

Available online at www.sciencedirect.com

Vision Research xxx (2005) xxx–xxx

**Vision
Research**

www.elsevier.com/locate/visres

Shape-specific perceptual learning in a figure-ground segregation task

Do-Joon Yi ^{a,*}, Ingrid R. Olson ^b, Marvin M. Chun ^c

^a Yale University, USA

^b University of Pennsylvania, USA

^c Yale University, USA

Received 28 July 2005; received in revised form 6 September 2005

Abstract

What does perceptual experience contribute to figure-ground segregation? To study this question, we trained observers to search for symmetric dot patterns embedded in random dot backgrounds. Training improved shape segmentation, but learning did not completely transfer either to untrained locations or to untrained shapes. Such partial specificity persisted for a month after training. Interestingly, training on shapes in empty backgrounds did not help segmentation of the trained shapes in noisy backgrounds. Our results suggest that perceptual training increases the involvement of early sensory neurons in the segmentation of trained shapes, and that successful segmentation requires perceptual skills beyond shape recognition alone.

© 2005 Published by Elsevier Ltd.

Keywords: Perceptual learning; Figure-ground; Segmentation; Shape; Search

1. Introduction

A fundamental step in object recognition is to segregate retinal images into figure and ground. Successful segmentation of an object can be achieved through bottom-up processing based on regional contrast defined by color, luminance, or texture. In a cluttered visual scene, however, segmentation benefits greatly from top-down information, such as expectancy or knowledge of contents. A classic example of this is Richard Gregory's Dalmatian dog image, which looks like a bunch of splotches to the naïve. Viewers who have previously experienced this degraded scene can promptly delineate the area corresponding to the dog against the background of similar black-and-white blobs. Such degradation is a common occurrence in everyday vision; due to occlusion, poor illumination, and viewpoint changes, object boundaries are frequently lost or corrupted by spurious edges. Thus, computational approaches to object recognition consider figure-ground segregation as a

highly challenging task that requires the aid of top-down knowledge (see Kersten, Mamassian, & Yuille, 2004).

One source of top-down information comes from perceptual training. Training can help human observers improve their ability to segment novel objects from cluttered visual scenes. For instance, training enabled observers to perceive a closed contour of similarly oriented Gabor elements in a background of Gabor elements (Kourtzi, Betts, Sarkheil, & Welchman, 2005) or to recognize a mechanochemical structure ('digital embryo') from a clutter of similar structures (Brady & Kersten, 2003). Interestingly, those learning effects do not simply reflect improvements in general-purpose perceptual skills; learning was typically specific to the trained objects and did not completely transfer to other untrained objects (see also, Kovacs, Kozma, Feher, & Benedek, 1999; Li & Gilbert, 2002). Such specificity is observed under a variety of conditions mimicking real world situations as well, such as when searching for a knife from an X-ray images of packed luggage (McCarley, Kramer, Wickens, Vidoni, & Boot, 2004) or when trying to diagnose an abnormality in medical X-rays (Sowden, Davies, & Roling, 2000).

Specificity has been extensively studied in perceptual learning under the basic assumption that the dependence of

* Corresponding author. Tel.: +1 203 432 9621; fax: +1 203 432 7172.
E-mail address: dojoon.yi@yale.edu (D.-J. Yi).

learning on specific stimulus parameters provides a behavioral probe into the underlying neuronal mechanisms (Karni & Bertini, 1997). In perceptual learning of fundamental dimensions of visual analysis, such as line orientation (Shiu & Pashler, 1992; Vogels & Orban, 1985), hyperacuity (Fahle & Edelman, 1993; Poggio, Fahle, & Edelman, 1992), motion (Ball & Sekuler, 1982, 1987; Vaina, Belliveau, des Roziers, & Zeffiro, 1998), texture (Ahissar & Hochstein, 1993, 1997; Karni & Sagi, 1991, 1993), and primitive shape (Sigman & Gilbert, 2000), the results, with some exceptions, revealed two types of specificity in perceptual learning: (1) improved performance restricted to the trained stimulus, and (2) specificity to the trained location. Because early sensory neurons respond to a specific range of stimuli and locations, such findings suggest that the cortical representation of the trained objects might be mediated by neural plasticity in early cortical areas (Gilbert, Sigman, & Crist, 2001; Karni & Bertini, 1997; but see Mollon & Danilova, 1996).

Both stimulus- and location-specificity may be characteristic of object recognition in cluttered scenes as well. So far, however, the evidence for such learning specificity in segmentation tasks has been limited. Although previous studies have shown that learning is specific to the trained shapes (Brady & Kersten, 2003; Kourtzi et al., 2005; McCarley et al., 2004; Sowden et al., 2000), none of these studies have tested whether segmentation learning is location-specific.

A related issue is the extent to which shape learning per se accounts for improved object recognition in cluttered scenes or whether learning affects segmentation processes as well. It is already known that, in an empty background, the perceptual learning of complex shapes are specific to the trained shapes and (or) to the trained locations (Dill & Fahle, 1997; Furmanski & Engel, 2000; Nazir & O'Regan, 1990). Thus, one might expect that a shape trained extensively in isolation might acquire a privileged representation that allows it to pop out of cluttered scenes with less effort, driving most of the improvement in segmentation tasks. Alternatively, efficient segmentation might require specialized perceptual skills beyond what is acquired during shape learning alone. In this regard, we wish to distinguish shape segmentation learning from shape learning throughout this paper; the former refers to training on shapes in noisy backgrounds whereas the latter refers to training on shapes in empty backgrounds.

Another feature of perceptual learning is that it typically evolves through extensive effort and prolonged time of repetition, in contrast to priming, which requires only a few exposures to take effect. Moreover, perceptual learning effects characteristically show a long, persistent time course spanning several days, weeks, and even months (Ball & Sekuler, 1987; Fiorentini & Berardi, 1981; Gilbert et al., 2001; Karni & Sagi, 1991, 1993). Along with stimulus- and location-specificity, the long-term durability in perceptual learning suggests that acquired visual skills are mediated by the rewiring of neural connections in early visual cortex

(Gilbert et al., 2001; Karni & Bertini, 1997). Given such theoretical significance, it would be informative to test the long-term retention of perceptual improvement after training in the shape segmentation task.

Hence, the current study explores perceptual learning for objects camouflaged in noise to examine how perceptual experience may influence figure-ground segregation. In our paradigm, we presented symmetric dot patterns (Dill & Fahle, 1997; Nazir & O'Regan, 1990) against a random dot background field.¹ Since the search field looked homogeneous without practice, subjects' detection performance improved as a function of learning achieved through perceptual training. We tested the stimulus-specificity, location-specificity, and time-course of perceptual learning. The results can be interpreted to constrain possible neural mechanisms for performance in this figure-ground segregation task.

2. Experiment 1: Perceptual learning of shape segmentation

We examined the nature of top-down processing in figure-ground segregation using a perceptual learning paradigm. Observers were trained on a set of shapes embedded in random dot backgrounds. Each specific shape was trained in a specific and separate location in the periphery. Each display was flashed for only a brief period of time. Because the task was very challenging, observers had to depend on perceptual skills that were acquired through extensive training on the task.

After training, we first tested the location specificity of shape segmentation learning by translating the trained shapes into untrained locations across the horizontal meridian. Topographic mapping studies of macaque and human V4 have reported that the receptive fields (RFs) of area V4 are confined to a single quadrant of the visual field while RFs in its anterior part, area TE, largely overlap across the vertical and horizontal midlines (Boussaoud, Desimone, & Ungerleider, 1991; Gattass, Sousa, & Gross, 1988; McKeefry & Zeki, 1997; Schiller & Lee, 1991). Based on this physiology, we may use psychophysics to constrain the neural locus of our learning effects: if performance on the trained shapes drops when the shapes are translated to a new quadrant within the same hemifield, then we may infer that shape-specific learning occurs at the level of V4 or earlier (e.g., Bar & Biederman, 1999; McAuliffe & Knowlton, 2000).

The second test condition examined orientation specificity by inverting trained shapes. The third and fourth test conditions examined shape specificity by introducing novel shapes. In addition, the last two tests were conducted across a one-month gap to measure long-term retention of shape

¹ Barlow and Reeves (1979) have used populations of static dots to study symmetry perception. All or a part of dots formed global symmetry across the two sides of the visual field in their study. In contrast, a symmetric dot pattern was locally embedded in a random dot background in our study.

164 learning. Both shape specificity and its long-term persist- 186
 165 tence may support the idea that the shape segmentation 187
 166 learning results from neural rewiring in relatively early cor- 188
 167 tical areas (Gilbert et al., 2001; Karni & Bertini, 1997). 189

168 2.1. Method 191

169 2.1.1. Observers 192

170 Three graduate and one undergraduate students volun- 193
 171 teered as observers. One observer (SP) had prior experience 194
 172 in a preliminary test two years prior to the present experi- 195
 173 ment. The others were naïve to the purposes of the experi- 196
 174 ment. 197

175 2.1.2. Apparatus and stimuli 200

176 A G4 Macintosh computer presented the stimuli and 201
 177 collected responses using MATLAB with Psychophysics 202
 178 Toolbox extensions (Brainard, 1997; Pelli, 1997). A 17-inch 203
 179 monitor was viewed from a viewing distance of 82 cm, fixed 204
 180 with a chin rest. 205

181 Each search display presented 240 white dots and a 206
 182 white fixation cross against a black background, subtend- 207
 183 ing $4.8^\circ \times 4.8^\circ$. Each dot was a 0.1° diameter disk, and the 208
 184 fixation cross was made of two 0.2° -long crosshairs. As 209
 185 illustrated in Fig. 1A, each of the four search targets con-

sisted of seven pairs of dots, which were symmetrically 186
 arranged along the vertical axis on an imaginary 7×7 grid, 187
 subtending $1.2^\circ \times 1.2^\circ$, with two or three dots per column 188
 and two dots per row. The central column of the pattern 189
 was left empty to prevent symmetric dots on each side from 190
 forming a cluster in the middle. To discourage observers 191
 from using the empty column as a target feature cue, a 192
 noise dot was added to a random position in this central 193
 column in each trial. Different observers performed the task 194
 with different symmetric target shapes. Distractors were 195
 also generated trial-by-trial using the same rule as targets 196
 except that the dots were arranged asymmetrically. The 197
 search items, both targets and distractors, were presented in 198
 the center of the visual quadrants with 1.7° center-to-fixa- 199
 tion eccentricity. 200

201 2.1.3. Design and procedure 202

The experiment consisted of a total of 18 sessions, two 203
 sessions per day. Observers conducted the first 16 sessions 204
 for 8–10 days and completed the last two sessions about a 205
 month after the sixteenth session (lag; 35 days for JJ and 206
 NT, 28 days for SP and RD). Each session was divided 207
 into four 80-trial blocks. Within each block, observers 208
 divided their attention into two diagonally opposite 209
 visual quadrants to search for target shapes, and they

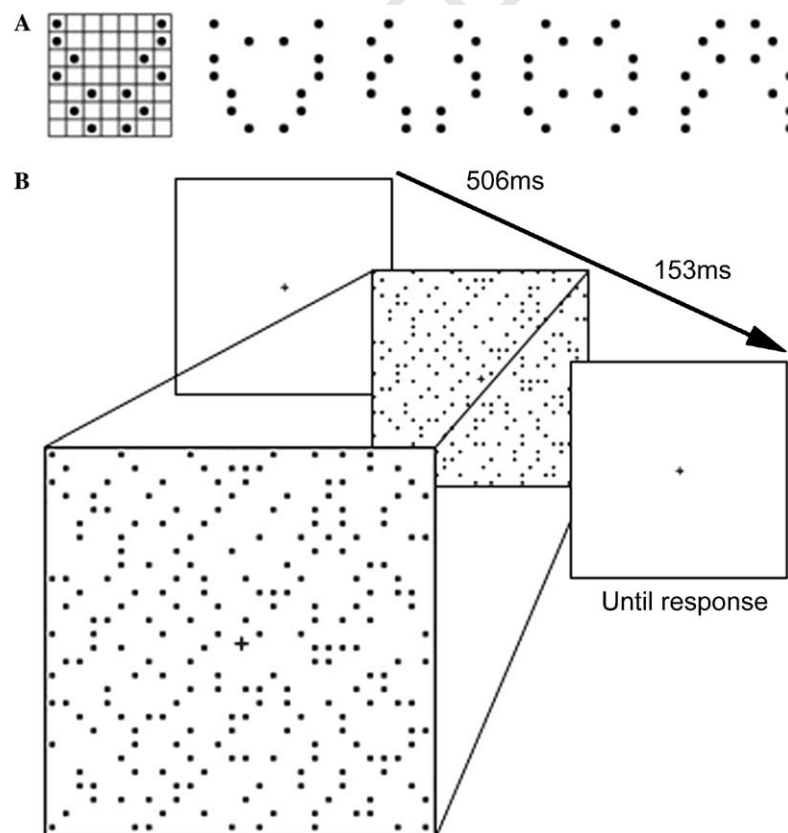


Fig. 1. Stimuli and procedure used in Experiment 1. (A) Examples of target shapes, each of which consisted of seven pairs of dots symmetrically positioned on an imaginary 7-by-7 grid. An additional dot was added to the middle column randomly in each trial. (B) Example of display procedure in a target present trial. In the enlarged random dot display, readers can find a symmetric target shape at the middle of the upper right visual quadrant. All stimuli are in scale and their colors were all white against a black background.

switched quadrants in alternating blocks. For example, an observer performed symmetry detection in the upper-right and the lower-left quadrants in odd numbered blocks and in the upper-left and the lower-right quadrants in even numbered blocks of a session. We restricted the possible number of target locations within a block to reduce the difficulty of the task. The order of target location pairs was counterbalanced across sessions and subjects. Each block contained 40 target-present trials that were randomly intermixed with the remaining 40 target-absent trials. Each target appeared 20 times within its respective quadrant. Thus, each block presented a target 20 times within each of the two attended quadrants. Accordingly, the following block presented two different targets in the other pair of quadrants. These two types of blocks were repeated once in the session. As such, each of the four targets was presented in its trained location for a total of 40 times within each session.

Observers practiced symmetry detection with a set of four targets in the first session and searched for a different target set (hereafter referred to as ‘old’ targets) in the next nine consecutive sessions. Note that, in these training sessions, each of the four old targets always appeared in a unique visual quadrant, such that specific shapes were trained in specific and separate locations (hereafter referred to as ‘old’ locations). In the remaining sessions, such ‘old-shape-old-location’ sessions (hereafter referred to as ‘old’ sessions) were interleaved with ‘test’ sessions to examine the specificity of learning. The first test session measured location specificity by translating each old target up or down into the opposite quadrant within the same hemifield. The second test session examined orientation specificity by inverting old targets in old locations. The third test session tested shape specificity by replacing old targets with a set of four new targets. Finally, two additional surprise sessions were conducted after a month delay to test long-term retention of perceptual learning. Observers performed an old session and a test session with a new target set. The order of test and old sessions in each day was counterbalanced between subjects.

Fig. 1B shows an example of the trial sequence. Each trial began with a central cross, which observers were instructed to fixate throughout the trial. After 506 ms, the search display was presented for 153 ms and followed by a blank display with the fixation cross. Observers then made a response in an unspeeded manner by pressing ‘k (labeled ●)’ when no target was detected, ‘j (labeled <)’ when a target was detected in the left quadrant, or ‘l (labeled >)’ when a target was detected in the right quadrant. Feedback was given for 306 ms after each response regarding its correctness or error type. The next trial was initiated 506 ms after the feedback disappeared. At the end of each block, performance was summarized in terms of hit rate, false alarm rate, and sensitivity (d'). A break was given every 40 trials, during which a pair of target locations in the next 40 trials were marked by two dimly outlined squares. At the beginning of each session, observers were given specific informa-

tion about the condition they would perform in the session.²

2.2. Results and discussion

We calculated two dependent measures for each session; one is the percentage of correct localization when targets were detected (e.g., responding ‘left’ for a upper left target). The other is sensitivity (d'), which was calculated by considering both correctly and incorrectly localized target responses as hits (e.g., responding either ‘left’ or ‘right’ for a upper left target). These two measures are plotted in Figs. 1A and B, respectively. It is clear that all observers gradually improved their performance for old targets during the first five days although the amount of improvement varied greatly across observers. To test the specificity of shape segmentation learning, we will focus our subsequent analyses on d' which is based on all types of responses. Since target localization was generally less accurate in test sessions than in same-day old sessions (Fig. 2A), the d' underestimates the difference between test vs. old sessions, making it a more stringent measure to test learning specificity. In addition, to parcel out the variability in the amount of learning across observers and across sessions, we calculated a specificity index (SI) for each test condition; $SI = 1 - (d' \text{ of a test session} / d' \text{ of the same-day old session})$. Fig. 2C shows the specificity indices, in which higher values indicate more specific learning. Each specificity index was submitted to a two-tailed one-sample t -test against 0 at a critical value of 0.05.

The specificity indices revealed that shape segmentation learning is significantly restricted to the trained location, orientation, and shape ($SI = 0.44$, $t = 4.604$ for location; $SI = 0.43$, $t = 3.822$ for orientation; $SI = 0.59$, $t = 4.864$ for shape; $SI = 0.55$, $t = 6.228$ for shape after a month). First of all, sensitivity deteriorated when the two trained locations in each hemifield exchanged their trained shapes. This demonstration of location specificity suggests that at least some aspects of shape segmentation learning may occur as early as area V4 where the receptive fields are confined to a single visual quadrant (Boussaoud et al., 1991; Gattass et al., 1988; McKeefry & Zeki, 1997; Schiller & Lee, 1991). It is worth noting, however, that the sensitivity after translation was still superior to that in the initial training sessions, indicating that shape segmentation learning had both location-specific and non-specific components. Location-specific components of training

² Observers’ knowledge of the specific aspects of the tests might enhance learning transfer because observers could imagine translating or inverting the trained shapes to assist their detection in the test sessions. Such a strategy should enhance transfer effects, serving as a stringent, conservative test of our interests in learning specificity. Without such instruction, observers might have been perceptually biased for the familiar training parameters, performing more poorly on the transfer tests than they are otherwise capable of, amplifying patterns of learning specificity. More importantly, our instructions did not convey any information about our hypotheses, leaving observers naïve to the predicted outcomes (except for one experienced observer, SP, whose results did not differ from the other observers).

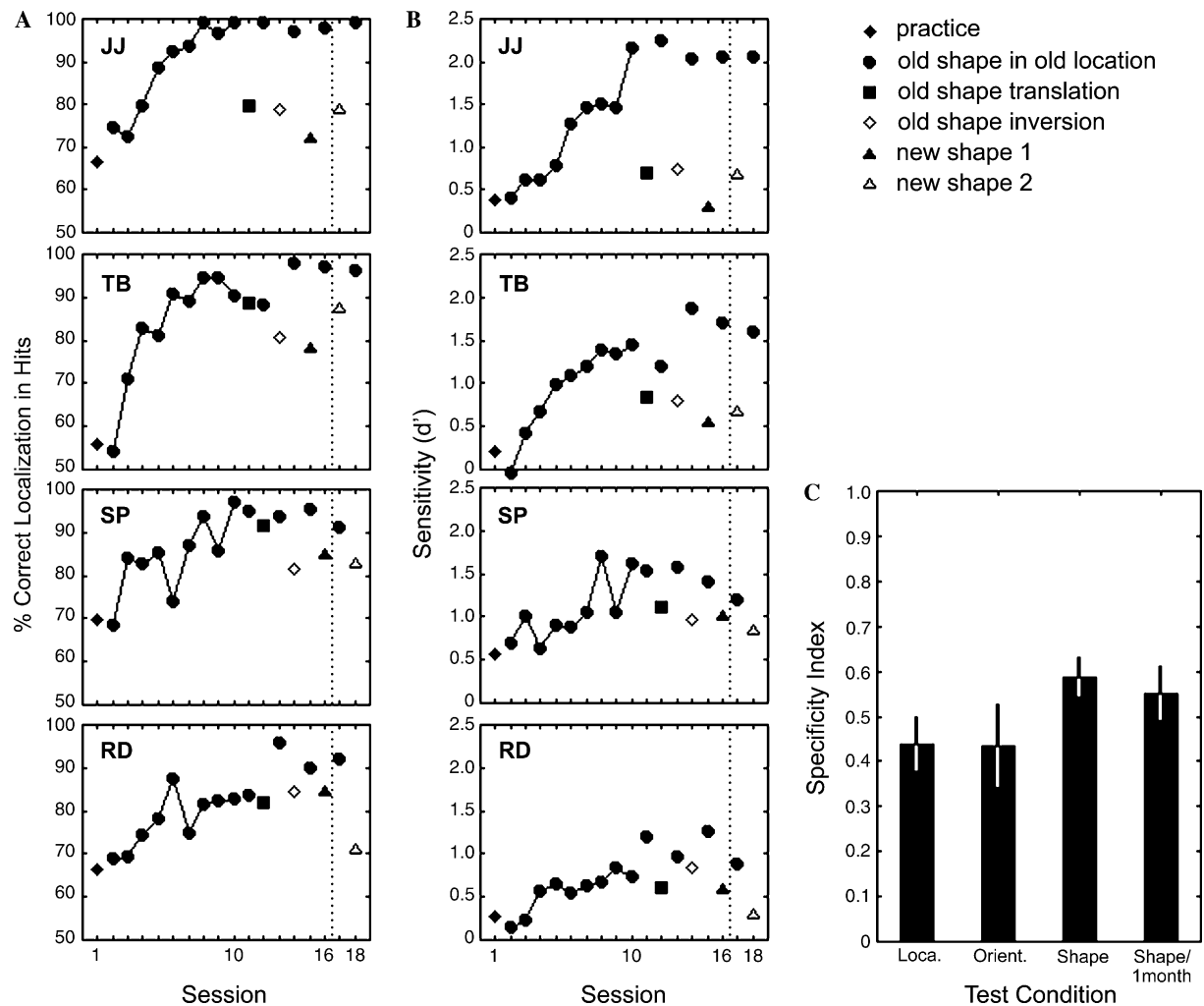


Fig. 2. Learning curve and specificity index in Experiment 1. (A) Percentage of correct localization of the targets for four individual observers when they reported target detection. (B) Sensitivity, d' , of the target detection. Dotted line represents one-month lag. (C) Specificity index. See Section 2.2 for calculation. Loca.: within-hemifield translation. Orient.: inversion. Shape: novel shapes. Shape/1 month: novel shapes after a one-month lag. Error bars indicate one standard error of the mean.

311 may include shape-specific information that would be
 312 dependent on neuronal representations in early retino-
 313 topic visual areas. Components of learning that were not
 314 location-specific may include general processes such as
 315 increased familiarity with the task and improvements in
 316 dividing visual attention. The observed partial specificity
 317 (or partial transfer) therefore implicates multiple loci of
 318 shape segmentation learning.

319 The hypothesis of early visual cortex participating in shape
 320 segmentation learning is further supported by the subsequent
 321 test conditions. The inverted versions of trained shapes were
 322 treated almost as novel shapes. A potential account for such
 323 orientation specificity is that inversion changed the location
 324 and orientation of local target features, which recruited new
 325 populations of early sensory neurons because these neurons
 326 are responsive to specific ranges of stimulus orientation within
 327 small RFs. An alternative (or complementary) account might
 328 be that trained shapes were processed configurally, which pro-
 329 duced an inversion effect as in face perception (Gauthier,
 330 Williams, Tarr, & Tanaka, 1998).

Especially important is the comparable specificity indices
 331 for the two novel shape conditions with a one-month
 332 lag between them. This result is supplemented by the fact
 333 that the two old sessions, one before and the other after the
 334 lag, were not different in d' ($d' = 1.60$ and 1.43 , respectively;
 335 paired $t = 2.087$, $p > .1$). Along with our evidence for loca-
 336 tion- and stimulus-specificity, this long-term retention effect
 337 supports the hypothesis that perceptual learning in this
 338 experiment was mediated by the rewiring of neural connec-
 339 tions in relatively early cortical areas (Gilbert et al., 2001;
 340 Karni & Bertini, 1997). The differences in specificity indices
 341 between pairs of test conditions were not significant
 342 (p 's > 0.15).
 343

344 Two aspects of the experiment ensured that observers
 345 fixated their eyes on the cross at the center of a search dis-
 346 play. On the one hand, the stimulus duration was too short
 347 for saccades to foveate any search item (Saslow, 1967). On
 348 the other hand, the target probability of one of two loca-
 349 tions in a block was low (25%), such that a 'sit-and-wait'
 350 strategy would be disadvantageous. If observers had

monitored only one of two target locations, they would have shown low sensitivity but very high percentage of correct localization in hit responses, a pattern not found in our data.

In sum, Experiment 1 demonstrated learning specificity both for trained shapes and for trained locations in our figure-ground segregation task. However, Experiment 1 did not clearly address exactly what observers learned during training. One hypothesis is that training helped form better shape representations, achieving finer matches between shape representations and the perceptual input. In addition, training may have helped additional processes involved in shape segregation learning, such as filtering out background noise. The alternative hypothesis is that training improved shape representations alone, such that shape matching would be sufficient for segmenting a shape from its background. In fact, several previous studies have reported both stimulus- or (and) location-specific dot pattern learning, similar to what we found for shape-segmentation learning here (Dill & Fahle, 1997; Nazir & O'Regan, 1990). Therefore, we conducted Experiment 2 to test whether perceptual improvement in segmentation was driven solely by shape learning alone or whether observers learned to segregate learned shapes from the noisy background as well.

3. Experiment 2: Shape learning vs. segmentation learning

We trained observers on target shapes without background dots in order to test whether shape learning alone could account for the shape segmentation improvement in Experiment 1. During training, observers performed the same detection task as in Experiment 1 between two diagonally opposite visual locations while ignoring the other locations in displays without background dots. The effect of shape learning on segmentation was tested by introducing background dots after training. If shape learning was sufficient for improved segmentation in Experiment 1, then the observers in this experiment should perform segmentation as well as the observers with the same amount of shape segmentation learning in Experiment 1.

In subsequent test conditions, we attempted to replicate the previous reports on symmetric dot pattern learning (Dill & Fahle, 1997; Nazir & O'Regan, 1990). Specificity was tested for location, orientation, and shape in the same order as in Experiment 1 to facilitate relevant comparisons.

The last test condition tested an alternative account for location specificity, based on the difference in eccentricity among local target features. When a target was presented in a peripheral location, the target feature closest to the fixation might be better represented than others. However, once the target is translated up or down into the other quadrant within hemifield, the feature that was previously closest to fixation is no longer the closest. Consequently, although the shape of the target and its eccentricity did not change, the eccentricity-dependent gradient of acuity might cause the target to be perceived differently, impairing target

recognition (Nazir & O'Regan, 1990). To test this possibility, we simply translated old targets into the opposite left or right quadrant across the vertical meridian. In previous studies of pattern recognition, translating a pattern did not impair recognition as far as the axes of a local shape symmetry and a global display symmetry between translation were parallel to each other (e.g., Dill & Fahle, 1999). Thus, if this account is correct, improved sensitivity after training should survive translation because the trained shape is globally symmetric across the vertical meridian, such that mirror images of the same features appear at the position closest to fixation.

3.1. Method

All methods were the same as in Experiment 1 except those described below. Four observers (two graduate students and two research associates) were newly recruited. Except during two test sessions, random dots were not shown in the background, leaving only four search items and a fixation on the screen. Observers searched for a target in two diagonally opposite locations within a block.

The experiment consisted of 16 sessions, two sessions per day. Observers conducted the first eight sessions for four consecutive days and, after a lag of two weeks (12 days for HC; 16 days for YX, AL, and JD), they completed the remaining eight sessions for four to five days. A set of four targets ('old' targets) was trained on for four sessions across the first two days and repeated in one of two daily sessions in the remaining days. Note that, during these sessions, each old target was trained in a specific location ('old' locations). The first test session was the same as the old sessions but with random dots in the background. The performance in the first test session was then compared with that in the second test session, which introduced a new set of targets with background dots. Testing resumed about two weeks later without the background dots. The third test session examined location specificity of shape learning by translating each old target up or down into a new quadrant within hemifield. The fourth test session measured orientation specificity by inverting old targets in their old locations. The fifth test session tested shape specificity by introducing a new set of targets. Finally, the sixth test session examined location specificity again but by translating old targets left or right into the opposite quadrant across hemifield.

3.2. Results and discussion

Figs. 3A and B show that when targets were presented without background dots, both correct localization and sensitivity substantially increased with only a few sessions of training. Observers in Experiment 2 achieved about twice greater d' than those in Experiment 1. As shown in Fig. 3C, we again calculated the specificity index for each test condition to achieve meaningful comparisons between conditions while parceling out the variability in the amount of learning across observers and sessions.

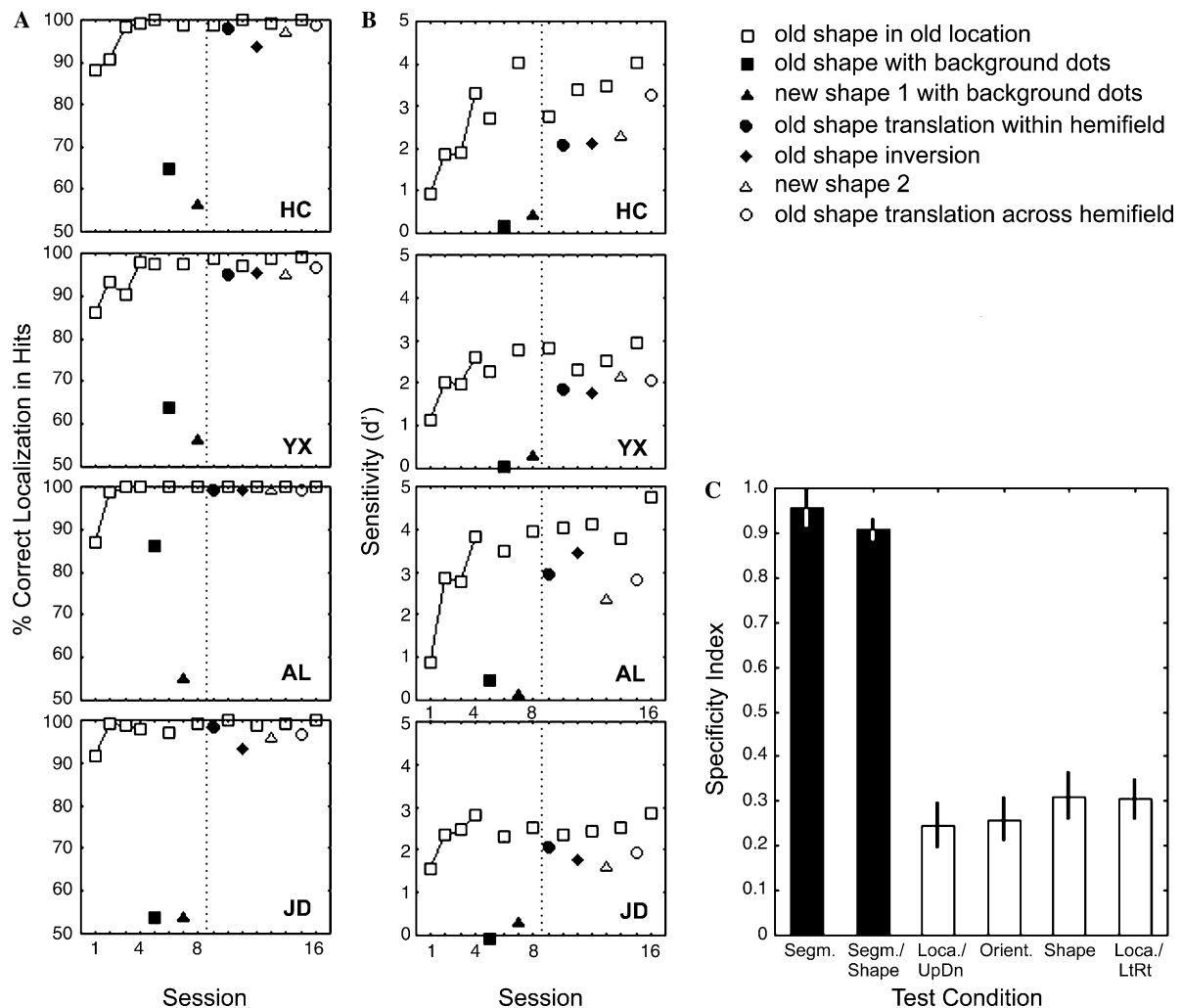


Fig. 3. Learning curve and specificity index in Experiment 2. (A) Percentage of correct localization of the targets for four individual observers when they reported target detection. (B) Sensitivity, d' , of the target detection. Dotted line represents a two-week lag. (C) Specificity index. See Section 2.2 for calculation. Segm.: segmentation of old shapes in a random dot background. Segm./Shape: segmentation of novel shapes in a random dot background. Loca./UpDn: within-hemifield translation. Orient.: inversion. Shape: novel shapes. Loca./LtRt: between-hemifield translation. Filled bar; noisy background condition. Empty bar; no-background condition. Error bars indicate one standard error of the mean.

459 The main analysis of interest was whether or not shape
 460 learning alone could lead to successful segmentation. The
 461 results showed that improved shape recognition did not
 462 transfer to the segmentation condition. Old targets with back-
 463 ground dots were detected no better than chance (mean
 464 $d' = 0.15$ for segmentation; $t = 1.303$, $p > .2$) and were compar-
 465 able to new targets both in d' (mean $d' = 0.29$ for segmenta-
 466 tion with new targets; paired $t = 0.844$, $p > .4$) and in specificity
 467 index (SI = 0.95 and 0.91, respectively; paired $t = 0.841$, $p > .4$).
 468 The lack of transfer to the segmentation condition is also
 469 demonstrated by Fig. 4, which plots mean d' s from both
 470 Experiment 1 and 2 up to the fifth session.³ Note that the d' in
 471 the segmentation condition of Experiment 2 was significantly
 472 lower than in the fifth session of Experiment 1 (mean $d' = 0.15$

vs. 0.83; independent samples $t = 4.980$, $p < 0.005$) and as low
 473 as in the second session of Experiment 1 (mean $d' = 0.15$ vs.
 474 0.30; independent samples $t = 0.743$, $p > .4$). Overall, these
 475 results suggest that shape learning alone cannot explain the
 476 improvement in segmentation in Experiment 1. Rather, success-
 477 ful segmentation might require additional perceptual
 478 skills, which are highly specific to the training context.
 479

Although not sufficient for segmentation, the shape learn-
 480 ing in Experiment 2 still revealed common signatures of percep-
 481 tual learning. As shown in Fig. 3B, long-term retention of
 482 shape learning was demonstrated by the old session right
 483 after the two-week lag. Sensitivity did not significantly change
 484 from the old session just before the lag (mean $d' = 2.99$ vs.
 485 3.32, respectively; paired $t = 1.031$, $p > .3$). In addition, as
 486 shown in Fig. 3C, improved performance was significantly
 487 limited to the trained location, orientation, and shape
 488 (SI = 0.25, $t = 5.654$ for location within hemifield; SI = 0.26,
 489 $t = 5.760$ for orientation; SI = 0.31, $t = 5.56$ for shape). Speci-
 490 ficity indices were generally lower in Experiment 2 than in
 491

³ The first session in Experiment 1 was practice and thus omitted here. In addition, the segmentation condition in Experiment 2 was conducted in the fifth session by two observers and in the sixth sessions by the other two observers.

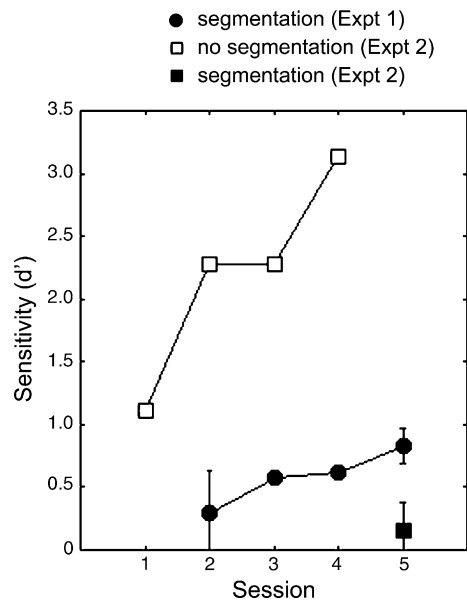


Fig. 4. Dissociation between shape learning and shape segmentation learning. See Section 3.2 for details.

be unique to our task. For example, during the earlier part of training, observers slowly accommodated to our task by deploying attention accurately on two separate target locations and by holding the fleeting representations of target candidates from a brief display. Informative instructions might have further aided partial transfer of skills. Observers might have also used high-level cognitive strategies tailored to each test condition, such as imagining mental rotation of trained shapes for the inversion condition.

4. General discussion

This study investigated how perceptual learning affects visual segregation. Experiment 1 demonstrated that perceptual learning of specific dot patterns facilitated figure segmentation from a random dot background. The learning effect was somewhat specific to trained locations and trained shapes. Improved segmentation performance and its partial specificity persisted for one month after training, implicating the increased involvement of sensory neurons in the segmentation of the trained shapes presumably through the rewiring of neural connections along the visual pathways (Gilbert et al., 2001; Karni & Bertini, 1997). Experiment 2 further showed that shape learning per se could not account for the improvement in shape segmentation. Shapes trained in empty backgrounds became nearly invisible to observers when random dots were introduced to the background, suggesting that segmentation requires further perceptual skills beyond shape recognition in isolation.

4.1. Shape recognition in a cluttered scene

An important feature of the present study was the use of a noisy background that effectively camouflaged the target from naïve observers. In this section, we will discuss which mechanisms may have contributed to shape-specific segregation learning in our task. First, we can reject the possibility that observers learned to detect symmetry per se. Although symmetry is a salient cue for shape recognition, bottom-up detection of symmetry alone cannot account for the learning improvements in this task. If observers were simply learning to detect low-level symmetry, they should have also improved on the novel symmetric shapes as much as the trained shapes. However, observers were much better at detecting trained symmetric shapes, suggesting that they were not just learning to detect symmetry.

Second, we can also distinguish our study from texture-segregation learning, which shows specific improvements for trained objects appearing in textured backgrounds (Karni & Sagi, 1991, 1993). Perceptual learning in these studies was also dependent on the structure of the textured background. In our experiments, the background dots had no consistent structure that would mediate such texture-specific learning.

Our experiments also rule out the hypothesis that shape learning alone could fully explain the shape-specific

Experiment 1, although the differences were not statistically significant (independent samples $t = 1.846$, $p > .1$ for location; $t = 1.434$, $p > .2$ for orientation; $t = 2.072$, $p > .08$ for shape).

Location specificity was also observed in the last condition in which old targets were translated across the vertical meridian. Since old targets preserved symmetry both locally and globally with this translation, the difference in eccentricity between local target features should not affect detection performance after such translation (e.g., Dill & Fahle, 1999). Nevertheless, the specificity index indicates that performance still deteriorated in untrained locations (SI = 0.30; $t = 6.734$) and the effect was comparable to the translation within hemifields (paired $t = 0.973$, $p > .4$). These results indirectly suggest that the location-specific effects observed within hemifield cannot be explained by asymmetries in which local features fall closest to fixation.⁴

Partial specificity (or partial transfer) observed for translation, inversion, and novel stimuli, along with comparable results in Experiment 1, suggests that some general skills acquired through training transferred to the test conditions. Given that complete learning specificity (no transfer) has been reported with similar dot patterns in empty backgrounds (Dill & Fahle, 1997), such transferable skills might

⁴ One potential problem with this control test is the involvement of hemifield change on top of quadrant change, which was inevitable in order to keep the axes of a local shape symmetry and a global display symmetry parallel to each other. This concern motivated us to train new group of observers with the same experimental methods except each target shape was symmetric around the horizontal axis. Translating such shapes up or down into the untrained quadrant kept the local and global axes parallel. Four observers trained in noisy backgrounds while the other four observers trained in empty backgrounds. Both groups of observers showed partial specificity that was nearly identical to the results reported here, indicating that the observed location specificity was not an artifact of the visual eccentricity gradient.

improvements in segregation. Training shapes in isolation did not later facilitate segmentation of the same shapes in noisy backgrounds. Since the observers in Experiment 2 achieved excellent detection performance after two days of training (accuracy mean = 93.4%, SD = 3.3%), this lack of transfer cannot be ascribed to incomplete shape learning. Rather, the dissociation between shape learning and shape segmentation learning implicates fundamental differences in perceptual skills acquired with or without a background noise.

One potential distinction between shape learning and shape segmentation learning is the degree of task difficulty. Shape segmentation was much more difficult when a background was noisy, as indicated by the large performance difference between the two experiments. It has been proposed that perceptual learning becomes more specific to the training parameters as task difficulty increases (Ahissar & Hochstein, 1997, 2004; for a recent review). Because difficult perceptual tasks require more accurate spatial representation of stimulus attributes, the improvement on the tasks depends on plasticity in early sensory neurons, which preserve fine-grained spatial resolution of the visual input. In our study, recognition of shapes in noisy backgrounds may require more accurate spatial information of dot arrays than recognition of the shapes in empty backgrounds. For example, initial processes of segmentation may include recognition of familiar dot formations from a target location. Then, based on their local coordinates, these dot formations can be pieced together to form a shape configuration. These effortful processes may be extensively trained during shape segmentation learning with background noise, while they may be bypassed during shape learning without background noise. In the absence of noise, the boundaries of target shapes are sharply delineated, and so configural information of the shapes can be easily extracted. Thus, rigorous localization of local features might not be as critical as it would be for the shapes camouflaged by background noise. As such, shape recognition may rely more on the early visual cortex to achieve accurate spatial representations when segmentation is challenged with background noise.

The distinction between shape learning and shape segmentation learning is, in fact, reminiscent of a critical question in object recognition; how does the brain learn to recognize a novel object in a cluttered scene when no top-down knowledge of the object is available to assist segmentation at the very beginning of the learning, an issue known as the ‘bootstrapped learning problem’ (Brady & Kersten, 2003)? In the simplest case, referred to as ‘opportunistic learning,’ one may expect that learning can proceed only when bottom-up information such as motion or color cues define the shape boundary. However, human observers can learn to identify a novel shape even when the shape is invisibly camouflaged at the beginning of the training (Brady & Kersten, 2003). One possibility is that observers learn to segregate covarying shape features based on statistical regularities (Fiser & Aslin, 2001). In the present study, shape

segmentation learning in Experiment 1 may be an example of bootstrapped learning,⁵ whereas shape learning in Experiment 2 without background noise may be an example of opportunistic learning. In the real world, these two types of learning should work together to establish object knowledge.

The observed dissociation in our study helps to understand how this cooperation is realized along the visual pathways; the hypothesis is that bootstrapped learning is driven by earlier visual cortical areas relative to opportunistic learning. This hypothesis is supported by a recent functional neuroimaging results (Kourtzi et al., 2005). Training for segmenting similarly oriented Gabor elements decreased the blood oxygenation-level dependent (BOLD) signals in high-level visual cortex when a salient bottom-up grouping cue was available (i.e., in a uniformly oriented Gabor field; opportunistic learning), but increased the BOLD signals in the both low- and high-level visual cortices when the bottom-up cue was poor (i.e., in a randomly oriented Gabor field; bootstrapped learning). Along with our study, these findings suggest that the way how shape learning is engraved in the brain can differ based on the level of noise in the background scene.

4.2. Specificity of shape segmentation learning

The theoretical importance of learning specificity resides in the general assumption that learning is mediated by early sensory cortex where neurons are selective and their receptive fields (RFs) are small (Fahle, 2005; for a review). In line with this assumption, we conjectured that, in our experiments, the representation of the trained shapes moved down along the visual pathways as the early sensory neurons became increasingly involved in localizing the oriented segments of dot arrays that defined the contours of the shapes in our task (Ahissar & Hochstein, 2004; Gilbert et al., 2001; Hochstein & Ahissar, 2002; Karni & Bertini, 1997). The test of within-hemifield translation helps to interpret the anatomical locus of learning in our experiments. Topographic mapping studies of monkeys (Bous-saoud et al., 1991; Gattass et al., 1988) and humans (McKeefry & Zeki, 1997) have documented that the RFs of neurons in areas V4 and earlier do not generally cross the vertical and horizontal meridians. The lack of complete transfer across the horizontal meridian (in both isolation and noisy background conditions), therefore, implicates the involvement of neurons at or earlier than area V4 (e.g., Bar & Biederman, 1999; McAuliffe & Knowlton, 2000).

This interpretation of specificity, however, deserves a note of caution since alternative accounts have been raised (Dill, 2002; Mollon & Danilova, 1996). Especially relevant here is the possibility that even identical stimulation might

⁵ Because target shapes were symmetric, our shape segmentation learning does not perfectly fit to the notion of bootstrapped learning. However, given that observers informally reported that some trained shapes looked asymmetric, our intuition predicts similar results with training on random dot patterns in noisy backgrounds.

674 result in a different neuronal response depending on which
675 subregion of a large RF is stimulated.⁶ In that case, neurons
676 beyond area V4 or TEO with RF's covering more than a
677 visual quadrant could mimic the observed location specific-
678 ity. Thus, the hypothesized involvement of early visual cor-
679 tex in shape segmentation learning requires converging
680 evidence. In this regard, functional neuroimaging evidence
681 of learning-dependent activity changes in early visual cor-
682 tex in humans is more decisive (Furmanski, Schluppeck, &
683 Engel, 2004; Kourtzi et al., 2005; Schiltz et al., 1999; Sch-
684 wartz, Maquet, & Frith, 2002).

685 4.3. Summary

686 In a cluttered visual scene, the mechanism of figure-
687 ground segregation benefits greatly from top-down
688 knowledge. It is one of the crucial functions of biological
689 vision, which is hardly matched by machine vision. In two
690 experiments, we demonstrated that such top-down
691 knowledge can be acquired through perceptual training.
692 We also reported several interesting properties of shape
693 segmentation learning. Especially, partial specificity in
694 our study supports the idea that shape segmentation from
695 a noisy background is achieved through multiple stages,
696 including lower levels of visual processing. Altogether,
697 our findings help understand how perceptual experience
698 is encoded from cluttered visual scenes and where such
699 learning is mediated by plasticity along the visual path-
700 way.

701 Acknowledgment

702 This research was supported by NIH Grant EY014193.

703 References

- 704 Ahissar, M., & Hochstein, S. (1993). Attentional control of early percep-
705 tual learning. *Proceedings of the National Academy of Sciences of the*
706 *United States of America*, 90(12), 5718–5722.
707 Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of
708 perceptual learning. *Nature*, 387(6631), 401–406.
709 Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual
710 perceptual learning. *Trends in Cognitive Sciences*, 8(10), 457–464.
711 Ball, K., & Sekuler, R. (1982). A specific and enduring improvement in
712 visual motion discrimination. *Science*, 218(4573), 697–698.
713 Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion
714 discrimination. *Vision Research*, 27(6), 953–965.
715 Bar, M., & Biederman, I. (1999). Localizing the cortical region mediating
716 visual awareness of object identity. *Proceedings of the National Acad-*
717 *emy of Sciences of the United States of America*, 96(4), 1790–1793.
718 Barlow, H. B., & Reeves, B. C. (1979). The versatility and absolute
719 efficiency of detecting mirror symmetry in random dot displays. *Vision*
720 *Research*, 19(7), 783–793.
721 Boussaoud, D., Desimone, R., & Ungerleider, L. G. (1991). Visual topogra-
722 phy of area TEO in the macaque. *Journal of Comparative Neurology*,
723 306(4), 554–575.

- Brady, M. J., & Kersten, D. (2003). Bootstrapped learning of novel objects. 724
Journal of Vision, 3(6), 413–422. 725
Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 726
433–436. 727
Dill, M. (2002). Specificity versus invariance of perceptual learning: The 728
example of position. In M. Fahle & T. Poggio (Eds.), *Perceptual learn-* 729
ing (pp. 219–231). Cambridge: The MIT press. 730
Dill, M., & Fahle, M. (1997). The role of visual field position in pattern-dis- 731
crimination learning. *Proceedings of the Royal Society of London.* 732
Series B: Biological Sciences, 264(1384), 1031–1036. 733
Dill, M., & Fahle, M. (1999). Display symmetry affects positional specific- 734
ity in same-different judgment of pairs of novel visual patterns. *Vision* 735
Research, 39(22), 3752–3760. 736
Fahle, M. (2005). Perceptual learning: Specificity versus generalization. 737
Current Opinion in Neurobiology, 15(2), 154–160. 738
Fahle, M., & Edelman, S. (1993). Long-term learning in vernier acuity: 739
Effects of stimulus orientation, range and of feedback. *Vision Research*, 740
33(3), 397–412. 741
Fiorentini, A., & Berardi, N. (1981). Learning in grating waveform dis- 742
crimination: Specificity for orientation and spatial frequency. *Vision* 743
Research, 21(7), 1149–1158. 744
Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher- 745
order spatial structures from visual scenes. *Psychological Science*, 746
12(6), 499–504. 747
Furmanski, C. S., & Engel, S. A. (2000). Perceptual learning in object recog- 748
nition: Object specificity and size invariance. *Vision Research*, 40(5), 749
473–484. 750
Furmanski, C. S., Schluppeck, D., & Engel, S. A. (2004). Learning strength- 751
ens the response of primary visual cortex to simple patterns. *Current* 752
Biology, 14(7), 573–578. 753
Gattass, R., Sousa, A. P., & Gross, C. G. (1988). Visuotopic organization 754
and extent of V3 and V4 of the macaque. *Journal of Neuroscience*, 8(6), 755
1831–1845. 756
Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training 'gree- 757
ble' experts: A framework for studying expert object recognition pro- 758
cesses. *Vision Research*, 38(15–16), 2401–2428. 759
Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of per- 760
ceptual learning. *Neuron*, 31, 681–697. 761
Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and 762
reverse hierarchies in the visual system. *Neuron*, 36(5), 791–804. 763
Karni, A., & Bertini, G. (1997). Learning perceptual skills: Behavioral 764
probes into adult cortical plasticity. *Current Opinion in Neurobiology*, 765
7(4), 530–535. 766
Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture dis- 767
crimination: Evidence for primary visual cortex plasticity. *Proceedings* 768
of the National Academy of Sciences of the United States of America, 769
88(11), 4966–4970. 770
Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. 771
Nature, 365(6443), 250–252. 772
Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as 773
Bayesian inference. *Annual Review of Psychology*, 55, 271–304. 774
Kourtzi, Z., Betts, L. R., Sarkheil, P., & Welchman, A. E. (2005). Distrib- 775
uted neural plasticity for shape learning in the human visual cortex. 776
Public Library of Science. Biology, 3(7), e204. 777
Kovacs, I., Kozma, P., Feher, A., & Benedek, G. (1999). Late maturation 778
of visual spatial integration in humans. *Proceedings of the National* 779
Academy of Sciences of the United States of America, 96(21), 12204– 780
12209. 781
Li, W., & Gilbert, C. D. (2002). Global contour saliency and local colinear 782
interactions. *Journal of Neurophysiology*, 88(5), 2846–2856. 783
McAuliffe, S. P., & Knowlton, B. J. (2000). Long-term retinotopic priming 784
in object identification. *Perception and Psychophysics*, 62(5), 953–959. 785
McCarley, J. S., Kramer, A. F., Wickens, C. D., Vidoni, E. D., & Boot, W. 786
R. (2004). Visual skills in airport-security screening. *Psychological Sci-* 787
ence, 15(5), 302–306. 788
McKeefry, D. J., & Zeki, S. (1997). The position and topography of the 789
human colour centre as revealed by functional magnetic resonance 790
imaging. *Brain*, 120(Pt. 12), 2229–2242. 791

⁶ We thank an anonymous reviewer of our earlier draft for raising this alternative account.

- 792 Mollon, J. D., & Danilova, M. V. (1996). Three remarks on perceptual
793 learning. *Spatial Vision*, 10(1), 51–58.
- 794 Nazir, T. A., & O'Regan, J. K. (1990). Some results on translation invari-
795 ance in the human visual system. *Spatial Vision*, 5(2), 81–100.
- 796 Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics:
797 transforming numbers into movies. *Spatial Vision*, 10(4), 437–442.
- 798 Poggio, T., Fahle, M., & Edelman, S. (1992). Fast perceptual learning in
799 visual hyperacuity. *Science*, 256(5059), 1018–1021.
- 800 Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the*
801 *Optical Society of America*, 57(8), 1030–1033.
- 802 Schiller, P. H., & Lee, K. (1991). The role of the primate extrastriate area
803 V4 in vision. *Science*, 251, 1251–1253.
- 804 Schiltz, C., Bodart, J. M., Dubois, S., Dejardin, S., Michel, C., Roucoux, A.,
805 et al. (1999). Neuronal mechanisms of perceptual learning: Changes in
806 human brain activity with training in orientation discrimination. *Neu-*
807 *roimage*, 9(1), 46–62.
- 808 Schwartz, S., Maquet, P., & Frith, C. (2002). Neural correlates of percep-
809 tual learning: A functional MRI study of visual texture discrimination.
Proceedings of the National Academy of Sciences of the United States of 810
America, 99(26), 17137–17142. 811
- Shiu, L. P., & Pashler, H. (1992). Improvement in line orientation discrimi- 812
nation is retinally local but dependent on cognitive set. *Perception and* 813
Psychophysics, 52(5), 582–588. 814
- Sigman, M., & Gilbert, C. D. (2000). Learning to find a shape. *Nature Neu-* 815
roscience, 3(3), 264–269. 816
- Sowden, P. T., Davies, I. R., & Roling, P. (2000). Perceptual learning of the 817
detection of features in X-ray images: A functional role for improve- 818
ments in adults' visual sensitivity? *Journal of Experimental Psychology:* 819
Human Perception & Performance, 26(1), 379–390. 820
- Vaina, L. M., Belliveau, J. W., des Roziers, E. B., & Zeffiro, T. A. (1998). 821
Neural systems underlying learning and representation of global 822
motion. *Proceedings of the National Academy of Sciences of the United* 823
States of America, 95(21), 12657–12662. 824
- Vogels, R., & Orban, G. A. (1985). The effect of practice on the oblique 825
effect in line orientation judgments. *Vision Research*, 25(11), 1679– 826
1687. 827