Behavioral/Systems/Cognitive

The Role of Top-Down Task Context in Learning to Perceive Objects

Yiying Song,[1,2] **Siyuan Hu,**[1,2] **Xueting Li,**[1] **Wu Li,**[1] **and Jia Liu**[1,2]

[1]State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China and [2]Graduate University of Chinese Academy of Sciences, Beijing 100049, China

In high-level perceptual regions of the ventral visual pathway in humans, experience shapes the functional properties of the cortex: the fusiform face area responds most strongly to faces of familiar rather than unfamiliar races, and the visual word form area (VWFA) is tuned only to familiar orthographies. But are these regions affected only by the bottom-up stimulus information they receive during learning, or does the effect of perceptual experience depend on the way that stimulus information is used during learning? Here, we test the hypothesis that top-down influences (i.e., task context) modulate the effect of perceptual experience on functional selectivities of the high-level visual cortex. Specifically, we test whether experience with novel visual stimuli produces a greater effect on the VWFA when those stimuli are associated with meanings (via association learning) but produces a greater effect on shape-processing regions when trained in a discrimination task without associated meanings. Our result supports this hypothesis and further shows that learning is transferred to novel objects that share parts with the trained objects. Thus, the effects of experience on selectivities of the high-level visual cortex depend on the task context in which that experience occurs and the perceptual processing strategy by which objects are encoded during learning.

Introduction

Substantial evidence indicates that the human visual system is plastic, modified by perceptual learning (Golby et al., 2001; Baker et al., 2007). After learning, the visual system becomes more sensitive to trained objects compared with untrained ones (Gauthier et al., 1999; Grill-Spector et al., 2000; Janzen and van Turennout, 2004; Sigman et al., 2005; Op de Beeck et al., 2006; Jiang et al., 2007; Weisberg et al., 2007). Here, we used functional magnetic resonance imaging (MRI) to examine whether the effect of perceptual experience on functional selectivities of the high-level visual cortex depends on the way that stimulus information is used during learning.

There is a consensus that no single cortical locus is exclusively responsible for object learning (Bukach et al., 2006; Kourtzi and DiCarlo, 2006). Rather, the learning-induced changes are distributed throughout the visual system (Op de Beeck and Baker, 2010). However, regarding the question of where and how object representations are constructed through learning, hypotheses of stimulus-driven learning and task-guided learning have different predictions (Kanwisher, 2001). The stimulus-driven learning hypothesis suggests that the prior preference of neurons biases the visual system to develop object representations in or near cortical regions whose preferred stimuli share similar visual features or physical properties (Dale et al., 1999; Jacobs, 1999) regardless of the task context within which objects are learned. However, different computational problems faced by the visual system in complex environments may recruit different neural substrates in processing the same object (Wong et al., 2009a,b). Therefore, object representations are likely developed in cortical regions that are intimately involved in tackling specific computational demands (Op de Beeck et al., 2006; Weisberg et al., 2007). That is, object representations are developed in cortical regions originally responsive to objects that are experience related but not necessarily visually similar to trained objects.

Although the top-down task context may recruit different neural substrates to construct new representations, processing strategies used to encode stimulus information during the learning may play a role in shaping the boundary of the tolerance of the representations to visual variations of trained objects. The holistic strategy processes the information on individual parts relatively inseparably from the information on relations between the parts (e.g., Tanaka and Farah, 1993; Gauthier and Tarr, 2002). Therefore, the learning would be specific to the trained objects and there would be no transfer of learning from trained objects to their variations. In contrast, the parts-based strategy processes objects as a combination of simple and independent parts (e.g., Marr, 1982; Biederman, 1987). Therefore, the individual parts of an object would be encoded when the whole object is learned, so a transfer of learning to the variants that share those parts would be expected (Riesenhuber and Poggio, 2000; Golcu and Gilbert, 2009).

To address these questions, we trained subjects to learn a set of novel objects either through visual association learning or through shape discrimination learning and investigated where in

the brain object representations were constructed and whether the learning effect transferred to their variations.

Materials and Methods

Subjects

Sixteen subjects (age 22–28; 9 females) participated in the experiment of visual association learning, and another 12 subjects (age 20–24; 8 females) participated in the experiment of shape discrimination learning. Four subjects from each group also participated in a behavioral discrimination task before and after the training. All subjects were right handed, had normal or corrected-to-normal visual acuity, and were native Chinese speakers who have studied English for at least 10 years. The functional MRI (fMRI) protocol was approved by the Institutional Review Board of Beijing Normal University, Beijing, China. Informed consent was obtained from all subjects before their participation.

General procedure

A set of novel objects were created (Fig. 1A). Two groups of subjects were trained in either a visual association task or a shape discrimination task. During the association learning, the objects were paired with English words and subjects were instructed to learn their associations (Fig. 1B). During the discrimination learning, subjects reported the identity of objects based on their shapes, which were briefly presented and then backwardly masked (Fig. 1C). After subjects' performance reached asymptotes, fMRI was used to characterize the cortical distribution of learning effects and the transfer of learning to variants of the trained objects by examining neural responses within the visual word form area (VWFA), which responds to the orthography of English words (Cohen et al., 2000, 2002), and the lateral occipital cortex (LO), a part of the lateral occipital complex (LOC) that is involved in processing the shape of general objects (Grill-Spector et al., 2000, 2001; Kourtzi and Kanwisher, 2001).

Stimuli

Dumbbell-shaped objects used in the training and fMRI scanning were constructed by connecting two component figures with a connection bar (Fig. 1A). The component figures were made of simple shapes (e.g., arc, line, circle, square) that were designed to avoid resemblance to everyday objects.

Sixteen objects used for training were created from 16 component figures. In particular, the 16 figures were arbitrarily divided into four groups, each of which contained four component figures. By combining any two figures in a group with a connection bar, four dumbbell-shaped objects were created (Fig. 1A, left) (for the whole set of stimuli, see supplemental Fig. 1, available at www.jneurosci.org as supplemental material). Note that once a figure was selected as a left (or right) component of a dumbbell-shaped object, it remained on the left (or right) in other combinations to ensure the uniqueness of component figures in each dumbbell-shaped object. Therefore, the dumbbell-shaped objects could not be learned based only on one component figure on either end. Instead, the configurations of the component figures had to be taken into account to acquire good performance on object recognition.

In addition to the to-be-trained objects, 16 variant objects were created by randomly combining a left component figure in one group and a right component in another group of the to-be-trained objects (Fig. 1A, middle). Therefore, the variants contained the same part components as the to-be-trained objects, but their configurations were novel.

Finally, another set of 16 dumbbell-shaped objects were constructed in the same way as the to-be-trained objects, but from a new set of component figures to serve as a nonexposed baseline condition (Fig. 1A, right).
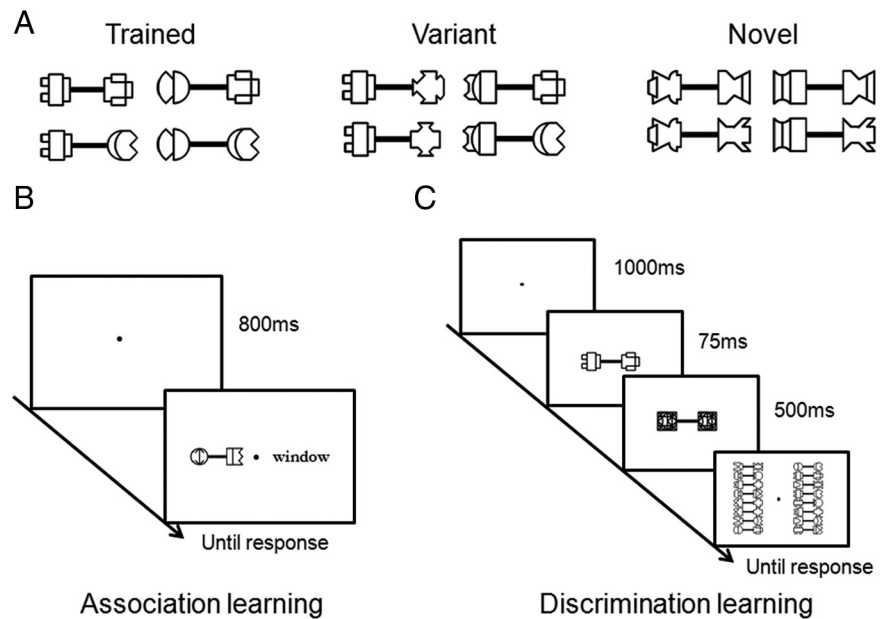

Figure 1. Stimuli and training procedure. **A**, Four dumbbell-shaped exemplars for each stimulus condition. **B**, Visual association learning. Subjects conducted an association judgment task to determine whether a stimulus pair was one of the actual paired associates. Each trial started with a blank screen, followed by a dumbbell-shaped object presented along with an English word. The display was presented until a response was made. **C**, Shape discrimination learning. Subjects learned to discriminate the shape of briefly presented dumbbell-shaped objects. Each trial started with a blank screen, followed by a dumbbell-shaped object briefly presented on the center of screen that was immediately masked. Subjects then selected the target among all 16 objects presented simultaneously on the screen.

Behavioral training

Visual association learning. A paired-associate association learning paradigm was used to train subjects to learn 16 paired associates (Fig. 1B). Subjects were first shown all paired associates to become familiar with their associations. Then, they conducted an association judgment task (two-alternative forced choice) to determine whether the stimulus pair was one of the actual paired associates. Each trial started with a blank screen for 800 ms, followed by a dumbbell-shaped object presented along with an English word. The display was presented until a response was made. Auditory feedback was given to indicate whether the response was correct (high pitch) or incorrect (low pitch). There were 480 trials (30 trials per object) in one training session, half of which contained correct paired associates. The behavioral training was terminated when the reaction time (RT) reached an asymptote (i.e., no significant decrease in RT in at least three consecutive sessions). This criterion was chosen so that a subject's end performance was equated both within and between the learning tasks. All subjects completed at least six training sessions (i.e., 180 trials per object), which took ~3–4 h in total for 3 successive days.

Shape discrimination learning. A shape discrimination learning paradigm was used to train subjects to learn the shape of 16 dumbbell-shaped objects (Fig. 1C). Each trial started with a blank screen for 1 s, followed by a dumbbell-shaped object briefly presented on the center of screen for 75 ms that was immediately masked for 500 ms. Subsequently, all 16 objects were presented simultaneously on the screen. Subjects reported the target object with a mouse click (i.e., a 16-alternative forced choice task). Auditory feedback was immediately provided. There were 80 trials per training session (i.e., 5 trials per object). The behavioral training ended when the accuracy reached an asymptote (i.e., no significant increase in accuracy in at least three consecutive sessions). All subjects completed at least 31 training sessions (i.e., 155 trials per object), which took ~8–10 h in total for 5–6 successive days.

fMRI scanning

Each subject participated in a single session consisting of the following: (1) two blocked-design functional localizer scans; and (2) three blocked-design experimental scans. The localizer scan consisted of English words, human frontal view faces, line-drawing objects, and scrambled line-

drawing objects. The experimental scan consisted of three types of the dumbbell-shaped objects: objects from behavioral training (Trained), variants of the trained objects (Variant), and nonexposed novel objects (Novel) (Fig. 1A). Each scan lasted 5 min and 15 s and consisted of sixteen 15 s blocks with five 15 s blocks of fixation periods interleaved. In each 15 s block of both localizer and experimental scans, 16 different exemplars of a given stimulus condition were shown in the center of the screen, each of which were presented for 400 ms followed by a blank interval of 500 ms.

During the experimental scan, subjects performed neither the association task nor the discrimination task that they performed in the behavioral training. Instead, they performed an identity one-back task (i.e., pressing a button whenever two identical images in a row repeated). This orthogonal task was designed to ensure that differences in cortical distribution of learning effects from these two task contexts, if observed, were not mainly from the differences in tasks performed during the scan. In addition, this task also ensured that subjects allocated attention equally to all types of stimulus conditions, regardless of whether they were familiar or not.

MRI data acquisition
Scanning was done on a 3T Siemens Trio scanner with an eight-channel, phase-arrayed coil at Beijing Normal University Imaging Center for Brain Research, Beijing, China. Thirty 2.8-mm-thick (20% skip) near axial slices were collected (in-plane resolution = 1.4 × 1.4 mm) and oriented parallel to each subject's temporal cortex to cover the inferior portion of the occipital lobes as well as the posterior portion of the temporal lobes, including the VWFA, fusiform face area (FFA), and LO. This relatively high in-plane resolution involves considerably less averaging across distinct neural populations and therefore reduces partial voluming so as to enable detection of small learning-induced changes in neural activity. T2*-weighted gradient-echo, echo-planar imaging procedures were used [repetition time (TR) = 3 s, echo time (TE) = 32 ms, flip angle = 90°]. In addition, MPRAGE, an inversion prepared gradient echo sequence (TR/TE/inversion time = 2.73 s/3.44 ms/1 s, flip angle = 7°, voxel size = 1.1 × 1.1 ×1.9 mm), was used to acquire three-dimensional structural images.

fMRI data analysis
Functional data were analyzed with the FreeSurfer functional analysis stream (CorTech Labs) (Dale et al., 1999; Fischl et al., 1999), froi (http://froi.sourceforge.net), and in-house Matlab code. After data preprocessing, including motion correction, intensity normalization, and spatial smoothing (Gaussian kernel, 4 mm full width at half maximum), voxel time courses for each individual subject were fitted by a general linear model. Each condition was modeled by a boxcar regressor matching its time course that was then convolved with a gamma function (delta = 2.25, tau = 1.25).

The method of individual subject analysis was used for characterizing the learning-induced neural changes because the size of learning effects may be small and the exact cortical loci representing the learned objects may vary across subjects. Therefore, in this study we first identified object-selective regions of interest (ROIs) separately for each subject from the localizer scan. The VWFA was defined as a set of contiguous voxels in the mid-fusiform gyrus in the left hemisphere that showed significantly higher responses to English words compared with line-drawing objects ($p < 0.01$, uncorrected for multiple comparisons). Face-selective and object-selective regions were defined in the same way by the contrasts of faces versus objects and objects versus scrambled objects, respectively. Because the occipital face area and the anterior subregion of the LOC, the posterior fusiform sulcus, could not be reliably localized across subjects, we focused here on the FFA and LO. Both the FFA and LO were bilaterally localized, but the VWFA was found only in the left hemisphere. Because the right FFA and LO are anatomically far from the VWFA in the left hemisphere, they were less critical than the left FFA and left LO in examining how the learning effect was distributed in the ventral visual cortex. Therefore, we focused on the spatial distribution of learning effects in the left hemisphere, with the results in the right hemisphere reported.

For the ROI analysis, percentage signal changes were extracted and averaged by condition across all experimental scans and all voxels within each subject's predefined ROIs, and this response was combined across subjects by averaging. Because fMRI responses typically lag 4–6 s after the neural response, the magnitude of the ROI activity was measured as the average percentage change in MR signal at the latencies of 6, 9, 12, 15, and 18 s (TR = 3 s, block length = 15 s) compared with a fixation as a baseline. Percentage signal changes, one per experimental condition per ROI per subject, were submitted to repeated-measures ANOVA, followed by *post hoc* pairwise two-tailed *t* tests.

Results
The role of task context in object learning
In the association learning, subjects were trained to learn the paired associates of English words and dumbbell-shaped objects. As expected, association training greatly shortened the RT in judging the paired associates (Fig. 2A, left). The RT in association judgment decreased monotonically from session 1 to session 5 ($F_{(4, 48)} = 21.10, p < 0.001$) and then reached an asymptote from session 6 to session 8 ($F_{(2, 30)} = 1.40, p > 0.05$). The mean accuracy remained >90% during the entire training period.

In the shape discrimination learning, another group of subjects were trained to discriminate briefly presented dumbbell-shaped objects. As shown on the left side of Figure 2B, the training significantly improved the behavioral performance, from a near-chance level of performance in accuracy in the first session (22%) to >85% after 26 sessions ($F_{(25, 275)} = 61.13, p < 0.001$). From the 27th session to 31st session, the accuracy reached an asymptote ($F_{(4, 44)} < 1$).

Four subjects from each group were further tested before and after the training with the same task used in the discrimination learning to examine the specificity of object learning. After the discrimination learning there was a significant interaction of stimulus (Trained versus Novel) by learning (before vs after) ($F_{(1, 3)} = 46.16, p < 0.01$), with the performance in discriminating the trained objects significantly higher than that for the novel objects after the training ($t_{(3)} = 4.93, p < 0.05$) but not before ($t_{(3)} = 1.81, p > 0.05$) (Fig. 2B, right). In contrast, the association learning did not selectively improve subjects' ability in discriminating the trained objects over the novel objects, as the interaction of stimulus by learning did not reach significance ($F_{(1, 3)} < 1$) (Fig. 2A, right). Thus, the association learning and discrimination learning posed different computational demands on the trained objects, resulting in subjects performing differently on the trained objects after the learning. Note that both stimulus exposure time (i.e., 180 trials per object in the association learning versus 155 trials in the discrimination learning on average) and subjects' end performance (i.e., no further improvement in performance) were approximately equated between two learning tasks.

To investigate the neural correlates of the task context on object learning, we examined the effect of perceptual experience on functional selectivities of the predefined ROIs after the learning. During the scan, a one-back task, rather than the tasks in behavioral training, was used. This orthogonal task was used to ensure that differences in cortical activation, if observed, resulted from the difference in the task context during the object learning and not from the difference in tasks performed during the scan. As expected, the behavioral performance in the one-back task was not significantly different between subject groups (supplemental Fig. 2, available at www.jneurosci.org as supplemental material).

To characterize the learning effect, the response to Trained versus Novel within the predefined ROIs (Fig. 3A) was compared. After the association learning, the magnitude of the response of each ROI was analyzed in a two-way ANOVA where the fact-

ors were object-selective cortical region (VWFA versus LO) and stimulus (Trained versus Novel). The main effects of both cortical region ($F_{(1, 15)} = 11.05, p < 0.01$) and stimulus ($F_{(1, 15)} = 5.57, p < 0.05$) were significant. More importantly, the amount of learning effects differed across cortical regions, suggested by a significant two-way interaction of cortical region by stimulus ($F_{(1, 15)} = 5.20, p < 0.05$). In fact, a significantly higher response to the trained objects (versus the novel objects) was found in the VWFA ($t_{(15)} = 3.11, p < 0.01$) but not in the LO ($t_{(15)} < 1$) (Fig. 3B). To further examine how the learning effect was distributed in the ventral visual cortex, we examined whether the left FFA, a cortical region that resides in the left fusiform gyrus near the VWFA, was shaped by the association learning. No learning effect was observed in the left FFA ($t_{(15)} < 1$) (Fig. 4A), and the interaction of the VWFA versus the left FFA by stimulus was significant ($F_{(1, 15)} = 10.35, p < 0.01$), suggesting that the learning-induced changes are restricted to specific regions and are not widespread across the visual system. Besides the object-selective regions in the left hemisphere, no learning effect was found in the right FFA ($t_{(15)} = -1.35, p > 0.05$) or right LO ($t_{(15)} < 1$).

A similar two-way ANOVA was performed for cortical responses after the discrimination learning. This ANOVA found a significant interaction of cortical region by stimulus ($F_{(1, 11)} = 7.44; p < 0.05$), again indicating that the amount of learning effect differs across ROIs. Neither the main effect of cortical region ($F_{(1, 11)} = 4.05, p > 0.05$) nor that of stimulus ($F_{(1, 11)} = 1.20, p > 0.05$) reached significance. Interestingly, the VWFA, which showed a higher response to the trained objects after the association learning, responded equally to the trained and novel objects after the discrimination learning ($t_{(11)} = -1.88, p > 0.05$). The LO, on the other hand, showed a significantly higher response to the trained objects ($t_{(11)} = 2.67, p < 0.05$) (Fig. 3C). However, no learning effect was observed in the right LO ($t_{(11)} = -1.43, p > 0.05$). The failure of observing the learning effect in the right LO might be caused by the hemispheric asymmetry in object perception, because the left hemisphere prefers objects with high-spatial frequency and the right hemisphere shows an opposite advantage (Sergent, 1983; Grabowska and Nowicka, 1996; Mercure et al., 2008). The stimuli in this study were high-contrast line-drawing objects that are preferred stimuli for the left hemisphere, whereas previous studies that reported the learning effect in the right LO used objects with smoothed surfaces in object discrimination tasks (Grill-Spector et al., 2000; Op de Beeck et al., 2006). In addition, we found that discrimination learning did not significantly induce a higher response to the trained objects in the FFA; instead, both left and right FFA responses to the trained objects were even lower than those of the novel objects ($p < 0.05$) (Fig. 4B).

In sum, experience with novel objects produced a greater effect on the VWFA when those stimuli were associated with words in the association learning, but it produced a greater effect on the LO when learned in the discrimination learning. The finding that the task context in object learning modulated the functional se-
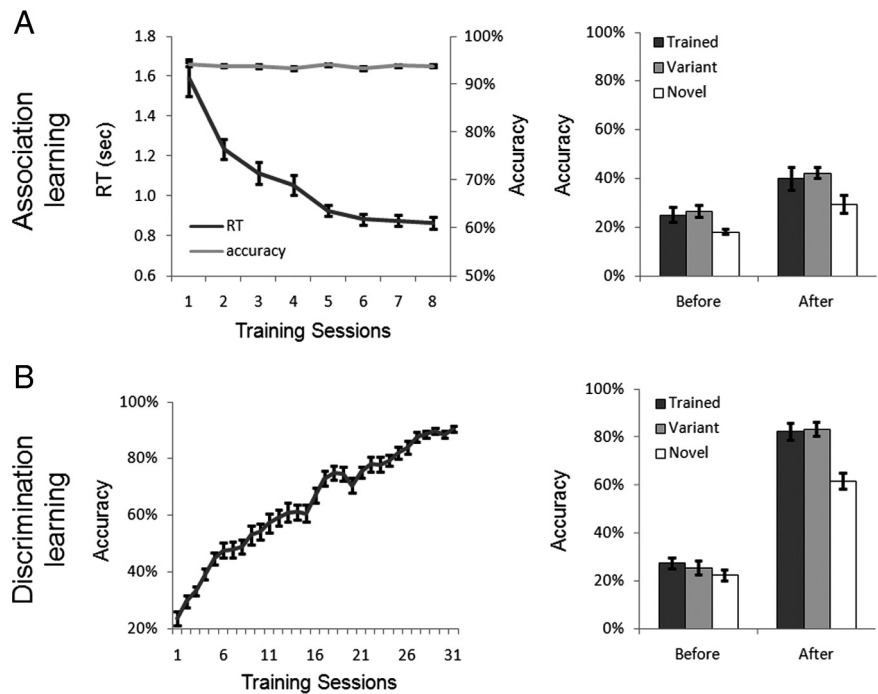

Figure 2. Behavioral training results. **A,** Association learning. Left, The RT monotonically decreased while the accuracy remained >90%. Note that the data from subjects who had fewer than eight training sessions were not included. Right, The association learning did not selectively increase subjects' ability in discriminating the trained objects. **B,** Discrimination learning. Left, The accuracy monotonically increased along with the number of training sessions. Right, After the learning, subjects' discrimination abilities on the trained objects were significantly improved over the novel objects and generalized to their variants. The error bars show the SEM of the behavioral performance across subjects.

lectivities of the high-level visual cortex is further supported by a significant three-way interaction of stimulus (Trained versus Novel), task context (association versus discrimination learning), and cortical region (VWFA versus LO) ($F_{(1, 26)} = 12.41, p < 0.005$). In addition, we also found a role of top-down task context in overall activation levels of the cortical regions revealed by a significant two-way interaction of cortical region by task context ($F_{(1, 26)} = 13.67, p < 0.005$). That is, after the association learning the overall activation level was smaller in the VWFA (versus LO), whereas after the discrimination learning the overall activation level in the LO was smaller (not significant). Although this interaction was based on a comparison between subjects, it implies that the task context may have different effects in modulating the overall activation level between cortical regions and the learning effect within a region. On the other hand, other factors such as stimulus exposure time and subjects' end performance might play a role as well, but they surely did not differ more than the top-down task context for this triple interaction. In fact, the analysis on individual differences in both exposure time and end performance suggests that they did not correlate with learning-induced neural changes observed in either the VWFA or the LO (supplemental Fig. 3, available at www.jneurosci.org as supplemental material).

Transfer of learning to the variants of the trained objects
The variants of the trained objects were included in the discrimination test before and after the discrimination learning to examine the transfer of learning from the trained objects to them (Fig. 2B, right). After the discrimination learning, the performance in discriminating the variants was equal to that for the trained objects ($t_{(3)} < 1$), and significantly higher than the novel objects ($t_{(3)} = 4.10, p < 0.05$) but not before ($t_{(3)} = 1.36, p > 0.05$). This result suggests
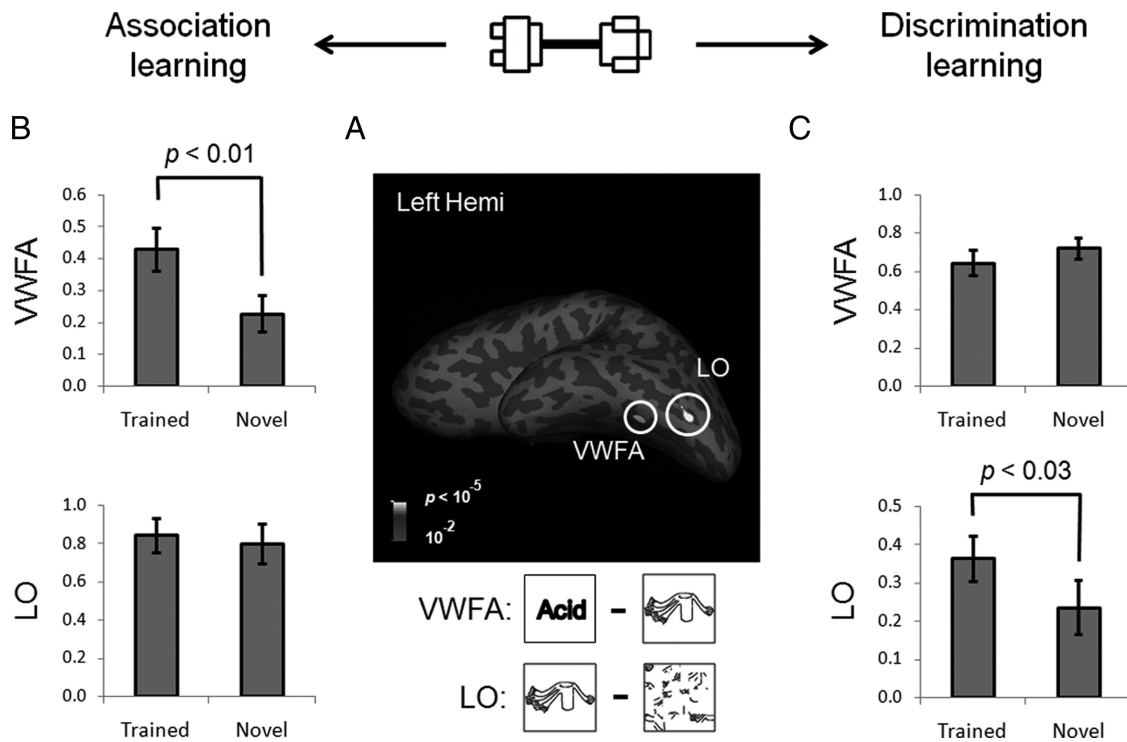
Figure 3. The learning effects in the VWFA and LO. **A**, Object-selective regions, the VWFA and LO, from the fMRI localizer scan in the left hemisphere (Left Hemi) of a typical subject shown on an inflated surface. Sulci are shown in dark gray and gyri in light gray. **B**, After the association learning, VWFA response to the trained objects was significantly higher than that of the novel objects, and no learning effect was observed in the LO. **C**, After the discrimination learning, the learning-induced change in blood oxygenation level-dependent (BOLD) responses was only observed in the LO, but not in the VWFA. The y-axis indicates the percentage signal change. The error bars show the SEM of the BOLD responses across subjects.
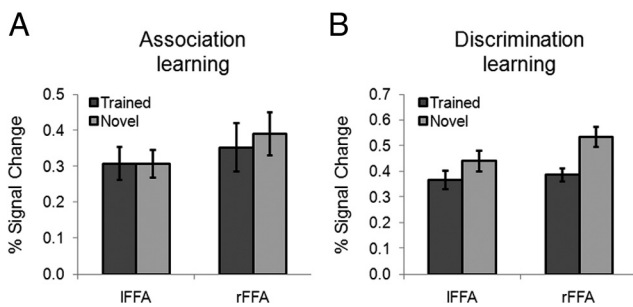

Figure 4. **A**, **B**, FFA responses to the trained objects and the novel objects after the association learning (**A**) and the discrimination learning (**B**). lFFA, Left FFA; rFFA, right FFA.

that the learning effect on the trained objects was generalized to the variants that shared the same parts with the trained objects. In line with the behavioral transfer of the learning, the LO responded to the variants as strongly as the trained objects ($t_{(11)} < 1$) and significantly more strongly than the novel objects ($t_{(11)} = 2.27, p < 0.05$), suggesting that the objects might be learned in a parts-based fashion (Fig. 5A, right).

To examine whether the parts-based processing strategy took place during the discrimination learning, we sorted the stimuli based on the part information they contained. Among all 16 target candidates shown in the display in each trial (Fig. 1C), one object was identical to the cue object (whole), two foil objects contained one part component of the target object (part), and the remaining distracters contained none (none) (Fig. 5A, left). If subjects' responses were based on one part component only, but not the two part components in combination, the subjects might have wrongly selected the foiled objects as the target. Therefore, the false alarm rate of selecting the foil objects as the target was

proportional to subjects' discrimination performance based on one part component. Because the chance of correctly selecting the target object based on one part component was 50%, the accuracy of correctly discriminating only one part component was two times of the false alarm rate. Similarly, the accuracy of correctly selecting the target based on the combination of two part components was the subtraction of the false alarm rate from subjects' total correct responses. As shown in the middle of Figure 5A, the performance based on only one part component first increased and then decreased along the learning (part), whereas the accuracy in discriminating the objects based on the combination of two parts was near 0% in the first session and then increased monotonically (whole). Thus, the processing of the trained objects in the discrimination learning was first based on one part component and then the combination of two components ($F_{(30, 330)} = 30.09, p < 0.001$). Note that the increased preference of using two part components in object discrimination does not necessarily imply the application of the holistic processing strategy. The contribution of each part component for the object discrimination may be additive (Macevoy and Epstein, 2009).

Similar parts-based processing strategy was observed in the association learning as well (Fig. 5B). At each association judgment, one of three types of dumbbell-shaped objects was presented: an object that was a correct associate of the presented English word (whole), an object that shared one part component with the target object (part), or an object that shared none (none) (Fig. 5B, left). If the parts-based strategy were at work, objects that contained only one part component would slow down the reaction time and decrease the accuracy because they contained conflict information for association judgment. This is exactly what is shown in the middle of Figure 5B, where the accuracy was decreased ($F_{(1, 12)} = 73.77, p < 0.001$) and the reaction time was
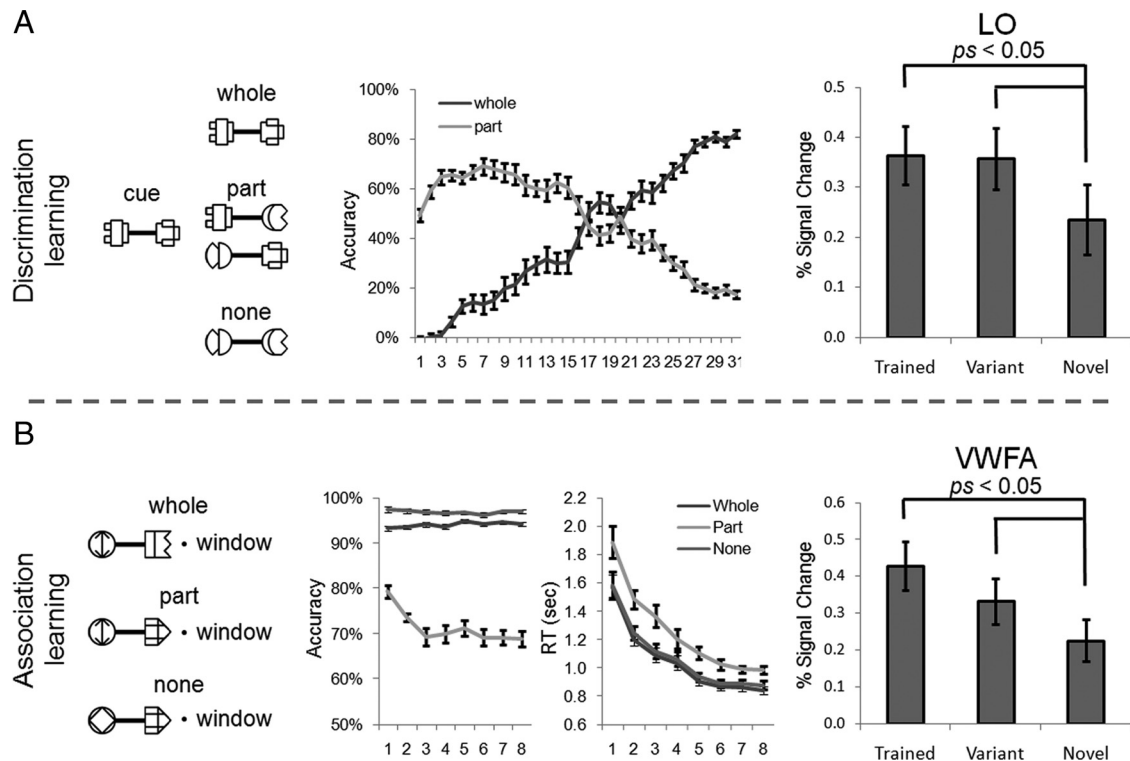
Figure 5. The parts-based processing strategy and the transfer of learning. **A**, Discrimination learning. Left, Subjects may match the cue object with the target object based on one part component or two part components in combination. Middle, The one-component-based performance (part) increased in early sessions and then decreased in the later sessions, whereas the two-component-based performance (whole) increased monotonically along the learning. Right, LO response to the variants that contained the part components of the trained objects was not different from the trained objects and was larger than the novel objects. **B**, Association learning. Left, The pairs for association judgment were three types: an object that was a correct associate of the presented English word (whole), an object that shared one part component with the target object (part), or an object that shared none (none). Middle, The association judgment on the incorrect pairs that shared one component was less accurate and slower. Right, VWFA response to the variants was as large as the trained objects and significantly larger than the novel objects.

delayed ($F_{(1, 12)} = 21.93$, $p < 0.001$) on objects that shared one part component with the target objects. In line with the parts-based processing strategy during the association learning, VWFA response to the variants that shared the part components with the trained objects was not significantly different from the trained objects ($t_{(15)} = 1.72$, $p > 0.05$) and was significantly larger than the novel objects ($t_{(15)} = 2.19$, $p < 0.05$) (Fig. 5B, right).

In sum, the learning effect observed in the VWFA and LO was transferred from the trained objects to their respective variants regardless of their behavioral relevance to the learning task. The behavioral learning curves further indicate that the transfer of the learning is caused by the parts-based strategy that was implicitly adopted during the learning.

Discussion

In this study, we investigated where and how learning-induced object representations were constructed in the visual system under different top-down task contexts. Subjects were trained to either associate a set of novel objects to English words or discriminate their shapes. We found that the association learning induced a larger VWFA response to the trained objects relative to the novel objects, whereas after the discrimination learning a larger response to the trained objects was only observed in the LO. The double dissociation of the VWFA and LO in response to the trained objects under different task contexts suggests that not only does the task context have differential learning effects on neurons within a cortical region, but it also determines which cortical regions are recruited in object learning. In addition, the transfer of learning to the variants that shared part components

with the trained objects was observed after both association and discrimination learning, suggesting that the trained objects were learned in a parts-based fashion.

Our study devised several novel designs to examine the role of task context in object learning. First, unlike most previous fMRI studies on object learning in which only one type of learning task was used, we presented two different computational problems in learning the same set of novel objects: association learning and discrimination learning. The physical properties and exposure time of the trained objects and subjects' end performance were approximately matched between learning tasks, so the difference in the spatial distribution of learning effects was mainly caused by the difference in top-down, task-relevant visual experience. Second, most previous studies included only trained and novel conditions, which is unlikely to illustrate how object learning takes place. Here, we included all three necessary conditions: the trained objects, the variants, and the novel objects. This design enabled us to directly examine how specific the object learning was and which perceptual processing strategy, parts-based or holistic, was used during the learning. Finally, although different tasks were used during learning, subjects performed an orthogonal task (i.e., the identity one-back task) in the scanner when learning effects on the visual system were mapped. The application of the orthogonal task largely avoids potentially confounding influences by the task itself and other possible confounding factors such as task difficulty or attention.

Although our study is one of the first ones to directly compare object learning under different task contexts with fMRI (see also Wong et al., 2009b), our results dovetail nicely with a number of

recent studies using either association or discrimination learning tasks in object learning. Neurophysiological studies on association learning in nonhuman primates have found that neurons in the inferior temporal cortex responded to both pictures of the paired associates after association learning regardless of the geometrical differences between them (Miyashita and Chang, 1988; Sakai and Miyashita, 1991; Erickson and Desimone, 1999; Erickson et al., 2000; Messinger et al., 2001; Schlack and Albright, 2007). Similarly, when novel objects were trained as navigation landmarks (Janzen and van Turennout, 2004), tools (Weisberg et al., 2007), or words (McCandliss et al., 1997; Hashimoto and Sakai, 2004; Sandak et al., 2004; Bitan et al., 2005; Breitenstein et al., 2005; Xue et al., 2006; Xue and Poldrack, 2007), regions involved in spatial navigation, manipulation of common tools, or word shape analysis, respectively, were activated. However, the learning effect observed in the VWFA may come from mental imagery instead, because through association subjects might have translated the trained objects to their corresponding words, which then activated the VWFA. Although intuitive, the mental imagery alternative was unlikely to account for our finding; VWFA response to the variants was as large as the trained objects, whereas the variants were unable to be veridically translated to any English words. In short, through association learning, object representations were developed in regions whose preferred stimuli were experientially associated with the new objects.

In parallel, neurophysiological and neuroimaging studies have also revealed that after discrimination learning the selectivity of neural responses to trained objects is increased in the inferior temporal cortex of monkeys (Kobatake et al., 1998; Baker et al., 2002) or in the LOC of humans (Grill-Spector et al., 2000; Op de Beeck et al., 2006; Jiang et al., 2007; Li et al., 2009). Recently, Op de Beeck et al. (2006) examined the spatial distribution of learning effects in shape discrimination and found that activated voxels for trained objects formed multiple discrete small clusters, many of which overlapped with the LOC. Similarly, our study also showed that, after the discrimination learning, the learning effect was observed in this shape analysis region but not in other specialized regions tested, such as the FFA or the VWFA.

The double dissociation of the VWFA and LO in response to the trained objects after learning under different top-down task contexts indicates that the size of the learning effect is not associated with the prior existing cortical responses to the to-be-trained objects (see also Op de Beeck et al., 2006). Learning does not simply produce a constant overall change of neural responses in the ventral system; instead, the spatial distribution of learning effects is largely restricted to specific regions. More importantly, it is not the responsive characteristics of these regions that justify its exclusive site for object learning. Rather, learning-induced object representations are constructed in regions that are posited to perform computations unique to the objects based on how they are learned. Thus, cortical regions of plasticity for object learning do not purely result from bottom-up visual processing; instead, this process appears to be dynamic and task dependent. The task-dependent plasticity in object learning may help explain why visually different objects (e.g., houses and scenes) are encoded in the same cortical region (e.g., parahippocampal place area) (Epstein and Kanwisher, 1998), possibly because they may be founded on the same computation (e.g., spatial navigation).

Interestingly, the learning-induced changes were not restricted to the trained objects. Instead, the learning was transferred to the variant objects that shared part components with the trained objects (see also Weisberg et al., 2007), suggesting that the trained objects were not encoded as a holistic entity but rather as a collection of relatively independent parts (Marr, 1982; Biederman, 1987; Golcu and Gilbert, 2009). In fact, the transfer of learning was not affected by the relation among components, because combinations of the part components in the variants were novel and task irrelevant. This finding is consistent with previous neurophysiological results showing that neurons are sensitive to simplified parts of objects (Desimone et al., 1984; Tanaka et al., 1991; Tsunoda et al., 2001), and it may explain why object-selective regions also respond to the parts of their preferred stimuli without veridical configurations, such as the FFA responsive to scrambled faces (Liu et al., 2009) and the VWFA responsive to pseudowords and consonant letter strings (Cohen et al., 2003; McCandliss et al., 2003; Vigneau et al., 2005; Baker et al., 2007; Vinckier et al., 2007). On the other hand, the representation did not generalize to visually similar objects that did not contain the parts of the train objects, suggesting that similarity in visual appearance is not sufficient for the transfer of learning. Together, our results suggest that the learning-induced representation is tolerant of a certain amount of deviance away from the trained objects through a parts-based process during the learning.

In sum, our findings help specify factors that shape the functional selectivities of the high-level visual cortex through learning. The top-down task context determines where object representations are constructed, whereas the parts-based perceptual process enables the representations tolerant to a broad range of variations through learning a small number of exemplars. In our study, several questions remain unaddressed that are important topics for future research. First, we examined learning effects in ROIs as most previous studies did; however, it is also important to characterize the spatial distribution of learning effects with a voxelwise whole-volume analysis so as to examine the interactive processing between different cortical regions, ranging from the bottom-up encoding object features in the occipitotemporal cortex to the top-down categorical object representations in the prefrontal cortex (Jiang et al., 2007). Second, although the representation is tolerant to a certain amount of deviance, the variants may not be coded as efficiently as the trained ones, even though the net magnitude of the responses was the same (Gillebert et al., 2009; Glezer et al., 2009). Finally, here we only directly manipulated the top-down task context (association versus discrimination learning). It is also plausible that the distribution of learning effects may be influenced by how informative the selectivity of regions is for to-be-trained stimuli as well (Op de Beeck and Baker, 2010). Consistent with this hypothesis, we found that the learning effect was restricted to the left hemisphere when high spatial-frequency stimuli were used, whereas previous studies using stimuli with smoothed surfaces reported a right-hemisphere advantage in object learning (e.g., Grill-Spector et al., 2000; Op de Beeck et al., 2006). Future studies that directly manipulate stimulus properties may help investigate this bottom-up mechanism in object learning.

References

Baker CI, Behrmann M, Olson CR (2002) Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. Nat Neurosci 5:1210–1216.

Baker CI, Liu J, Wald LL, Kwong KK, Benner T, Kanwisher N (2007) Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. Proc Natl Acad Sci U S A 104:9087–9092.

Biederman I (1987) Recognition-by-components: a theory of human image understanding. Psychol Rev 94:115–147.

Bitan T, Manor D, Morocz IA, Karni A (2005) Effects of alphabeticality, practice and type of instruction on reading an artificial script: an fMRI study. Brain Res Cogn Brain Res 25:90–106.

Breitenstein C, Jansen A, Deppe M, Foerster AF, Sommer J, Wolbers T, Knecht S (2005) Hippocampus activity differentiates good from poor learners of a novel lexicon. Neuroimage 25:958–968.

Bukach CM, Gauthier I, Tarr MJ (2006) Beyond faces and modularity: the power of an expertise framework. Trends Cogn Sci 10:159–166.

Cohen L, Dehaene S, Naccache L, Lehericy S, Dehaene-Lambertz G, Henaff MA, Michel F (2000) The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. Brain 123:291–307.

Cohen L, Lehericy S, Chochon F, Lemer C, Rivaud S, Dehaene S (2002) Language-specific tuning of visual cortex? Functional properties of the visual word form area. Brain 125:1054–1069.

Cohen L, Martinaud O, Lemer C, Lehericy S, Samson Y, Obadia M, Slachevsky A, Dehaene S (2003) Visual word recognition in the left and right hemispheres: anatomical and functional correlates of peripheral alexias. Cereb Cortex 13:1313–1333.

Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9:179–194.

Desimone R, Albright TD, Gross CG, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. J Neurosci 4:2051–2062.

Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. Nature 392:598–601.

Erickson CA, Desimone R (1999) Responses of macaque perirhinal neurons during and after visual stimulus association learning. J Neurosci 19:10404–10416.

Erickson CA, Jagadeesh B, Desimone R (2000) Clustering of perirhinal neurons with similar properties following visual experience in adult monkeys. Nat Neurosci 3:1143–1148.

Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. Neuroimage 9:195–207.

Gauthier I, Tarr MJ (2002) Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. J Exp Psychol Hum Percept Perform 28:431–446.

Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC (1999) Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. Nat Neurosci 2:568–573.

Gillebert CR, Op de Beeck HP, Panis S, Wagemans J (2009) Subordinate categorization enhances the neural selectivity in human object-selective cortex for fine shape differences. J Cogn Neurosci 21:1054–1064.

Glezer LS, Jiang X, Riesenhuber M (2009) Evidence for highly selective neuronal tuning to whole words in the "visual word form area." Neuron 62:199–204.

Golby AJ, Gabrieli JD, Chiao JY, Eberhardt JL (2001) Differential responses in the fusiform region to same-race and other-race faces. Nat Neurosci 4:845–850.

Golcu D, Gilbert CD (2009) Perceptual learning of object shape. J Neurosci 29:13621–13629.

Grabowska A, Nowicka A (1996) Visual-spatial-frequency model of cerebral asymmetry: a critical survey of behavioral and electrophysiological studies. Psychol Bull 120:434–449.

Grill-Spector K, Kushnir T, Hendler T, Malach R (2000) The dynamics of object-selective activation correlate with recognition performance in humans. Nat Neurosci 3:837–843.

Grill-Spector K, Kourtzi Z, Kanwisher N (2001) The lateral occipital complex and its role in object recognition. Vis Res 41:1409–1422.

Hashimoto R, Sakai KL (2004) Learning letters in adulthood: direct visualization of cortical plasticity for forming a new link between orthography and phonology. Neuron 42:311–322.

Jacobs RA (1999) Computational studies of the development of functionally specialized neural modules. Trends Cogn Sci 3:31–38.

Janzen G, van Turennout M (2004) Selective neural representation of objects relevant for navigation. Nat Neurosci 7:673–677.

Jiang X, Bradley E, Rini RA, Zeffiro T, Vanmeter J, Riesenhuber M (2007) Categorization training results in shape- and category-selective human neural plasticity. Neuron 53:891–903.

Kanwisher N (2001) Faces and places: of central (and peripheral) interest. Nat Neurosci 4:455–456.

Kobatake E, Wang G, Tanaka K (1998) Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. J Neurophysiol 80:324–330.

Kourtzi Z, DiCarlo JJ (2006) Learning and neural plasticity in visual object recognition. Curr Opin Neurobiol 16:152–158.

Kourtzi Z, Kanwisher N (2001) Representation of perceived object shape by the human lateral occipital complex. Science 293:1506–1509.

Li S, Mayhew SD, Kourtzi Z (2009) Learning shapes the representation of behavioral choice in the human brain. Neuron 62:441–452.

Liu J, Harris A, Kanwisher N (2009) Perception of face parts and face configurations: an fMRI study. J Cogn Neurosci 22:203–211.

Macevoy SP, Epstein RA (2009) Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. Curr Biol 19:943–947.

Marr D (1982) Vision: a computational investigation into the human representation and processing of visual information. San Francisco: Freeman.

McCandliss BD, Posner MI, Givòn T (1997) Brain plasticity in learning visual words. Cogn Psychol 33:88–110.

McCandliss BD, Cohen L, Dehaene S (2003) The visual word form area: expertise for reading in the fusiform gyrus. Trends Cogn Sci 7:293–299.

Mercure E, Dick F, Halit H, Kaufman J, Johnson MH (2008) Differential lateralization for words and faces: category or psychophysics? J Cogn Neurosci 20:2070–2087.

Messinger A, Squire LR, Zola SM, Albright TD (2001) Neuronal representations of stimulus associations develop in the temporal lobe during learning. Proc Natl Acad Sci U S A 98:12239–12244.

Miyashita Y, Chang HS (1988) Neuronal correlate of pictorial short-term memory in the primate temporal cortex. Nature 331:68–70.

Op de Beeck HP, Baker CI (2010) The neural basis of visual object learning. Trends Cogn Sci 14:22–30.

Op de Beeck HP, Baker CI, DiCarlo JJ, Kanwisher NG (2006) Discrimination training alters object representations in human extrastriate cortex. J Neurosci 26:13025–13036.

Riesenhuber M, Poggio T (2000) Models of object recognition. Nat Neurosci [Suppl] 3:1199–1204.

Sakai K, Miyashita Y (1991) Neural organization for the long-term memory of paired associates. Nature 354:152–155.

Sandak R, Mencl WE, Frost SJ, Rueckl JG, Katz L, Moore DL, Mason SA, Fulbright RK, Constable RT, Pugh KR (2004) The neurobiology of adaptive learning in reading: a contrast of different training conditions. Cogn Affect Behav Neurosci 4:67–88.

Schlack A, Albright TD (2007) Remembering visual motion: neural correlates of associative plasticity and motion recall in cortical area MT. Neuron 53:881–890.

Sergent J (1983) Role of the input in visual hemispheric asymmetries. Psychol Bull 93:481–512.

Sigman M, Pan H, Yang Y, Stern E, Silbersweig D, Gilbert CD (2005) Top-down reorganization of activity in the visual pathway after learning a shape identification task. Neuron 46:823–835.

Tanaka JW, Farah MJ (1993) Parts and wholes in face recognition. Q J Exp Psychol A 46:225–245.

Tanaka K, Saito H, Fukada Y, Moriya M (1991) Coding visual images of objects in the inferotemporal cortex of the macaque monkey. J Neurophysiol 66:170–189.

Tsunoda K, Yamane Y, Nishizaki M, Tanifuji M (2001) Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. Nat Neurosci 4:832–838.

Vigneau M, Jobard G, Mazoyer B, Tzourio-Mazoyer N (2005) Word and non-word reading: what role for the visual word form area? Neuroimage 27:694–705.

Vinckier F, Dehaene S, Jobert A, Dubus JP, Sigman M, Cohen L (2007) Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. Neuron 55:143–156.

Weisberg J, van Turennout M, Martin A (2007) A neural system for learning about object function. Cereb Cortex 17:513–521.

Wong AC, Palmeri TJ, Gauthier I (2009a) Conditions for facelike expertise with objects: becoming a Ziggerin expert–but which type? Psychol Sci 20:1108–1117.

Wong AC, Palmeri TJ, Rogers BP, Gore JC, Gauthier I (2009b) Beyond shape: how you learn about objects affects how they are represented in visual cortex. PLoS One 4:e8405.

Xue G, Poldrack RA (2007) The neural substrates of visual perceptual learning of words: implications for the visual word form area hypothesis. J Cogn Neurosci 19:1643–1655.

Xue G, Chen C, Jin Z, Dong Q (2006) Language experience shapes fusiform activation when processing a logographic artificial language: an fMRI training study. Neuroimage 31:1315–1326.