \ll 0.001$ , mean accuracy = 97.55 %, s.e.m = 0.58 %). However, in Experiment 2A observers did not recruit texture as a cue to 3D shape ( $t(9) = 0.47$ ,  $p = 0.65$ , mean accuracy = 50.88 %, s.e.m. = 1.88 %). Observers showed a small bias to perceive the thicker gratings as closer ( $t(9) = 4.16$ ,  $p < 0.01$ , mean accuracy = 54.63 %, s.e.m. = 1.1 %), consistent with previous findings that low spatial frequencies are perceived as closer than high spatial frequencies (Gibson, 1950).

Figure 6A shows mean percent perceived as trained for the observers on training and test trials as a function of training trials for Experiment 2B. These data are summarized as Z-scores in Figure 6B. Again, the disambiguating cues were effective at determining observers' percepts on training trials ( $t(7) = 20.22$ ,  $p \ll 0.001$ , mean accuracy = 98.33 %, s.e.m = 0.49 %). The effect of training can be seen in either of two ways. First, Figure 6B shows that on average, training was effective: the average subject was biased in the direction of training, and because training was counterbalanced between two groups of subjects, this result cannot be due to overall bias to see thick gratings as closer. Looked at this way, the effect of training was significant: observers' percepts on ambiguous test trials were now biased in favor of the texture-configuration correlation used on the trainings trials ( $t(7) = 4.05$ ,  $p < 0.01$ , mean = 57.31 %, s.e.m. = 1.79 %). This learned bias was stronger ( $t(3) = 3.5$ ,  $p < 0.05$ ) when the surface with thicker gratings was configured as the closer surface ( $t(3) = 7.42$ ,  $p < 0.01$ , mean accuracy = 61.5 %, s.e.m. = 1.51 %) as compared to when the surface with thicker gratings was configured as the farther surface ( $t(3) = 3.23$ ,  $p < 0.05$ , mean = 53.13%, s.e.m. = 0.97 %).

However, it is probable that the average subject had a small bias to see thick gratings as closer. The data can be modeled as the combination of two effects: an overall bias in that direction, and an effect of learning that was opposite in direction for the two subgroups (Figure 6C and 6D). Probit analysis with terms for overall bias and an effect of learning (Backus, 2009) show an overall bias (in z-score units) of 0.11 (99 % CI: [0.08, 0.13], computed by resampling from the set of 8 subjects) and an effect of learning of 0.19 (99 % CI: [0.15, 0.22]). These effects are significantly different from 0 (chance) at  $p \ll 0.001$  level (assuming normally distributed data). Thus, while observers did show an overall bias to perceive thicker gratings as closer, there can be no question whether the training was effective. In Experiment 2B, a texture-contingent bias in favor of the training contingency was learned. Thus, the visual system came to use (i.e. recruited) the texture cue for a new purpose that the texture cue did not have before: texture came to (weakly) specify which face of the object was closer. A similar analysis for Experiment 2A revealed an overall bias

of 0.12 (99 % CI: [0.08, 0.16]) but not a significant effect of learning (0.02; 99% CI: [-0.02, 0.06]).

## 5. Discussion

The current study was designed to examine whether motion is critical for the learning of contingent biases, i.e. cue-recruitment, since previous demonstrations used stimuli containing motion. In Experiment 1, we used stationary stimuli and observed strong learned biases contingent on stimulus location. Motion was not a critical signal for acquiring this contingent bias. In Experiment 2, we showed that the visual system can learn biases for perceiving 3D shapes contingent on surface texture, and the learning was weaker than in Experiment 1.

Bistable stimuli have proved extremely useful in studying the cue-recruitment phenomenon due to the dichotomous nature of possible percepts, which makes them easy for the observer to report. In Experiment 1, we found that a strong location contingent bias, that lasts at least until the following day, can be learned for nonmoving stimuli just as they can for moving stimuli (Haijiang et al., 2006, Harrison & Backus, 2010a, Harrison & Backus, 2010b, Jain & Backus, 2010). These cue recruitment experiments used briefly presented stimuli and previous studies have shown that a similar perceptual stabilization can occur when ambiguous bistable stimuli are presented intermittently without explicit training and has been linked with perceptual memory (Brascamp, Knapen, Kanai, Noest, van Ee & van den Berg, 2008, Brascamp, Pearson, Blake & van den Berg, 2009, Carter & Cavanagh, 2007, Maier, Wilke, Logothetis & Leopold, 2003). The perceptual alternations in these studies have been explained based on separate perceptual biases at stimulus onset and during continuous viewing, with these two processes operating on different time scales (Brascamp et al., 2008, Pastukhov & Braun, 2008). The learning observed in this as well as previous cue-recruitment studies examining location contingent bias exhibit a time-scale much longer than those examined in the bistable perception literature, lasting at least until the next day and resistant to reverse training. However, it is possible that this perceptual stabilization plays a critical role in acquisition of the strong location contingent biases observed cue-recruitment studies (Jain et al., 2010).

Previous cue recruitment experiments with moving stimuli found larger effects (approximately 1.25 times when expressed in terms of z-scores) than we found in Experiment 1. This difference cannot be attributed to a bias to perceive the stimuli in one configuration over the other: observers' bias to perceive the stimuli as an open book facing towards them was not different from the bias to see the book opening away. Alternatively, Test stimuli might have been inherently weak in supporting unambiguous percepts, thereby reducing any possible effect. This alternative seems unlikely, however, as close to 100% of training trials were seen as specified by cues, and observers did not indicate being unsure of appearance on Test trials: very few trials were reported as "missed" (1, 2, and 18 trials on only 3 out of 20 sessions, respectively). The stimuli used in the current experiment differed from those used in previous experiments examining location-contingent bias in more ways than just removal of motion signals. Some of the key differences were stimulus size, distance between training locations, and the observers' task. Some or all of these factors could have affected the strength of learning. Of all these factors, the distance between training locations is perhaps least likely to have reduced the strength of learning, given the spread of spatial influence for bistable stimuli fell to chance at much smaller distances (Knapen, Brascamp, Adams & Graf, 2009) than the ones used in this study (2 deg vs 11.4 deg). Harrison and Backus (2012) found no evidence that size or rotation speed affected the strength of learning of object shape as cue to rotation direction, suggesting that smaller stimulus size in the current experiment is not very likely to have strongly affected the

learning. Further, previous findings with moving stimuli have proven to be highly robust across diverse conditions and experimental designs, so it seems likely that absence of motion was a key factor in reducing the magnitude of the effect.

Why would a location-contingent bias be learned more strongly for moving stimuli? One argument is based on the observation that associative learning can be implemented by change in the response of neurons that jointly encode multiple properties (Barlow, 1990, Braddick, Campbell & Atkinson, 1978, Harrison & Backus, 2012). For example, the location-contingent bias for rotation direction of a Necker cube (Haijiang et al., 2006) could be attributable to long-term potentiation of neurons in area MT (Harrison & Backus, in press), because neurons there jointly encode disparity and motion direction (DeAngelis, Cumming & Newsome, 1998). Thus, it may be that neurons in MT that encode both depth (for example crossed disparity) and motion (for example rightward) are easily trained to respond more strongly to (rightward) motion, resulting in greater perceived (near) depth for (rightward) moving stimuli. Neurons in area MT can be trained to respond to new visual cues (Schlack & Albright, 2007), so it would not be surprising to see changes in relative responsiveness to motion signals. It is unclear whether neurons in the superior lateral occipital (SLO) region or the inferotemporal (IT) region that encode the shapes of static objects (Grill-Spector, Kushnir, Hendler, Edelman, Itzhak & Malach, 1998, Kanwisher, Woods, Jacoboni & Mazziotta, 1997, Murray, Olshausen & Woods, 2003), contain any such population of neurons that could so easily bias apparent shape simply by responding more strongly to a low-level feature of the stimulus (such as MT neurons can do for motion).

Learning outcomes differed between Experiments 2A and 2B. In Experiment 2A we observed no learning, but in Experiment 2B, observers' percept of an ambiguous Necker cube became weakly contingent on the surface-texture. There were three differences between the methods of Experiment 2A and 2B. First, the disambiguating binocular cues were presented for the entire duration of the training trial in Experiment 2A as opposed to only for the first 120 ms in Experiment 2B. Second, the task demands did not require the observers to process the surface texture in Experiment 2A whereas it was necessary to do so in Experiment 2B. Third, stimuli were presented 9.4 degrees above the fixation mark in Experiment 2A, while they were presented at the fixation during Experiment 2B.

Some or all of these factors may have caused learning in Experiment 2B. In previous studies, training with low-information stimuli, such as with monocular cues only or short pulses of binocular disparity, resulted in stronger learned bias (Backus et al., 2011, Di Luca et al., 2010, Harrison et al., 2011, van Dam & Ernst, 2010), perhaps because resolving the ambiguity in stimuli that are difficult to resolve facilitates the learning of new, correlated cues that can help with disambiguation in the future (Harrison et al., 2011). The second and third differences, namely, the texture-dependent task and foveal presentation, respectively, could have caused differences in attention, because the task required processing the surface texture, in the central visual field. Attention is known to affect perceptual learning. For example, Ahissar and Hochstein (1993) found no improvement in task performance for task-irrelevant features while performance on task-relevant features improved within the same stimuli. Top-down influences, presumably mediated by attention, have also been reported in the form of task-dependent changes in neuronal response in a perceptual learning task (Crist, Li & Gilbert, 2001, Li, Piech & Gilbert, 2004). Mukai et al. (2011) found stronger learning at an attended location than at an unattended location. Further, imaging studies have shown that learning was associated with stronger initial activation of the attentional network (Mukai, Kim, Fukunaga, Japee, Marrett & Ungerleider, 2007). In an associative learning paradigm, similar to the current experimental design, researchers found that the learning of location-contingent bias for rotation direction of a Necker cube at the attended location was three times as strong as bias learned at the unattended location (Backus & Fuller, 2010).



Attention to texture in Experiment 2B could have been further enhanced by presenting stimuli in central vision, the habitual locus of attention within the visual field. Therefore, it seems likely that the learning observed in Experiment 2B occurred in part because observers attended to and processed the surface texture more thoroughly in Experiment 2B.

In Experiment 2B, we found a stronger texture-contingent bias in the group that was trained with thicker gratings as close than the group that was trained with thinner gratings as close. There are two possible explanations for this finding. Observers might have started with no bias to perceive thicker gratings as closer. In this case, the learning was stronger for the group that was trained with thicker gratings as closer compared to the group trained with thinner gratings as closer. Alternatively, observers could have had a weak preexisting bias to perceive thicker gratings as closer, and this bias then was enhanced for the group that was trained with thicker gratings as closer and weakened (to the point of reversal, in fact) for the group trained with thinner gratings as closer. Our experimental results cannot distinguish between these two possibilities, however, given the known relationship between spatial frequency and depth order (Gibson, 1950), the latter seems the more likely scenario. This finding is similar in nature to another cue-recruitment study where Jain and Backus (2010) found stronger learning in the group that was trained with a contingency that was consistent with a possibly pre-existing bias than in the group that was trained with a contingency that was opposite to the pre-existing bias.

Two recent studies examined related phenomena. Ernst (2007) found that two arbitrarily chosen signals, one tactile and one visual, were combined by observers during a discrimination task after the signals were put into correlation with one another during training. Discrimination performance was particularly disrupted for stimuli that contained “anti-correlated” signals (i.e. opposite covariation as compared to training). Ernst’s study showed that an associative learning paradigm caused a change in the perceptual system’s implicit belief, or Bayesian prior, for the signals’ contingency (joint probability). Our study differs structurally from that of Ernst, where training was conducted using a discrimination task at threshold with feedback. Because observers operated at threshold, appearance was by design only partially reliable as an indicator for the stimulus differences that observers were instructed to detect. Unlike our study, it is not clear in Ernst (2007) how the training affected the appearance of the stimulus, or if it did, whether tactile appearance, visual appearance, or both were affected.

Second, it is not only direct cues (as we used here), but also “ancillary” cues, that might in principle be recruited (Backus, 2011). In fact, cue recruitment has already been demonstrated for non-moving stimuli in the case of an ancillary cue. An ancillary cue is a signal that specifies the reliabilities of other cue(s) for some property of a scene, but does not itself directly depend on the property (Landy, Maloney, Johnston & Young, 1995). For example, binocular relative disparity becomes less reliable relative to perspective cues as viewing distance increases, so vergence eye posture is an ancillary cue that specifies the reliability of disparity as an indicator of surface slant (Backus & Banks, 1999). Ancillary cue recruitment occurred in a study by Seydell, Knill, and Trommershäuser (2010). In that study, observers judged the slants of stereoscopically presented diamonds and ellipses. The aspect ratio of the figure within the image, dubbed the “figural compression” cue, provided a second slant cue that was manipulated separately from the stereoscopic cue. When the figures’ aspect ratios (in the simulated world) had high variation, figural compression had lower statistical reliability and figural compression was appropriately given relatively less weight by observers during slant judgment. Critically, observers were able to recruit the shape category of the figure, diamond or ellipse, as an ancillary cue that specified the statistical reliability of the figural compression cue. Thus, when judgments of diamonds and ellipses were intermixed, and diamonds but not ellipses had highly variable aspect ratios,

observers relied less on figural compression when judging the slants of the diamonds, and vice versa. Of course, the fact that an ancillary cue can be learned does not strongly predict whether direct cues can be learned, as demonstrated in the current study. Learning to use a signal as a cue that specifies when another signal is reliable is qualitatively different from learning that a signal is itself a useful cue that varies with, and therefore indicates, the state of an environmental property.

## 6. Conclusion

Strong location contingent biases that affect apparent 3D shape can be learned in the absence of motion signals. We were able to find conditions -- monocular disambiguation and attention to relevant features -- under which the visual system learned to use an abstract property, surface texture, to interpret 3D configuration. These findings extend the generality of cue recruitment as a form of perceptual learning. Training stimuli need not contain motion for cue-contingent biases to be acquired, nor do test stimuli need to contain motion for a learned cue-contingent bias to affect appearance.

## Acknowledgments

This work was supported by NIH grant R01-EY 013988. We thank B. Hunter McFadden and Martha Lain for assistance in running the experiment.

## References

- Ahissar M, Hochstein S. Attentional control of early perceptual learning. *Proceedings of the National Academy of Sciences of the United States of America*. 1993; 90(12):5718–5722. [PubMed: 8516322]
- Attneave F. Multistability in perception. *Scientific American*. 1971; 225(6):63–71. [PubMed: 5116412]
- Backus BT. The Mixture of Bernoulli Experts: a theory to quantify reliance on cues in dichotomous perceptual decisions. *J Vis*. 2009; 9(1):6, 1–19.
- Backus, BT. Recruitment of new visual cues for perceptual appearance. In: Trommershauser, J.; Körding, K.; Landy, MS., editors. *Sensory Cue Integration*. Oxford, UK: Oxford University Press; 2011.
- Backus BT, Banks MS. Estimator reliability and distance scaling in stereoscopic slant perception. *Perception*. 1999; 28(2):217–242. [PubMed: 10615462]
- Backus BT, Fuller S. Attention mediates learned perceptual bias for bistable stimuli. *Journal of Vision*. 2010; 10(7):1106.
- Backus BT, Haijiang Q. Competition between newly recruited and pre-existing visual cues during the construction of visual appearance. *Vision Research*. 2007; 47(7):919–924. [PubMed: 17303207]
- Backus BT, Jain A, Fuller S. Cue recruitment for extrinsic signals after training with low-information stimuli. *Journal of Vision*. 2011; 11(11):983.
- Barlow, HB. A theory about the functional role and synaptic mechanism of visual after-effects. In: Blakemore, CB., editor. *Vision: coding and efficiency*. Cambridge: Cambridge University Press; 1990.
- Blake R, Logothetis NK. Visual competition. *Nat Rev Neurosci*. 2002; 3(1):13–21. [PubMed: 11823801]
- Born RT, Bradley DC. Structure and function of visual area MT. *Annu Rev Neurosci*. 2005; 28:157–189. [PubMed: 16022593]
- Braddick, O.; Campbell, F.; Atkinson, J. Channels in Vision: Basic Aspects. In: Held, R.; Leibowitz, HW.; Teuber, HL., editors. *Handbook of Sensory Physiology*. Vol. VIII. 1978. p. 3-38.
- Brascamp JW, Knapen TH, Kanai R, Noest AJ, van Ee R, van den Berg AV. Multi-timescale perceptual history resolves visual ambiguity. *PLoS One*. 2008; 3(1):e1497. [PubMed: 18231584]

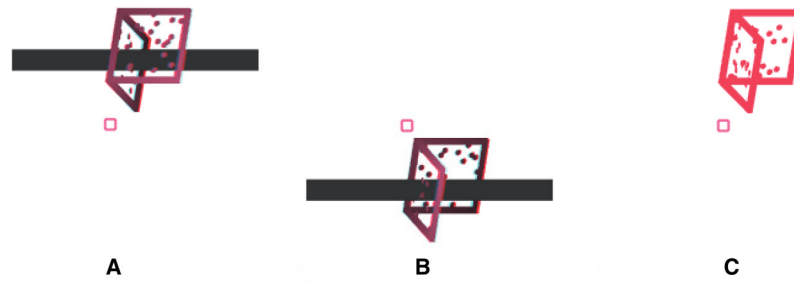
- Brascamp JW, Pearson J, Blake R, van den Berg AV. Intermittent ambiguous stimuli: Implicit memory causes periodic perceptual alternations. *Journal of Vision*. 2009; 9(3):1–23.
- Brunswik, E. Perception and the representative design of psychological experiments. Berkeley, CA: University of California Press; 1956. p. 92p. 123-131.
- Carter O, Cavanagh P. Onset rivalry: brief presentation isolates an early independent phase of perceptual competition. *PLoS One*. 2007; 2(4):e343. [PubMed: 17406667]
- Champion RA, Adams WJ. Modification of the convexity prior but not the light-from-above prior in visual search with shaded objects. *J Vis*. 2007; 7(13):10, 11–10. [PubMed: 17997638]
- Crist RE, Li W, Gilbert CD. Learning to see: experience and attention in primary visual cortex. *Nature neuroscience*. 2001; 4(5):519–525.
- DeAngelis G, Cumming B, Newsome W. Cortical area MT and the perception of stereoscopic depth. *Nature*. 1998; 394:677–680. [PubMed: 9716130]
- Di Luca M, Ernst MO, Backus BT. Learning to use an invisible visual signal for perception. *Current Biology*. 2010; 20(20):1860–1863. [PubMed: 20933421]
- Doshier BA, Sperling G, Wurst SA. Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Res*. 1986; 26(6):973–990. [PubMed: 3750879]
- Ernst MO. Learning to integrate arbitrary signals from vision and touch. *J Vis*. 2007; 7(5):7, 1–14. [PubMed: 18217847]
- Gibson, JJ. The perception of the visual world. Boston: Houghton Mifflin; 1950.
- Grill-Spector K, Kushnir T, Hendler T, Edelman S, Itzhak Y, Malach R. A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human brain mapping*. 1998; 6(4):316–328. [PubMed: 9704268]
- Haijiang Q, Saunders JA, Stone RW, Backus BT. Demonstration of cue recruitment: Change in visual appearance by means of Pavlovian conditioning. *Proc Natl Acad Sci U S A*. 2006; 103:483–488. [PubMed: 16387858]
- Harrison SJ, Backus BT. Disambiguating Necker cube rotation using a location cue: What types of spatial location signal can the visual system learn? *Journal of Vision*. 2010a; 10(6:23):1–15.
- Harrison SJ, Backus BT. Uninformative visual experience establishes long term perceptual bias. *Vision Res*. 2010b; 50(18):1905–1911. [PubMed: 20600232]
- Harrison SJ, Backus BT. Associative learning of shape as a cue to appearance: A new demonstration of cue recruitment. *Journal of Vision*. 2012; 12(3)
- Harrison SJ, Backus BT, Jain A. Disambiguation of Necker cube rotation by monocular and binocular depth cues: Relative effectiveness for establishing long-term bias. *Vision Research*. 2011; 51(9): 978–986. [PubMed: 21335023]
- Helmholtz, Hv. Treatise on Physiological Optics. Vol. III. New York: Optical Society of America; 1910/1925.
- Jain A, Backus BT. Experience affects the use of ego-motion signals during 3D shape perception. *J Vis*. 2010; 10(14)
- Jain A, Fuller S, Backus BT. Absence of cue-recruitment for extrinsic signals: sounds, spots, and swirling dots fail to influence perceived 3D rotation direction after training. *PLoS One*. 2010; 5(10):e13295. [PubMed: 20949047]
- Kanai R, Moradi F, Shimojo S, Verstraten FA. Perceptual alternation induced by visual transients. *Perception*. 2005; 34(7):803–822. [PubMed: 16124267]
- Kanwisher N, Woods RP, Jacoboni M, Mazziotta JC. A locus in human extrastriate cortex for visual shape analysis. *Journal of Cognitive Neuroscience*. 1997; 9:133–275. e86a774f-c71e-2344-c8d3-e248860fa620.
- Knapen T, Brascamp J, Adams WJ, Graf EW. The spatial scale of perceptual memory in ambiguous figure perception. *J Vis*. 2009; 9(13):16, 11–12.
- Landy MS, Maloney LT, Johnston EB, Young M. Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research*. 1995; 35(3):389–412. [PubMed: 7892735]
- Li W, Piech V, Gilbert CD. Perceptual learning and top-down influences in primary visual cortex. *Nat Neurosci*. 2004; 7(6):651–657. [PubMed: 15156149]

- Maier A, Wilke M, Logothetis NK, Leopold DA. Perception of temporally interleaved ambiguous patterns. *Current biology : CB*. 2003; 13(13):1076–1085. [PubMed: 12842006]
- Mukai I, Bahadur K, Kesavabhotla K, Ungerleider LG. Exogenous and endogenous attention during perceptual learning differentially affect post-training target thresholds. *Journal of Vision*. 2011; 11(1)
- Mukai I, Kim D, Fukunaga M, Japee S, Marrett S, Ungerleider LG. Activations in visual and attention-related areas predict and correlate with the degree of perceptual learning. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2007; 27(42):11401–11411. [PubMed: 17942734]
- Murray SO, Olshausen BA, Woods DL. Processing shape, motion and three-dimensional shape-from-motion in the human cortex. *Cerebral cortex*. 2003; 13(5):508–516. [PubMed: 12679297]
- Pastukhov A, Braun J. A short-term memory of multi-stable perception. *J Vis*. 2008; 8(13):7, 1–14. [PubMed: 19146337]
- Saito H, Yukie M, Tanaka K, Hikosaka K, Fukada Y, Iwai E. Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *J Neurosci*. 1986; 6(1):145–157. [PubMed: 3944616]
- Schlack A, Albright TD. Remembering Visual Motion: Neural Correlates of Associative Plasticity and Motion Recall in Cortical Area MT. *Neuron*. 2007; 53:881–890. [PubMed: 17359922]
- Seydell A, Knill DC, Trommershauser J. Adapting internal statistical models for interpreting visual cues to depth. *Journal of Vision*. 2010; 10(4):1, 1–27. [PubMed: 20465321]
- Sinha P, Poggio T. Role of learning in three-dimensional form perception. *Nature*. 1996; 384(6608):460–463. [PubMed: 8945472]
- Tanaka K, Saito H. Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *J Neurophysiol*. 1989; 62(3):626–641. [PubMed: 2769351]
- van Dam LC, Ernst MO. Preexposure disrupts learning of location-contingent perceptual biases for ambiguous stimuli. *Journal of Vision*. 2010; 10(8):15. [PubMed: 20884590]

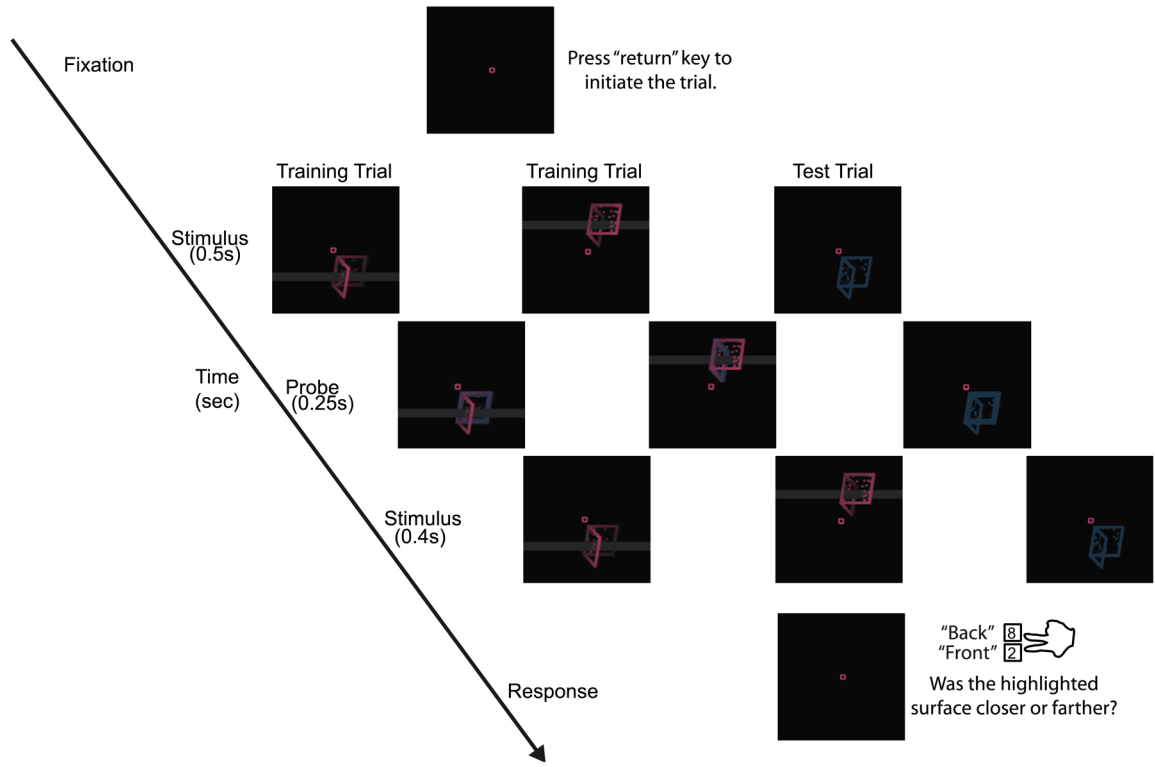
### Highlights

- Earlier studies of cue recruitment used motion, so findings could be idiosyncratic.
- Location contingent bias for 3D shape was acquired in the absence of motion.
- Surface-texture was also recruited as a cue that disambiguated 3D shape.

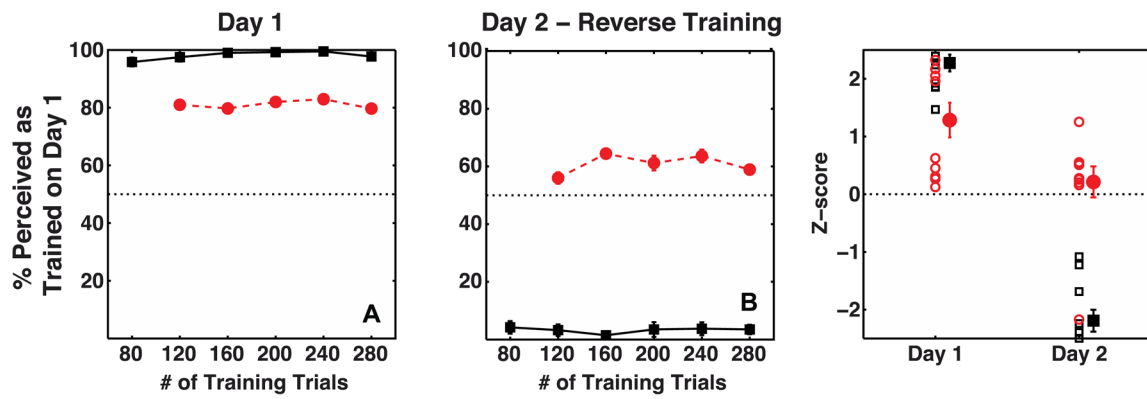




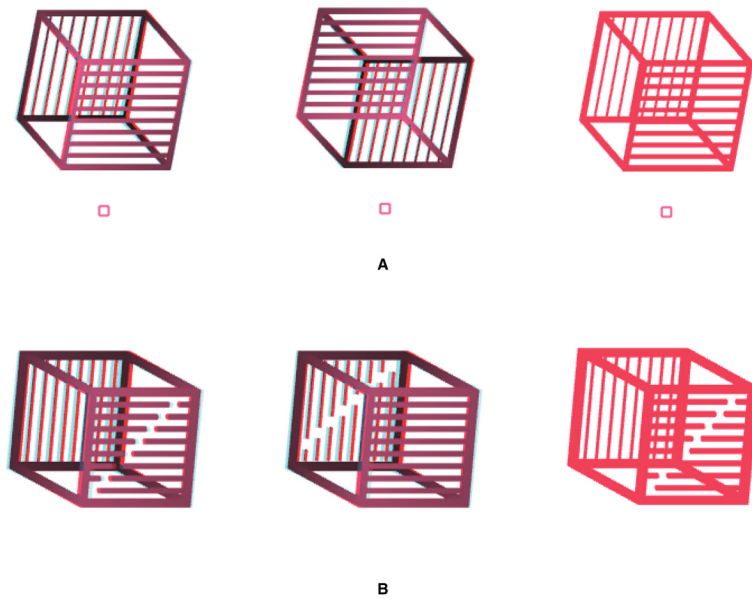
**Figure 1.** Stimuli used in Experiment 1. Panels A and B depict the two configurations of the disambiguated stimuli presented on training trials. Training stimuli were presented in stereo using anaglyph glasses. Panel C shows a typical ambiguous stimulus presented on test trials. Test stimuli were presented monocularly. The white background used here is for illustration purposes only; the stimuli were presented on black background during the experiments.



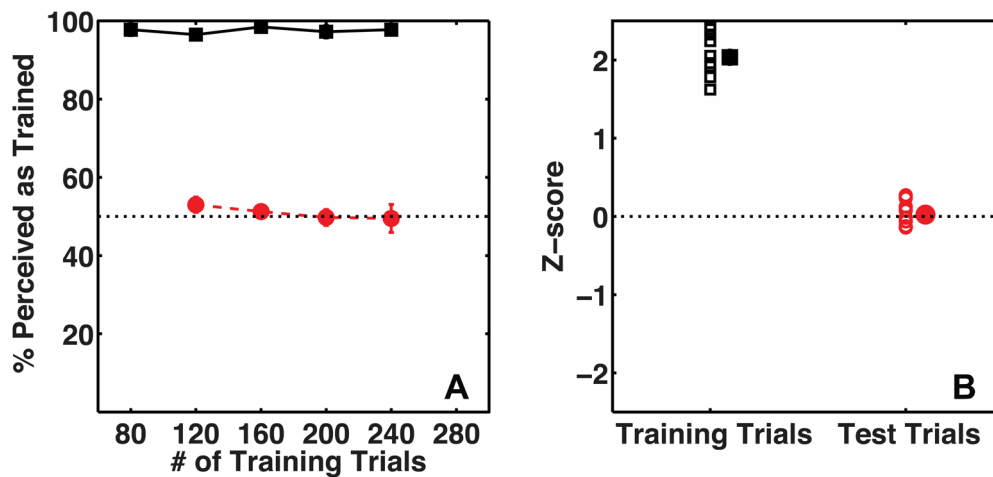
**Figure 2.** Structure of three typical trials (two training trials and one test trial) in Experiment 1. Like training trials, test trials appeared both above and below fixation. One face, chosen randomly, was highlighted during the 0.25 sec probe.



**Figure 3.** Average percent perceived as trained on Day 1 by observers (N=10) on Training trials (black squares) and Test trials (red circles). Panels A and B show data as a function of the number of training trials on Day 1 and Day 2, respectively. Panel C shows the same data as a Z-score measure for the entire session, with a separate marker for each subject. The mean of the data is plotted with a slight horizontal shift to prevent individual data being occluded. The error bars show the standard error of the mean.

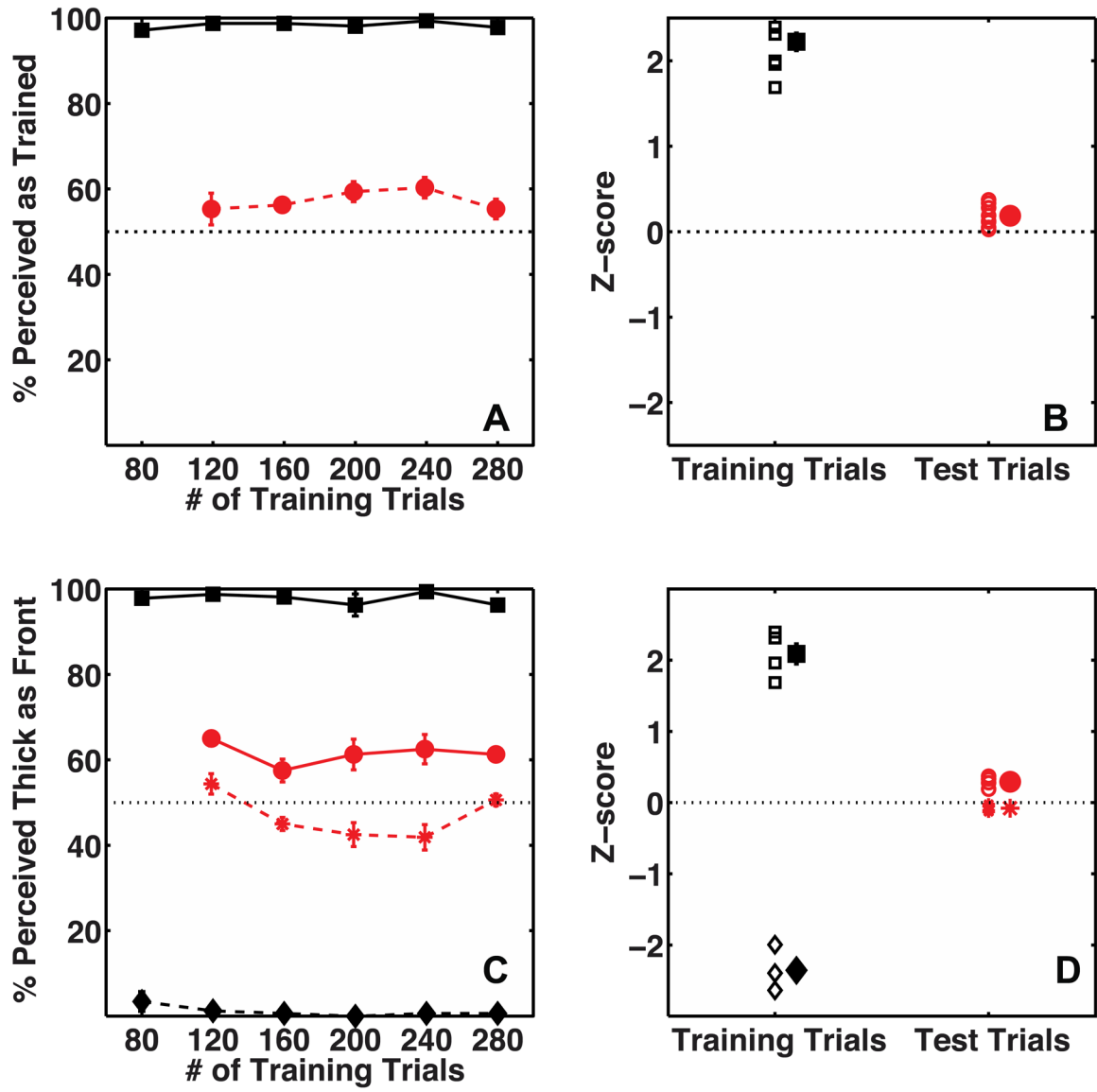


**Figure 4.** Examples of binocular disambiguated training trials and monocular ambiguous test trials as used in Experiment 2A (top row) and Experiment 2B (bottom row). The white background used here is for illustration purposes only; the stimuli were presented on black background in the experiments.



**Figure 5.** Average percent perceived as trained by observers (N=10) on Training trials (black squares) and Test trials (red circles). Panel A shows data as a function of the number of training trials. Panel B shows the same data as a Z-score measure for the entire session, with a separate marker for each subject. The mean of the data is plotted with a slight horizontal shift to prevent individual data being occluded. The error-bars show the standard error of the mean.





**Figure 6.**

**A)** Average percent perceived as trained by observers (N=8) on Training trials (black squares) and Test trials (red circles) as a function of the number of training trials. **B)** The same data is shown as a Z-score measure for the entire session, with a separate marker for each subject. **C)** Average percentage of trials seen as having thick gratings as closer observers (N=8) on Training trials and Test trials as a function of the number of training trials. The black squares and red circles represent observers trained with thick gratings as front contingency while the black diamond and red stars represent observers train with thin gratings as front contingency. **D)** The same data is shown as a Z-score measure for the entire session, with a separate marker for each subject. In Panels B and D the mean of the data is plotted with a slight horizontal shift to prevent individual data being occluded. The error bars show the standard error of the mean.