Associative learning of shape as a cue to appearance: A new demonstration of cue recruitment

Sarah J. Harrison
Department of Psychology, Royal Holloway University of London, Egham, UK
Graduate Center for Vision Science, SUNY College of Optometry, USA, & SUNY Eye Institute, USA

Benjamin T. Backus

The perceived rotation direction of a wire-frame Necker cube at stimulus onset can be conditioned to be dependent on retinal location (B. T. Backus & Q. Haijiang, 2007; S. J. Harrison & B. T. Backus, 2010a). This phenomenon was proposed to be an example of the visual system learning new cues to visual appearance, by adaptation in response to new experiences. Here, we demonstrate recruitment of a new cue, object shape, for the appearance of rotating 3D objects. The cue was established by interleaving ambiguous and disambiguated instances of two shapes, cubes and spheres, at the same retinal location. Disambiguated cubes and spheres rotated in opposite directions. A significant bias was consequently introduced in the resolution of ambiguity, whereby the proportions of ambiguous shapes perceived as rotating clockwise differed, in the direction predicted by their disambiguated counterparts. This finding suggests that training led the visual system to distinguish between the two shapes. The association of rotation direction and shape was only achieved when monocular depth cues were used to depict rotation in depth; shapes disambiguated by binocular disparity did not lead to recruitment of the shape cue. We speculate that this difference may be the consequence of a difference in the neural pathways by which the disambiguating cues act. This new instance of the cue recruitment effect opens possibilities for further generalization of the phenomenon.

Keywords: learning, motion—3D, depth, perceptual bias, bistable perception, structure from motion


Introduction

Many adaptive capabilities of the visual system are well known and are usually demonstrated either through an improvement in discrimination or detection performance (e.g., Fine & Jacobs, 2002; Seitz & Dinse, 2007) or by negative aftereffects on appearance (color aftereffects, for example). These types of adaptation are thought to promote greater efficiency in the use of low-level visual signals and heightened sensitivity to deviations from the norm (Barlow, 1990; Webster, Werner, & Field, 2005), respectively. More recently however, a very different third type of adaptation has been demonstrated, whereby prolonged intermittent presentation of a stimulus establishes a bias to resolve ambiguous visual input in accord with previous experience, at stimulus onset (e.g., Brascamp et al., 2008; Haijiang, Saunders, Stone, & Backus, 2006; Klink et al., 2008; Leopold, Wilke, Maier, & Logothetis, 2002; Long & Moran, 2007; Noest, van Ee, Nijs, & van Wezel, 2007; Pastukhov & Braun, 2008). Rather than exhibiting a short-lived negative aftereffect, which typically follows sustained viewing of a visual feature, the visual system becomes resistant to change, with regard to interpretation of the stimulus, i.e., a kind of positive aftereffect. Concurrently, it was established that the stabilized percept could be trained to be dependent on a previously unrelated stimulus parameter or “signal,” in an associative learning paradigm (Haijiang et al., 2006). That is, a bias was learned to interpret visual input in the same way as on previous occurrences of that signal. This signal, therefore, was said to have been “recruited” by the visual system as a cue to visual appearance.

The original demonstration of cue recruitment used a perceptually bistable Necker cube stimulus, which can be perceived as rotating in either direction at stimulus onset: Cubes with rotation direction disambiguated by depth cues were presented either above or below fixation. Interleaved presentation of ambiguous cubes at the same two locations resulted in the perceptual outcome of ambiguous cubes adopting the same location—rotation contingency as specified by disambiguated cubes. The recruitment of retinal location as an independent cue to perceived rotation direction was demonstrated not only for ambiguous stimuli (Haijiang et al., 2006) but also for stimuli that combine the new cue and other depth cues (Backus & Haijiang, 2007).

The cue recruitment effect, as described above, was proposed to be an example of the visual system learning...
new cues to visual appearance, through experience of association with long-trusted cues (cf. Backus & Haijiang, 2007; Berkeley, 1709). In this way, it was suggested that the visual system could adapt in the face of new experiences, perhaps with the function of becoming optimized for rapid and efficient processing of visual input. However, a lack of further examples of associative learning of this type could lead to concerns about the generality of the cue recruitment finding and the appropriate framework within which to interpret it.

In the experiments that follow, we used the same conditioning paradigm as previously (Backus & Haijiang, 2007; Haijiang et al., 2006; Harrison & Backus, 2010a, 2010b), with the aim of recruiting object shape as a new cue to object rotation direction. We interleaved ambiguous and disambiguated instances of two rotating shapes at the same retinal location, with disambiguated instances of the two shapes assigned opposite rotation directions. The extent to which the perceptual outcome for ambiguous instances was the same as for their disambiguated counterparts of the same shape was our measure of cue recruitment. It has previously been demonstrated that perceptual stabilization of an ambiguous rotating 3D shape transfers between different shapes (Maier, Wilke, Logothetis, & Leopold, 2003) and that the retinal location cue transfers between rotating objects with different shapes (van Dam & Ernst, 2010). However, neither of these findings addresses our question of whether shape itself can be trained as a cue.

To anticipate our findings, we demonstrate that significant recruitment of the shape cue can be achieved, using the combined monocular depth cues of occlusion, self-occlusion, depth from shading, and haze to disambiguate shape–rotation direction in the conditioning stimulus. Follow-up experiments rule out possible confounds for this finding of a shape-contingent rotation bias. We therefore conclude that the effect is a new case of associative learning of a cue for visual appearance and provide a possible reason why it was observed only when using monocular depth cues during training.

**General methods**

**Hardware and software**

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

**Shape stimuli**

Simulated rotating shape 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. Subjects were seated so as to be comfortable at a viewing distance of 1.0 m from the display screen, which was also the distance to the center of the stimulus as specified by binocular vergence demand and accommodative demand.

Two combinations of depth cues were used in the following series of experiments to control the direction of rotation of disambiguated shapes (Figure 1). The first cue set was binocular disparity and occlusion (“disparity plus occlusion”), a combination of cues that had been highly effective in the original demonstrations of cue recruitment (Backus & Haijiang, 2007; Haijiang et al., 2006). Shapes were presented with geometrically correct disparity (as described in the Methods section of Experiment 1). The occlusion cue consisted of a central column around which the shape rotated. The column was a 2-dimensional vertical strip of 4.0-cm width. Far portions of the shape were occluded as they moved around the back of the column, whereas closer portions of the shape were visible in front of the column.

The second cue set consisted of monocular depth cues only (“monocular only”), and, like “disparity plus...
occlusion;” this cue set has been demonstrated to be effective in conditioning a retinal location cue (Harrison, Backus, & Jain, 2011). Cube frame edges depicted both a depth-from-shading cue through use of a directional light source and an internal occlusion cue. The impact of the light source was further enhanced by use of the Vizard “fog” function, which simulates the effect of haze or “aerial perspective,” causing the contrast between adjacent faces of a cube edge to decrease with distance (as simulated by the other depth cues). The central column occlusion cue was once again present.

Ambiguous shapes 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. All shapes were presented using orthographic projection so that perspective cues did not indicate front and back of the cube in the test case.

Two 3-dimensional shapes were used in the following experiments: a wire-frame cube and a wire-frame sphere. The sphere was constructed from three wire-frame great circles that interlocked on a common central axis. Each circle was offset from the others by a rotational angle of 60 degrees, giving the appearance of a “beach ball” with six segments. The dimensions and parameters associated with each shape are described prior to each experiment.

All shapes contained 25 randomly placed dots on each transparent face or surface segment, which stabilized appearance on ambiguous instances as a single rigid rotating body. All shapes rotated from a starting orientation at stimulus onset such that their front and back edges were vertical and coincident at the center of the image (45 degrees of yaw). The roll and pitch angles, which determine whether the shape appeared to be viewed from above or below at stimulus onset, were either both +25 or both −25 degrees (Figure 1). These two configurations are associated with different motions of the cube edges at stimulus onset, which can be a confounding cue as to the direction of rotation (Grossmann & Dobbins, 2006). Hence, the viewing angle at stimulus onset, “above” or “below,” was balanced across both training and test trials as per the methodology introduced by Haijiang et al. (2006).

**Experimental procedure**

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, which was done with a key press, and to then maintain fixation. On initiation of each trial, a rotating shape appeared centered 12.0 degrees below the fixation marker. Simultaneously, a comparison dot repeated short cycles of horizontal motion through the fixation marker.

The 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 (key presses “2” and “8,” respectively). Subjects were instructed that we were interested in the accuracy, not the speed, of their responses. 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 training trials and monocularly on test trials. The shape and comparison dot remained on the screen for a minimum of 1.5 s and a maximum of 6.0 s; the subject’s response terminated the presentation at any time after 1.5 s. As subjects responded immediately upon forming a clear rotation percept, and their response terminated the stimulus presentation (unless less than 1.5 s had passed since stimulus onset), it was extremely rare for a reversal of percept to occur. Each session consisted of 480 trials and lasted approximately 30 min depending on the pace of the individual subject.

Due to random assignment of dot direction, the measure of interest, perceived direction of rotation, was not correlated with the key-press response. None of the subjects who took part in the following experiments reported being aware that rotation direction was dependent on stimulus shape even though shape correlated perfectly with rotation direction on disambiguated trials.

**Subjects**

Subjects were adult members of the public who were recruited from the local area (New York City) via Craig’s List and were paid for their time. Subjects’ vision was normal or corrected to normal with non-bifocal lenses. Stereoscopic vision was assessed using the TNO Stereacuity test; subjects were required to have a minimum stereocuity of 240 s of arc. However, stereocuity in static images is not always an indicator of stereocuity in dynamic images such as the rotating cube stimulus used here (Rouse, Tittle, & Braunstein, 1989). Accordingly, our critical measure of subjects’ suitability for the experiment, in terms of both stereocuity and task comprehension, was their performance on training trials on the first day of testing. In the course of the following five experiments, six subjects, who met other criteria but did not reach performance levels on disambiguated trials of 90% or more for each shape individually, on Day 1, were excluded (2 each from Experiments 2, 4, and 5). A total of 40 subjects completed the experiments in the study.

**Analysis**

The percent of shapes seen as rotating in the direction specified by their disambiguated counterpart, for each of
the two shapes, was transformed into a z-score, “zCube” and “zSphere,” i.e., we used a probit (inverse cumulative normal) transformation (Backus, 2009; Dosher, Sterling, & 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 1 non-conforming response within the 120 observations of each ambiguous shape within a session or 2 non-conforming responses when considering all 240 ambiguous trials. In Experiments 4 and 5, where results for each of the 8 stimulus parameter combinations are depicted separately, a limiting z-score of ±1.834 is used, corresponding to 1 non-conforming response within 30 observations. Where comparisons of z-scores for cubes and spheres were made, non-parametric tests were used for consistency throughout, due to the bimodal distribution of z-scores in some experiments.

For each subject, z-scores for the two shapes were summed, giving a single “zDiff” measure of the extent to which perceived rotation differed between the two shapes. In the case of ambiguous shapes therefore, zDiff is a measure of training-induced bias, which is independent of any global, preexisting bias. As will be seen, the zDiff measure is remarkably robust to intersubject variability in the absolute magnitude of the two component z-scores. Zdiffs approximated a normal distribution and so are analyzed using parametric tests.

Where analysis indicated that a shape–rotation bias was adopted on Day 1 ambiguous trials, subjects were tested on a second consecutive day to assess the extent of the bias through counterconditioning. On Day 2, subjects viewed the opposite shape–rotation contingency to that viewed on Day 1. Day 2 zDiffs were assessed relative to the shape–rotation contingency depicted on Day 1.

Experiment 1

In this first experiment, we attempted to train object shape as a cue to rotation direction, using a conditioning paradigm that was very similar to the previously demonstrated recruitment of a retinal location cue (Harrison & Backus, 2010a). Two simulated volumetric shapes, a sphere and a cube, were presented in a randomly interleaved fashion at a single retinal location. For half of the instances of each shape, rotation direction was disambiguated by the disparity-plus-occlusion cue set; the other half were ambiguous, as detailed in the General methods section.

Methods

The edges of the wire-frame cube were solid rectangular parallelepipeds with length of 20.0 cm and width and breadth of 0.3 cm. The edges of the wire-frame circles were curved parallelepipeds with a square cross-section of 0.5 cm width and breadth of 0.3 cm. The diameter of the sphere was 30.0 cm; this diameter produces a sphere of approximately equal spatial extent (or “radius”) as the cube, and so the shapes were perceived as approximately equal “size.” However, due to the yaw of cube stimuli, the maximum front-to-back distance of the cube (during rotation) was slightly less than that of the sphere, and this led to a smaller maximum front-to-back disparity (0.9 vs. 1.1 degrees of visual angle). Both shapes rotated about a vertical axis at a rate of 45 degrees s⁻¹.

On Day 1, subjects were trained under either “cube-clockwise (CW), sphere-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 blocks of 8 trials that contained all combinations of trial type (disambiguated/ambiguous), shape type (cube/sphere), and shape viewpoint (above/below), as detailed in the General methods section. The first block of trials was constrained to present the 8 possible trials in a disambiguated–ambiguous sequence, alternating between shape types and shape viewpoints. The order of presentation of these first 8 trials was counterbalanced across subjects.

Results

Z-scores for cubes and spheres, for all 8 subjects, are presented in Figure 2. All subjects perceived the vast majority of ambiguous stimuli as rotating in the same direction, regardless of shape, resulting in a strongly bimodal distribution of z-scores. Either cubes were seen according to the rotating direction depicted on disambiguated instances (positive z-score) and spheres were seen as rotating oppositely to disambiguated instances and hence in the same direction as cubes (negative z-score) or vice versa. Hence, the average zDiff (zCube + zSphere) was not significantly different from zero (t(7) = 0.92, p = 0.39, 95% CI = [−0.55, 0.24]; Figure 15). (Dividing the ambiguous trials into 6 consecutive blocks each containing 20 cube and 20 sphere instances revealed that there were no significant changes in perceptual outcome through the course of the testing session.) The perceived direction of rotation on ambiguous trials was not, over all subjects, significantly dominated by the direction depicted by disambiguated instances of one shape or the other (zCube vs. zSphere; Wilcoxon signed rank Z = −0.14, p = 0.89).

Discussion

Clearly, shape was not recruited as a cue to rotation direction. Instead, the visual system “chose” one rotation direction and resolved all ambiguous instances accordingly, as if the two shapes were indistinguishable to the
bias mechanism. This result is not entirely surprising given the known transference of bias between shapes (van Dam & Ernst, 2010). This pattern of visual response has also been seen when attempting to train spatiotopic (i.e., allocentric or world) location as a cue (Harrison & Backus, 2010a), where the visual system adopted the same outcome on nearly all ambiguous instances at a given retinal location, regardless of spatiotopic location.

We have previously conjectured that rotation in depth depicted by binocular disparity-plus-occlusion and monocular-only cue sets activates neural populations that are at least partially overlapping but that the two cue sets have different adaptation strengths either because they exert their influence through different neural pathways or due to contributions by non-overlapping neural populations (Harrison et al., 2011). This led us to consider whether using different cue sets for disambiguation of the two shapes might enable recruitment of the shape cue.

Experiment 2

Methods

In Experiment 2, disambiguated cubes used the monocular-only cue set described in the General methods section. The edges of the cube were solid rectangular parallelepipeds with length of 20.0 cm (outside edges) and increased width and breadth of 2.0 cm so as to allow the use of shape-from-shading, self-occlusion, and haze cues to depth. Ambiguous cube stimuli had the same dimensions as disambiguated cubes but contained no depth cues. Ambiguous and disambiguated spheres were depicted exactly as in Experiment 1; disambiguated spheres still employed the “disparity-plus-occlusion” cue set, and both they and their ambiguous counterparts had wire-frame edges. The trial sequence was identical to Experiment 1, with the order of the first 8 trials balanced across subjects, and the remaining 472 trials randomly interleaved within balanced blocks of 8 trials.

Results

Z-scores for cubes and spheres, for all 8 subjects, are presented in Figure 3. The average zDiff was very slightly above zero (zDiff = 0.30, t(7) = 2.51, p = 0.04, 95% CI = [0.017, 0.58]; Figure 15), which suggests a small shape–rotation bias. Unlike in Experiment 1, the perceived direction of rotation on ambiguous trials (both cubes and spheres) was now strongly dominated by that depicted by disambiguated cubes (zCube vs. zSphere; Wilcoxon signed rank Z = −2.38, p = 0.02)—note that the direction of rotation of cubes was counterbalanced across subjects.
As before, there were no significant changes in perceptual outcome through the course of the testing session.

One plausible explanation for a rotation bias being established in the same direction as that depicted by disambiguated cubes is that subjects may have viewed them for longer: If disambiguation of rotation direction by our monocular-only cue set required a longer viewing time than did disambiguation by our disparity-plus-occlusion cue set, then one could conjecture that it was not the disambiguating cues but instead the difference in the length of stimulus presentation that led to the dominance of the cubes. However, there was no significant difference in presentation times between disambiguated cubes and spheres (cube mean presentation time = 2.29 s, SEM = 0.11 s; sphere mean display time = 2.27 s, SEM = 0.14; paired difference t(7) = 0.47, p = 0.65). However, because stimuli remained visible for a minimum of 1.5 s even when subjects responded sooner, potentially creating a situation where stimuli were present but unattended following a fast response, we also verified that there were no differences in reaction time (cube mean RT = 2.17 s, SEM = 0.16 s; sphere mean RT = 2.12 s, SEM = 0.19 s; paired difference t(7) = 0.50, p = 0.64). In fact, the lack of significant differences also extended to a 2-way comparison that included test trials.

Discussion

The ability of the cube to dominate the perceived rotation direction is of interest. This did not occur when both shapes were disambiguated by the disparity-plus-occlusion cue set, and so we conclude that the critical factor in the cubes’ influence was the use of the monocular-only cue set to disambiguate rotation direction. An alternative possibility, that the bias is a result of differences in the total luminance and contrast between stimuli, caused by the different widths of edge elements in monocular-only and disparity-plus-occlusion stimuli, has previously been ruled out (Harrison et al., 2011). Despite finding a positive zDiff, it would be premature to conclude that shape was successfully established as a cue to rotation direction: The magnitude of the effect was very small, much smaller than that seen previously for a retinal location cue (Harrison & Backus, 2010a). Hence, it is possible that this result represents a chance finding of significance. Nevertheless, the strong influence of monocularly disambiguated stimuli led us to try a further stimulus variation in our attempt to recruit shape as a cue.

Experiment 3

Methods—Day 1

In Experiment 3, both disambiguated cubes and spheres used the monocular-only cue set described in the General methods section. Cube dimensions were as described for Experiment 2. Spheres had an outside diameter of 30.0 cm, as previously, but wire-frame edges had increased width and breadth, like the cube edges, of 2.0 cm. All ambiguous stimuli had the same dimensions as their disambiguated shape counterparts but contained no depth cues. The trial sequence was identical to Experiment 1, with the order of the first 8 trials balanced across subjects, and the remaining 472 trials randomly interleaved within balanced blocks of 8 trials.

Results—Day 1

Z-scores for cubes and spheres, for all 8 subjects, are presented in Figure 4. The overall perceived direction of rotation was no longer dominated by the direction depicted by one or the other disambiguated shape (zCube vs. zSphere; Wilcoxon signed rank Z = −0.84, p = 0.40). Again, there was evidence of a small shape–rotation bias; zDiff was significant (zDiff mean = 0.48, SEM = 0.16, t = 3.0, p = 0.021; Figure 15) and slightly larger than in Experiment 2.

Methods—Day 2

A learned shape–rotation bias was again present and was slightly larger in magnitude than seen in Experiment 2. We assessed the longevity of the bias by counterconditioning
subjects on a second, consecutive day. On Day 2, subjects viewed disambiguated shapes with the opposite shape-rotation contingency to that seen the previous day. If little long-term learning occurred on Day 1, then subjects should show a bias to perceive ambiguous stimuli on Day 2 according to the shape-rotation contingency of Day 2 disambiguated stimuli (negative zDiff). However, if Day 1 experience caused learning, subjects’ perception of ambiguous stimuli on Day 2 should reflect a bias in the direction of the shape-rotation contingency experienced on Day 1 (positive zDiff).

Results—Day 2

Z-scores for cubes and spheres, for all 8 subjects on Day 2, are presented in Figure 5. On Day 2, zDiff was significantly negative (zDiff mean = $-0.49$, SEM = 0.19, $t(7) = 2.54$, $p = 0.04$; Figure 16) and was not significantly different from that which would be predicted if equal but opposite learning of the shape-rotation bias occurred on Day 2 as had occurred on Day 1 ($t(7) = 0.03$, $p = 0.97$, 95% CI = $[-0.54, 0.56]$). As before, the perceptual outcomes for ambiguous spheres and cubes were equally likely to match that of their disambiguated counterparts (zCube vs. zSphere; Wilcoxon signed rank $Z = -0.28$, $p = 0.78$).

Discussion

Only a small shape-rotation bias was learned on Day 1, which was not strong enough to be retained on Day 2. However, as in Experiment 2, all but one subject had a positive zDiff measure of trained bias on Day 1. Hence, it is increasingly likely that this is a non-trivial result, reflecting recruitment of the shape cue.

One could ask whether the result is due to a confound such as the two shapes occupying non-overlapping regions of the visual field and hence activating marginally different neuronal populations. Similarly, a slight difference in peak, or mean, retinal velocity, due to the different shapes having matched angular velocity, could be hypothesized to cause different sets of neurons to be activated by each of the two shapes. We suggest that this is unlikely, as the spatial extent and speed of rotation of each shape were identical to those used in Experiment 1, where no shape-rotation bias was learned. The difference between the two experiments is in the type of disambiguating cues used: Experiment 1 used the disparity-plus-occlusion cue set, whereas Experiment 3 used the monocular-only cue set. The type of disambiguating depth cues was the critical factor in recruitment of shape as a cue to rotation direction; this idea will be discussed further in the General discussion section.

In Experiments 4 and 5, we provide further evidence that our result is not due to methodological or stimulus confounds. In Experiment 4, different sizes and rotational speeds were used for disambiguated and ambiguous instances of both shapes. If the result obtained in Experiment 3 is due to differences in these low-level properties between the two shapes rather than the different shapes themselves, then using a mixture of shape sizes and speeds should disrupt the effect. In Experiment 5, we rule out the possibility that our manipulation of the starting sequence of trials contributed to our result.

Experiment 4

Methods

In Experiment 4, we verify that differences in the low-level properties of size or retinal velocity do not underlie the bias observed in Experiment 3. For instance, parts of the sphere impinged on retinotopic locations that the cube did not, due to the cube’s non-uniform spatial extent. A difference in mean retinal velocities could likewise lead to non-identical activation of neurons selective for these low-level visual properties.

Disambiguated cubes and disambiguated spheres both used the monocular-only cue set described in the General methods section. However, two different sizes of each shape were used as were two different speeds of rotation.

![Figure 5](image-url)
Our aim was not to evaluate the impacts of size and speed on the bias. Rather, our logic was that introduction of different sizes and speeds would disrupt learning of the bias if it was dependent on differences in these low-level properties but would not disrupt learning of the bias if stimulus shape was the cue to rotation direction that was being learned by the visual system.

Large cubes had outside edges of length 23.0 cm, and small cubes had edges of length 17.0 cm. Large spheres had an outside diameter of 36.0 cm, and small spheres had an outside diameter of 26.0 cm. For all shapes, edges had width and breadth of 2.0 cm. All ambiguous stimuli had the same dimensions as their disambiguated shape counterparts but contained no depth cues. Both shapes, in both their disambiguated and ambiguous forms, were presented rotating at two speeds, 35 and 55 degrees s\(^{-1}\).

Shapes had either “above” or “below” viewpoint at stimulus onset. Combining all possible parameters resulted in a total of 32 possible stimuli.

Previous anecdotal observations of strong adapting effects of larger or faster shapes led us to devise a starting sequence of stimuli that was intended to encourage recruitment of the shape cue by delaying presentation of a “large and fast” pairing. Shapes were presented in disambiguated–ambiguous pairs, for the first eight presentations. A small, fast instance of the first shape was presented, followed by a large, slow instance of the second shape, then a small, slow version of the first shape, and finally a large fast instance of the second shape. The first two pairs were presented in the “above” viewpoint and the second two pairs in the “bottom” viewpoint. The assignment of the cube and the sphere to the “first”

![Sample stimuli: Large and small cubes and spheres, disambiguated by the monocular-only cue set. (Images are cropped and do not depict the full extent of the display screen.)](image-url)
and “second” positions and the shape–rotation contingency were counterbalanced across subjects. The next 24 presentations were drawn randomly from the remaining 24 (out of 32) possible stimuli. The remaining 448 trials were randomly interleaved within balanced blocks of 32 trials that contained all combinations of trial type, shape type, above/below shape viewpoint, shape size, and shape–rotation speed.

Results—Day 1

$Z$-scores for cubes and spheres, for 8 subjects, are shown in Figure 7. Day 1 $z$Diff was significantly different from zero (mean = 0.74, $SEM = 0.14$, $t(7) = 5.31$, $p = 0.001$, 95% CI = [0.41, 1.06]; Figure 15). As before, the perceptual outcomes for ambiguous spheres and cubes were equally likely to match that of their disambiguated counterparts ($z$Cube vs. $z$Sphere; Wilcoxon signed rank $Z = -0.28$, $p = 0.78$).

While our experimental design was not intended to have the statistical power to explore the factors of shape size and speed, it is of interest to look at the extent to which the rotation bias was adopted by each of the eight types of ambiguous stimulus. Figure 8 shows $z$-scores separated by stimulus type, for each of the eight subjects. There is no reason to suspect that the rotational bias was adopted more strongly by some types of stimulus than others. However, our logic does not depend on this evidence, only on the lack of disruption of the bias (Figure 15). For instance, even if the bias had been observed predominantly in one stimulus—say large fast cubes—a hypothetical mechanism capable of learning a low-level rotational bias specific to only the “large and fast” stimuli would need to be highly idiosyncratic.

Figure 7. $Z$-scores for perceived rotation of ambiguous spheres and cubes for 8 subjects on Day 1 in Experiment 4. A positive $z$-score means that shapes were seen rotating in the same direction as their disambiguated counterparts, whereas a negative $z$-score means that shapes were seen as rotating in the opposite directions.

Figure 8. $Z$-scores for perceived rotation of ambiguous spheres and cubes (small slow, SS; small fast, SF; big slow, BS; big fast, BF) on Day 1 in Experiment 4. Each graph shows data for an individual subject. The order of subjects is the same as that in Figure 7. A positive $z$-score means that shapes were seen rotating in the same direction as their disambiguated counterparts, whereas a negative $z$-score means that shapes were seen as rotating in the opposite directions.
Results—Day 2

Subjects received counterconditioning on Day 2, viewing the reverse shape–rotation contingency to that shown on Day 1. Z-scores for cubes and spheres, for 8 subjects, are shown in Figure 9. Day 2 zDiff was not significantly different from zero (mean = −0.035, SEM = 0.19, 95% CI = [−0.48, 0.41]; Figure 16) but was significantly less negative than would be expected if the shape–rotation bias was established to the same extent on Day 2 as it had been on Day 1 (t(7) = −3.41, p = 0.011), i.e., the retention of bias was significant. Cubes and spheres were equally likely to be seen as rotating in the same direction as their Day 1 disambiguated counterparts, whereas a negative z-score means that shapes were seen as rotating in the opposite directions.

Experiment 5

Methods

Finally, in Experiment 5, we departed from the strict structure of the starting sequence of trials used in all previous experiments and instead randomized the order of stimulus presentation from the very start of the session. As in Experiment 4, two different sizes of each shape, and two different rotation speeds, were used. Together with the two trial types (training, test), two shapes (cube, sphere), and two shape viewpoints (above, below), this
resulted in 32 possible parameter combinations. Trials were drawn randomly from balanced blocks containing one of each of the 32 possible trials. No other ordering was imposed. This meant that it was possible for a trial sequence to start with several ambiguous trials before a disambiguated stimulus was presented or several instances of one shape before the other shape was presented. Deviation from a strictly balanced sequence could potentially lead to later perceptual outcomes adopting the outcome seen on the first ambiguous presentations (Harrison & Backus, 2010b) or to the rotation direction depicted by one shape or the other becoming dominant, as we have previously observed that a bias can be very rapidly adopted, indeed, within the first few trials (e.g., Harrison & Backus, 2010a; van Dam & Ernst, 2010). Removing the strict structure of the starting sequence therefore tests the robustness of the shape recruitment to these possible effects.

Results—Day 1

Z-scores for cubes and spheres separately, for 8 subjects, are presented in Figure 11. ZDiff was significantly different from zero (mean = 0.70, SEM = 0.21, t(7) = 3.28, p = 0.014; Figure 15). There was a non-significant trend for ambiguous cubes to be more likely than ambiguous spheres to be seen as rotating in the direction depicted by their unambiguous counterparts (zCube vs. zSphere; Wilcoxon signed rank Z = −1.82, p = 0.07). The cube and sphere z-scores are shown broken down into their individual stimulus types in Figure 12.

Results—Day 2

Subjects received counterconditioning on Day 2, viewing the reverse shape–rotation contingency to that shown

Figure 11. Z-scores for perceived rotation of ambiguous spheres and cubes for 8 subjects on Day 1 in Experiment 5. A positive z-score means that shapes were seen rotating in the same direction as their disambiguated counterparts, whereas a negative z-score means that shapes were seen as rotating in the opposite directions.

Figure 12. Z-scores for perceived rotation of ambiguous spheres and cubes (small slow, SS; small fast, SF; big slow, BS; big fast, BF) on Day 1 in Experiment 5. Each graph shows data for an individual subject. The order of subjects is the same as that in Figure 11. A positive z-score means that shapes were seen rotating in the same direction as their Day 1 disambiguated counterparts, whereas a negative z-score means that shapes were seen as rotating in the opposite directions.
on Day 1. Results from Day 2 counterconditioning are shown in Figure 13. Day 2 zDiff was not significantly different from zero ($t(7) = 0.64, p = 0.54, 95\% \text{ CI} = [-0.39, 0.22]$; Figure 16) but was significantly less negative than that expected if the shape–rotation contingency was adopted to the same extent on Day 2 as it had been on Day 1 ($t(7) = -2.42, p = 0.05$), suggesting a marginal retention of bias. There was a non-significant trend for ambiguous cubes to be more likely than ambiguous spheres to be seen as rotating in the direction depicted by their unambiguous counterparts ($z_{\text{Cube}}$ vs. $z_{\text{Sphere}}$; Wilcoxon signed rank $Z = -1.26, p = 0.21$). The cube and sphere $z$-scores are shown broken down into their individual stimulus types in Figure 14.

On Day 1 of Experiment 5, it could be suggested that the shape–rotation bias was due primarily to the perceptual outcomes of ambiguous cubes; there was a non-significant trend for cubes to adopt their trained rotation to a greater extent than did spheres. However, note that even if we did conclude that the bias was predominantly acquired by one of the two shapes, this scenario would still constitute learning of the shape cue providing $z$Diff was significant, as the visual system must be distinguishing between the shapes in order for there to be a difference in their acquired biases. Due to the well-established percept stabilization effect for both spheres and cubes in isolation (see Introduction section) and the clearly demonstrated ease with which the learned bias transfers between shapes (Maier et al., 2003; van Dam & Ernst, 2010), we do not consider it a viable possibility that rotation biases are independently acquired (or not acquired) from disambiguated instances to ambiguous instances of each shape.

With respect to the observed trend itself, we feel that a very mundane explanation may suffice: Combining data...

![Figure 13](image13.png)

**Figure 13.** Z-scores for perceived rotation of ambiguous spheres and cubes for 8 subjects on Day 2 in Experiment 5. The order of subjects is the same as that in Figure 11. A positive $z$-score means that shapes were seen rotating in the same direction as their Day 1 disambiguated counterparts, whereas a negative $z$-score means that shapes were seen as rotating in the opposite directions.

![Figure 14](image14.png)

**Figure 14.** Z-scores for perceived rotation of ambiguous spheres and cubes (small slow, SS; small fast, SF; big slow, BS; big fast, BF) on Day 2 in Experiment 5. Each graph shows data for an individual subject. The order of subjects is the same as that in Figure 13. A positive $z$-score means that shapes were seen rotating in the same direction as their Day 1 disambiguated counterparts, whereas a negative $z$-score means that shapes were seen as rotating in the opposite directions.
Experiments 3 and 4 revealed that the shape depicted on the first trial was significantly more likely to have the higher $z$-score ($t(14) = 2.17$, $p = 0.047$). The random ordering of trials in Experiment 5 in fact resulted in four out of eight subjects “correctly” perceiving cube rotation on the first trial (irrespective of whether it was disambiguated or ambiguous stimulus, bearing in mind that perceptual outcomes from ambiguous trials can establish strong biases also; Harrison & Backus, 2010b). In comparison, only one subject correctly perceived a sphere on their first trial. Of the remaining three subjects, two had cubes as their first disambiguated shape (with earlier trials being “incorrectly” perceived). While our experiment was not designed to explore the complex interplay between the sequence of trial types and their perceptual outcomes in establishing perceptual bias, it would appear that in Experiment 5 cubes may have had a “head start”: Subjects’ perceptual experiences earlier in the sequence of trials were likely highly influential in establishing the measured bias.

Comparing results across the three experiments that used only monocular depth cues to disambiguate rotation direction (Experiments 3, 4, and 5; 1-way ANOVAs), we found no significant difference in Day 1 $z$Diff ($F(2) = 0.65$, $p = 0.53$), Day 2 $z$Diff ($F(2) = 2.01$, $p = 0.15$), or the change in $z$Diff from Day 1 to Day 2 ($F(2) = 0.18$, $p = 0.84$). We conclude that the shape–rotation bias established in Experiment 3 was not dependent on low-level stimulus differences nor was it due to a confound introduced by the starting sequence of trials. Having established that there were no significant differences between the three above experiments, we pooled the results to increase statistical power: We found a small but highly significant Day 1 $z$Diff (mean = 0.64, $SEM = 0.10$, $t(23) = 6.46$, $p < 0.01$) and a non-significantly negative Day 2 $z$Diff (mean = -0.20, $SEM = 0.10$, $t(23) = -1.94$, $p = 0.06$, 95% CI = [-0.42, 0.01]). Day 2 $z$Diff, despite being negative, was significantly different from that which would be expected if Day 2 shape–rotation contingency was adopted to the same extent as on Day 1 ($t(23) = -3.03$, $p < 0.01$). Overall results from Experiments 1–5 ($z$Diffs) are shown in Figures 15 and 16 (Days 1 and 2, respectively). Considering all three experiments together, cubes and spheres were equally likely to be seen as rotating in the same direction as their disambiguated counterparts, both on Day 1 (zCube vs. zSphere; Wilcoxon signed rank $Z = -1.40$, $p = 0.16$) and on Day 2 (Wilcoxon signed rank $Z = -0.33$, $p = 0.74$). A parametric test of zCube vs. zSphere (which is in fact appropriate for the aggregated data) likewise found no significant difference (Day 1, $t(23) = 1.65$, $p = 0.11$; Day 2, $t(23) = 0.742$, $p = 0.93$).

**General discussion**

Using an associative learning paradigm, we have demonstrated recruitment of object shape as a cue for object rotation in bistable simulated 3D shapes. Recruitment of the shape cue was an effect of small magnitude...
but was consistent across observers. This finding demonstrates a generalization of the “cue recruitment effect” (Haijiang et al., 2006)—whereby the visual system learns through contingent exposure to a previously unrelated cue that the cue is predictive of some parameter of an object’s visual appearance—and lends support to the functional interpretation of cue recruitment as a means by which the visual system adapts to the environment.

The cue recruitment effect was originally demonstrated through association of retinal location with rotation direction in the bistable Necker cube; disambiguated instances of cubes were presented rotating in opposite directions at two locations, and the perceptual outcome of interleaved ambiguous cubes rapidly adopted the depicted location–rotation contingency. Here, through an analogous procedure of interleaving ambiguous and disambiguated instances of two shapes that rotated in opposite directions at a single retinal location, we establish a significant shape-contingent bias, which is evident in the perceived rotation direction at onset of the ambiguous objects. Within the cue recruitment paradigm, this constitutes recruitment of shape as a cue to rotation direction (appearance) of the objects. Notably, the shape cue was only recruited when associative learning was conducted using monocular depth cues to specify rotation direction in disambiguated shapes and not when using binocular disparity as a depth cue. Under these conditions, the positive bias was found in twenty-three out of twenty-four subjects (Experiments 3, 4, and 5), with the robustness of the effect underscored by its existence despite high intersubject variability in the two component z-scores. Together with previous findings of only weak recruitment of a translation cue (Haijiang et al., 2006), weak recruitment of vertical disparity (Di Luca, Ernst, & Backus, 2010), and no recruitment of a sound cue and a variety of other visual cues (Jain, Fuller, & Backus, 2010), this suggests that there may be constraints on what cues can be recruited, and we discuss this below.

First though, how can one explain the functionality of an adaptation that appears to be a “positive aftereffect,”” as opposed to the well-documented alternating of percepts that occurs when such bistable stimuli as used here are presented with shorter ISIs (Klink et al., 2008; Leopold et al., 2002; Noest et al., 2007)? Environmental factors that affect the direction of aftereffects following adaptation have previously been discussed (Backus, Garrigan, Haijiang, & Balasubramanian, 2006). Here, one possibility is that longer interstimulus intervals are statistically indicative of an object onset (an exposure to the observer) that is likely to be a new, independent, event. Under these circumstances, we would suggest that it is initially most beneficial to the visual system to rapidly interpret the object in accordance with similar, previously experienced, events rather than evaluate other possible interpretations. In contrast, at shorter interstimulus intervals, one could postulate that the visual system is “behaving” as if the stimulus had never been removed (e.g., perhaps very transiently occluded), in the sense that, statistically, the presentation is likely a continuation of the same stimulus occurrence. Hence, percept switching occurs (as it does during prolonged viewing of such stimuli, e.g., Leopold et al., 2002). In this case, the percept switching may be related to the classical negative aftereffect, thought to be a by-product of inhibitory mechanisms designed to enhance sensitivity to the onset of stimuli that differ from that previously presented. These seemingly contradictory positive and negative aftereffects could potentially be accommodated within a Bayesian framework, such as that presented by Stocker and Simoncelli (2006) where negative aftereffects result from a change in the likelihood function and positive aftereffects result from a change in the prior.

Contingent effects are generally recognized to require joint encoding, at some level, of the parameters to be linked (Barlow, 1990; Braddick, Campbell, & Atkinson, 1978). Previously, we demonstrated that a bias for rotation in depth could be trained to be strongly dependent on retinal location (Haijiang et al., 2006; Harrison & Backus, 2010a). We proposed that this bias was instantiated in hMT+, where local elements of our disambiguated cubes could cause adaptation in neurons that are jointly tuned to direction of motion and disparity (analogous to macaque MT: DeAngelis, Cumming, & Newsome, 1998; DeAngelis & Uka, 2003; Maunsell & Van Essen, 1983a, 1983b). This would then lead to altered responses to leftrightward- and rightward-moving elements in ambiguous cubes. Critically, MT is retinotopically mapped, and so the state of adaptation could be specific to retinal location, explaining the location-dependent bias that we observed. Additional support for hMT+ as the neural locus of the bias is found in the neural response of macaque MT (Bradley, Chang, & Andersen, 1998; Dodd, Krug, Cumming, & Parker, 2001; Grunewald, Bradley, & Andersen, 2002) and human MT+ (Brouwer & van Ee, 2007), which have been shown to predict the perceived direction of motion in structure-from-motion stimuli.

Our subsequent finding that, when training a retinal location bias, disambiguation by a monocular-only cue set led to a stronger bias than disambiguation by disparity and occlusion (Harrison et al., 2011) led us to postulate that the bias for rotation direction could be instantiated entirely in MT but that activation and therefore adaptation in MT were different in response to monocular depth cues than to binocular disparity. However, this explanation would not suffice for the results presented here, as we know of no evidence that MT encodes 3D object shape or indeed any aspect of shape. Additionally, the rotation bias has been shown to transfer between shapes presented at a single retinal location, when trained using binocular disparity (van Dam & Ernst, 2010).

The results we present here lead us to favor an alternative explanation, whereby the perceptual outcome in the ambiguous Necker cube is a function of neural activity in higher cortical areas also, such as regions in parietal and...
frontal cortices that are activated by imagery of 2D rotation (Seurinck, de Lange, Achten, & Vingerhoets, 2010) and regions such as lateral occipital cortex that are activated when volume is attributed to 2D images (Moore & Engel, 2001). These areas might be “trained” when using shapes disambiguated by the monocular-only cue set but not when using a cue set containing binocular disparity, as the strong bottom-up signals provide local evidence of depth and direction of motion without the necessity for further inference. It is the similarity between our monocularly disambiguated and ambiguous stimuli, in terms of the requirement for the visual system to make inferences, which would result in their greater neural overlap for representation of rotation direction. Indeed, our hypothesis is not dissimilar from saying that ambiguous and monocularly disambiguated stimuli have “more in common” than do our ambiguous and disparity-plus-occlusion stimuli.

In a related vein, it is interesting to note the findings of Adams, Kerrigan, and Graf (2010), who found that recalibration of the shading prior occurred when conflicting visual disparity information was presented intermittently after stimulus onset but not when present continuously from stimulus onset. Adams et al. highlighted the importance of an error signal in recalibration, which never occurred when the (strongly unambiguous) disparity information was present at stimulus onset. However, their findings can also be interpreted within our slightly broader framework that it is the inferential, or decisional, stages of perception that are influential in learning, even in the absence of cue conflict such as in our experimental paradigm. In both their and our paradigms, little inference was required when binocular disparity was present at stimulus onset; the perceptual outcome was simply “steamrollered” by the strong bottom-up depth signal, resulting in lower levels of learning.

Of course, one could propose that the type of learning we are eliciting here does not require that the associated properties be jointly encoded in the same neural population. However, this flies in the face of current understanding of contingent effects and leaves unanswered the question of how such learning could be mechanistically implemented.

Hence, we hypothesize that associative learning as demonstrated here will only be found between properties that can be jointly encoded by individual neurons. If two properties (here, shape and rotation direction) are primarily represented in different neural loci, then it will be difficult to train their association, because other properties that are strongly jointly encoded (here, retinal location and rotation direction, in hMT) will become associated to a far greater extent. However, if the neural locus of the unwanted association can be “de-emphasized,” as was achieved here by using monococular depth signals to side step strong bottom-up adaptation of hMT, learned association of other parameters that are only weakly related in terms of neural representation may be possible.

Conclusion

We have demonstrated recruitment of a new cue, object shape, for the appearance of rotating 3D objects. We propose that this is an instance of the cue recruitment effect, first illustrated through a learned rotation bias that was contingent on retinal location (Backus & Haijiang, 2007; Haijiang et al., 2006) or object translation direction (Haijiang et al., 2006) and more recently vertical disparity (Di Luca et al., 2010). These are suggested to be examples of the visual system learning new cues to visual appearance, by adaptation in response to new experiences. The association of rotation direction and shape seen in the present study was only achieved when monocular depth cues were used to depict rotation in depth; shapes disambiguated by binocular disparity did not lead to recruitment of the shape cue. The finding of a new example of cue recruitment, together with the means by which it was found, opens possibilities for further generalization of the phenomenon.

Acknowledgments

This research was supported by Grants NSF BCS-0810944 and NIH R01 EY 013988-07.

Commercial relationships: none.
Corresponding author: Sarah J. Harrison.
Email: sarah.harrison@rhul.ac.uk.
Address: Royal Holloway, University of London, Egham Hill, Egham, Surrey TW20 0EX, UK.

References


