
ELSEVIER
VISION
RESEARCH
An International Journal for Functional Aspects of Vision
L
R

Vision Research 86 (2013) 107–114

Contents lists available at SciVerse ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

The effect of spatial frequency on perceptual learning of inverted faces

Adélaïde de Heering *, Daphne Maurer

McMaster University, Hamilton, Ontario, Canada

ARTICLE INFO

Article history:
Received 16 December 2011
Received in revised form 22 April 2013
Available online 3 May 2013

Keywords:
Faces
Learning
Training
Spatial frequency
Inverted
Adults

ABSTRACT

We investigated the efficacy of training adults to recognize full spectrum inverted faces presented with different viewpoints. To examine the role of different spatial frequencies in any learning, we also used high-pass filtered faces that preserved featural information and low-pass filtered faces that severely reduced that featural information. Although all groups got faster over the 2 days of training, there was more improvement in accuracy for the group exposed to full spectrum faces than in the two groups exposed to filtered faces, both of which improved more modestly and only when the same faces were shown on the 2 days of training. For the group exposed to the full spectrum range and, to a lesser extent, for those exposed to high frequency faces, training generalized to a new set of full spectrum faces of a different size in a different task, but did not lead to evidence of holistic processing or improved sensitivity to feature shape or spacing in inverted faces. Overall these results demonstrate that only 2 h of practice in recognizing full-spectrum inverted faces presented from multiple points of view is sufficient to improve recognition of the trained faces and to generalize to novel instances. Perceptual learning also occurred for low and high frequency faces but to a smaller extent.

1. Introduction

Perceptual learning refers to an increase in the ability to extract information from the environment, as a result of practice and experience (Gibson, 1969; for other definitions see Ball & Sekuler, 1987; Fiorentini & Berardi, 1981; Schoups, Vogels, & Orban, 1995). It has been demonstrated for simple stimuli such as gratings (Ball & Sekuler, 1987; Fahle, Edelman, & Poggio, 1995; Fiorentini & Berardi, 1981; Karni & Sagi, 1991; McKee & Westheimer, 1978; Poggio, Fahle, & Edelman, 1992; Schoups, Vogels, & Orban, 1995) and for complex visual stimuli such as shapes and objects (Furmanski & Engel, 2000; Gold, Bennett, & Sekuler, 1999; Rubin, Nakayama, & Shapley, 1997; Sigman & Gilbert, 2000; Yi, Olson, & Chun, 2006). Improvement is often specific to the stimuli used during training (for reviews, see Levi & Li, 2009; Sagi & Tanne, 1994). For example, practice with feedback improves accuracy on a spatial frequency discrimination task, but changing the spatial frequency of the target by an octave, or its orientation by 90° abolished these effects (Fiorentini & Berardi, 1981). Specificity was also found after training on the discrimination of the direction of motion, on the perception of contour, and on figure-ground segmentation (Ball & Sekuler, 1987; Fiorentini & Berardi, 1981; Rubin, Nakayama, & Shapley, 1997; Sigman & Gilbert, 2000; Yi, Olson, & Chun, 2006). The results for more complex stimuli such as objects are mixed. Some authors found improvement restricted to the trained set of objects such as triangles of a particular size and orientation (Sigman & Gilbert, 2000) whereas others showed that improvement in recognition of common grey-scaled objects transferred almost completely across changes in image size (Furmanski & Engel, 2000).

Perceptual learning had also been used to explore the plasticity of domains in which adults have expertise, such as face processing. Practice with feedback over several days significantly improves accuracy for recognizing the identity of upright faces despite the fact that before training adults had had a lifetime of exposure to that category of stimuli (e.g., Dolan et al., 1997; Gold, Sekuler, & Bennett, 2004; Gold, Bennett, & Sekuler, 1999; Hussain, Sekuler, & Bennett, 2009a, 2009b). Recently, Hussain, Sekuler, and Bennett (2011) also showed that this improvement was maintained on retests approximately 1 year after training. The effects of training with upright faces have sometimes been found to transfer to novel faces (Jim/Anti-Jim, Bi et al., 2010; another twin picture, Robbins & McKone, 2003), and sometimes not (Hussain, Sekuler, & Bennett, 2009b, 2011). In contrast, the training effects do transfer to novel points of view (Dwyer et al., 2009), changed illumination (Moses, Ullman, & Edelman, 1995), and changes in size and visual field (Bi et al., 2010).

Adults' poorer processing of inverted faces than of upright faces (Yin, 1969) is typically attributed to limited exposure to this face category (e.g., Rossion, 2009). A few studies have examined whether increased exposure – through training – can improve adults' discrimination of inverted faces (Bi et al., 2010; Dwyer

* Corresponding author. Address: Institute of Research in Psychology (IPSY) & Institute of Neuroscience (IoNS), Université Catholique de Louvain, 10 Place du Cardinal Mercier, 1348 Louvain-la-Neuve, Belgium. Fax: +32 10 47 37 74.
E-mail address: adelaide.deheering@uclouvain.be (A. de Heering).

0042-6989/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.visres.2013.04.014

et al., 2009; Hussain, Sekuler, & Bennett, 2009b; Laguesse et al., 2012; Moses et al., 1995; Robbins & McKone, 2003). All demonstrated that training with inverted faces is effective but to a lesser extent than what is observed for upright faces when the latter were used for comparison (Bi et al., 2010; Dwyer et al., 2009; Hussain, Sekuler, & Bennett, 2009b; Moses et al., 1995; Robbins & McKone, 2003). From these studies, evidence of generalization to novel inverted faces is mixed: Hussain, Sekuler, and Bennett (2009b) found limited evidence for it while Laguesse et al. (2012) showed a significant decrease of the face inversion effect after training with inverted faces even though novel face identities were used at post-test. The authors attributed their success to the length of the challenging training they used (2 weeks), the large number of faces they presented during training (30 faces), the different depth-rotated views of the training faces and their inclusion of a pre-test and a of post-test composed of novel face identities.

In the present study, we also attempted to enhance the training effects for inverted faces by discouraging the learning of specific instances and instead encouraging the development of an effective processing strategy that could be generalized to new instances. Specifically, we trained one group of participants with multiple faces, each of which was presented from a number of points of view. In addition, we examined whether spatial frequency filtering influenced learning. To this end, we presented a second group of participants with high spatial frequency faces that emphasize the featural information that adults can use almost as efficiently in processing inverted as upright faces (e.g., Collishaw & Hole, 2000; Maurer, Le Grand, & Mondloch, 2002; Rossion, 2008, 2009) and a third group of participants with low spatial frequency faces that de-emphasize those features to encourage the use of more global information of the type that adults use efficiently for upright but not inverted faces (e.g., Goffaux & Rossion, 2006; Maurer, Le Grand, & Mondloch, 2002; Rossion, 2008, 2009). Although some studies showed that adults use the same mid-spatial frequencies to process upright and inverted faces (Boutet, Collin, & Faubert, 2003; Gaspar, Sekuler, & Bennett, 2008; Watier, Collin, & Boutet, 2010; Willenbockel et al., 2010), it has also been demonstrated that holistic/global face perception is supported by low spatial frequencies in adults (Goffaux & Rossion, 2006). Based on the latter evidence, we expected high-pass and low-pass filtering to selectively encourage the learning of featural or of holistic/configural strategies that might in turn affect differently the patterns of generalization.

The training paradigm was based on the short regime used by Hussain, Sekuler, and Bennett (2009b) to induce improvements with full-spectrum inverted faces. Specifically, participants were trained over 2 days to view a face then find it among 10 facial images. In order to test transfer of training, half the participants were trained with the same 10 faces on the second day of training, and half, with a set of 10 new faces. Unlike Hussain, Sekuler, and Bennett (2009b), the target face varied across 7 different viewpoints, while the choice faces were always presented in full-front view. This variation was introduced to encourage the learning of a general strategy, rather than specific images.

To further explore the extent of learning and its generalization to novel exemplars, participants were also tested before and after the 2 days of training on 4 tasks composed of full spectrum faces not used during training and of a different size than the trained faces: a simultaneous face matching task (Task 1), a delayed face matching task (Task 2), a composite task that measures holistic face processing (Task 3) and the Jane task that measures sensitivity to differences in the shape of features and their spacing (Task 4). Changes from pre-test to post-test in the trained groups were compared to those obtained from a control group that was tested twice at the same intervals but without intervening training. Based on the previous study by, Hussain, Sekuler, and Bennett (2009b) using the same training paradigm, we expected that generalization to novel instances of inverted faces would be unlikely (Task 1–Task 2). We also thought that any enhancement in holistic processing (Task 3), or in sensitivity to feature spacing (Task 4), would be most likely after training on low spatial frequency faces because of its emphasis on global processing and that any enhancement in featural processing (Task 4) would be most likely after training to high spatial frequency faces because of its emphasis on featural information.

1.1. Methods

1.1.1. Participants

The sample consisted of 64 participants between the ages of 18 and 30 years ($X = 21$; SD = 2.7) who participated either for remuneration or for credit in a psychology course. All had normal or corrected-to-normal vision. Specifically, their linear letter acuity (Lighthouse Visual Acuity Chart) was at least 20/20, they showed fusion at near on the Worth four-dot test and they had stereo acuity of at least 40 arcsec on the Titmus test. Sixteen participants were assigned to each of the 3 training groups and 16 to the control condition and not trained at all.

1.1.2. Procedure

The Research Ethics Board of McMaster University approved the study. Informed written consent was obtained from all participants prior to testing and they received a debriefing form at the end of the experiment. Participants came to the lab for 1 h on 4 consecutive days. On the first and last day, they completed the pre-test and post-test, respectively. On the second day and third day, except for the control group, they received training with feedback on inverted faces.

1.1.2.1. Pre-test and post-test. Participants were seated in a dark room 100 cm from a Dell Trinitron P1140 computer screen (51 cm diagonally) controlled either by a Mac Mini running on OSX 10.4.2 (Tasks 1 and 2) or a PowerMac G4 cube running on OS.9.2.1 (Tasks 3 and 4). Stimulus presentation was controlled by Superlab (version 4.0.7b for Tasks 1 and 2 and version 1.77 for Tasks 3 and 4). Stimuli always consisted of grey scale images of faces. Accuracy (% correct responses) and reaction times (ms) were recorded.

The order of the task was counterbalanced across participants but remained identical for each participant from the pre-test to the post-test.

Tasks 1 and 2 were adapted from Busigny and Rossion (2010) to test participants' recognition abilities for faces presented across different viewpoints. They were AB-X tasks in which participants used the mouse to click on the face with the same identity as the target face presented at the top of the screen from two ¾ profile faces presented at the bottom of the screen (Task 1) or on another screen (Task 2). In Task 1, a trial started with a fixation cross presented for 100 ms and was followed by 3 faces (the target face, the matching face, and the distractor) presented simultaneously. The trial ended with participant's response and was followed by the next trial after a 100 ms inter-stimulus interval (ISI). In this task, stimuli subtended approximately 5.7° by 7.1° of visual angle from the testing distance of 100 cm. Task 2 was the same except that the target face disappeared after 100 ms, and following a 500 ms delay, the matching and distractors face appeared and remained on the screen until the participant's response. In this task, each face subtended approximately 7.1° by 9.2° of visual angle from the testing distance of 100 cm. In both tasks, there were 60 trials per block.

Task 3 used the composite face effect, originally described by Young, Hellawell, and Hay (1987) and Hole (1994), to measure holistic face processing. We used a variant of the task and more

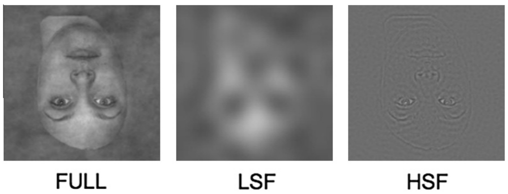

Fig. 2. Example of a full spectrum, a low spatial frequency filtered (<5 cpi) and a high spatial frequency filtered (>24 cpi) face stimulus.

specifically the stimuli and procedure of Le Grand et al. (2004) in which participants are shown two faces sequentially and must decide if their top halves are the same or different. In separate blocks, the face parts were either misaligned (to offset interference from holistic processing) or aligned with each other (to encourage holistic processing). Stimuli in the aligned condition were 5.6° by 8° of visual angle, from a distance of 100 cm. Stimuli in the misaligned condition were 8.4° by 8° of visual angle, from a distance of 100 cm. Each block started with 4 practice trials without feedback followed by an intermixed series of trials with faces sharing an identical top half ($n = 24$) or different top halves ($n = 24$). On every trial, the bottom halves of the faces were different. Holistic processing makes it difficult to ignore the bottom halves in the aligned condition, leading to errors and increased reaction times on same trials. Each trial started with a fixation cross appearing in the middle of the screen. After the participant pushed the space bar to start the trial, a face appeared centrally for 200 ms, followed by a 300 ms ISI, and a second face for 200 ms. Participants' response (same or different top half) ended the trial and triggered the fixation cross. The order of the aligned and misaligned blocks was randomized across participants but always the same for each participant in the pre-test and the post-test.

Task 4, the Jane task originally developed by Mondloch, Le Grand, and Maurer (2002) was created to test participants' ability to detect featural manipulations (i.e., Jane's eyes and mouth replaced by the facial features of different females) and spacing manipulations (i.e., Jane's eyes moved up, down, closer together, or farther apart together with her mouth moved up or down). All stimuli were 5.7° by 9.1° of visual angle, from a testing distance of 100 cm. The experimenter initiated the task by introducing Jane and her sisters who look alike but are different people. The featural and spacing/configural manipulations were presented in blocks of 30 trials respectively, with the correct response 'same' for half the trials in each block. The order of the blocks was randomized across participants but always the same for each participant in the pre-test and the post-test. A trial started with a fixation cross. When participants pressed the spacebar, it brought up the first face centrally on the screen for 200 ms directly followed by a 300 ms ISI. Then the second face appeared in the center of the screen until the participant signaled whether the two faces were the same or different. There were three practice trials before the test.

1.1.2.2. Training. Stimuli were inverted greyscale digitized photographs of two sets of 10 Caucasian women aged 17–25 years taken from 7 viewpoints: frontal, turned 45° to the left and to the right, turned 90° to the left and to the right (profile views), looking straight up, and looking straight down (Fig. 1). All models had minimal make-up and neutral expressions and they wore caps to conceal their external features such as hair and ears. Target faces and choice faces were centered on a uniform 250 × 250 pixels grey background and subtended 3.4° by 3.4° of visual angle when viewed from the training distance of 66 cm.

These stimuli were Fourier transformed using ideal band-pass filters to generate the low spatial frequency (LSF) and the high spatial frequency (HSF) versions of each full-spectrum (FULL) facial identity (Fig. 2). Specifically, for the LSF faces, we filtered the original images with a band-pass ideal filter, with the low and high cutoffs placed at 0.0001 and 5 cycle per image [cpi], respectively (1 cycle per degree [cpd] when viewed at 66 cm). We chose these cutoffs with the objective of eliminating featural information and simulating the range of spatial frequencies newborns use to acquire information about faces, namely 0–1 cpd (Acerra, Burnod, & de Schonen, 2002; Banks & Bennett, 1991; de Heering et al., 2008). When exposure to this range is not received during early infancy—because of bilateral congenital cataracts—normal face expertise later fails to develop (e.g., de Heering & Maurer, 2012; Le Grand et al., 2001, 2004). For the HSF faces, we used a band-pass ideal filter placed at 24 cpi (4.8 cpd when viewed at 66 cm) and 100 cpi (no visible energy remains after this threshold) that preserved featural information that adults can use when processing inverted faces (e.g., Maurer, Le Grand, & Mondloch, 2002). After filtering, the mean luminance of the FULL, LSF and HSF faces was of 81 cd/m^2, 81 cd/m^2 and 57 cd/m^2 and their mean RMS contrast was of 0.20, 0.18 and 0.04, respectively.

Participants were trained in a dark room while seated 66 cm away from a Dell Trinitron P1140 computer screen. A Mac mini (Apple) with Superlab 4.0.7b controlled the presentation of the stimuli. Accuracy (% of correct responses) and reaction times (ms) were recorded. Training occurred during 2 sessions of approximately 1 h spread over 2 consecutive days. On the second day of training, half of the participants were trained with the same faces as on the first day (Group 1) and half with a new set of faces (Group 2), in each case with the same spatial frequency filtering. Each training session started with 10 practice trials followed by a total of 630 trials displayed in 9 blocks of 70 trials. If they wished, participants could take a break between the blocks. A training trial started with the presentation of a central fixation cross (100 ms), then of a blank interval (100 ms) followed by an inverted face (500 ms) randomly chosen from the 7 viewpoints (each viewpoint of each trained face was presented an equal number of times over the course of the experiment). After a second 100 ms blank interval, 10 full front inverted faces were displayed in two rows of 5 face images (Fig. 3). Their spatial frequency content matched the target. Participants were asked to decide as fast and as accurately as possible which of the 10 faces had the same identity as the one presented previously by clicking on it with the mouse. The location of the 10 faces was constant across trials. Auditory feedback was provided after each response (high- and low-pitched tones for correct and incorrect responses, respectively), and the next trial started after the feedback.

2. Results

2.1. Pre-test

We ran separate ANOVAs to examine whether the performance of the groups differed during the pre-test. For tasks 1 and 2, the

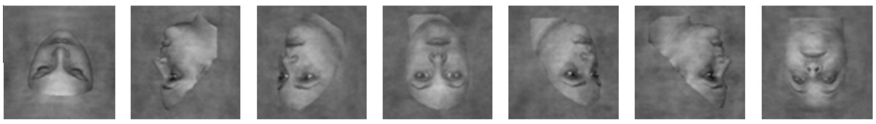

Fig. 1. Example of an inverted face stimulus from the full spectrum condition presented with 7 different viewpoints (looking down, turned 90° to the left, turned 45° to the left, frontal, turned 45° to the right, turned 90° to the right and looking up).

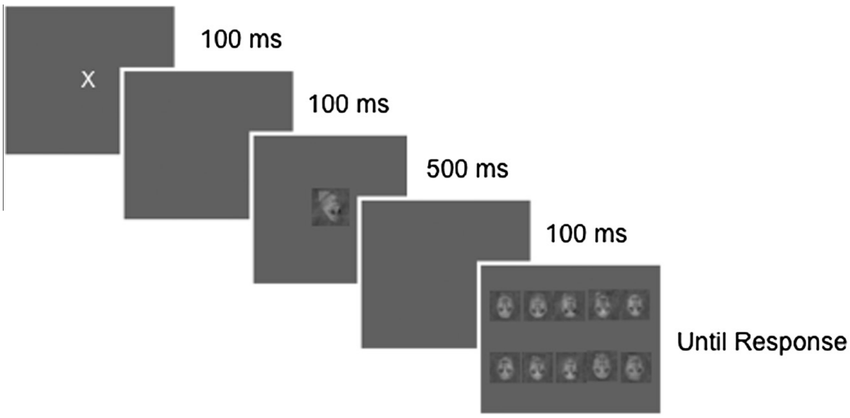

Fig. 3. Temporal sequence of stimuli during training.

only factor was group (FULL vs. LSF vs. HSF vs. Control). For Tasks 3 and 4, we included the additional within-subject factor of alignment (aligned vs. misaligned) or block (feature vs. spacing differences) in the analyses, respectively.[1]

For accuracy, there was no difference among the groups for Task 1 ($F(3,63) = .369$, $p = .776$) or Task 2 ($F(3,62) = 1.211$, $p = .314$). For Task 3, there was no main effect of the alignment of the face parts, a pattern suggesting, as expected, no composite face effect for inverted faces ($F(1,60) = 1.397$, $p = .242$), no interaction between the composite effect and the group ($F(3,60) = .894$, $p = .450$) and no main effect of the group ($F(3,60) = .145$, $p = .932$). For Task 4, there was a main effect of the testing block ($F(1,60) = 19.260$, $p < .0001$), participants being overall better to match inverted faces based on their features (76%; SD = 11) than based on the spacing between their features (68%; SD = 13). Conversely, there was no interaction between the group and the testing block ($F(3,60) = .585$, $p = .627$) and crucially, no main effect of group ($F(3,60) = 1.662$, $p = .185$).[2]

For median correct reaction times (ms), there was no effect of group for Task 1 ($F(3,62) = 1.317$, $p = .277$) and Task 2 ($F(3,63) = .926$, $p = .434$). For Task 3, there was no main effect of the alignment of the face parts, a pattern suggesting no composite face effect for inverted faces ($F(1,60) = .148$, $p = .702$), no interaction between the composite effect and the group ($F(3,60) = .775$, $p = .513$) and no main effect of group ($F(3,60) = .157$, $p = .925$). For Task 4, there was, as for accuracy, a main effect of the testing block ($F(1,60) = 4.241$, $p = .044$), participants being overall faster to match inverted faces based on their features (810 ms; SD = 180) than based on the spacing between their features (856 ms; SD = 264). Conversely, there was no interaction between the group and the testing block ($F(3,60) = 1.135$, $p = .342$) and no main effect of group ($F(3,60) = 1.135$, $p = .342$).[3]

2.2. Training

The data from the 18 blocks (2 days × 9 blocks of 70 trials) were collapsed into 6 blocks of 210 trials. Blocks 1–3, therefore, contained the data collected on the first day of training and Blocks 4–6 contained the data collected on the second day of training.

A repeated measures ANOVA on accuracy (% correct) with the training block (1–6) as the within-subject variable and the trained group (FULL, LSF, HSF) and subgroup (G1 trained with same faces on Day 2; G2 trained with different faces on Day 2) as the between-subjects variables indicated a main effect of the training block ($F(3.81,159.88) = 36.136$, $p < .0001$), with participants improving significantly from the beginning to the end of training (Fig. 4A). There was also an interaction between the trained group and the training block ($F(7.61,159.88) = 14.761$, $p < .0001$) because the difference between the trained groups was larger at the end than at the beginning of training (Fig. 4A). This hypothesis was confirmed by multiple t-tests with Bonferroni comparisons (alpha = .05/3 = .017) on the amount of improvement from Block 1 to Block 6 for each of the 3 trained groups. Participants improved by 20% when they were trained on full-spectrum faces during the training, which was significantly more than the 7% improvement observed in the HSF group ($p < .0001$) and the 1% improvement of the LSF group ($p < .0001$). The amount of improvement also differed significantly between the HSF group (7%) and the LSF group (1%) ($p = .034$). Finally there was a triple interaction among the training block, the trained group, and the subgroup ($F(7.62,159.88$[4]$) = 4.509$, $p < .0001$).

As a follow-up to this 3-way interaction, we performed repeated measures ANOVAs on accuracy (% correct) for each trained group (FULL, LSF, HSF) separately with the training block (1–6) as the within-subject variable and the subgroup (G1, G2) as the between-subjects variable. When there was an interaction with group, we further decomposed the ANOVA by applying planned comparison t-tests to test whether there was significant improvement from Block 1 to Block 6 for each (sub)group., adjusted by Bonferroni correction (alpha = .05/2 = .025).

For the FULL group, there was no main effect of subgroup ($F(1,14) = .129$, $p = .725$) and no interaction between the training block and the subgroup ($F(3.81,53.28) = 2.081$, $p = .099$) (Fig. 4A). The effects of training block in this condition replicated those in the original analysis. For the HSF group, there was an interaction between the training block and the subgroup ($F(3.327,46.58) = 2.733$, $p = .049$). Group 1, the group which saw the same faces on both days, improved significantly by 12% from the beginning to the end of the training (Block 6 vs. Block 1: $t(7) = -3.745$, $p = .007$). On the contrary, Group 2, which was exposed to a novel set of faces on Day 2, did not improve significantly (Block 6 vs. Block 1: $t(7) = -1.013$, $p = .345$) (Fig. 4A). Finally for the LSF group, there was also an interaction between the training block and the

[1] Participants were also tested with upright faces, but since we were primarily interested in generalization to novel examples of inverted faces, the few significant differences are indicated in footnotes.

[2] For upright trials, we did not find any significant effect of group on participants' accuracy, except for Task 4 where the control group's advantage for featural trials over spacing trials was significantly larger than the difference shown by the group later trained on HSF inverted faces (two-tailed Dunnett t-tests: $p = .015$).

[3] As for inverted trials, we did not find any significant effect of group on participants' median correct reaction times (ms) for upright faces (ps > .05).

[4] We used a Huynh–Feldt correction when the Mauchly's test of sphericity indicated that the data were not normally distributed.

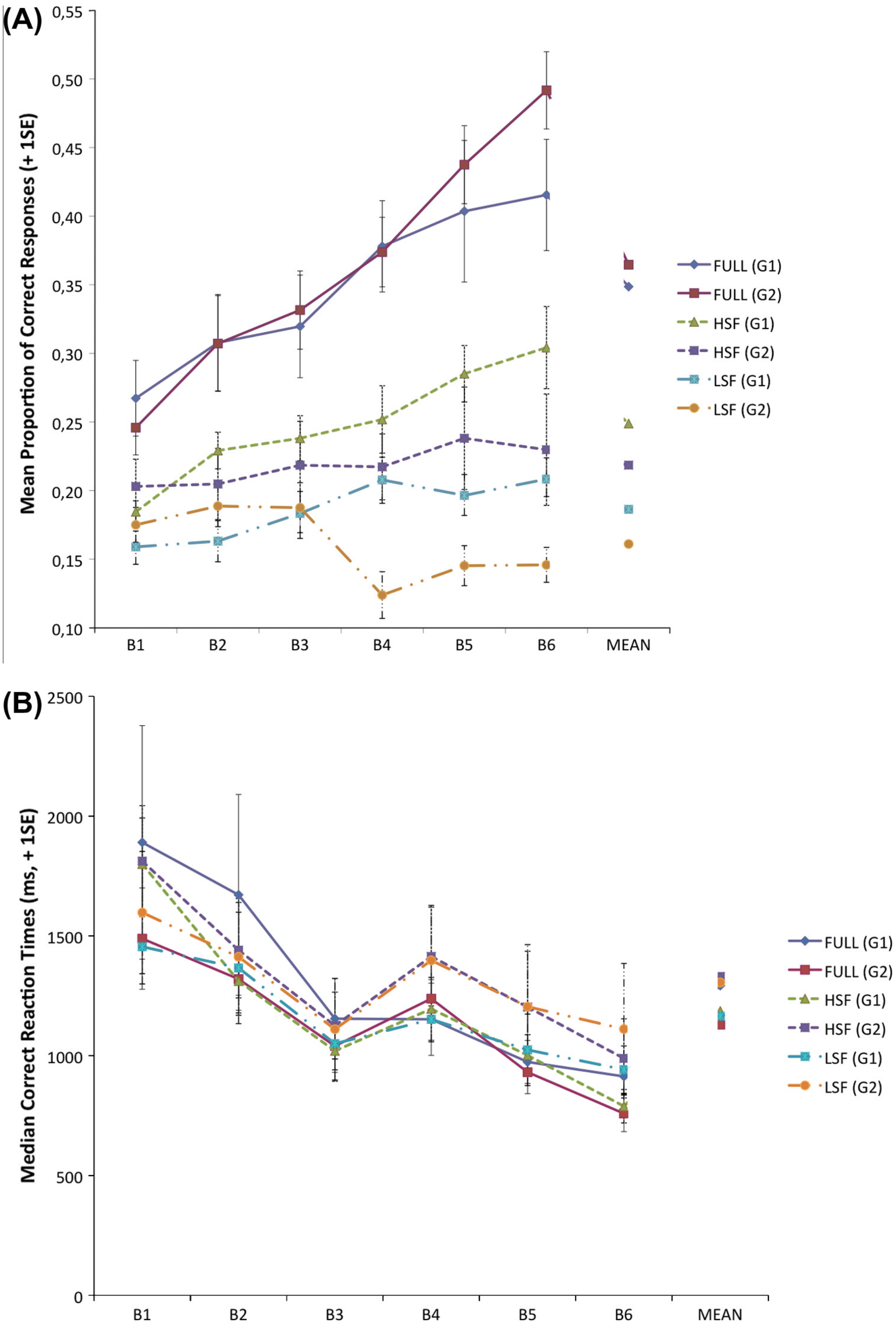

Fig. 4. Participants' mean proportion of correct responses (A) and median correct reaction times (B) during the 6 training blocks (Day 1: Blocks 1–3; Day 2: Blocks 4–6), according to the training group (FULL; HSF; LSF) and the subgroup (Group 1; Group 2) they were assigned to. Chance level is at 10%. Bars represent between-subjects standard errors.

subgroup ($F(5,70) = 12.419$, $p < .0001$): participants in Group 1 improved significantly by 5% from the beginning to the end of training (Block 6 vs. Block 1: $t(7) = -3.031$, $p = .019$), unlike participants in Group 2 whose accuracy did not improve significantly over the training (Block 6 vs. Block 1: $t(7) = 2.067$, $p = .078$) (Fig. 4A).

We also performed a repeated measure ANOVA on participants' median reaction times (ms) on correct trials with the training block (1–6) as the within-subject variable and the trained group (FULL, LSF, HSF) and the subgroup (G1, G2) as the between-subjects variables. There was a main effect of the training block ($F(2.193,92.11) = 34.587$, $p < .0001$) because participants got faster overall over training (Fig. 4B). There were no other main effects or interactions.

2.3. Post-test

For the post-test, we first calculated the trained participants' difference scores between pre-test and post-test based on their accuracy (%; POST-PRE). Then we performed an ANOVA with the trained group (FULL, LSF, HSF) and the subgroup (G1 exposed to same faces

Table 1
Participants' mean proportion of correct responses in the upright and the inverted condition of the 4 tasks used at pre-test and at post-test, according to the group they were assigned to (FULL, LSF, HSF, Control). Participants' composite face effect (CFE) was calculated as the ratio between their mean proportion of correct responses with aligned and misaligned trials ((M − A)/(M + A)).

		Full		LSF		HSF		Control	
		PRE	POST	PRE	POST	PRE	POST	PRE	POST
Task 1									
UP	Accuracy – all trials	0.88	0.94	0.90	0.92	0.94	0.95	0.91	0.93
INV	Accuracy – all trials	0.80	0.85	0.81	0.83	0.81	0.87	0.79	0.82
Task 2									
UP	Accuracy – all trials	0.86	0.91	0.86	0.88	0.90	0.93	0.88	0.89
INV	Accuracy – all trials	0.73	0.83	0.78	0.78	0.76	0.84	0.76	0.77
Task 3									
UP	Same trials (CFE)	−0.17	−0.10	−0.15	−0.07	−0.10	−0.12	−0.20	−0.09
INV	Same trials (CFE)	−0.05	−0.03	0.01	−0.01	0.03	−0.02	−0.01	−0.03
UP	Different trials (CFE)	−0.01	−0.03	0.01	0.01	−0.02	−0.02	−0.02	−0.02
INV	Different trials (CFE)	0.01	0.02	0.01	0.01	−0.01	−0.01	0.00	0.03
Task 4									
UP	Featural trials	0.91	0.89	0.87	0.88	0.91	0.91	0.85	0.84
INV	Featural trials	0.78	0.82	0.74	0.77	0.80	0.84	0.72	0.74
UP	Configural trials	0.80	0.80	0.83	0.80	0.82	0.84	0.72	0.74
INV	Configural trials	0.66	0.72	0.67	0.71	0.73	0.78	0.66	0.69

on Day 2; G2 exposed to different faces on Day 2) as between subject variables in order to determine whether the subgroups differed from each other. All groups had higher accuracy on the post-test than on the pre-test (see Table 1). We added the alignment of the face (aligned vs. misaligned) and the block (feature vs. spacing differences) as additional within subject variables for Tasks 3 and 4, respectively. There was no main effect or interaction involving the subgroup factor (*p*s > .05). We, therefore, collapsed across the 2 subgroups for each trained group and contrasted the accuracy of the trained groups and of the untrained control group in a new series of ANOVAs. The goal here was to assess whether the amount of improvement in the 4 tasks was greater in any of the trained groups than in the untrained control group. There was a main effect of group in the delayed face matching task (Task 2: $F(3,63) = 4.620$, $p = .006$), but not in the other tasks (*p*s > .05). No other main effect or interaction reached significance (*p*s > .05).[5] One-tailed Dunnett *t*-tests were then used to assess whether the amount of improvement was greater in any of the trained group compared to the untrained control group in Task 2. They indicated greater improvement in accuracy in the FULL group ($X = 10\%$, SE = .03; $t(30) = 2.951$, $p = .007$) and in the HSF group ($X = 8\%$, SE = .05; $t(29) = 2.873$, $p = .038$) than in the control group ($X = 1\%$, SE = .02). Conversely, the LSF group ($X = 1\%$, SE = .01; $t(30) = -.020$, $p = .759$) did not improve more than the control group (Fig. 5).

The same analyses on the trained participants' difference scores between pre-test and post-test calculated based on their median correct reaction times (ms; PRE-POST) suggested, as for accuracy, no main effect or interaction involving the group or the subgroup variable (*p*s > .05). The same pattern was obtained when the data were collapsed across subgroups and the untrained control group was added to the analyses (*p*s > .05).[6]

3. Discussion

In this study, we investigated the efficacy of training adults for only 2 h spread across two consecutive days to recognize inverted faces with different viewpoints and its generalization to new instances. When training involved full spectrum faces, participants improved on average by 20%, whether or not the faces changed on the second day of training (G1 = 27%; G2 = 24%). Nevertheless, their accuracy was still low (around 46%) at the end of training even when they were in Group 1 and had only 10 inverted faces to learn that remained the same across the 2 days of training (for similarresults with full spectrum inverted faces, see Hussain, Sekuler, & Bennett, 2009a, 2009b; Laguesse et al., 2012). As in Laguesse et al. (2012) after 2 weeks of training, the FULL group also showed evidence of generalization: they improved on the second day of training even when the faces were different from the ones learned on the previous day and they improved more than the control group on the delayed matching of a new set of full spectrum inverted faces of a different size and presented from different points of view (Task 2). The lack of generalization in the putatively easier task - the simultaneous task - cannot be explained by a ceiling effect (mean pre-test accuracy of 80%) but may have arisen because the control group improved by 3% between the pre- and post-test (vs. 1% on Task 2) and/or because the structure of the task (simultaneous discrimination) differed from the structure of training (delayed matching to sample), which was shared by Task 2 and by the training and post-test chosen by Laguesse et al. (2012). That pattern of restricted generalization is consistent with evidence that when adults make visual discriminations, their perceptual and memory representations are jointly affected by the spatial frequency content of the stimuli and the structure of the task (e.g., Lalonde & Chaudhuri, 2002; Nemes et al., 2011). Our results for the FULL group nevertheless contrast with those of Hussain, Sekuler, and Bennett (2009b) who reported (almost) no improvement from baseline in adults trained for approximately the same time with full spectrum inverted faces that changed between the first and second day, as they did for one of our groups. This discrepancy may have arisen among other things from our study presenting the faces from different point of views during training as in Laguesse et al. (2012), a variation that may have forced the observers to learn to extract invariants in addition to the pictorial information in the faces.

Participants trained with high spatial frequency and low spatial frequency inverted faces improved but to a lesser extent than those exposed to the full range and only when they saw the same 10 faces on the 2 days of training. The improvement was greater with high spatial frequency (HSF) faces than with low spatial frequency (LSF) faces, despite the fact that overall luminance and the mean contrast of the HSF stimuli was lower than those of the LSF stimuli.

[5] The same analyses performed on participants' difference scores for upright trials indicated no main effect or interaction involving the group factor (ps > .05).

[6] The same analyses performed on participants' difference scores for upright trials indicated no main effect or interaction involving the group factor (ps > .05).

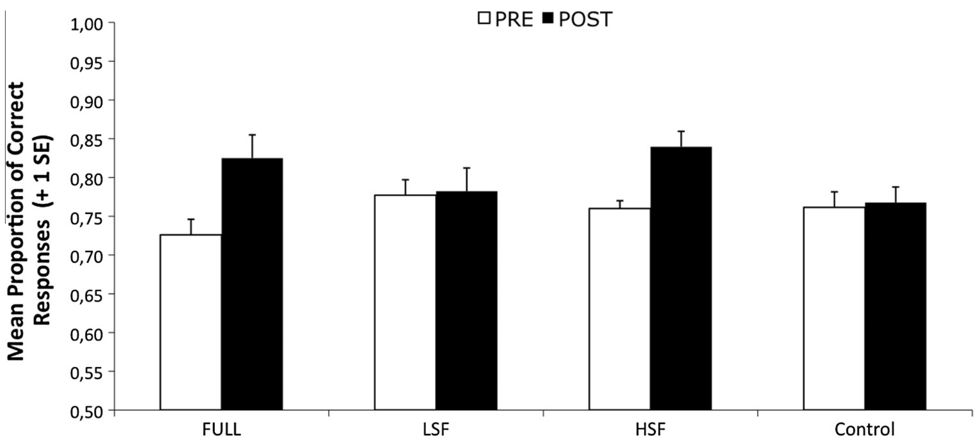

Fig. 5. Improvement in accuracy (% correct) in the inverted condition of the delayed face matching task (Task 2) before and after training for each group. Chance level is at 50%. Bars represent between-subjects standard errors.

It seems that participants shown HSF faces changed during the first day of training from perceiving the stimuli as nearly indistinguishable (accuracy near chance) to perceiving the specific exemplars as psychologically separated, or differentiated (Goldstone, 1998). However they did not learn any more general strategy during the first day of training that they could apply to new instances presented on the second day of training. Nevertheless, by the end of the second day of training participants of the HSF group appeared to have developed a skill that they could apply to full spectrum inverted faces during the post-test since they showed more improvement than the control group in Task 2 (FULL = 10% vs. HSF = 8% vs. Control = 1%). Our guess is that these participants slowly improved in their ability to process the features of inverted faces, the edges of which were more conspicuous after filtering, and they subsequently could transfer that ability to unfiltered faces (for similar results see Robbins & McKone, 2003). If this explanation is correct, it is puzzling that the featural condition of the Jane task did not reveal any differential improvement, unless one considers that the task does not only assess featural face processing but also the ability to extract surface reflectance cues on which the subjects were not trained (McKone & Yovel, 2009). Perhaps this part of the post-test was also not sensitive enough to detect a greater change in the high spatial frequency group than in other groups because the strategy this group learned depended on extracting shape information from not just the eyes and the mouth but also from other features, such as the nose, the shape of which was held constant in the task.

Participants trained with low spatial frequency faces showed evidence of only weak learning, even though the faces they viewed had the same luminance as the full spectrum faces and similar contrast: as for the other groups, their reaction times decreased during training and they improved in accuracy by only 5% across training and only if the faces remained the same on Day 2. However, there was no evidence of generalization on any of the tasks of the post-test because their accuracy and reaction times did not differ significantly from those of the untrained control group. These patterns suggest that they were unable to pick up a significant amount of global information conveyed by this range of spatial frequencies although they were encouraged to do so by the nature of the stimuli. Nor were they able to develop or improve a general global strategy that could be transferred to full spectrum faces.

Overall, our results show that only 2 h of training on a set of full spectrum inverted faces presented from multiple views is effective and can support generalization to a new set of full spectrum inverted faces. Similar evidence was found in pigeons who, after being trained to recognize objects from multiple viewpoints, showed robust performance with novel views of the trained objects (Soto, Siow, & Wasserman, 2012). Our results also indicate that perceptual learning can occur for low spatial frequency and high spatial frequency faces but only if their number is limited. Finally, the transfer of learning observed in participants exposed to high spatial frequency faces suggests that the information they convey, namely featural information, may be especially important for perceptual learning of inverted faces.

Future studies might investigate the effect of training for a longer period of time with the filtered and non-filtered inverted faces used in this study. We hypothesize that participants exposed to low spatial frequency faces could take more time than those exposed to high or the full range of spatial frequencies to reach their asymptotic level of accuracy but then would show better generalization because of the importance of holistic processing for face recognition (Rossion, 2008, 2009). It would also be interesting to explore the exact range of low and high spatial frequencies needed to benefit from a 2-h training procedure when a different set of inverted faces is involved on the second day of training. One could indeed expect different effects of training when the range of low spatial frequencies is broader than the very restricted range used in the current study (see for example, Goffaux & Rossion, 2006; Morrisson & Schyns, 2001; but see Gaspar, Sekuler, & Bennett, 2008; Willenbockel et al., 2010) and, with more effective low frequency training, one might also expect transfer to the composite face task and the Jane spacing task, which would provide evidence of increased holistic/configural face processing.

Acknowledgments

This research was supported by a grant from the Canadian Natural Sciences and Engineering Research Council to D.M. (9797). We would like to thank Mayu Nishimura and Xiaoqing Gao for help with the creation of the stimuli used for the training as well as Nimasha Weliwitigoda and Sally Stafford for helping to test the participants.

References

Acerra, F., Burnod, Y., & de Schonen, S. (2002). Modelling aspects of face processing in early infancy. *Developmental Science, 5*, 98–117.

Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research, 27*, 953–965.

Banks, M. S., & Bennett, P. J. (1991). Anatomical and physiological constraints on neonatal visual sensitivity and determinants of fixation behaviour. In M. J. S. Weiss & P. R. Zelazo (Eds.), *Newborn attention: Biological constraints and the influence of experience* (pp. 177–217). Norwood, NJ: Ablex.

Bi, T., Chen, N., Weng, Q., He, D., & Fang, F. (2010). Learning to discriminate face views. *Journal of Neurophysiology, 104*, 3305–3311.

Boutet, I., Collin, C., & Faubert, J. (2003). Configural face encoding and spatial frequency information. *Perception & Psychophysics, 65*, 1078–1093.

Busigny, T., & Rossion, B. (2010). Acquired prosopagnosia abolishes the face inversion effect. *Cortex, 46*, 965–981.

Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception, 29*, 893–909.

de Heering, A., & Maurer, D. (2012). Face memory deficits in patients deprived of early visual input by bilateral congenital cataracts. *Developmental Psychobiology*. http://dx.doi.org/10.1002/dev.21094.

de Heering, A., Turati, C., Rossion, B., Bulf, H., Goffaux, V., & Simion, F. (2008). Newborns' face recognition is based on spatial frequencies below 0.5 cycles per degree. *Cognition, 106*, 444–454.

Dolan, R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S., et al. (1997). How the brain learns to see objects and faces in an impoverished context. *Nature, 389*, 596–599.

Dwyer, D. M., Mundy, M. E., Vladeanu, M., & Honey, R. C. (2009). Perceptual learning and acquired face familiarity: Evidence from inversion, use of internal features, and generalization between viewpoints. *Visual Cognition, 17*, 334–355.

Fahle, M., Edelman, S., & Poggio, T. (1995). Fast perceptual learning in hyperacuity. *Vision Research, 35*, 3003–3013.

Fiorentini, A., & Berardi, N. (1981). Learning in grating waveform discrimination: Specificity for orientation and spatial frequency. *Vision Research, 21*, 1149–1158.

Furmanski, C. S., & Engel, S. A. (2000). Perceptual learning in object recognition: Object specificity and size invariance. *Vision Research, 40*, 473–484.

Gaspar, K., Sekuler, A. B., & Bennett, P. J. (2008). Spatial frequency tuning of upright and inverted face identification. *Vision Research, 48*, 2817–2826.

Gibson, E. J. (1969). *Principles of perceptual learning and development*. East Norwalk, CT, US: Appleton-Century-Crofts.

Goffaux, V., & Rossion, B. (2006). Faces are "spatial" – Holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1023–1039.

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research, 39*, 3537–3560.

Gold, J. M., Sekuler, A. B., & Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science, 28*, 167–207.

Goldstone, R. B. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585–612.

Hole, G. J. (1994). Configural factors in the perception of unfamiliar faces. *Perception, 23*, 65–74.

Hussain, Z., Sekuler, A. B., & Bennett, P. J. (2009a). How much practice is needed to produce perceptual learning? *Vision Research, 49*, 2624–2634.

Hussain, Z., Sekuler, A. B., & Bennett, P. J. (2009b). Perceptual learning modifies inversion effects for faces and textures. *Vision Research, 49*, 2273–2284.

Hussain, Z., Sekuler, A. B., & Bennett, P. J. (2011). Superior identification of familiar visual patterns a year after learning. *Psychological Science, 22*, 724–730.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *PNAS, 88*, 4966–4970.

Laguesse, R., Dormal, G., Biervoye, A., Kuefner, D., & Rossion, B. (2012). Extensive visual training in adulthood significantly reduces the face inversion effect. *Journal of Vision, 12*, 1–13.

Lalonde, J., & Chaudhuri (2002). Task-dependent transfer of perceptual to memory representations during delayed spatial frequency discrimination. *Vision Research, 42*, 1759–1769.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Early visual experience and face processing. *Nature, 410*, 890.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2004). Impairment in holistic face processing following early visual deprivation. *Psychological Science, 15*, 762–768.

Levi, D. M., & Li, R. W. (2009). Improving the performance of the amblyopic visual system. *Philosophical Transactions of the Royal Society London B: Biological Sciences, 364*, 399–407.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*, 255–260.

McKee, S. P., & Westheimer, G. (1978). Improvement in Vernier acuity with practice. *Perception and Psychophysics, 24*, 258–262.

McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing faces, and sometimes not? Toward a new theory of holistic processing. *Psychonomic Bulletin Review, 16*, 778–797.

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception, 31*, 553–566.

Morrisson, D. J., & Schyns, P. G. (2001). Usage of spatial scales for the categorization of faces, objects and scenes. *Psychonomic Bulletin & Review, 8*, 454–469.

Moses, Y., Ullman, S., & Edelman, S. (1995). Generalization to novel images in upright and inverted faces. *Perception, 25*, 443–461.

Nemes, V. A., Whitaker, D., Heron, J., & McKeefry, D. J. (2011). Multiple spatial frequency channels in human visual perceptual memory. *Vision Research, 51*, 2331–2339.

Poggio, T., Fahle, M., & Edelman, S. (1992). Face perceptual learning in visual hyperacuity. *Science, 256*, 1018–1021.

Robbins, R., & McKone, E. (2003). Can holistic processing be learned for inverted faces. *Cognition, 88*, 79–107.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica, 128*, 274–289.

Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: The perceptual field hypothesis. *Acta Psychologica, 132*, 300–312.

Rubin, N., Nakayama, K., & Shapley, R. (1997). Abrupt learning and retinal size specificity in illusory-contour perception. *Current Biology, 7*, 461–467.

Sagi, D., & Tanne, D. (1994). Perceptual learning: learning to see. *Current Opinion in Neurobiology, 4*, 195–199.

Schoups, A. A., Vogels, R., & Orban, G. A. (1995). Human perceptual learning in identifying the oblique orientation: Retinotopy, orientation specificity and monocularity. *Journal of Physiology, 483*, 797–810.

Sigman, M., & Gilbert, C. D. (2000). Learning to find a shape. *Nature Neuroscience, 3*, 264–269.

Soto, F. A., Siow, J. Y. M., & Wasserman, E. A. (2012). View-invariance learning in object recognition by pigeons depends on error-driven associative learning processes. *Vision Research, 62*, 148–161.

Watier, N. N., Collin, C. A., & Boutet, I. (2010). Spatial-frequency thresholds for configural and featural discrimination in upright and inverted faces. *Perception, 39*, 502–513.

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., et al. (2010). Does face inversion change spatial frequency tuning? *Journal of Experimental Psychology: Human Perception and Performance, 36*, 122–135.

Yi, D. J., Olson, I. R., & Chun, M. M. (2006). Shape-specific perceptual learning in a figure-ground segregation task. *Vision Research, 46*, 914–924.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141–145.

Young, A. W., Hellawell, A., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*, 747–759.