

# Demonstration of cue recruitment: Change in visual appearance by means of Pavlovian conditioning

Qi Haijiang\*, Jeffrey A. Saunders†, Rebecca W. Stone‡, and Benjamin T. Backus\*†‡§

\*Bioengineering Graduate Group, †Department of Psychology, and ‡Neuroscience Graduate Group, University of Pennsylvania, 3401 Walnut Street, C-Wing, 302-C, Philadelphia, PA 19104-6228

Edited by Dale Purves, Duke University Medical Center, Durham, NC, and approved November 9, 2005 (received for review August 5, 2005)

Until half a century ago, associative learning played a fundamental role in theories of perceptual appearance [Berkeley, G. (1709) *An Essay Towards a New Theory of Vision* (Dublin), 1st Ed.]. But starting in 1955 [Gibson, J. J. & Gibson, E. J. (1955) *Psychol. Rev.* 62, 32–41], most studies of perceptual learning have not been concerned with association or appearance but rather with improvements in discrimination ability. Here we describe a “cue recruitment” experiment, which is a straightforward adaptation of Pavlov’s classical conditioning experiment, that we used to measure changes in visual appearance caused by exposure to novel pairings of signals in visual stimuli. Trainees viewed movies of a rotating wire-frame (Necker) cube. This stimulus is perceptually bistable. On training trials, depth cues (stereo and occlusion) were added to force the perceived direction of rotation. Critically, an additional signal was also added, contingent on rotation direction. Stimuli on test trials contained the new signal but not the depth cues. Over 45 min, two of the three new signals that we tested acquired the ability to bias perceived rotation direction on their own. Results were consistent across the eight trainees in each experiment, and the new cue’s effectiveness was long lasting. Whereas most adaptation aftereffects on appearance are opposite in direction to the training stimuli, these effects were positive. An individual new signal can be recruited by the visual system as a cue for the construction of visual appearance. Cue recruitment experiments may prove useful for reexamining of the role of experience in perception.

classical conditioning | perceptual learning | bistable perception

A “visual percept” is the mental representation that one becomes consciously aware of when one’s eyes are open. It specifies the sizes and locations of surfaces and objects, surface properties such as color and visual texture, and the recognized identities of objects (1, 2). To reliably construct percepts from visual signals, the visual system must exploit the statistical relationships between the signals it measures, which are caused by regularities in the world. Percepts are computed automatically, as demonstrated by visual illusions that persist even when one knows they are illusory (3). Here we describe a simple test of the empiricists’ proposal that the visual system actively monitors and refines its functions for mapping signals from the world onto perceptual appearances, consistent with tracking changes in the meanings of those signals over time (3–10). In particular, we tested whether the visual system detects and utilizes new signals as “cues” that it relies upon to build percepts after exposure to novel correlations between input signals.

There is already considerable evidence that the rules that map signals onto percepts can change with experience. For example, it has been reported that experience can affect the following: (i) whether a surface with low luminance is seen as painted dark or merely in a shadow (11); (ii) the mapping between moving 2D images and the 3D representations they evoke (12, 13); (iii) the perceived distance of a U.S. coin, depending on whether an image of fixed size depicts a dime or half-dollar [with the dime’s image appearing closer because a dime’s true size is smaller (14)]; and (iv) the disambiguation of perceptually bistable images (such as might be seen as either

a dog or a chef), which can be made to depend on the spatial position at which the images are presented (15). It is also well established that, during development, visual experience is necessary for the system to learn how to interpret orientation, stereo, motion, and other visual cues (16).

It also seems to be the case that our perceptual systems make remarkably few errors, and perception is in many ways nearly optimal (17). Did evolution endow us with a mechanism to track contingencies between signal measurements (7), so that new signals might be used during perception? Brunswik (3) used the term “ecological validity” to describe a cue’s correlation with some property of the world, and he argued that the visual system tracks changes in validity, as would be needed to optimally resolve ambiguities through probabilistic inference. Our experiments address this issue.

We use the words “cue recruitment” to describe two things: (i) an experimental design and (ii) a type of learning. This same distinction must be made for other types of classical conditioning: one must be careful to distinguish the conditions under which the learning occurs from what (if anything) was learned. In a cue recruitment experiment, a new signal is put into correlation with existing cues during training, and the experimenter measures whether the new signal comes to elicit the same perceptual response that the existing cues elicit. If it does, we describe the learning as cue recruitment by the visual (or other perceptual) system. As noted earlier, we are concerned here specifically with perceptual appearance not visually guided behavior in general.

Cue recruitment is a simple form of associative learning. Why might one expect to see cue recruitment in perception? Because the organism cannot know the true state of the world, it must rely entirely on innate knowledge (6) and patterns of activity in its signal measurements (4). A Bayesian system that learns about the environment must therefore track correlations between signals (18). A simple form of such learning is to recruit a new signal to elicit an existing response (at some level within the nervous system) after noting the temporal co-occurrence of activity caused by the signal and activity that represents the response (9, 19). The contemporary view is that learning by association is a method for acquiring knowledge about the world (20), and in principle this learning could include knowledge about how visually measured signals are related to states of the world.

In a laboratory setting, an arbitrarily chosen visual signal can be put into artificial correlation with cues that reliably evoke some perceptual attribute (21). The perceptual attribute that we paired with the new signal was direction of rotation for a 3D wire-frame cube in a movie (Fig. 1 and Fig. 5, which is published as supporting information on the PNAS web site). The movie supported perceived rotation in either direction, but only one direction of rotation

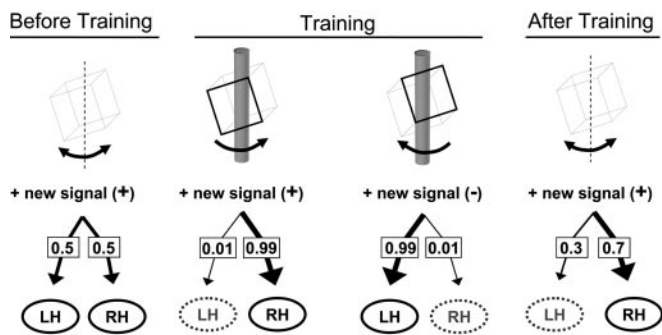
Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: POSN, position of the stimulus within the display (either above or below the center of the display screen); TRANSL, direction (up or down) of the stimulus’ translational motion. SOUND, pitch (high or low) of a two-tone auditory stimulus.

§To whom correspondence should be addressed. E-mail: backus@psych.upenn.edu.

© 2005 by The National Academy of Sciences of the USA



**Fig. 1.** Experimental paradigm to study cue recruitment. Before training, an ambiguous stimulus was equipotential, and a new signal had no effect. During training, stereo and occlusion cues specified the direction of rotation on each trial, and two values of the new signal (–) and (+) were presented in correlation with the two directions of rotation. After training, the new signal disambiguated the rotation [as shown for the (+) signal; the symmetric case for the (–) signal is not shown]. LH, left-handed perceived rotation; RH, right-handed perceived rotation. Typical probabilities for each of these perceptual outcomes are shown in the boxes.

was typically seen on any given trial. Our measure of cue recruitment was the percentage of trials on which the cube was seen to rotate in the direction specified by the new cue.

Our experiments have three features that are methodologically important for reasons that may not be immediately obvious. First, the new signals we used were suprathreshold and easily perceived in their own right, quite apart from any learned effects they had on the targeted perceptual attribute (i.e., rotation direction). This high salience of the novel signals ensured that changes in appearance were not caused by improvement in the visual system's ability to extract the signal itself. Second, the targeted perceptual attribute was also trivially easy to perceive, which was important in making the instructions easy to follow, as we did not want trainees to search for alternative strategies based on the appearance of some other perceptual attribute (such as the new signal itself). A third important feature of our experimental method was that the task required judging appearance relative to an additional, randomized signal (a dot that moved right or left). This indirect task prevented simple response bias (i.e., an association between the new signal and a particular button) from interfering with the trainees' use of button presses to report on appearance.

## Methods

On training trials, preestablished depth cues (stereo and occlusion) were added to the display to disambiguate the perceived rotation, and critically, a new signal was also added. This new signal had one of two values, (+) or (–), depending on the direction of rotation indicated by the preestablished cues. Thus the new signal was contingent on the direction of rotation, so if the system can learn from experience (for example by using Bayes' rule), then the new signal should become a cue for direction of rotation. Every 11th trial was a test trial, on which the preexisting cues were removed, leaving only the new signal to disambiguate the percept.

**Training Cues.** The experimental manipulation was to expose trainees to novel correlations between previously unrelated signals and the direction of 3D rotation of the cube. The three signals tested were as follows: (i) top vs. bottom of screen position (POSN), (ii) up vs. down of translational motion (TRANSL), and (iii) high-pitched vs. low-pitched sound (SOUND). Movies 1–9, which are published as supporting information on the PNAS web site, describe the stimuli in greater detail. To anticipate the results, reliable changes in perception (in the predicted direction) were obtained for POSN and TRANSL but not for SOUND.

**Apparatus and Display.** Red–green stereo anaglyph images (22) were presented by rear-projecting from an LP350 DLP projector (InFocus, Wilsonville, OR) onto  $166 \times 125$ -cm region of a projection screen. The trainee was seated 200 cm away, with eye level aligned to the center of the display area.

**Stimuli, Task, and Procedure.** The stimuli were 8.33-s animations showing a wire-frame cube, covered with random dots, rotating in depth around a vertical axis. A separate horizontally moving probe dot was shown at the top or middle of the display, moving either left or right. The cube was oriented in space so that one of its diagonals was perpendicular to the axis of rotation. The cube was simulated to have 28.9-cm edges (diagonal length, 50 cm), subtending  $14^\circ$  of visual angle. The moving probe dot subtended  $0.74^\circ$ . The rate of rotation of the cube was  $0.754 \text{ rad/s}$  for all three experiments.

There were two types of trials, training trials and test trials. On test trials, orthographic projection was used to render the rotating cube, and images were presented only to the right eye. Under these conditions, structure-from-motion information is ambiguous, leading to a bistable percept of either left-handed or right-handed 3D rotation. On training trials, two depth cues were added to disambiguate the direction of 3D rotation. First, disparities were added to the binocular images, consistent with one of the two 3D interpretations. Second, an opaque cylinder (10 cm by 122 cm) was added to lie along the rotation axis of the cube, so it provided static and dynamic occlusion cues to depth order.

All displays contained a 4-cm-wide fixation square ( $1.15^\circ$ ), positioned at the center of the screen for the POSN and TRANSL experiments or 58 cm above the center for the SOUND experiment. Trainees were instructed to fixate on the square. Eye position was not monitored because the logic of the experiment does not require that fixation was accurate.

The task was to judge whether the probe dot moved in the same direction as the front or the back of the rotating cube and to press 2 or 8 on a keypad, respectively. This task ensured that the trainee's button presses were uncorrelated with the new signal over the course of the session, to prevent motor response bias from contributing to the measured effects. This task felt natural and was easy to do.

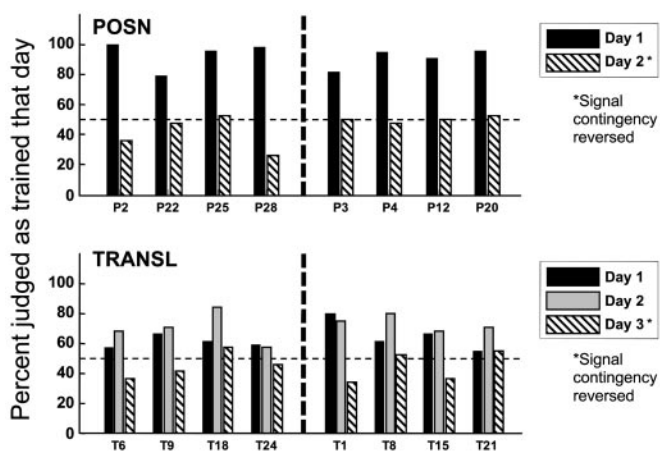
Feedback (a visually presented smiling or frowning cartoon face) was provided immediately after the trainee responded to indicate whether the response was correct. On test trials, the feedback was always positive, regardless of the response. On training trials, if the response was incorrect, a delay of 6 s before the next trial was imposed as a penalty. We instructed trainees that it was possible that the direction of rotation could appear to flip in the middle of a trial, and instructed them to indicate whether such a flip occurred by first pressing 0, and then 2 or 8, according to their initial impression. Based on this measure, spontaneous reversals appeared to be rare, occurring on between 0% and 1.5% of trials. These trials were excluded from analysis.

Sessions lasted 45–60 min, during which trainees responded to a total of 440–470 trials. The trials were arranged in a repeating sequence of 10 training trials followed by 1 test trial. Trials were self-paced, with built-in breaks every 50 trials. The cue-rotation pairing was counterbalanced across trainees, to control for the (unlikely) possibility that trainees might have a preexisting bias to perceive left- or right-handed rotation contingent on the trained signal.

**Trainees (Subjects).** Trainees who performed multiple sessions did so on consecutive days. All trainees were naive to the purposes of the experiment and were paid for participating. Trainees gave informed consent in accordance with a protocol approved by the Institutional Review Board panel of the University of Pennsylvania.

See *Supporting Methods*, which is published as supporting information on the PNAS web site, for additional methods.





**Fig. 4.** Persistence of learning into the next day. The data plotted are the percentage of trials on which the perceived rotation direction agreed with the new signal, according to its contingency that day. Each pair (Upper) or triplet (Lower) of bars represents data from one trainee. A bar height  $>50\%$  means the cue was effective in the predicted direction. If trainees started each day in the same state, bar heights would be the same on all days for a given trainee, which was not the case. Instead, continued training on day 2 (TRANSL) resulted in additional bias, and reversed training on day 2 (POSN) or day 3 (TRANSL) resulted in less bias (in the new predicted direction) as compared with that seen on day 1. *T* tests for the between-day differences (across trainees) were as follows: (i) for POSN,  $H_0$  (null hypothesis): day 1 = day 2 was rejected at  $P < 0.001$ ; (ii) for TRANSL,  $H_0$ : day 1 = day 2 was rejected at  $P < 0.05$ ; and (iii) for TRANSL,  $H_0$ : day 2 = day 3 was rejected at  $P < 0.001$ . Only those trainees who completed all sessions of their experiment are shown.

it is possible that these signals are special. Additional experiments are clearly needed to measure cue recruitment for a variety of additional signals and perceptual attributes to determine the generality of these findings.

Second, whether the chosen signals are special or not, if the total amount of learning in a cue recruitment experiment is small (i.e., if training provides only a small cue-contingent bias in appearance), then a sensitive measure of learning is needed. It was for this reason, following the logic of Wallach and Austin (15), that we used a perceptually bistable stimulus to demonstrate changes in appearance. For such a stimulus, the learning needed only to bias the percept by an amount that was large relative to whatever factors normally resolve the competition between the two perceived forms.

Third, contingent adaptation aftereffects (CAA) are ubiquitous phenomena in visual experiments (24–27). In our experiments, training stimuli contained motion that was correlated with binocular disparity. In this situation, a CAA would be expected to bias the appearance of binocular zero-disparity stimuli in a direction opposite to the effect of cue recruitment. We avoided this potential problem by using monocular stimuli for which the disparity was underlined rather than binocular stimuli with zero disparity. In other words, test stimuli were ambiguous because binocular disparity was undefined, rather than because disparities were set to zero as in previous experiments (28, 29). Nawrot and Blake (28) have shown that this stereomotion CAA can also be eliminated by training with stimuli in which rotation direction is specified by pictorial cues (such as occlusion) alone. However, we were unable to take this approach because perceived rotation on training trials was not always forced in the correct direction when we used only occlusion as the depth cue.

In principle, removal of the preexisting depth cues from the test stimuli could also help because their presence, even if their values were such that they did not bias perception, would be expected to reduce the relative effectiveness of a newly learned cue as a consequence of weighted averaging with those cues (30). Reducing the number of cues reduces the number of factors involved in

constructing the percept, thereby increasing the weight given to the new cue.

What actually changed in the nervous system to allow the new cue to bias the percept? It is often quite challenging to determine which internal variables are associated during learning to account for a change in responses to stimuli (31). However, in our experiment, a direct association between the new cue and either stereo or occlusion can be ruled out. The role of these preexisting cues was to disambiguate the wire-frame cube, but perceived rotation direction also depended on whether the near part of the cube moved left or right. Thus, training the observer with static stereo and occlusion cues would not have led to bias in perceived rotation direction in our experiment, because there would be no physical difference between the two types of training stimulus. Thus, training must have affected perceived rotation direction in some other way, such as eliciting activity in neurons that represent both disparity and motion or directly modifying a higher-level neural representation of rotation in depth. In the case of the POSN cue, a promising start would be to separately measure learning that depends on retinal position and learning that depends on position in the world, as these two signals were confounded in the present study.

Why did the learning rates differ for the three different signals we used? Learning rates in our experiment were presumably under the control of the system, because all of the relevant signals were suprathreshold. Our working hypothesis is that the system will learn rapidly or slowly, according to its (implicit) internal belief as to whether learning is appropriate in the given situation. Differences in learning rates have been found in other domains. For example, rats learn to associate illness with flavors more easily than with lights or sounds (ref. 32; see also refs. 33 and 34). Failure to learn the sound cue in our experiment may reflect a similar predisposition to not learn certain correlations. The sounds we used may have been judged by the system as not plausibly bearing on the cube's rotation direction, perhaps because the auditory and visual stimuli did not have simultaneous onsets. Simultaneous sounds can bias visual percepts (35, 36), and for perception, cues that are mutually relevant normally would be measured by the system at the same time, so learning mechanisms should exploit this fact.

The results are notable for lasting into the next day (in fact, during pilot experiments, two of the authors could not be easily retrained even after several weeks). The long-lasting nature of these effects distinguishes them from most negative adaptation aftereffects (which are short-lived). An exception, however, is the McCollough effect (37), a negative adaptation aftereffect that can last many days. The McCollough effect has also been described in instances of associative learning (38). To induce the McCollough effect, two images are viewed in alternation for several minutes: vertical black-and-red stripes alternate every few seconds with horizontal black-and-green stripes. After this training period, vertical black-and-white stripes look black-and-greenish, and horizontal black-and-white stripes look black-and-pink.

Learning in the McCollough effect appears to be limited to specific signals such as orientation and spatial frequency that are known stimulus features for early visual coding (39, 40). Negative adaptation aftereffects are expected for processes of recalibration: in this case, the system sticks to its original belief that color and orientation in the world are uncorrelated (41–43) and modifies itself in the direction that would serve, under that assumption, to eliminate bias. If so, the McCollough effect is long-lasting because the training stimuli are treated by the system as strong evidence of the need for recalibration; after this recalibration occurs, it could take a long time for recalibration by natural stimuli to remove the last traces of bias induced by the training stimuli.

More generally, the visual system has two functionally distinct problems: keeping itself calibrated and tracking the meanings of signals from the world (e.g., cue recruitment). Persistent exposure to stimuli with signals in novel correlations should be expected to elicit adaptations in both of these functions. We

would associate negative aftereffects with recalibration and positive aftereffects with changes in signal utilization. Many instances of recalibration are known; is there any evidence of the latter, excepting the experiments we report here? A number of other phenomena are closely related to cue recruitment in that they directly demonstrate a positively correlated change in appearance, as a result of learned associations under laboratory control. These phenomena include the following:

- Likely previous instances of cue recruitment. Fieandt (3, 11) used classical conditioning to change the perceived color (albedo) of a surface, as measured in a matching task: at first the surface appeared to be well lit and painted dark, but through experience it came to appear in a shadow and painted light. Fieandt also reported that a new visual cue (a button) and a new sound cue (a bell) were both somewhat effective in modulating how the stimulus was perceived. These early studies reported that learned contextual cues under experimental control could bias appearance. Howells (44) reported that a tone became effective at biasing perceived color, as measured in matching experiments.
- Changes in cue weights. If two visual cues are in conflict and tactile information consistently agrees with one of them during training, then that visual cue will be given greater weight (45, 46). Note that no new cues are recruited in this case.
- Modulation of cue weights. A third visual cue (or “context”) can be recruited to modulate the weights of the two conflicting visual cues, and some third cues are learned faster than others (47). This modulation demonstrates recruitment of an “auxiliary cue” (30).
- Changes in the assumptions used to interpret sense data. Experience with a visual scene can systematically bias how shading cues are used to infer the shapes of new 3D objects in the scene (48).
- 3D interpretations of images. A given 2D image can be made to evoke two or more 3D interpretations, as a result of repeated association between the image and information that specifies the 3D object (12, 13). This effect is clearly associative learning. It differs from cue recruitment in that the laws that relate 3D shapes to their 2D projections are largely known by the visual system before the experiment and because no new signal acquired the ability to control the percept (became a cue) on a trial-by-trial basis. It would be interesting to train on both of two 3D shapes, contingent on some third signal (such as color); we would expect the new signal to become a cue that controls, on a trial-by-trial basis, which of two 3D shapes is seen.
- Disambiguation of a bistable percept of object identity. Presentation position can bias the perceived identity of an ambiguous figure after association between position and unambiguous (but configurally similar) figures (15). We interpret this result as cue recruitment. The authors did not determine whether the learning was long lasting.
- Conditioned afterimages. It has been reported that visual experiences (in the absence of visual input) can be elicited as conditioned responses to tones after pairing images with tones (49). This finding would demonstrate a conditioned visual response but not one that occurs during normal vision.

These findings, together with our own, suggest that appearance can be modified by experience, and indeed that perception is stable not because the visual system is stable, but because it continually adapts itself to a world that is stable. Associative learning appears to be a mechanism by which it does so. Associative learning is significant for a wide range of human responses, including motor behaviors, glandular secretions and other physiological responses, emotional responses, verbal learning, category learning, judgments of event-relatedness, reasoning, and a variety of effects in social psychology (38, 50). Historically, one of the first uses of associative learning as a concept was to explain perception (4). Why then was it given relatively little attention in the past 50 years?

The failure to find classical conditioning of perceptual responses during the 1950s was interpreted as indicative of the perceptual system’s high degree of sophistication: It was concluded that classical conditioning did not affect perception because simple stimulus-to-response mappings do not give an appropriate description of the perception (51). However, the view that classical conditioning is implemented by a direct mapping from stimuli to responses has since been rejected in favor of the view that learning in classical conditioning represents change in the organism’s representation of contingency (52).

Gibson (53) defined “perceptual learning” as follows:

Any relatively permanent and consistent change in the perception of a stimulus array after practice or experience with this array will be considered perceptual learning.

This definition matches the generic meaning of the words perceptual learning. However, in the absence of positive experimental findings (51), many researchers assumed that perceptual learning did not include associative learning. In 1955, there was no straightforward way to challenge Gibson and Gibson (54) when they denied that changes in the utilization of suprathreshold signals, as took place in our experiments, might have importance for perception.<sup>†</sup> Since then, many researchers have defined perceptual learning in terms of discrimination performance (55–59), which excludes any effects of associative learning on appearance. A recent textbook on perceptual learning (60) states:

Contrary to associative learning, perceptual learning does not bind together two processes that were separated but improves discrimination between stimuli that could not be discriminated before the learning.

Current textbooks on perception certainly discuss appearance, but they make no reference to associative learning.

Brunswik’s (3) theory, which includes conditioned learning of the ecological validities of cues, has found wide application in the field of judgment and decision making (61), and other aspects of his theory are starting to find use in perception and computer vision (62–64). Many others in the modern literature, besides Brunswik, have argued that associative learning must play a role in perception. Hebb (9) argued explicitly for such a role and developed ideas for implementing it neurally. Barlow (25) noted the “astonishingly deep knowledge of the normal patterns of associated activation our visual system possesses and automatically uses,” and argued that perceptual learning must include mechanisms for detecting statistical regularities in the visual environment (see also ref. 65). Indeed, it is difficult to see how one’s visual system could learn this statistical structure without the ability to detect novel correlations among signal measurements. Wallach (66) proposed that the calibration of perceptual estimators occurs through association. Purves and colleagues have documented many instances in which the magnitude of a perceptual illusion is well correlated with a statistic in natural scenes, which strongly suggests a role for associative learning in the construction of appearance (67). Baron (68) and Fiser and Aslin (69–71) described several instances of learned associations between visually presented signals that improved recognition performance, although it is not possible to determine whether these improvements were mediated by changes in appearance. Wallis and Bülhoff (72) reviewed a variety of evidence in favor of associative learning in object recognition, and Geisler and Diehl (73) discussed the evolution of facultative adaptations for perception, i.e., mech-

<sup>†</sup>Gibson and Gibson (54) suggested that, “Perceptual learning, then, consists of responding to variables of physical stimulation not previously responded to.” And later, “True perceptual learning experiments are limited to those concerned with discrimination,” (meaning that the only controlled perceptual learning experiments to date were discrimination experiments).

anisms that allow the organism to adapt appropriately to the statistics of its environment during its lifetime.

A useful distinction between cue recruitment and learning to discriminate may be the limiting factor on the learning rate. For cue recruitment, the rate at which an organism learns to use a signal as a new cue must necessarily reflect the organism's implicit or explicit belief about whether current observations of correlation have predictive value about what the signal will mean in the future. It is hard to see why learning to discriminate would be limited by this factor, given that the cost of learning is presumably much smaller than the gains to be had from improved performance. More likely, learning to discriminate proceeds at the fastest rate possible and is limited by the system's ability to create a better template (to extract a feature through proper combination of many relevant sensors, so as to achieve a better signal-to-noise ratio). Detection is not a limiting factor in cue recruitment because the new signal is already suprathreshold: Learning can occur within very few trials, if the system so chooses.

A second useful distinction may be that cue recruitment entails a change in the utilization of visual signals, without any necessary increase in the number of percepts the organism is capable of constructing. This difference is easiest to see when a small set of stimuli is either differentiated, giving rise to more percepts (as when developing an appreciation for wine), or else simply remapped, so that a given stimulus comes to look like something else (as with our rotating cubes).

Cue recruitment and learning to discriminate might interact in interesting ways. In some cases, learning to discriminate is greatly facilitated by brief practice with easy stimuli (74, 75), and subjective reports are that the practice causes previously indiscriminable stimuli to look different (74). These observations lead one to suspect that some internally represented cue became associated

with a perceptual attribute during viewing of the easy stimuli, which generalized to the more difficult stimuli.

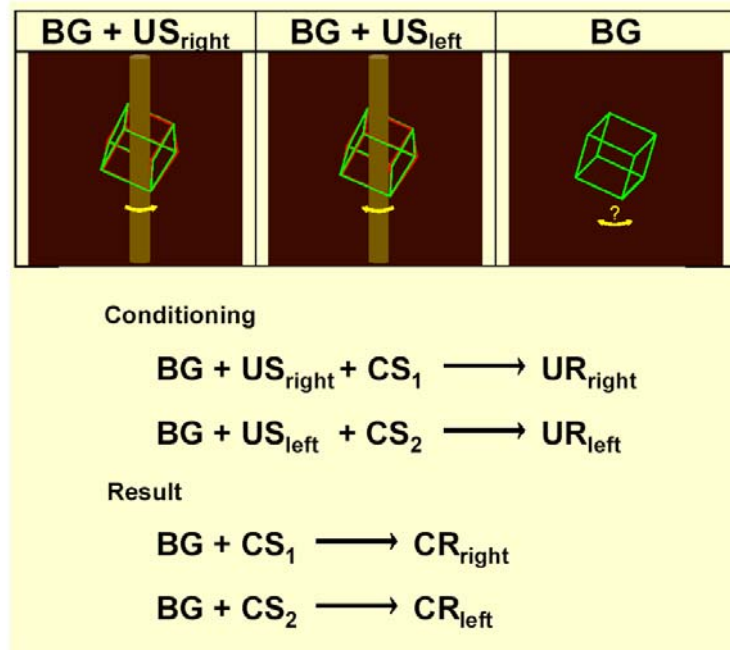
Our experiments confirmed two cases of cue recruitment for which the learning rate was nonzero. When combined with previous theoretical and experimental work, our results provide strong evidence for associative learning in the processes that construct percepts. It remains to be seen just how widely such learning can be shown to occur; the current work provides only a start. Some interesting implications are that learning in perceptual mechanisms can now be studied in isolation, independent of other learning that might also occur at other stages during visually guided behavior, and that intelligent computer vision systems may have to be equipped with mechanisms for learning new associations between signals.

In summary, cue recruitment experiments provide an additional tool for studying perceptual appearance, by measuring the rate of learning at which an arbitrarily chosen signal comes to control an arbitrarily chosen perceptual attribute. Until many such rates have been measured and a pattern of results established, the presumption should probably return to what it was 50 years ago: That associative learning [what Bishop Berkeley (4) called the "habitual or customary connexion between two sorts of ideas"] is of ongoing importance for perception, even in adults.

We thank Jonathan Baron, David Brainard, Jacob Nachmias, Larry Palmer, Virginia Richards, Martin Seligman, and other readers from the University of Pennsylvania for their comments on drafts of this manuscript. We thank Clark Ohnesorge for bringing the work of Egon Brunswik to our attention, Marianne Promberger for helping us read von Fieandt (1936), and Richard Pater for discussion and comments. This work was supported by National Institutes of Health Grants R01-EY013988 and P30-EY001583 and by the University Research Foundation at the University of Pennsylvania.

1. von Helmholtz, H. (1910) *Handbuch der Physiologischen Optik*, trans. Optical Society of America (1925) *The Perceptions of Vision*, Treatise on Physiological Optics (Optical Society of America, New York), Vol. 3 (German).
2. Gibson, J. J. (1979) *The Ecological Approach to Visual Perception* (Houghton Mifflin, Boston).
3. Brunswik, E. (1956) *Perception and the Representative Design of Psychological Experiments* (Univ. of California Press, Berkeley), pp. 92, 96, 123–131.
4. Berkeley, G. (1709) *An Essay Towards a New Theory of Vision* (Dublin), 1st Ed.
5. Condillac, E. B. A. d. (1754) *Traite on Sensations*, trans. Philip, F. (1982) *A Treatise on Systems, A Treatise on Sensations*, Philosophical Works of Etienne Bonnot, Abbe de Condillac (Erlbaum, Hillsdale, NJ), Vol. 1.
6. Kant, I. (1781) *Critique of Pure Reason*; trans. Meiklejohn, J. M. D. (1990) in *Critique of Pure Reason*, Philosophical Classics (Dover, New York).
7. von Helmholtz, H. (1878) *The Facts of Perception*; reprinted (1971) in *Selected Writings of Hermann Helmholtz*, ed. Kahl, R. (Wesleyan Univ. Press, Middletown, CT).
8. James, W. (1890) *The Principles of Psychology*; reprinted (1983) (Harvard Univ. Press, Cambridge, MA), pp. 722–912.
9. Hebb, D. O. (1949) *Organization of Behavior* (Wiley, New York).
10. Ames, A. J. (1953) in *Vision and Action*, ed. Ratner, S. (Rutgers Univ. Press, New Brunswick, NJ), pp. 251–274.
11. von Fieandt, K. (1936) *Archiv für die Gesamte Psychologie* **96**, 467–495.
12. Wallach, H., O'Connell, D. N. & Neisser, U. (1953) *J. Exp. Psychol.* **45**, 360–368.
13. Sinha, P. & Poggio, T. (1996) *Nature* **384**, 460–463.
14. Epstein, W. (1965) *Am. J. Psychol.* **78**, 120–123.
15. Wallach, H. & Austin, P. (1954) *Am. J. Psychol.* **67**, 338–340.
16. Simmons, K. (1993) *Early Visual Development: Normal and Abnormal* (Oxford Univ. Press, New York).
17. Kersten, D., Mamassian, P. & Yuille, A. (2004) *Annu. Rev. Psychol.* **55**, 271–304.
18. Knill, D. C. & Richards, W. (1996) *Perception as Bayesian Inference* (Cambridge Univ. Press, Cambridge, U.K.).
19. Pavlov, I. P. (1927) *Conditioned Reflexes* (Oxford Univ. Press, Oxford).
20. Rescorla, R. A. (2003) *Span. J. Psychol.* **6**, 185–195.
21. Smedslund, J. (1955) *Multiple Probability Learning* (Oslo Univ. Press, Oslo).
22. Mulligan, J. B. (1986) *Perception* **15**, 27–36.
23. Gallistel, C. R., Fairhurst, S. & Balsam, P. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 13124–13131.
24. Mayhew, J. E. & Anstis, S. M. (1972) *Percept. Psychophys.* **12**, 77–85.
25. Barlow, H. (1990) *Vision Res.* **30**, 1561–1571.
26. Durgin, F. H. & Proffitt, D. R. (1996) *Spat. Vis.* **9**, 423–474.
27. Blaser, E. & Domini, F. (2002) *Vision Res.* **42**, 273–279.
28. Nawrot, M. & Blake, R. (1991) *Percept. Psychophys.* **49**, 230–244.
29. Bradley, D. C., Chang, G. C. & Andersen, R. A. (1998) *Nature* **392**, 714–717.
30. Landy, M. S., Maloney, L. T., Johnston, E. B. & Young, M. (1995) *Vision Res.* **35**, 389–412.
31. Rescorla, R. A. (1980) *Pavlovian Second-Order Conditioning: Studies in Associative Learning* (Erlbaum, Hillsdale, NJ).
32. Garcia, J. & Koelling, R. A. (1966) *Psychon. Sci.* **4**, 123–124.
33. Wilcoxon, H. C., Dragoin, W. B. & Kral, P. A. (1971) *Science* **171**, 826–828.
34. Seligman, M. E. P. (1970) *Psychol. Rev.* **77**, 406–418.
35. Sekuler, R., Sekuler, A. B. & Lau, R. (1997) *Nature* **385**, 308 (lett.).
36. Ecker, A. J. & Heller, L. M. (2005) *Perception* **34**, 59–75.
37. McCollough, C. (1965) *Science* **149**, 1115–1116.
38. Siegel, S. & Allan, L. G. (1996) *Psychon. Bull. Rev.* **3**, 314–321.
39. Sigel, C. & Nachmias, J. (1975) *Vision Res.* **15**, 829–836.
40. Allan, L. G., Siegel, S., Kulatunga-Moruzi, C., Eissenberg, T. & Chapman, C. A. (1997) *Percept. Psychophys.* **59**, 1327–1334.
41. Dodwell, P. C. & Humphrey, G. K. (1990) *Psychol. Rev.* **97**, 78–89.
42. Allan, L. G. & Siegel, S. (1993) *Psychol. Rev.* **100**, 342–346; discussion 347–350.
43. Bedford, F. L. (1995) *Cognition* **54**, 253–297.
44. Howells, T. H. (1944) *J. Exp. Psychol.* **34**, 87–103.
45. Ernst, M. O., Banks, M. S. & Bühlhoff, H. H. (2000) *Nat. Neurosci.* **3**, 69–73.
46. Atkins, J. E., Fiser, J. & Jacobs, R. A. (2001) *Vision Res.* **41**, 449–461.
47. Jacobs, R. A. & Fine, I. (1999) *Vision Res.* **39**, 4062–4075.
48. Adams, W. J., Graf, E. W. & Ernst, M. O. (2004) *Nat. Neurosci.* **7**, 1057–1058.
49. Davies, P., Davies, G. L. & Bennett, S. (1982) *Perception* **11**, 663–669.
50. Lieberman, D. A. (2000) *Learning: Behavior and Cognition* (Wadsworth, Stamford, CT), 3rd Ed.
51. Drever, J. (1960) *Annu. Rev. Psychol.* **11**, 131–160.
52. Rescorla, R. A. (1988) *Am. Psychol.* **43**, 151–160.
53. Gibson, E. J. (1963) *Annu. Rev. Psychol.* **14**, 29–56.
54. Gibson, J. J. & Gibson, E. J. (1955) *Psychol. Rev.* **62**, 32–41.
55. Sagi, D. & Tanne, D. (1994) *Curr. Opin. Neurobiol.* **4**, 195–199.
56. Fahle, M. & Morgan, M. (1996) *Curr. Biol.* **6**, 292–297.
57. Doshier, B. A. & Lu, Z. L. (1999) *Vision Res.* **39**, 3197–3221.
58. Gibson, E. J. & Pick, A. D. (2000) *An Ecological Approach to Perceptual Learning and Development* (Oxford Univ. Press, New York).
59. Fine, I. & Jacobs, R. A. (2002) *J. Vis.* **2**, 190–203.
60. Fahle, M. (2002) in *Perceptual Learning*, eds. Fahle, M. & Poggio, T. (MIT Press, Cambridge, MA), pp. ix–xii.
61. Hammond, K. R. & Stewart, T. R. (2001) *The Essential Brunswik* (Oxford Univ. Press, New York).
62. Martin, D., Fowlkes, C., Tal, D. & Malik, J. (2001) *Proc. Int. Conf. Comput. Vis.* **2**, 416–423.
63. Elder, J. H. & Goldberg, R. M. (2002) *J. Vis.* **2**, 324–353.
64. Geisler, W. S. & Kersten, D. (2002) *Nat. Neurosci.* **5**, 508–510.
65. Barlow, H. (2001) *Behav. Brain Sci.* **24**, 602–607; discussion 652–671.
66. Wallach, H. (1985) *Am. Psychol.* **40**, 399–404.
67. Purves, D. & Lotto, R. B. (2003) *Why We See What We Do: An Empirical Theory of Vision* (Sinauer, Sunderland, MA).
68. Baron, J. (1974) *Can. J. Psychol.* **28**, 37–50.
69. Fiser, J. & Aslin, R. N. (2001) *Psychol. Sci.* **12**, 499–504.
70. Fiser, J. & Aslin, R. N. (2002) *J. Exp. Psychol. Learn. Mem. Cognit.* **28**, 458–467.
71. Fiser, J. & Aslin, R. N. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 15822–15826.
72. Wallis, G. & Bühlhoff, H. (2002) in *Perceptual Learning*, eds. Fahle, M. & Poggio, T. (MIT Press, Cambridge, MA), pp. 299–315.
73. Geisler, W. S. & Diehl, R. L. (2002) *Philos. Trans. R. Soc. London B* **357**, 419–448.
74. Rubin, N., Nakayama, K. & Shapley, R. (1997) *Curr. Biol.* **7**, 461–467.
75. Ahissar, M. & Hochstein, S. (2004) *Trends Cogn. Sci.* **8**, 457–464.

## Supporting material for Haijiang et al. ("Cue recruitment") PNAS 2006 (online 2005)



**Fig. 5.** Conditioning procedure (experimental design). The experiments can be described as classical conditioning experiments, as follows. The background (BG) included the testing apparatus and the ambiguous projection of a rotating cube. The unconditioned stimuli (US<sub>right</sub> and US<sub>left</sub>) were stereo and occlusion cues, which were added to the display so as to make it consistent with left-handed or right-handed rotation, respectively. The unconditioned response was the perceived direction of rotation (left-handed or right-handed rotation, UR<sub>right</sub> and UR<sub>left</sub>). The conditioned stimuli were the two values of the new signal (CS<sub>1</sub> and CS<sub>2</sub>). For the POSN experiment, CS<sub>1</sub> and CS<sub>2</sub> were stimulus positions (at the top and bottom of the display screen, respectively); for the TRANSL experiment, CS<sub>1</sub> and CS<sub>2</sub> were movement in an upward or downward direction; and for the SOUND experiment, CS<sub>1</sub> and CS<sub>2</sub> were a high- and low-pitched tone sequence. Before training, BG elicits UR<sub>right</sub>- or UR<sub>left</sub>-independent of whether the trial also contains CS<sub>1</sub> or CS<sub>2</sub>. After training, BG + CS elicits right-handed perceived rotation (CR<sub>right</sub>) or left-handed perceived rotation (CR<sub>left</sub>), depending on the contingency between the CS and the US during training.

Supporting Movie 1

**Movie 1.** Training trial with "top" POSN cue. The rotation direction of the cube is specified by stereo cues (red–green anaglyph for green filter over right eye) and occlusion cues (vertical post). In this movie, these cues specify that the front of the cube moves leftward. The box is a fixation marker. The moving dot is part of the trainee's task (see *Stimuli, Task, and Procedure*). The cube and occluder were horizontally centered on the screen, and positioned vertically 33 cm above or below the center of the screen. On

training trials, the cube's vertical position was correlated with the direction of 3D rotation. The initial rotational pose of the cube on training trials was randomized. The probe dot moved horizontally at a speed of 19 cm/s ( $5.4^\circ/s$ ) in the plane of the screen, along the screen's horizontal midline ( $\pm 36$ -cm range).

Supporting Movie 2

**Movie 2.** Training trial with "bottom" POSN cue. Stereo and occlusion specify that the front of the cube moves rightward.

Supporting Movie 3

**Movie 3.** Test trial with "bottom" POSN cue. Stereo and occlusion cues have been removed and rotation direction is ambiguous (green image is visible only to right eye). The cube had a constant neutral starting pose on test trials.

Supporting Movie 4

**Movie 4.** Training trial with "downward" TRANSL cue. The cube and occluder were presented 30 cm to the right or left of the center. While rotating, the cube moved vertically along its axis of rotation at a speed of 8.4 cm/s ( $2.4^\circ/s$ ), starting from a random vertical position and "wrapping" around when it moved off the screen. On training trials, the direction of vertical translation was correlated with the direction of 3D rotation. (The stimulus appeared on the left or right side of the screen, unrelated to 3D rotation.) The initial pose of the cube was randomized on both training trials and test trials. The probe dot moved horizontally at a speed of 7.9 cm/s ( $2.3^\circ/s$ ) along the screen's horizontal midline. To prevent the dot from overlapping with the moving cube, it was presented on the opposite side of the display. Its starting position was chosen randomly, between 0 and 30 cm from the center.

Supporting Movie 5

**Movie 5.** Training trial with "upward" TRANSL cue. See Movie 4 legend for details.

Supporting Movie 6

**Movie 6.** Test trial with "downward" TRANSL cue. See Movie 4 legend for details.

Supporting Movie 7

**Movie 7.** Training trial with "high-pitched" SOUND cue. In the SOUND experiment, the cube was always presented at the center of the screen. A sound cue was also presented, starting 600 ms before the appearance of the cube and continuing for 2 s. It consisted of a repeated two-tone sequence, either increasing low-pitched (300 ms at 200 Hz followed by 400 ms at 400 Hz) or decreasing high-pitched (300 ms at 2800 Hz followed by 400 ms at 2500 Hz). The tone sequence was correlated with the 3D rotation of the cube. The initial



pose of the cube was randomized on training trials and was always the same (a left–right symmetric pose) on test trials. The probe dot moved at 19 cm/s (5.4°/s) along a horizontal axis 60 cm above the center of the screen. Its initial position was chosen randomly ( $\pm 36$ -cm range).

#### Supporting Movie 8

**Movie 8.** Training trial with "low-pitched" SOUND cue. See Movie 7 legend for details.

#### Supporting Movie 9

**Movie 9.** Test trial with "high-pitched" SOUND cue. See Movie 7 legend for details.

### **Supporting Methods**

**Apparatus and Display.** The projector had a pixel resolution of  $1,024 \times 768$  and a refresh rate of 60 Hz. To present binocular images, we used red/green glasses and color anaglyph images for the animation frames. The anaglyphs were generated so that the ghosting because of incomplete binocular separation consisted primarily of differences in background chromaticity rather than in luminance (1). Both edges and points were antialiased, using a custom-generated look-up table to compute blended anaglyph color values for partially filled pixels. To render stereoviews, a constant interocular separation of 6.2 cm was assumed.

**Stimuli.** On test trials, orthographic projection caused the cubes to appear slightly distorted in shape on test trials. Unlike the test trials, the images on training trials were perspective projections. On half of the training trials, we reversed the perspective cue so that the training object had the appearance (and was geometrically correctly displayed to be) a distorted cube. This was done so that test trials would not stand out as being different from training trials (otherwise, only test-trial stimuli would appear distorted). Under these conditions, the perceived 3D-rotation direction was reliably in the direction specified by disparity and occlusion. This outcome was confirmed in the data: Responses on training trials (after initial practice of 30 trials) were as specified by the stimulus (median, 99% correct; range, 91–100% correct).

**Trainees.** Thirty-six trainees volunteered for the study. Of these volunteers, 8 failed to pass the screening test (for stereo vision and making  $>90\%$  correct responses on training trials) and were not tested further. Of the remaining 28 trainees, 9 participated in the POSN study, 11 in the TRANSL study, and 8 in the SOUND study. In the POSN study, 8 of the trainees performed two sessions, and in the TRANSL study, 8 of the trainees performed three sessions.

1. Mulligan, J. B. (1986) *Perception* **15**, 27–36.