Learning to Discriminate Face Views

Taiyong Bi, Nihong Chen, Qiujie Weng, Dongjun He and Fang Fang

J Neurophysiol 104:3305-3311, 2010. First published 14 July 2010; doi:10.1152/jn.00286.2010

J Neurophysiol 104: 3305–3311, 2010.
First published July 14, 2010; doi:10.1152/jn.00286.2010.

Learning to Discriminate Face Views

Taiyong Bi,* Nihong Chen,* Qiujie Weng,* Dongjun He, and Fang Fang
Department of Psychology and Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, China

Submitted 22 March 2010; accepted in final form 7 July 2010

Bi T, Chen N, Weng Q, He D, Fang F. Learning to discriminate face views. *J Neurophysiol* 104: 3305–3311, 2010. First published July 14, 2010; doi:10.1152/jn.00286.2010. Although perceptual learning of simple visual features has been studied extensively and intensively for many years, we still know little about the mechanisms of perceptual learning of complex object recognition. In a series of seven experiments, human perceptual learning in discrimination of in-depth orientation of face view was studied using psychophysical methods. We trained subjects to discriminate face orientations around a face view (i.e., 30°) over eight daily sessions, which resulted in a significant improvement in sensitivity to the face view orientation. This improved sensitivity was highly specific to the trained orientation and persisted up to 6 mo. Different from perceptual learning of simple visual features, this orientation-specific learning effect could completely transfer across changes in face size, visual field, and face identity. A complete transfer also occurred between two partial face images that were mutually exclusive but constituted a complete face. However, the transfer of the learning effect between upright and inverted faces and between a face and a paperclip object was very weak. These results shed light on the mechanisms of the perceptual learning of face view discrimination. They suggest that the visual system had learned how to compute face orientation from face configural information more accurately and that a large amount of plastic changes took place at a level of higher visual processing where size-, location-, and identity-invariant face views are represented.

INTRODUCTION

Perceptual learning is the phenomenon that training can improve sensory feature discrimination and object recognition (Fahle and Poggio 2002). It has been observed in all sensory modalities and has been studied intensively in past decades because of its close links to cortical plasticity (Fahle 2005; Gilbert et al. 2001). In vision, the majority of perceptual learning studies focus on experience-dependent changes taking place at early processing stages in the visual system. It has been shown that training can improve performance in discriminating many elementary visual features, including contrast (Yu et al. 2004), orientation (Schoups et al. 1995), spatial phase (Berardi and Fiorentini 1987), stereoacuity (Fendick and Westheimer 1983), hyperacuity (Fahle and Edelman 1993), motion direction (Ball and Sekuler 1987), and texture (Karni and Sagi 1991). One of the central questions in perceptual learning is its specificity and generalization (transfer), which have profound implications for the underlying neural mechanisms (Gilbert et al. 2001). Perceptual learning in low-level vision is usually characterized by its specificity to trained visual attributes like retinal position, orientation, or eye. These characteristics point to early visual cortical areas as its neural substrate, where the visual topography is precisely mapped, the receptive fields are small, orientation tuning is sharp, and monocular neurons exist. Indeed, recent single-unit and brain imaging studies confirmed this idea and showed that the locus of cortical changes accompanying perceptual learning could be as early as in V1 (Furmanski et al. 2004; Pourtois et al. 2008; Schoups et al. 2001; Schwartz et al. 2002; but see also Ghose et al. 2002).

Perceptual learning also occurs with more complex visual stimuli such as shape, object, and face, as shown by several psychophysical studies. (Furmanski and Engel 2000; Gold et al. 1999; Sigman and Gilbert 2000). Parallel to low-level vision, the effect of object recognition training was also specific to the trained set of objects (Sigman and Gilbert 2000) but showed a higher degree of generalization than low-level vision (e.g., complete transfer across changes in object size; Furmanski and Engel 2000), which is in accordance with the size-invariant object representation in high-level visual cortex (Ito et al. 1995). Single-unit studies in monkeys have shown that training on recognition and discrimination of objects could induce changes in the strength and object selectivity of neuronal responses in the inferior temporal visual cortex (Baker et al. 2002; Freedman et al. 2006; Sigala and Logothetis 2002). Functional MRI (fMRI) studies in human subjects found that object recognition training could either enhance cortical responses in object-selective areas (Grill-Spector et al. 2000; Kourtzi et al. 2005; Tarr and Gauthier 2000) or change the spatial distribution of activity to trained objects in extrastriate cortex (Op de Beeck et al. 2006).

Although perceptual learning in high-level vision is receiving increased attention from researchers, relative to perceptual learning in low-level vision, its characteristics remain largely unknown and its neural mechanisms are still elusive. Here, we carried out seven experiments to systematically study the characteristics of perceptual learning in face view discrimination. *Experiment 1* established the basic learning protocol and showed that, after training, performance in face view discrimination could significantly improve in an orientation-specific and long-lasting manner. In *experiments 2–7*, we measured the transfer of the perceptual learning across changes in face size, face part, visual field, face identity, in-plane orientation, and object category. With the measurements of specificity and transfer, we can infer what is learned during training and the cortical locus and the neural mechanism of the face view learning (Fahle 2005).

We believe that face view learning is a good starting point to study perceptual learning in high-level vision. First, significant advance has been made in understanding how face view is represented in the visual cortex. In monkey, face view–selective neurons clustering in the superior temporal sulcus (STS) and inferior temporal cortex (IT) have been well studied (De

* T. Bi, N. Chen, and Q. Weng contributed equally to this work.
Address for reprint requests and other correspondence: F. Fang, Dept. of Psychology and Key Lab. of Machine Perception (Ministry of Education), Peking Univ., Beijing 100871, China (E-mail: ffang@pku.edu.cn).

was randomized. Their spatial positions were randomly distributed within a $6.2 \times 6.2°$ area whose center was coincident with the fixation point, with a constraint that these two face views were separated by $\geq 1.5°$ of visual angle. Subjects were asked to make a two-alternative-forced-choice (2-AFC) judgment of the orientation of the second face relative to the first face (left or right). A high-pitched tone was provided after a wrong response, and the next trial began 1 s after response. The θ varied trial by trial and was controlled by the QUEST staircase to estimate subjects' face view discrimination threshold (75% correct).

averaged as a measure of subjects' discrimination performance and plotted as a function of orientation. Note that subjects were randomly selected to be trained at either −30 or +30°. Because training at the two orientations induced a similar learning effect, for the sake of presentation simplicity, the discrimination performance functions for subjects trained at −30° were flipped horizontally and averaged together with the functions for subjects trained at +30°. Subjects' performance improvement at an orientation was calculated as (pretraining threshold – posttraining threshold)/pretraining threshold × 100%. To measure the time course of the training effect (learning curve), discrimination thresholds from 25 QUEST staircases in a daily training session were averaged and plotted as a function of training day. Learning curves were fitted with a power function (Jeter et al. 2009).

To quantify the transfer of training between the trained and the test stimuli, transfer index was defined as the ratio of performance improvement with the test stimulus and that with the trained stimulus. Performance improvement with the trained stimulus over eight daily training sessions was calculated as (1st day threshold – 8th day threshold)/1st day threshold × 100%. The test stimulus here had the same orientation as the trained stimulus. Paired *t*-test and independent-samples *t*-test were carried out for within-subject comparisons and between-subject comparisons, respectively.

RESULTS

Perceptual learning in face view discrimination

In *experiment 1*, we first measured subjects' face view discrimination thresholds at seven orientations of −90, −60, −30, 0, 30, 60, and 90° (Fig. 1, *A* and *B*). Subjects practiced for 8,000 trials during eight daily training sessions on face view discrimination at the orientation of 30°. Throughout the training course, their discrimination thresholds gradually decreased, which resulted in a 36% performance improvement (Fig. 1*C*). After training, we measured thresholds at the seven orientations again.

Before training, subjects had a significant lower threshold (better performance) at 0° than the thresholds at other orientations [all $t(7) > 4.7$, $P < 0.01$; gray line in Fig. 1*D*], which is consistent with the claim that 3D symmetric shapes are discriminated more efficiently than asymmetric ones (Liu and Kersten 2003). After training, the threshold at 30° was comparable to that at 0° [$t(7) = 0.45$, $P = 0.67$] and was significantly lower than those at other orientations [all $t(7) > 4.4$, $P < 0.01$; black line in Fig. 1*D*]. We calculated the percent improvement in discrimination performance after training. The improvement at the trained orientation of 30° was 44%, which was significantly higher than those (about or <10%) at other orientations [black line in Fig. 1*E*; all $t(7) > 4.7$, $P < 0.01$]. These results suggest an orientation-specific perceptual learning in face view discrimination.

To examine the persistence of the learning effect, we measured the discrimination thresholds 1 and 6 mo after training (red and green lines in Fig. 1*D*). Relative to the performance before training, the percent improvements in discrimination performance at the trained orientation of 30° was 34% 1 mo after training and 33% 6 mo after training (red and green lines in Fig. 1*E*). This means that the learning effect was long-lasting, and 75% of the effect was kept after a half-year break. It is noteworthy that the learning effect persisted in an orientation-specific manner. The improvements at the untrained orientations were around or <10%, which were significantly lower than the improvement at 30° [all $t(7) > 3.5$, $P < 0.01$, except the marginal significance at 90° 6 mo after training, $t(7) = 2.1$, $P = 0.077$]. The long-lasting orientation-specific perceptual learning in face view discrimination was quite robust and consistent across individual subjects.

Transfer of the face learning effect

Experiments 2–7 were designed to study the transfer of training from trained stimuli to test stimuli. The test stimuli were always faces. The trained stimuli shared more or less properties with the test stimuli. Similar to *experiment 1*, subjects underwent eight daily training sessions to discriminate views of the trained stimulus at the orientation of 30°. Before and after training, we measured subjects' face view discrimination thresholds with the test stimulus at the seven orientations of −90, −60, −30, 0, 30, 60, and 90°.

Experiment 2 studied how a size change from the trained stimulus to the test stimulus could affect the transfer of learning (Fig. 2*A*). The area of the test stimulus was four times that of the trained stimulus. The 8-day training resulted in a 31% improvement in discrimination performance with the trained stimulus at the orientation of 30°. Relative to the performance before training, the improvement after training with the test stimulus at 30° was 33%, significantly higher than the improvements at the untrained orientations [all $t(7) > 2.5$, $P < 0.05$].

Experiment 3 investigated the effect of face part change on the transfer of learning. The trained and the test stimuli were mutually exclusive, and they constituted a complete face (Fig. 2*B*). Throughout the training course, subjects' discrimination thresholds gradually decreased, which resulted in a 38% performance improvement with the trained face. Relative to the pretraining performance, the improvement after training with the test face at 30° was 41%, significantly higher than the improvements at the untrained orientations [all $t(5) > 3.1$, $P < 0.05$].

In *experiment 4*, the trained face was presented in the lower visual field, and the test face was presented in either the lower or the upper visual fields (Fig. 2*C*). The objective of the experiment was twofold. First, *experiments 1–3* showed that face view learning could take place around the fixation point. Here we examined if the learning could occur at a more eccentric area, e.g., the lower or the upper visual field. Second, we were interested in the effect of visual field change on the transfer of learning. Similar to *experiments 1–3*, eight training sessions in the lower visual field led to a 27% performance improvement. Before and after training, we measured face view discrimination thresholds in both the lower and the upper visual fields. Subjects' performance improvements at the trained orientation were 33 and 32% in the lower and the upper visual fields, respectively, both of which were orientation specific [lower visual field: all $t(5) > 2.6$, $P < 0.05$, dotted line; upper visual field: all $t(5) > 3.3$, $P < 0.05$, solid line].

Experiment 5 examined the effect of face identity change on the transfer of learning (within-category transfer). The trained face was Anti-Jim and the test face was Jim (Fig. 2*D*). They were a face/anti-face pair, which lay at the two ends of a face identity trajectory (Leopold et al. 2001). After eight daily training sessions, subjects' performance with the trained face improved by 37%. We also found that the learning transfer was orientation specific because the improvement with the test stimulus at 30° was 41%, significantly higher than those at the untrained orientations [all $t(6) > 2.5$, $P < 0.05$].

In *experiments 6* and *7*, the test stimulus was a face and the trained stimuli were an inverted face (Fig. 2*E*) and an M-like paperclip (Fig. 2*F*). Here we examined how face inversion and object category change affected the learning transfer. Similar to the training effect with an upright face, training with the inverted face and the paperclip at the orientation of 30° also improved subjects' discrimination performance by 36 and 45%, respectively. However, compared with *experiments 2–5*, the performance improvements with the test face at 30° were weak (13 and 22%). These improvements were not orientation-specific because there was no significant difference between the trained orientation and the untrained orientations [inverted face: all $t(6) < 1.9$, $P > 0.12$; paperclip: all $t(7) < 2.3$; $P > 0.07$].

To quantify the transfer of training between the trained and the test stimuli, the transfer index was calculated as the ratio of performance improvement with the test stimulus and that with the trained stimulus (Fig. 3). A large index means that a large amount of the training effect has been transferred to the test stimulus; in other words, the performance improvement with the test stimulus can be largely attributed to the training effect. The transfer indices in *experiments 1–5* were 1.23, 1.12, 1.06, 1.27, and 1.18. There was no significant difference among them [$F(6,50) = 1.176$, $P = 0.336$]. Note that, in *experiment 1*, the test and the trained stimuli were the same. These results suggest a complete transfer from the trained stimulus to the test stimulus in *experiments 2–5*. Why are the indices larger than 1? This is because the threshold measurement before training also led to some learning effect. The transfer indices in *experiments 6* and *7* were 0.37 and 0.54, respectively. There was no significant difference between them ($t = 0.85$, $P = 0.41$). However, they were significantly lower than the indices in *experiments 1–5* [$F(8,65) = 5.616$, $P < 0.001$], which suggests a partial or weak transfer in *experiments 6* and *7*.

DISCUSSION

We see faces from various viewing angles every day. Face view perception informs us not only about a person's identity but also about his/her social attention. Even from a small face view change, we could infer changes in their current goals and intentions (Nummenmaa and Calder 2009). Can our ability of face view perception (discrimination) be improved with training? In this study, a series of seven experiments was conducted to address this question and to investigate the characteristics of perceptual learning of face view discrimination. *Experiment 1* showed that training led to a significant improvement in sensitivity to face view orientation. The improvement was highly specific to the trained orientation and lasted up to 6 mo. In *experiments 2–5*, we found that the orientation-specific learning effect completely transferred across changes in face size, visual field, and face identity. A complete transfer also occurred between two partial face images that were mutually exclusive but constituted a complete face. However, the transfers were weak between an upright face and an inverted face and between a paperclip object and a face, as shown in *experiments 6* and *7*. It should be noted that, in most experiments, only one face stimulus was used. These conclusions can be further strengthened if more face stimuli were used.

Face view learning exhibited two important characteristics of perceptual learning: specificity and persistence (Liu 1999; Sasaki et al. 2010). Training at 30° had a very weak effect on the discrimination performance at other orientations, even at 0 and 60°. It could be argued that subjects' view sensitivity at 0° (cardinal orientation) was already very high, leaving little room for improvement. To rule out the explanation, we trained two subjects using the same procedure as that for 30°. The training resulted in a 48% performance improvement at 0°, comparable to the training effect at 30°. It has been reported that some face neurons in STS responded symmetrically to left and right views (De Souza et al. 2005). This would predict that training at 30° should also lead to a higher performance improvement at $-30°$ than at 0, ± 60, and $\pm 90°$. However, we did not find such an effect, which indicates that training might have a very weak or little influence on these STS neurons. The benefits of perceptual learning with visual features are usually long-lasting, persisting for up to 2 yr (Karni and Sagi 1993). In high-level vision, training effects with a shape/object identification task could last 1 mo (Furmanski and Engel 2000; Sigman and Gilbert 2000). Here, we expanded these results by showing that the orientation-specific face view learning could last $\leq$6 mo.

Using a visual search or identification task, past studies indicated that face/shape recognition is subject to perceptual learning (Furmanski and Engel 2000; Golcu and Gilbert 2009; Hussain et al. 2009; Sigman and Gilbert 2000). However, few of them studied the characteristics of high-level visual perceptual learning as comprehensively as this study. There are two similar findings in previous studies and ours. One is the complete transfer across a change in face size, in agreement with the finding that object learning was insensitive to image size (Furmanski and Engel 2000). The other is the weak transfer from an upright face to its vertical inversion. Hussain et al. (2009) also found that face identification training was largely specific to the in-plane orientations of trained faces. Both studies suggest that the neural codings of upright faces and inverted faces are quite different. However, a significant contrast between past studies and ours is whether learning was restricted to the area where the stimulus was trained. Our study showed a complete transfer from the lower visual field to the upper visual field. Such a transfer between visual fields was not found in other studies (Cox and DiCarlo 2008; Dill and Fahle 1997; Nazir and O'Regan 1990). Two noteworthy points can explain the discrepancy. First, in these studies, subjects were trained to identify simple shapes that can be coded by retinotopic areas (e.g., V2 and V4; Hedge and Van Essen 2007). Second, these shapes were trained at a fixed position in the visual field. The face stimuli in our study were randomly presented in a small area, which might subserve the spatial transfer of the learning effect.

Where does face view learning occur in the visual cortex? In the human visual cortex, there are three face-selective areas: OFA, STS, and FFA (Fang et al. 2007; Haxby et al. 2000), which are the possible loci of the learning. Also, a recent study (Sigman et al. 2005) suggested that retinotopic areas played an important role in high-level visual perceptual learning with a shape identification task. By measuring the transfer of learning from the trained face to other visual stimuli more or less resembling it, we can infer where face view learning took place in the visual system and what was learned during training. First, the complete transfers across changes in face size, face part, and visual field could rule out early retinotopic areas and OFA. Because the receptive fields of neurons in early retinotopic areas are small (Smith et al. 2001), they are sensitive to stimulus changes in size, local feature, and position in the visual field. Although OFA is at a higher position in the visual processing hierarchy than early retinotopic areas, it is still retinotopically organized (see a review by Wandell et al. 2007)and is sensitive to changes in face part (Pitcher et al. 2007). Thus the properties of early retinotopic areas and OFA do not support the complete transfers we observed. Second, the complete transfer of the learning from a face to its anti-face suggests that the learning effect is identity invariant, which resonates with the finding that most view-selective face neurons in macaque STS are not sensitive to identity (Perrett et al. 1992). Using fMRI adaptation, Fang et al. (2007) showed that both STS and FFA in human visual cortex could encode face views. However, Grill-Spector et al. (2004) showed that FFA, but not STS, is selective for face identity. Thus STS is more likely to be the cortical locus of face view learning than FFA. Third, the weak transfers from the trained face to its in-plane inversion and to the paperclip object can rule out an alternative explanation that the face view discrimination learning is a general 3D orientation discrimination learning and takes place at cortical areas coding 3D orientation. Hinkle and Connor (2002) reported that some neurons in macaque area V4 carry robust signals for 3D orientation defined by binocular disparity. Although the orientation of the face views was not defined by disparity, the finding by Hinkle and Connor does suggest the alternative explanation for our observation. A key prediction of this 3D orientation learning explanation is that it should not depend on the object representation in the visual system, as long as the trained and the test stimuli are in the same 3D orientation. In our study, however, only weak transfers were

GRANTS

This work was supported by the National Natural Science Foundation of China (Projects 30870762, 90920012, and 30925014) and the Ministry of Science and Technology of China (2010CB833903).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

REFERENCES

Andrews TJ, Ewbank MP. Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *NeuroImage* 23: 905–913, 2004.

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

Ball K, Sekuler R. Direction-specific improvement in motion discrimination. *Vision Res* 27: 953–965, 1987.

Berardi N, Fiorentini A. Interhemispheric transfer of visual information in humans: spatial characteristics. *J Physiol* 384: 633–647, 1987.

Bi T, Su J, Chen J, Fang F. The role of gaze direction in face viewpoint aftereffect. *Vision Res* 49: 2322–2327, 2009.

Cox DD, DiCarlo JJ. Does learned shape selectivity in inferior temporal cortex automatically generalize across retinal position? *J Neurosci* 28: 10045–10055, 2008.

De Souza WC, Eifuku S, Tamura R, Nishijo H, Ono T. Differential characteristics of face neuron responses within the anterior superior temporal sulcus of macaques. *J Neurophysiol* 94: 1252–1266, 2005.

Desimone R, Albright TD, Gross CG, Bruce C. Stimulus-selective properties of inferior temporal neurons -328-9-ophysiol