Does Extensive Training at Individuating Novel Objects in Adulthood Lead to Visual Expertise? The Role of Facelikeness

Aliette Lochy[1], Friederike G. S. Zimmermann[1,2], Renaud Laguesse[1], Verena Willenbockel[2,3], Bruno Rossion[1,4], and Quoc C. Vuong[2]

Abstract

■ Human adults have a rich visual experience with seeing human faces since birth, which may contribute to the acquisition of perceptual processes that rapidly and automatically individuate faces. According to a generic visual expertise hypothesis, extensive experience with nonface objects may similarly lead to efficient processing of objects at the individual level. However, whether extensive training in adulthood leads to visual expertise remains debated. One key issue is the extent to which the acquisition of visual expertise depends on the resemblance of objects to faces in terms of the spatial configuration of parts. We therefore trained naive human adults to individuate a large set of novel parametric multipart objects. Critically, one group of participants trained with the objects in a "facelike" stimulus orientation, whereas a second group trained with the same objects but with the objects rotated 180° in the picture plane into a "nonfacelike" orientation. We used a fast periodic visual stimulation EEG protocol to objectively quantify participants' ability to discriminate untrained exemplars before and after training. EEG responses associated with the frequency of identity change in a fast stimulation sequence, which reflects rapid and automatic perceptual processes, were observed over lateral occipital sites for both groups before training. There was a significant, albeit small, increase in these responses after training but only for the facelike group and only to facelike stimuli. Our findings indicate that perceived facelikeness plays a role in visual expertise and highlight how the adult perceptual system exploits familiar spatial configurations when learning new object categories. ■

INTRODUCTION

Human adults interact with many people during their daily social activities. Not surprisingly, the face is an important cue for individuating both familiar and unfamiliar people. Despite the fact that individual faces are visually similar to each other, adults nevertheless can quickly and accurately recognize faces at an individual level (e.g., to correctly greet "Sue" or "Sally"). By comparison, objects from other categories are typically recognized at a basic (e.g., "dog" or "car") or subordinate (e.g., "beagle" or "corvette") level (e.g., Jolicoeur & Milliken, 1989; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) but rarely at the individual level. Indeed, expertise in face recognition has been defined as a capacity to rapidly and automatically individuate face exemplars at the perceptual level (Tanaka & Curran, 2001; Carey, 1992).

The adult perceptual processing efficiency for individual faces has been attributed in part to an innate sensitivity to a "face template" present at birth (Turati, Simion, Milani, & Umiltà, 2002; Morton & Johnson, 1991) and

in part to lifelong exposure to faces coupled with the need to individuate faces for social interactions (Mondloch, Maurer, & Ahola, 2006; Carey, 1992). There is extensive evidence from neuroimaging and intracerebral recordings that face recognition in human adults relies predominantly on a right-lateralized network of brain regions in the ventral occipitotemporal cortex that respond preferentially to faces (e.g., Jonas et al., 2016; Rossion, Hanseeuw, & Dricot, 2012; Haxby, Hoffman, & Gobbini, 2000; Allison, Puce, Spencer, & McCarthy, 1999; Kanwisher, McDermott, & Chun, 1997; Sergent, Ohta, & MacDonald, 1992). More recently, researchers have used a powerful fast periodic visual stimulation (FPVS) EEG protocol to objectively quantify rapid and automatic perceptual responses to face individuation over posterior occipitotemporal sites (Liu-Shuang, Norcia, & Rossion, 2014; Rossion & Boremanse, 2011). These individual discrimination responses are reduced by stimulus manipulations that affect face perception such as picture plane inversion or contrast reversal (for a review, see Rossion, 2014).

According to an influential generic visual expertise hypothesis, extensive visual experience with exemplars from nonface object categories can lead to the recruitment of similar perceptual processes and neural structures used

[1]University of Louvain, [2]Newcastle University, [3]Goethe University Frankfurt, [4]Centre Hospitalier Regional Universitaire de Nancy

to individuate faces (Tanaka & Curran, 2001; Tarr & Gauthier, 2000; Carey, 1992; Diamond & Carey, 1986). There is behavioral and neural evidence for this hypothesis in natural experts, such as dog judges (Tanaka & Curran, 2001; Diamond & Carey, 1986), car hobbyists (Rossion, Collins, Goffaux, & Curran, 2007; Xu, 2005; Gauthier, Curran, Curby, & Collins, 2003; Gauthier, Skudlarski, Gore, & Anderson, 2000), and bird watchers (Xu, 2005; Tanaka & Curran, 2001; Gauthier et al., 2000). However, there are many uncontrolled factors that contribute to visual recognition performance after years of real-world experience with different categories (e.g., age of onset of interest in the category, duration and richness of the visual experience). Moreover, most natural experts generally do not individuate exemplars from the category of interest but recognize them at a subordinate level (e.g., bird species or cars models).

To better control for these factors, a specific version of the visual expertise hypothesis states that training to individuate novel 3-D objects in experimental settings can lead to behavioral changes and increased activation in face-selective brain regions during adulthood (Tarr & Gauthier, 2000). The most widely used set of novel objects for this purpose are Greebles (Gauthier & Tarr, 1997). After approximately 10 hr of training with Greebles, participants become more sensitive to the spatial relationships between Greeble parts (Gauthier & Tarr, 1997, 2002). This configural processing is presumed to be a behavioral hallmark of individual face recognition expertise (Rossion, 2013; Tanaka & Farah, 1993; Sergent, 1984). Behavioral effects of training-induced changes in configural processing were supported by neuroimaging results. First, BOLD activity in the face-selective brain region of the mid-fusiform gyrus (i.e., the so-called "fusiform face area" [FFA]; Kanwisher et al., 1997) increased when participants discriminated untrained Greebles in the scanner (Gauthier & Tarr, 2002; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999). Second, BOLD activity in the right FFA positively correlated with behavioral measures of expertise with Greebles (Gauthier & Tarr, 2002). Similar behavioral and/or neural results were found using other novel objects and individual-level training protocols (e.g., Wong, Palmeri, & Gauthier, 2009; Wong, Palmeri, Rogers, Gore, & Gauthier, 2009). However, the findings in support of the visual expertise hypothesis have been contested, as a number of studies failed to find increases in configural sensitivity or neural activity in the FFA after extensive training or natural expertise with nonface objects, including Greebles (e.g., Brants, Wagemans, & Op de Beeck, 2011; Robbins & McKone, 2007; Rhodes, Byatt, Michie, & Puce, 2004; for a review, see McKone, Kanwisher, & Duchaine, 2007). Moreover, training-induced increase in BOLD activation has been found in object-selective regions but not in FFA (Op de Beeck, James, DiCarlo, & Kanwisher, 2006).

There are several potential reasons for the discrepancy in findings across different expertise studies, particularly those that train adults to individuate novel nonface objects. First, most studies using novel objects have largely dismissed the potential role of "facelikeness" in the acquisition of visual expertise in adulthood. That is, human adults may exploit preexisting face templates to recognize nonface objects that physically resemble faces, particularly if these objects have parts arranged in a similar spatial configuration as facial parts (Vuong et al., 2016; Brants et al., 2011; Biederman & Kalocsai, 1997; see Xu, 2005, for a similar argument for bird and car expertise). Indeed, several studies have shown that complex visual stimuli that resemble faces can elicit BOLD activity in brain regions associated with face recognition (e.g., Liu et al., 2014; Davidenko, Remus, & Grill-Spector, 2012; Churches, Baron-Cohen, & Ring, 2009; Hadjikhani, Kveraga, Naik, & Ahlfors, 2009). Second, there are additional stimulus-related issues with the novel objects used, particularly for Greebles (and, to some extent, for other novel object sets as well; see Vuong et al., 2016, for a critique). Individual Greebles can have one of five possible body shapes only, and each Greeble has a unique set of distinctive parts. Consequently, individual Greebles may be learned and recognized on the basis of a single diagnostic part or on the independent processing of these parts. Hence, even individuals with acquired prosopagnosia who have severe face recognition deficits are able to learn Greebles as well as control participants (Rezlescu, Barton, Pitcher, & Duchaine, 2014). The physical and perceptual similarities between Greeble exemplars have also not been quantified. By physical similarity, we mean similarity based on physical measurements of the stimuli (e.g., comparing pixel values of images of objects). Taken together, potential differences in how observers across different Greeble training studies relied on single parts, or the degree to which the observers perceived the facelikeness of the Greebles or were affected by the perceptual (dis)similarity between Greebles, may account for the reported differences in neural findings even when identical training paradigms and Greebles were used (compare Brants et al., 2011, and Gauthier et al., 1999).

Last, previous training studies do not provide evidence that participants individuate objects more rapidly and automatically at the perceptual level after training than before training. First, the behavioral tasks used to gauge "expertise" effects after training (e.g., Wong, Palmeri, & Gauthier, 2009; Wong, Palmeri, Rogers, et al., 2009; Gauthier & Tarr, 1997, 2002) do not necessarily tap perceptual processes. Behavioral performance on these tasks may be influenced by postperceptual mechanisms at attentional or decisional stages as objects are typically presented unmasked and for a long duration (500 msec or longer; see Harel, Kravitz, & Baker, 2013). Second, the fMRI paradigms used to assess training-induced neural changes typically average BOLD activity across several seconds so that any changes in activity may again not necessarily reflect perceptual processes related to object individuation. Using EEG, Rossion, Kung, and Tarr (2004)

showed that training with asymmetric Greebles led to a decrease in the amplitude of the N170 component in response to a face when it was concurrently presented with a Greeble. This decrease in the N170 response to a face was also observed when it was presented concurrently with another face (Jacques & Rossion, 2004). Although this finding indicates that training with novel objects leads to competition with face processing at an early perceptual stage, there is no direct evidence that the Greebles themselves were processed rapidly and automatically at the individual level. Other EEG studies with Greebles have found training-induced changes to the N170 component in response to individual Greebles after training (Cao, Jiang, Li, Xia, & Floyd, 2015; Rossion, Gauthier, Goffaux, Tarr, & Crommelinck, 2002). In these studies, however, participants attended to the Greebles to perform a task. Thus, despite their wide use, it is unclear whether behavioral and neural changes after training with Greebles reflect the acquisition of rapid and automatic perceptual processes related to visual expertise. Most importantly, it remains unclear whether the changes were due to their facelikeness (Brants et al., 2011; Biederman & Kalocsai, 1997) or other stimulus factors.

The goal of the current study was to test whether extensive training with novel parametric multipart 3-D objects in adults leads to the acquisition of visual expertise. This remains an important open question given the issues raised earlier. Here, we provide the most stringent test of the specific visual expertise hypothesis to date by addressing several key issues: (1) We directly compare the acquisition of visual expertise for "facelike" and "nonfacelike" novel objects equated on physical and perceptual similarities (Vuong et al., 2016); (2) we use parameterized novel objects that have similar body shapes and parts (i.e., no diagnostic local cues to identity); and (3) we use a powerful EEG protocol that allows us to objectively quantify rapid and automatic perceptual processes related to object individuation before and after training, rather than relying on behavioral measures that may be contaminated by attentional and decisional processes (but see Harel, 2016; Harel et al., 2013; and Harel, Gilaie-Dotan, Malach, & Shlomo, 2010, for the view that such nonperceptual processes may be important for visual expertise). By "rapid," we mean that objects can be individuated in less than 200 msec (i.e., a single glance) in line with electrophysiological evidence that sensitivity to individual faces emerges around this time (i.e., at the level of the N170 component; Jacques, d'Arripe, & Rossion, 2007; Heisz, Watter, & Shedden, 2006). By "automatic," we mean that objects can be individuated even if the objects themselves are irrelevant for the task. For example, the objects may be processed with little to no attention (Schneider & Chein, 2003; Wojciulik, Kanwisher, & Driver, 1998). Furthermore, we train participants for approximately 28 hr (compared with approximately 10 hr typically used in previous studies) over a 3-week period to

provide them with the greatest opportunity to acquire expertise with the individual objects in a laboratory context. Last, we use a greater variety of training tasks compared with previous studies to maintain participants' motivation and interest during the entire training period.

Combining FPVS and Novel Parametric Multipart Objects

Figure 1A presents an example array of the objects used in the current study. The individual objects were generated by varying parameters that defined the 3-D shape of the body and parts, which allowed us to quantify their physical and perceptual similarities (Vuong et al., 2016). Our objects capture some physical aspects of faces—particularly, a systematic spatial configuration of parts. Importantly, these objects are perceived to be more facelike in one stimulus orientation and more nonfacelike in the other orientation (i.e., rotated 180° in the picture plane; Figure 1B). Using the same sets of stimuli, we recently reported that facelikeness facilitated performance on a demanding delayed matching task across viewpoint changes but that the orientation manipulation did not affect low-level discrimination thresholds and higher-level perceptual similarity judgments (Vuong et al., 2016). Participants in that study rated objects in the facelike orientation to be more "facelike" ($M = 2.5$, $SE = 0.9$) than objects in the nonfacelike orientation ($M = 6.1$, $SE = 0.9$; with 1 being very facelike and 7 being not facelike at all).

Two groups of participants underwent an extensive training program at individuating the novel objects over a 3-week period. One group was trained to individuate the objects from only the facelike stimulus orientation, whereas the other group was trained with the same objects but from only the nonfacelike stimulus orientation. For both groups, we used an FPVS protocol with EEG recordings before and after training. In this protocol, the same individual stimulus is presented at a rapid periodic rate of approximately 6 stimuli/sec (visual stimulation frequency). New individual stimuli, different from the repeated stimulus and from each other, are presented on every fifth stimulus in the stimulation sequence (individual discrimination frequency). Participants were presented with separate 60-sec sequences containing new exemplars of the novel objects in either the facelike or nonfacelike stimulus orientation. They responded to color changes of a fixation cross superimposed on the images during each sequence. Individual discrimination can be objectively indexed by the neural response at the individual discrimination frequency. Importantly, this specific FPVS protocol has been optimized to measure rapid and automatic perceptual responses to face individuation (Dzhelyova & Rossion, 2014; Liu-Shuang et al., 2014; Rossion, 2014).

We make two hypotheses. First, if extensive training with novel objects leads to the acquisition of visual expertise, we
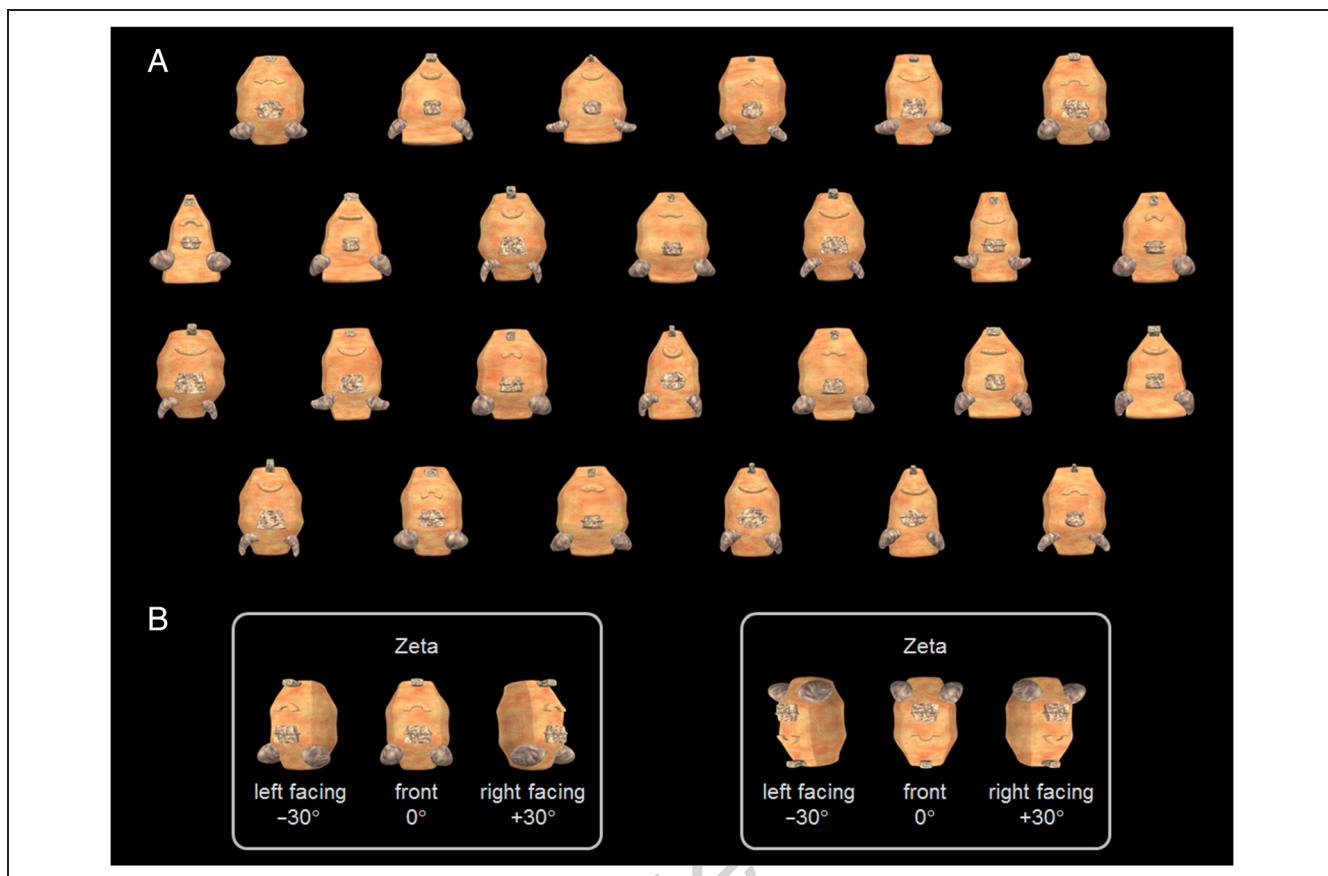
Figure 1. The novel, parametric, multipart 3-D objects used in the current study. (A) Example of a complete set of 26 objects shown in the nonfacelike orientation. (B) During training, each object was assigned a name label (e.g., "Zeta") and learned from three viewpoints. The images on the left illustrate "Zeta" in the nonfacelike stimulus orientation, and the images on the right illustrate "Zeta" in the facelike stimulus orientation.

expect to observe an increase of the neural response at the individual discrimination frequency for these objects in the trained stimulus orientation but not in the untrained orientation after training (relative to a pretraining baseline). Second, if the acquisition of visual expertise depends on prior familiarity with a stimulus configuration (i.e., here, facelikeness), we expect to observe an increase of the neural response at the individual discrimination frequency for participants in the facelike group but not for participants in the nonfacelike group. Given our focus on rapid and automatic perceptual processes, we do not necessarily expect a tight relationship between our neural and behavioral measures, because the latter could reflect general learning of the novel objects (i.e., overall performance improvement after training). Rather, our hypotheses are related to training-specific neural changes to these perceptual processes.

METHODS

Participants

Thirty right-handed volunteers participated in the study. All were native French speakers with normal or corrected-to-normal vision. They were naive to the purpose of the study and had never seen the stimuli before. Before volunteers provided written consent, they read an information sheet describing the testing and training procedures, time commitment, and reimbursement for their time. They were informed that they could stop their participation at any point without any consequences, but all volunteers completed the full experiment. The ethics were approved by the Newcastle University Ethics Committee. Before the start of training, participants completed a computerized version of the Benton Face Recognition Test (Benton, Sivan, Hamsher, Varney, & Spreen, 1983) and the Cambridge Face Memory Test (Duchaine & Nakayama, 2006). Performance of the two groups did not differ on either of the face tasks. Table 1 presents participant information for each group.

Stimuli

The procedure to generate our novel parametric multipart objects has been described in detail elsewhere (Vuong et al., 2016) and will be only briefly described here. The objects were mono-oriented and roughly symmetric about their vertical axis and had a similar spatial configuration of parts (see Figure 1A). Each object was composed of a large central body, three small parts

Table 1. The Number of Women, Mean Age (in Years), and Mean Number of Correct Items for Each Face Test Averaged across Participants in the Facelike and Nonfacelike Groups

	Facelike Group	Nonfacelike Group
Number of women	11	11
Age	22.3 (3.5)	21.6 (2.9)
Benton Face Recognition Test	45.0 (3.9)	45.0 (6.5)
Cambridge Face Memory Test	54.6 (9.3)	55.1 (8.0)

The standard deviation is presented in parentheses. Overall, the two groups did not differ in age or performance on the two face tests.

attached to it that defined its front (i.e., 0° view; see Figure 1B), and two small parts attached to its "side." The body and parts were defined by 19 parameters, such as the shape of the cross-section (from circle to square), amount of bending, and amount of tapering.

We first created 12 prototypes by fixing the values of the parameters to maximize the difference between prototypes. We then paired the 12 prototypes in all possible combinations (i.e., 66 pairs) and "morphed" the body and each of the parts at two levels (25% and 75%), which yielded 132 unique objects. The objects were rendered from three viewpoints (−30° ["left facing"], 0° ["front facing"], and +30° ["right facing"]; see Figure 1B) against a uniform black or gray background as 500 × 500 pixel images. We used the 25% and 75% morphs to create four object sets of 33 images each and randomly split the 25% morphs to create Sets 1 and 2 and the 75% morphs to create Sets 3 and 4. We arbitrarily selected 26 stimuli in each set to be experimental objects and the remaining to be practice objects. Each participant was randomly assigned one set for the training sessions and a different set for the testing sessions. The four sets were used equally often on training and testing sessions in each group.

Apparatus

The tasks were run using the Pychtoolbox (MATLAB; Mathworks, Inc., Natick, MA; Kleiner et al., 2007; Brainard, 1997; Pelli, 1997), E-Prime (Psychology Software Tools, Pittsburgh, PA), or JavaScript. For the behavioral tasks, participants were seated in a dimly lit room at a viewing distance of approximately 56 cm from a computer monitor (screen resolution of 1280 × 1024 pixels with a frame rate of 60 Hz). The objects were presented on a black background. The participants wore headphones to receive auditory feedback (when presented) and to reduce distraction.

During EEG recording, participants were comfortably seated in a dimly lit and sound-attenuated room at a viewing distance of 1 m from a computer monitor (screen resolution of 1280 × 1024 pixels with a frame rate of 120 Hz). The objects for the EEG recording were presented on a gray background. EEG was acquired at 512 Hz using a 128-channel Biosemi Active II system (Biosemi, Amsterdam, The Netherlands), with electrodes including standard 10–20 system locations as well as additional intermediate positions (www.biosemi.com). Two additional electrodes (Common 252 Mode Sense active electrode and Driven Right Leg passive electrode) were used as reference and ground electrodes, respectively. Eye movements were monitored using four electrodes placed at the outer canthi of the eyes and above and below the right eye.

Training Sessions

Table 2 shows the training program used in the current study, during which participants learned to individuate 26 objects from three different viewpoints. All 26 objects were introduced at once. Each object was assigned a novel four- to five-letter name beginning with a different letter of the alphabet (e.g., Aklo, Bifa, Camar, …, Zoti; most of these names were taken from the names used for the Greebles in Gauthier & Tarr, 1997). Participants completed fourteen 1- to 2-hr sessions in total, which were spread out over 3 weeks.

Training Tasks

Familiarization tasks. There were two tasks to familiarize participants with the 26 objects in their training set: card sorting and inspection. For the card sorting task (Rossion, Schiltz, Robaye, Pirenne, & Crommelinck, 2001), we printed each object from each viewpoint on 9.7 × 9.7 cm black laminated cards. Participants were given a shuffled set of cards, and their task was to match the three views of each object as accurately as possible with no time pressure.

In the inspection task, participants viewed different combinations of objects together from the same viewpoint. There were three blocks in this task. On each block, participants were first shown an array of eight or nine objects from one of the viewpoints with their name written underneath. They were instructed to carefully study the objects and then to press the space bar to proceed. Objects for each block were randomly selected from the 26 objects without replacement.

Naming tasks. These tasks required participants to name each object as quickly and accurately as possible. There were two versions of this task: naming and duration threshold. On each trial, participants were presented with a fixation cross for 500 msec followed by an object at the center of the screen. They responded by pressing the key corresponding to the first letter of the object's name as quickly and accurately as possible. Each object was presented from the three viewpoints within a block, and the presentation order was randomized. On feedback

Table 2. The Extended Training Program Used in the Current Study

Session	Training Task	Feedback	Trials	Total Trials
1	Card sorting	No	78	
	Visual search: pic–pic	Auditory	78	
	Inspection (Block 1)	No	13	
	Inspection (Block 2)	No	13	
	Inspection (Block 3)	No	12	
	Naming with response	No	78	
	Verification: pic first	Auditory	156	
	Naming	Visual	78	
	Naming	No	78	584
2–4	Visual search: pic–pic	Auditory	78	
	Inspection (Block 1)	No	13	
	Inspection (Block 2)	No	13	
	Inspection (Block 3)	No	12	
	Naming with response	No	78	
	Verification: pic first	Auditory	156	
	Naming	Visual	156	
	Naming	No	78	584
5–7	Verification: RISE	Auditory	104	
	Verification: pic first	Auditory	156	
	Visual search: name–pic	Auditory	78	
	Naming	Visual	156	
	Naming	No	78	
	Duration threshold	No	78	650
8	Card sorting	No	78	
	Matrix scanning	Visual	26	
	Verification: pic first	No	156	
	Verification: name first	No	156	
	Naming	Visual	78	
	Naming	No	78	
	Duration threshold	No	78	650
9	Matrix scanning	Visual	26	
	Verification: pic first	No	156	
	Verification: name first	No	156	
	Naming	Visual	78	
	Naming	No	78	
	Duration threshold	No	78	572

Table 2. (*continued*)

Session	Training Task	Feedback	Trials	Total Trials
10	Matrix scanning	Visual	26	
	Verification: pic first	Auditory	156	
	Naming	Visual	156	
	Verification: RISE	No	104	
	Naming	No	78	
	Duration threshold	No	78	598
11	Verification: RISE	No	104	
	Verification: name first	Auditory	156	
	Naming	Visual	156	
	Naming	No	78	
	Duration threshold	No	78	572
12	Matrix scanning	Auditory	26	
	Verification: name first	No	156	
	Naming	Visual	156	
	Verification: RISE	No	104	
	Naming	No	78	
	Duration threshold	No	78	598
13	Verification: RISE	No	104	
	Verification: pic first	No	156	
	Naming	Visual	156	
	Naming	No	78	
	Duration threshold	No	78	572
14	Card sorting	No	78	
	Verification: name first	No	156	
	Verification: pic first	No	156	
	Naming	Visual	156	
	Naming	No	78	
	Duration threshold	No	78	702
				8,550

For each session, the training tasks are presented in the order that they were run, along with the type of feedback and number of trials. Auditory feedback was given in the form of a 500-Hz tone for incorrect responses. Visual feedback was given in the form of a visual display of the object with its correct response.

blocks, if participants responded incorrectly, they were shown the object again with the correct name written underneath and were required to press the correct key to continue.

For the duration threshold task, we used the QUEST procedure (Watson & Pelli, 1983) to adjust the stimulus duration so that participants' naming performance

remained at approximately 82% correct. Incorrect responses increased the stimulus duration, whereas correct responses decreased the stimulus duration. The initial stimulus duration was set to 500 msec for Sessions 1 and 10 and to the duration threshold achieved at the end of the previous session for all other sessions. The objects were shown only from the 0° view in this task, and no feedback was provided.

Verification tasks. These tasks required participants to match an object and its name label. There were three versions of this task: name–picture verification, picture–name verification, and Random Image Structure Evolution (RISE) verification. On each trial, participants were presented with a fixation cross for 500 msec, followed by a blank screen for 300 msec, a target stimulus (name or picture) for 1000 msec, another blank screen for 500 msec, and, finally, a probe stimulus (name or picture). The participants' task was to decide whether the target and probe stimuli matched (i.e., were the same identity) and to press the corresponding key as quickly and accurately as possible. On feedback blocks, if participants responded incorrectly, they heard a 1500-Hz tone for 500 msec. The 26 objects were presented from each viewpoint once for match and nonmatch trials. On nonmatch trials, we randomly paired an object at each view with a nonmatching name label.

We also used the RISE (Sadr & Sinha, 2004) procedure so that participants could learn to match objects and name labels despite visual noise. In the RISE procedure, we generated movies in which the Fourier phase on each frame linearly transitioned from 0% (pure visual noise) to 100% (fully visible image) in 120 frames at the rate of 30 frames/sec (4 sec in total). The trial structure was the same as in the name–picture matching task, except that the probe image was replaced by the RISE movie.

Search tasks. For the last type of training task, participants searched for target objects in a variety of ways. There were three versions of this task: picture–picture search, name–picture search, and matrix scanning (Wong, Palmeri, & Gauthier, 2009). For the first two search tasks, the trial structure was as follows: On each trial, participants were presented with a fixation cross for 500 msec, followed by a blank screen for 500 msec, a target stimulus (name or picture from the 0° viewpoint) for 1000 msec, another blank screen for 500 msec, and, finally, a circular search array of six objects that always contained the target. All the images in the search array were shown from either the +30° or −30° viewpoint, randomly determined on each trial. The participants' task was to click on the target object in the array. If participants selected the incorrect object, they heard a 1500-Hz tone for 500 msec.

The matrix scanning task was adapted from Wong, Palmeri, and Gauthier (2009). In brief, on each trial, participants saw a 4 × 7 matrix of different objects shown from various viewpoints (randomly determined for each object in the matrix). They scanned the matrix from left to right, top to bottom, searching for target objects until they reached the end of the matrix. They then pressed the space bar and typed the first letter of the last target's name. If participants responded incorrectly, they heard a 1500-Hz tone for 500 msec.

Testing Sessions

There were two testing sessions. The first session occurred 1–3 days before training (pretesting), and the second session occurred 1–3 days after training (posttesting). In each session, we tested participants on a demanding four-alternative forced-choice (4AFC) delayed matching task (Vuong et al., 2016; Laguesse, Dormal, Biervoye, Kuefner, & Rossion, 2012), and we separately acquired EEG data using the FPVS protocol. Both the behavioral and EEG tests occurred on the same day, with the order counterbalanced across participants. For each participant, the same object set was used on both testing sessions, and this set was different from the one used during the training sessions. Participants were also tested on other behavioral and EEG tasks with different types of stimuli, but these additional data are not reported here. Note that, for the pretesting session, the 30 participants in this study formed a subset of the participants tested in Vuong et al. (2016).

4AFC Delayed Matching Task during Testing Sessions

In the 4AFC delayed matching task, participants were shown a target stimulus followed by four probe stimuli. Their task was to select which probe stimulus matched the target stimulus. Each trial began with a white fixation cross at the center of the screen for 500 msec, followed by a blank screen for another 500 msec, a target stimulus presented for 500 msec, another 500-msec blank screen, and four probe stimuli arranged in a 2 × 2 matrix. Participants pressed the key corresponding to the probe that matched the target (ignoring viewpoint changes) as quickly and accurately as possible. On each trial, the target and probe stimuli were always shown from the same stimulus orientation (facelike or nonfacelike), and the four probe stimuli were always shown from the same viewpoint. To avoid image matching and to increase the difficulty of the task, the target and probe stimuli were always shown from different viewpoints (e.g., target stimulus at 0° and all probe stimuli at +30°, or target stimulus at +30° and all probe stimuli at −30°). We randomly selected five of the six possible view combinations for each Stimulus × Orientation condition and for each participant to reduce the overall duration of the experiment. There were 260 trials (26 stimuli × 2 stimulus orientations × 5 repetitions). There were 20 practice trials with feedback (i.e., a 1500-Hz tone occurred when an incorrect response was made).
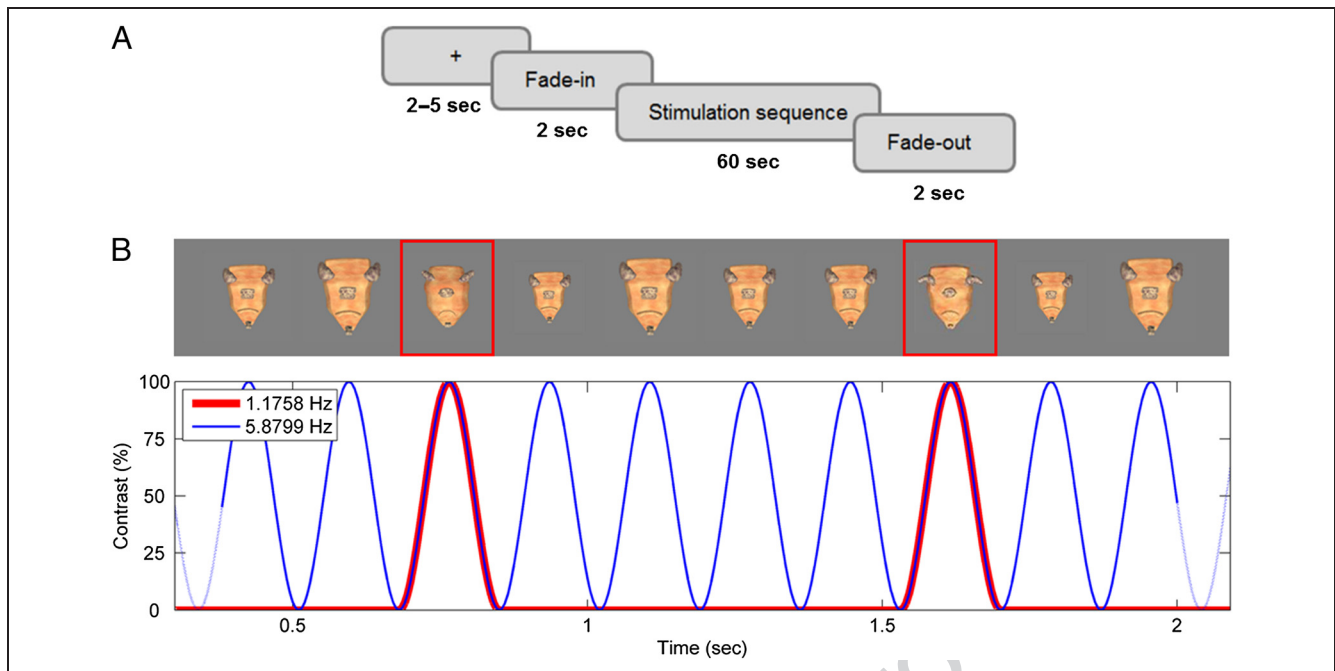
Figure 2. The FPVS protocol used in the current study. (A) Each trial consisted of a fixation period, an initial fade period, the main stimulation sequence, and an end fade period. (B) The time course of the main stimulation sequence. The same image was displayed with sinusoidal contrast modulation from 0% to 100% contrast at 5.8799 Hz during the 60-sec sequence (blue line). Identity changes occurred at 1.1758 Hz (red line) with a new exemplar replacing the repeated image on every fifth presentation (red squares). The images were randomly scaled to different sizes per presentation.

EEG Recording during Testing Sessions

During the EEG recording, participants were presented with FPVS sequences while they performed a fixation task (color-change monitoring). This task served to maintain a constant level of attention throughout the sequence. Figure 2 illustrates the FPVS protocol used. Each trial began with a black fixation cross on a gray background presented for a randomly determined duration between 2 and 5 sec. The cross was followed by a 2-sec "fade" period, then the main 60-sec stimulation sequence, and then another 2-sec "fade" period (Figure 2A). During the main stimulation sequence (Figure 2B), the same image was periodically presented by sinusoidally modulating that image's contrast from a gray background (0% contrast) to full (100%) contrast and back in 166 msec (reaching full contrast in 83 msec). This gave rise to a frequency of 5.8799 Hz—here referred to as the "visual stimulation" frequency. In addition, a new image was presented on every fifth item in the sequence. This gave rise to an additional stimulation frequency of 1.1758 Hz— here referred to as the "individual discrimination" frequency. There were 70 complete cycles and 14 identity changes in total. During the "fade" periods, the sinusoidal contrast modulation was additively combined with a linear contrast modulation in which the contrast increased from 0% to 100% in 2 sec (or vice versa). To reduce low-level image changes driving the EEG responses at the two frequencies, images randomly changed size on each presentation between 80% and 120%.

There were eight EEG trials for approximately 8 min of EEG recording. Four of the trials presented objects in the facelike stimulus orientation, and the other four trials presented the objects in the nonfacelike stimulus orientation. One of the 26 objects in the testing set was randomly selected as the repeated image for each trial. The remaining 25 objects served as exemplars introduced on every fifth stimulus (randomly selected without replacement for that trial). All objects were presented from the frontal view. The conditions were presented in random order. For the fixation task, the central cross changed color from blue to red for 200 msec, independently of the stimulus manipulation (six changes randomly distributed throughout the 60-sec sequence). Participants pressed the space bar whenever they detected this change. There was a break between each trial to ensure low-artifact EEG signals. Participants responded very accurately on this task (average = 97–98%).

EEG Analysis

Preprocessing. EEG analyses were carried out using Letswave (Version 5; nocions.webnode.com/letswave). After bandpass filtering between 0.1 and 100 Hz, the EEG data were segmented to include 2 sec before and after each sequence, resulting in 64-sec segments (−2 to 62 sec). Data files were then downsampled to 250 Hz to reduce file size and data processing time. Artifact-ridden or noisy channels were replaced using linear interpolation. All channels were rereferenced to the common average.

The EEG recordings were then segmented again from stimulation onset until 59.5828 sec, corresponding to 70 complete cycles at 1.1758 Hz (14,883 bins).

Frequency domain analysis. The four 60-sec stimulation sequences for the facelike and nonfacelike stimulus orientations were separately averaged in the time domain for each participant to reduce EEG activity that is not phase-locked to the stimulus. A fast Fourier transform was then applied to the averaged segments, and amplitude spectra were extracted for all channels (square root of the sum of squares of the real and imaginary parts divided by the number of data points). Because we used a long time window (59.5828 sec), the frequency analysis yielded spectra with a high-frequency resolution (1/59.5828 sec = 0.0168 Hz), providing a high signal-to-noise ratio (SNR; Regan, 1989) at the frequencies of interest (i.e., 1.1758 and 5.8799 Hz and their harmonics).

To correct for EEG noise, the amplitude spectrum for each participant and condition was corrected by subtracting the average amplitude of the 20 surrounding bins (10 bins on each side, excluding the immediately adjacent bins) from the amplitude at each frequency bin (Retter & Rossion, 2016). The resulting baseline-subtracted (SBL) response at each frequency bin is expressed in microvolts (µV). On the basis of the grand-averaged amplitude spectrum for each condition, z scores (difference between amplitude at the frequency of interest and mean amplitude of 20 surrounding bins divided by the standard deviation of the 20 surrounding bins) were estimated for posterior electrodes of interest to assess the significance of the response at 1.1758 and 5.8799 Hz and their corresponding harmonics. z Scores larger than 3.1 ($p < .001$) were considered significant. We then selected a similar number of harmonics (the maximal number of consecutive significant harmonics) across conditions averaged over training group, testing session, and stimulus orientation. For the individual discrimination frequency, we found significant responses from 1.1758 to 7.0548 Hz (from the first to sixth harmonics, excluding the visual stimulation frequency, over posterior electrodes). For the visual stimulation frequency, we found significant responses from 5.8799 to 35.2794 Hz (from the first to sixth harmonics). Finally, the SBL responses from significant harmonics were summed per participant (Retter & Rossion, 2016), separately for the two frequencies. To define ROIs for our analyses, we ranked electrodes according to their SBL response amplitudes averaged over training group, testing session, and stimulus orientation and selected the four electrodes common to both the facelike and nonfacelike groups with the largest amplitude in both the right-hemisphere (PO8,

Figure 3. Mean RTs for different tasks used during the 14 training sessions per training group. Some tasks were performed during a subset of the sessions, whereas the naming without feedback task was performed on every session. Both groups generally improved with training, and there were overall no significant differences between groups with a few small exceptions and trends (***$p < .001$; *$p < .07$).
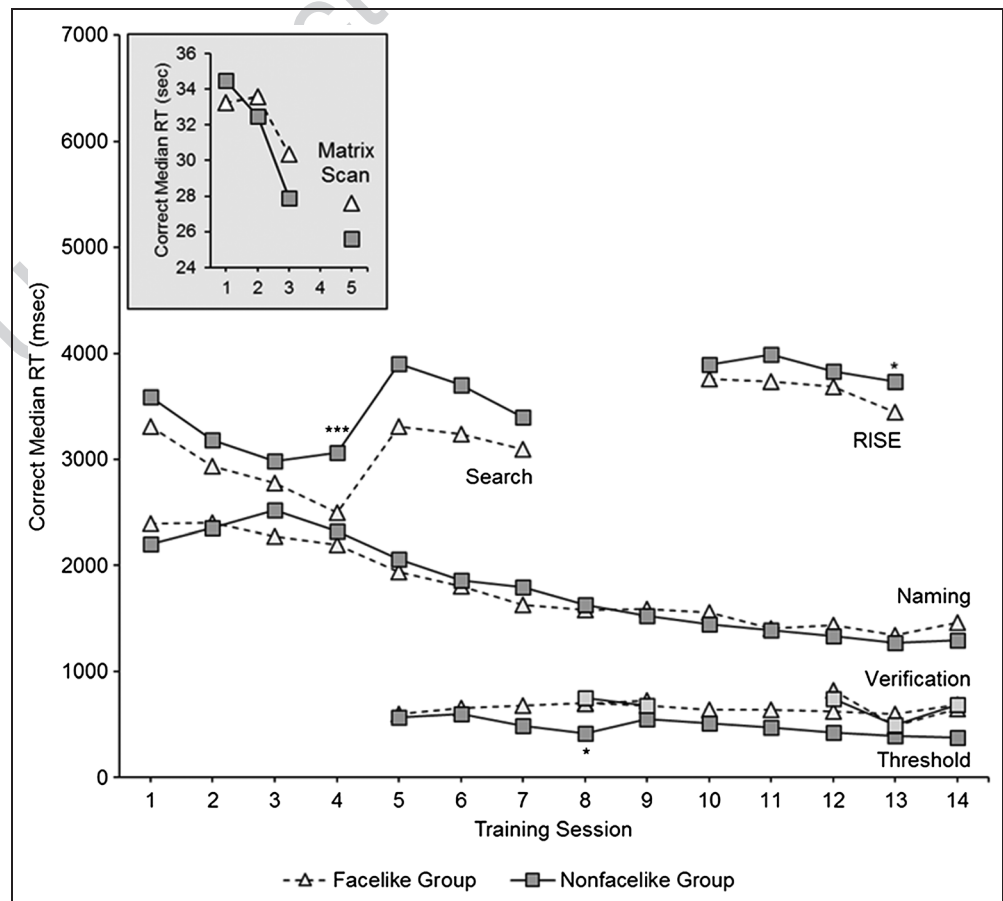
Table 3. Mean Proportion Correct (and *SEM*) Averaged across Participants in the Facelike and Nonfacelike Groups for the Different Training Tasks in Each Session

Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Naming														
Facelike	0.21	0.43	0.60	0.69	0.73	0.78	0.82	0.82	0.85	0.87	0.88	0.90	0.90	0.90
	0.03	0.06	0.06	0.06	0.06	0.05	0.05	0.04	0.03	0.03	0.03	0.02	0.03	0.03
Nonfacelike	0.19	0.46	0.63	0.74	0.80	0.81	0.84	0.87	0.87	0.82	0.90	0.92	0.92	0.93
	0.03	0.05	0.06	0.05	0.05	0.04	0.04	0.04	0.04	0.07	0.03	0.02	0.02	0.02
Threshold														
Facelike					0.69	0.74	0.78	0.79	0.81	0.84	0.85	0.86	0.85	0.86
					0.05	0.05	0.04	0.03	0.03	0.02	0.02	0.02	0.03	0.02
Nonfacelike					0.72	0.77	0.77	0.78	0.82	0.85	0.86	0.88	0.87	0.86
					0.05	0.04	0.06	0.06	0.03	0.02	0.02	0.02	0.02	0.01
RISE														
Facelike										0.81	0.87	0.85	0.88	
										0.09	0.04	0.07	0.05	
Nonfacelike										0.94	0.92	0.94	0.93	
										0.01	0.01	0.01	0.01	
Matrix scan														
Facelike								0.54	0.61	0.66		0.75		
								0.06	0.07	0.07		0.06		
Nonfacelike								0.50	0.66	0.72		0.75		
								0.08	0.06	0.05		0.05		
Verification														
Facelike								0.89	0.88			0.88	0.86	0.82
								0.04	0.05			0.08	0.08	0.10
Nonfacelike								0.93	0.94			0.96	0.96	0.95
								0.01	0.01			0.01	0.01	0.01
Search														
Facelike	0.79	0.80	0.82	0.86	0.82	0.86	0.88							
	0.01	0.01	0.02	0.01	0.05	0.03	0.04							
Nonfacelike	0.77	0.79	0.84	0.85	0.87	0.90	0.92							
	0.02	0.02	0.02	0.01	0.03	0.03	0.03							

Overall, the two groups did not differ in performance during the training sessions.
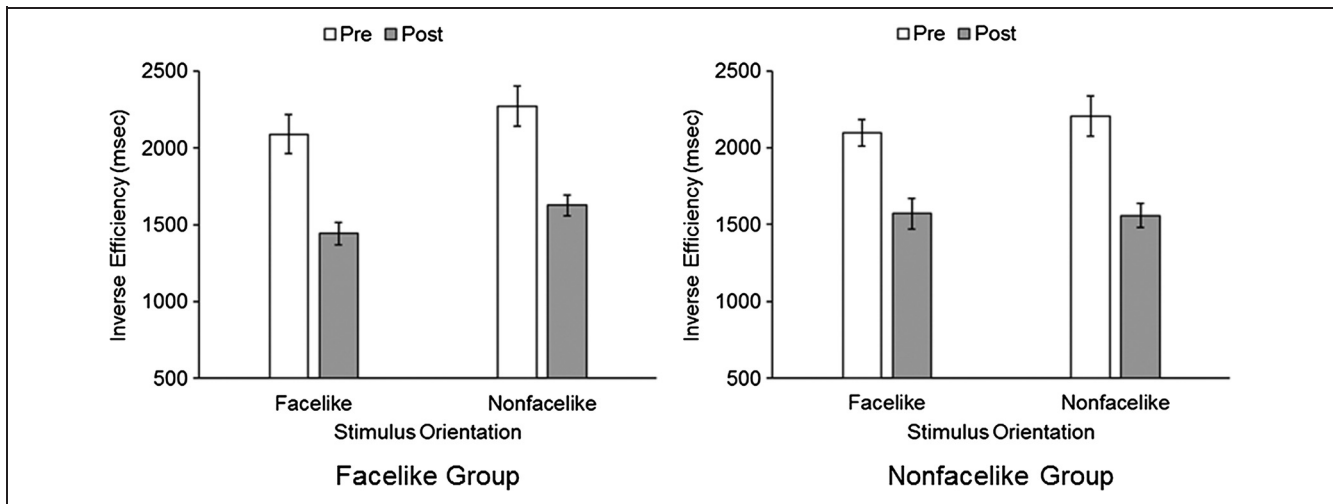
Figure 4. Mean inverse efficiency scores (correct median RT/proportion correct, in milliseconds) of the 4AFC delayed matching task for the facelike (left) and nonfacelike (right) groups as a function of testing session and stimulus orientation. The error bars in this and subsequent graphs represent the *SEM*.

PO10, PPO6, and POO6) and left-hemisphere (PO7, PO9, PPO5, and POO5) ROI. These are homologous electrodes. The SBL responses from the electrodes in each hemisphere were then averaged to define the right and left ROIs.

We used the SBL responses for all statistical analyses. To control for outliers, we first identified responses across all observations that were greater than 1.96 *SD*s above the grand mean (i.e., 95% confidence interval). These outliers were replaced by the participant's mean across the remaining conditions for that participant. For the individual discrimination frequency, there were 240 observations (15 participants × 2 training groups × 2 testing sessions × 2 stimulus orientations × 2 hemispheres) with 12 outliers. For the visual stimulation frequency, there were 120 observations (15 participants × 2 training groups × 2 testing sessions × 2 stimulus orientations) with six outliers. Note that, for visualization purposes, we used the SBL responses (including outliers) for scalp topographies and baseline-divided (SNR) responses (including outliers) for amplitude spectra.

RESULTS

Training Results

For the training data, we focused on the following tasks: naming without feedback, duration threshold, matrix scanning, picture–picture search, name–picture search, picture verification without feedback, name verification without feedback, and RISE verification. Figure 3 presents the correct median RTs, and Table 3 presents the proportion-correct data. For each training task and training session, we conducted a *t* test to compare the RTs and proportion-correct data between the facelike and nonfacelike groups. Overall, there were no significant differences in performance between the two groups for most comparisons. Importantly, the naming without feedback task allowed us to gauge learning more directly as it was the last task in each training session. Both groups responded progressively more quickly and more accurately across sessions, reaching an asymptote around Session 11 for both correct RTs and proportion correct. As evident in Figure 3 and Table 3, there were no significant differences between the two groups for any session (*t*s < 1.35).

Testing Results

4AFC Delayed Matching Task

Figure 4 presents the inverse efficiency scores in the 4AFC delayed matching task (correct median RT/proportion correct; Townsend & Ashby, 1983) as a function of training group, testing session, and stimulus orientation.

Table 4. Mean Proportion Correct and RT from the 4AFC Delayed Matching Task as a Function of Training Group, Stimulus Orientation, and Testing Session

	Proportion Correct		Correct Median RT (msec)	
	Pre	Post	Pre	Post
Facelike group				
Facelike	0.80 (0.01)	0.88 (0.01)	1678 (104)	1256 (56)
Nonfacelike	0.77 (0.02)	0.85 (0.02)	1751 (104)	1380 (54)
Nonfacelike group				
Facelike	0.78 (0.01)	0.89 (0.01)	1645 (71)	1395 (80)
Nonfacelike	0.77 (0.02)	0.90 (0.01)	1680 (85)	1396 (72)

The *SEM* is presented in parentheses.

Table 4 presents the mean and *SEM* for proportion correct (chance = 0.25) as well as correct RT in the different conditions. The inverse efficiency scores were submitted to a $2 \times 2 \times 2$ mixed ANOVA with Training group (facelike, nonfacelike) as a between-participant factor and Testing session (pretesting, posttesting) and Stimulus orientation (facelike, nonfacelike) as within-participant factors. There were main effects of Testing session, $F(1, 28) = 68.21, p < .001, \eta_p^2 = .71$, and Stimulus orientation, $F(1, 28) = 16.94, p < .001, \eta_p^2 = .38$, but no effect of Training group ($F < 1.0$). Participants responded more efficiently with facelike stimuli ($M = 1799$ msec, $SE = 58$ msec) compared with nonfacelike stimuli ($M = 1915$ msec, $SE = 62$ msec), and they responded more efficiently after training ($M = 1549$ msec, $SE = 54$ msec) than before training ($M = 2165$ msec, $SE = 82$ msec). There was also a significant interaction between Training group and Stimulus orientation, $F(1, 28) = 5.92, p = .02, \eta_p^2 = .17$. Participants in the facelike group responded more efficiently to the facelike stimuli ($M = 1765$ msec, $SE = 86$ msec) compared with the nonfacelike stimuli ($M = 1949$ msec, $SE = 92$ msec), $t(14) = 6.72, p < .001$. By comparison, participants in the nonfacelike group responded equally efficiently to both facelike ($M = 1834$ msec, $SE = 79$ msec) and nonfacelike ($M = 1881$ msec, $SE = 84$ msec) stimuli, $t(14) < 1.0$.

For comparison with our EEG results (see below), we also analyzed the inverse efficiency scores separately for the facelike and nonfacelike groups. For this analysis, the data were submitted to a 2×2 ANOVA with Testing session and Stimulus orientation as within-participant factors. For the facelike group, there were main effects of Testing session, $F(1, 28) = 37.86, p < .001, \eta_p^2 = .65$, and Stimulus orientation, $F(1, 28) = 45.14, p < .001$,

$\eta_p^2 = .76$. Participants responded more efficiently with facelike compared with nonfacelike stimuli. They also responded more quickly after training ($M = 1534$ msec, $SE = 67$ msec) than before training ($M = 2180$ msec, $SE = 128$ msec). By comparison, for the nonfacelike group, there was only a main effect of Testing session, $F(1, 28) = 30.58, p < .001, \eta_p^2 = .69$, with participants responding more efficiently after training ($M = 1564$ msec, $SE = 85$ msec) than before training ($M = 2151$ msec, $SE = 102$ msec). Overall, participants responded more efficiently for facelike stimuli compared with nonfacelike stimuli (see also Vuong et al., 2016). However, there was no behavioral change specific to the trained stimulus orientation for either group after training.

FPVS Responses in the EEG Recordings

Figure 5A shows the SNR spectrum averaged across groups and conditions. There is a clear individual discrimination response at 1.1758 Hz and its harmonics, which reflects the onset and offset of different identities in the sequence. There is also a clear visual stimulation response at 5.8799 Hz (and its harmonics, only the first shown at 11.7598 Hz), which reflects the onset and offset of each visual stimulus in the sequence. Figure 5B shows the SBL topography of the posterior scalp averaged across conditions for the facelike and nonfacelike groups, showing bilateral occipitotemporal responses for the individual discrimination response and medial occipital responses for the visual stimulation response.

Individual discrimination responses. Figure 6A shows the posterior scalp topographies as a function of testing

Figure 5. Summary of EEG responses. (A) The EEG spectrum of the response (SNR) over the right occipito-temporal electrode (PO8) averaged across groups and conditions. There are clear responses at the individual discrimination frequency (1.1758 Hz and its harmonics) and at the visual stimulation frequency (5.8799 Hz and its harmonics; only the first harmonic is shown). (B) The posterior scalp topography for the individual discrimination response (left) and visual stimulation response (right) averaged across conditions for the facelike and nonfacelike groups.
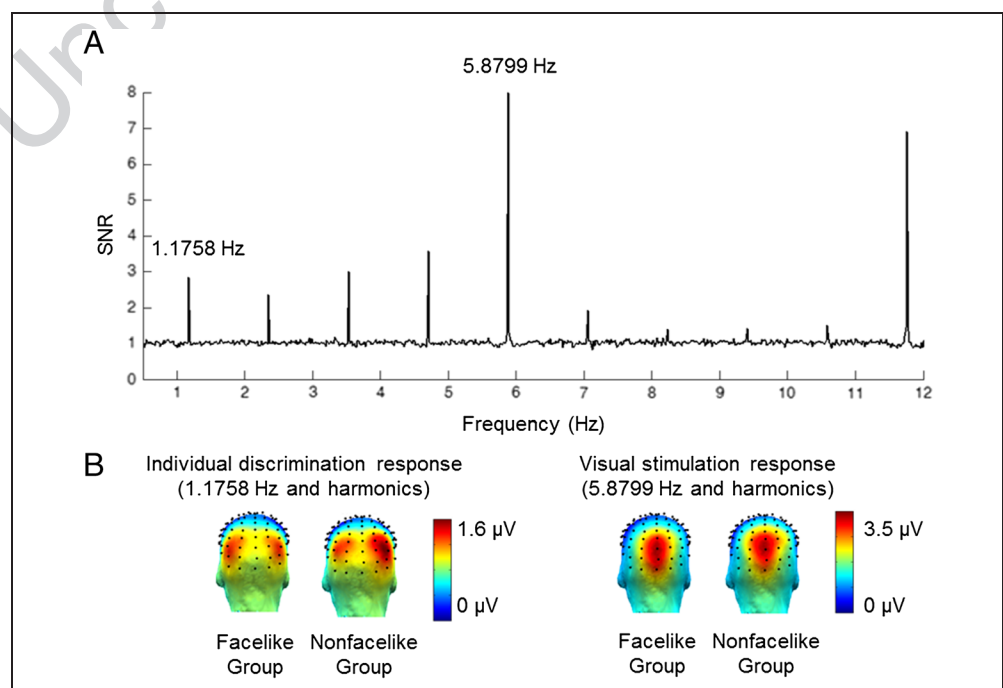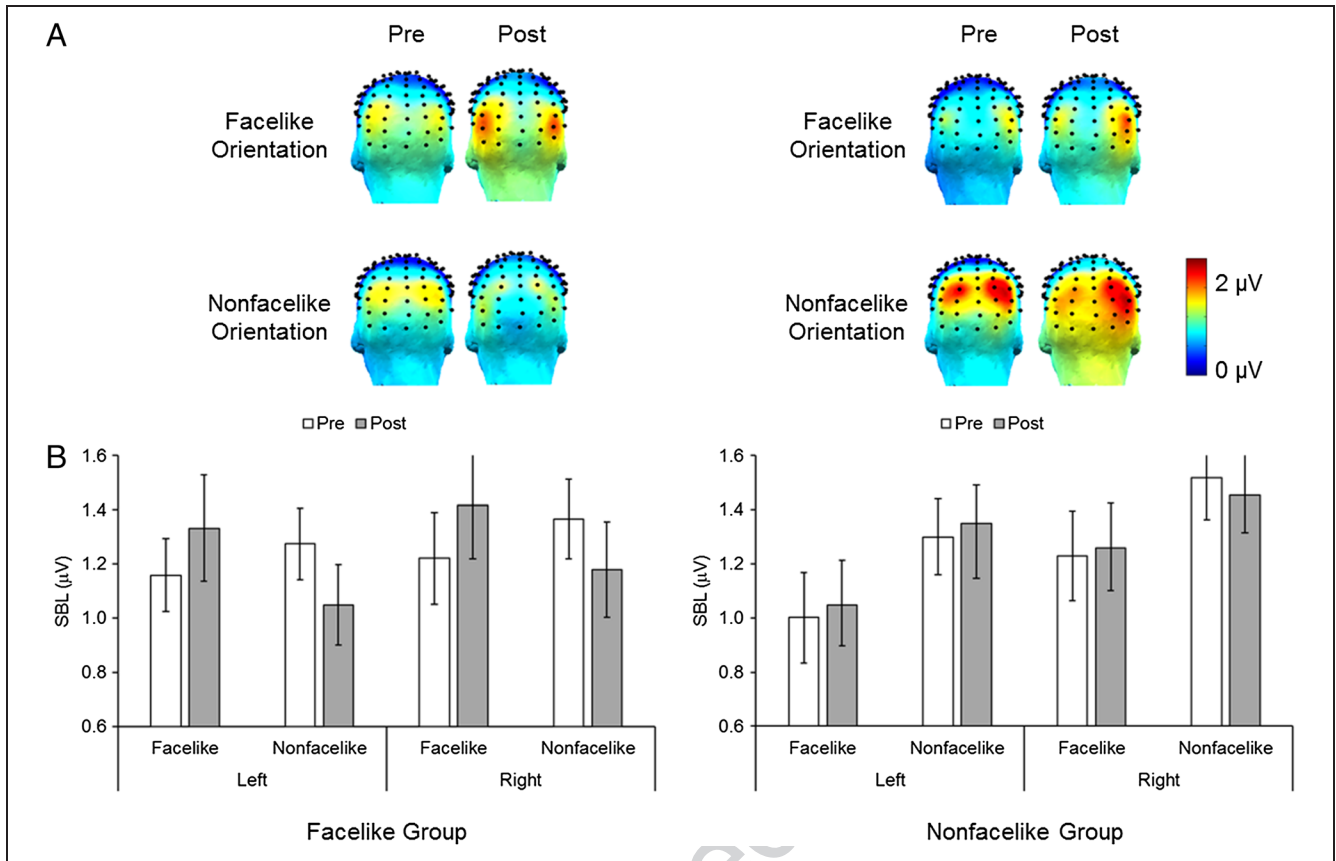
Figure 6. Individual discrimination responses in the EEG. (A) The posterior scalp topographies displaying the mean SBL amplitude at the individual discrimination frequency (1.1758 Hz and significant harmonics) for the facelike (left) and nonfacelike (right) groups as a function of testing session and stimulus orientation. Note that these topographies include the outlier responses. (B) Mean SBL amplitude for the facelike (left) and nonfacelike (right) groups as a function of testing session, stimulus orientation, and hemisphere (left ROI: PO7, PO9, PPO5, and POO5; right ROI: PO8, PO10, PPO6, and POO6).

session and stimulus orientation, separately for the facelike and nonfacelike groups. Figure 6B depicts the mean individual discrimination responses across the different conditions. We submitted the responses to an omnibus $2 \times 2 \times 2 \times 2$ mixed ANOVA with Training group as a between-participant factor and Testing session, Stimulus orientation, and Hemisphere as within-participant factors. There was no significant effect of Training group, Testing session, or Stimulus orientation ($ps > .15$). There was a main effect of Hemisphere, $F(1, 28) = 4.35, p = .046, \eta_p^2 = .13$, with a larger individual discrimination response in the right hemisphere ($M = 1.32$ μV, $SE = 0.09$ μV) compared with the left hemisphere ($M = 1.19$ μV, $SE = 0.09$ μV).

Importantly, there was a significant Training group × Stimulus orientation interaction, $F(1, 28) = 7.32, p = .01, \eta_p^2 = .21$, and a Testing session × Stimulus orientation interaction, $F(1, 28) = 4.81, p = .037, \eta_p^2 = .15$. There was also a trend for a three-way interaction between Training group, Testing session, and Stimulus orientation, $F(1, 28) = 2.91, p = .09, \eta_p^2 = .09$. No other interactions were significant ($ps > .38$). ERP studies show that the N170 component can increase to the trained stimulus orientation after training with novel objects (e.g., Cao

et al., 2015; Rossion et al., 2002, 2004). To further explore these interactions and given these previous ERP findings, we conducted separate 2×2 ANOVAs for each training group with Testing session and Stimulus orientation as within-participant factors. For the facelike group, there was only a significant Testing session × Stimulus orientation interaction, $F(1, 14) = 6.97, p = .02, \eta_p^2 = .33$. Post hoc t tests showed that there was a larger response to the facelike stimuli ($M = 1.37$ μV, $SE = 0.19$ μV) compared with the nonfacelike stimuli ($M = 1.11$ μV, $SE = 0.16$ μV) after training, $t(14) = 2.31, p = .04$, but responses did not differ significantly between the two stimulus orientations before training (facelike: $M = 1.19$ μV, $SE = 0.15$ μV; nonfacelike: $M = 1.29$ μV, $SE = 12$ μV), $t(14) < 1.0$. By comparison for the nonfacelike group, there was no significant Testing session × Stimulus orientation interaction ($F < 1.0$). There was a main effect of Stimulus orientation, $F(1, 14) = 7.53, p = .02, \eta_p^2 = .35$, with a larger response to the nonfacelike stimuli ($M = 1.41$ μV, $SE = 0.11$ μV) compared with the facelike stimuli ($M = 1.13$ μV, $SE = 0.09$ μV). No other main effects or interactions were significant ($ps > .35$). Overall, only participants in the facelike group showed a change in the individual discrimination response for the trained facelike stimulus orientation

after training. This finding provides evidence that extensive training can lead to training-induced neural changes to rapid and automatic perceptual processing but only if the novel objects resemble faces.

Participants in the facelike and nonfacelike groups showed learning as measured by their improved performance on the learning tasks (particularly naming) and on the 4AFC task after training. To assess whether there were any neural changes related to this general learning, we conducted a "whole-scalp" analysis on the individual discrimination response averaged across all electrodes, excluding the eight posterior electrodes that were analyzed in the ROI analysis. These other electrodes likely reflect neural responses outside the ventral occipitotemporal and lateral occipital cortex. We submitted the data to an omnibus ANOVA with Training group as a between-participant factor and Testing session and Stimulus orientation as within-participant factors. There was no main effect of Testing session, $F(1, 28) = 0.90$, $p = .35$, although there was a slight decrease of the amplitude of the individual discrimination response for both groups during the posttesting session compared with the pretesting session. For the facelike group, there was a decrease of 0.04 μV (pretesting: $M = 0.52$ μV, $SE = 0.09$ μV; posttesting: $M = 0.48$ μV, $SE = 0.05$ μV). For the nonfacelike group, this decrease was 0.06 μV (pretesting: $M = 0.56$ μV, $SE = 0.06$ μV; posttesting: $M = 0.50$ μV, $SE = 0.05$ μV). Testing session did not interact with any of the other factors (all Fs < 1).

Similar to our findings with the ROI analysis, we found a significant interaction between Training group and Stimulus orientation, $F(1, 28) = 13.84$, $p < .001$, $\eta_p^2 = .33$. For the facelike group, the individual discrimination response did not differ between the facelike ($M = 0.54$ μV, $SE = 0.08$ μV) and nonfacelike ($M = 0.47$ μV, $SE = 0.06$ μV) stimuli, $t(14) = 1.65$, $p = .12$. By comparison for the nonfacelike group, the response was larger

for the nonfacelike ($M = 0.64$ μV, $SE = 0.06$ μV) than facelike ($M = 0.43$ μV, $SE = 0.04$ μV) stimuli, $t(14) = 3.37$, $p = .005$. No other effects or interactions were significant. Thus, there was no evidence of neural changes related to general learning. Moreover, the larger response to the nonfacelike stimuli in the nonfacelike group appears on the whole scalp, not only in the posterior electrodes reflecting perceptual processes, so is unlikely related to any form of learning (although we do not know why there is this difference for the nonfacelike group, but not for the facelike group).

Visual Stimulation responses. Figure 7A and B show the mean visual stimulation response as a function of testing session and stimulus orientation, separately for the facelike and nonfacelike groups. We submitted the responses to a 2 × 2 × 2 mixed ANOVA with Training group as a between-participant factor and Testing session and Stimulus orientation as within-participant factors. There were no main effects of Training group or Testing session (ps > .14). There was a main effect of Stimulus orientation, $F(1, 28) = 12.94$, $p = .001$, $\eta_p^2 = .32$, but this effect was qualified by a significant Training group × Stimulus orientation interaction, $F(1, 28) = 5.00$, $p = .03$, $\eta_p^2 = .15$. For the facelike group, there was no significant difference in the responses to the facelike stimuli ($M = 3.15$ μV, $SE = 0.30$ μV) or the nonfacelike stimuli ($M = 3.39$ μV, $SE = 0.35$ μV), $t(14) = 1.10$, $p = .29$. For the nonfacelike group by comparison, responses were larger for the nonfacelike stimuli ($M = 4.10$ μV, $SE = 0.40$ μV) compared with the facelike stimuli ($M = 3.02$ μV, $SE = 0.41$ μV), $t(14) = 3.72$, $p < .001$. When the responses were analyzed separately for the facelike and nonfacelike groups, there was a marginally significant main effect of Testing session for the facelike group, $F(1, 14) = 4.51$, $p = .052$, $\eta_p^2 = .24$, showing a slight increase in response amplitude (pretesting: $M = 3.06$ μV, $SE = 0.27$ μV; posttesting: $M = 3.46$ μV,
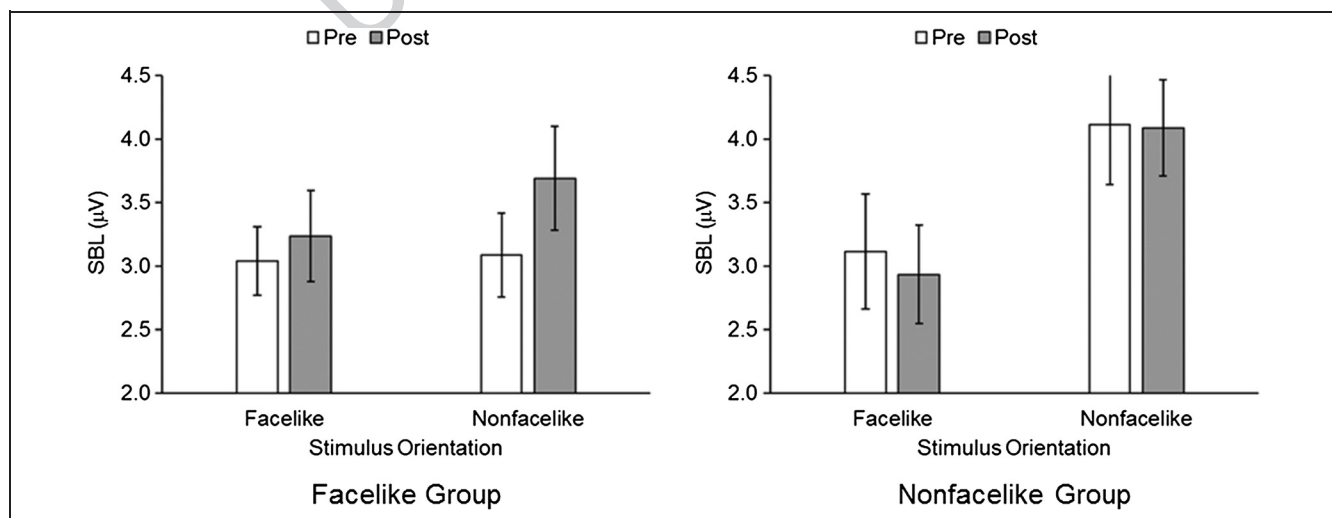

Figure 7. Mean SBL amplitude at the visual stimulation frequency (5.8798 Hz and significant harmonics) over the Oz site for the facelike (left) and nonfacelike (right) groups as a function of testing session and stimulus orientation.

$SE = 0.37$ μV). There was also a significant effect of Stimulus orientation for the nonfacelike group, $F(1, 14) = 13.81$, $p < .002$, $\eta_p^2 = .50$, showing higher response amplitudes for the nonfacelike stimuli (facelike: $M = 3.02$ μV, $SE = 0.41$ μV; nonfacelike: $M = 4.10$ μV, $SE = 0.40$ μV). All other effects were not significant ($ps > .12$). Thus, in contrast to the individual discrimination response, there were no changes in the visual stimulation response specific to the trained stimulus orientation for either group (i.e., no Testing session × Stimulus orientation interaction for either group).

DISCUSSION

In the current study, we tested whether extensive experience individuating exemplars of a novel object category would lead to the acquisition of visual expertise. Two groups of adult participants underwent extensive training (14 training sessions over 3 weeks, approximately 28 hr in total) with a well-controlled set of novel, parametric, multipart 3-D objects (Vuong et al., 2016). The two groups differed only in whether the participants were trained with the objects in a facelike or nonfacelike stimulus orientation. We used a novel approach in the domain of visual expertise—the FPVS protocol—and recorded EEG responses to a rapid stimulation sequence to evaluate whether participants showed evidence of training-induced changes to rapid and automatic object individuation, as these are important aspects of visual expertise (Tanaka & Curran, 2001). Consistent with our hypotheses, we found a training-induced modulation of the individual discrimination response over lateral occipital sites in both hemispheres for the facelike group: EEG responses at the individual discrimination frequency (1.1758 Hz) were stronger for the facelike (trained) than nonfacelike (untrained) stimulus orientation after training, whereas responses to the two orientations did not differ beforehand. In contrast, for the nonfacelike group, there were no changes to the individual discrimination response over the same sites after training. We found that changes to the EEG responses at the visual stimulation frequency (5.8799 Hz) over the Oz site were not specific to the trained stimulus orientation for participants in both the facelike and nonfacelike groups. Thus, training-induced neural changes were specific to the visual discrimination of objects at the individual level rather than reflecting a general response to visual stimulation by these objects.

The increase in the individual discrimination response for the facelike group after training occurred under very demanding presentation conditions: (1) The stimuli were only available for a single glance because of the fast stimulus presentation rate (approximately 6 stimuli/sec), (2) there was both forward and backward masking of individual exemplars in the sequence, and (3) participants' attention was directed at the fixation cross. Furthermore, this increase in the neural response was observed for un-

trained exemplars, which did not have name labels. Our study thus provides the first clear evidence for a training-induced increase in rapid and automatic object individuation in adulthood. Importantly, this increase, although small, was dependent on the perceived facelikeness of the individual objects with respect to the spatial configuration of object parts. Indeed, we previously showed that naive participants rated the facelike stimulus orientation as more "facelike" than the nonfacelike stimulus orientation (Vuong et al., 2016). Although previous studies have shown evidence for training-induced neural changes in ERP markers of early perceptual processes (e.g., the N170 component; Cao et al., 2015; Rossion et al., 2002, 2004), there were limitations in the stimulus set used (Vuong et al., 2016). Furthermore, slow EEG paradigms do not necessarily measure neural responses that reflect rapid and automatic visual discrimination at the individual level.

Although we found neural changes for the facelike group after extensive training, we did not find a larger individual discrimination response in the right compared with the left hemisphere after training for this group. Previous studies consistently found larger individual discrimination responses in the right hemisphere for face stimuli (e.g., Liu-Shuang, Torfs, & Rossion, 2016; Liu-Shuang et al., 2014). Thus, our results suggest that brain regions and neural mechanisms different from those that process faces were recruited for the objects after training. That is, brain regions and neural mechanisms involved in individual face recognition do not necessarily underlie visual expertise in general. We also found significant individual discrimination responses over bilateral occipital sites in both groups to both stimulus orientations before any training. This finding likely reflects the fact that our novel objects were perceptually distinctive enough for participants to detect a change in object identity during the stimulus sequence before training and even when their attention was directed at the fixation cross. To some extent, it should not be surprising that there is rapid and automatic perceptual processing of perceptually discriminable objects irrespective of training. Our results, however, clearly show that extensive training can enhance this perceptual discriminability. Critically, this increased perceptual discriminability was only observed after training with the objects from the facelike but not the nonfacelike stimulus orientation, although the training stimuli were identical in both groups except for a 180° rotation in the picture plane. Last, we found that participants in the nonfacelike group showed a stronger neural response to individual changes in the nonfacelike stimuli than in the facelike stimuli at the pretesting session but this difference was also present at the posttesting session. We do not know why this group showed this difference between the two stimulus orientations. However, and critically for our hypotheses, there are two observations to note. First, this stronger response to the nonfacelike stimuli was not specific to perceptual processes, as

we found it across the whole scalp. Second, it was not specific to identity changes, as we also found a stronger response to the nonfacelike stimuli for the visual stimulation frequency.

The results from the training tasks showed that participants in the facelike and nonfacelike groups learned the objects at the individual level to the same degree so differences in the individual discrimination responses between the two groups are unlikely related to potential differences in learning. Participants in both groups named trained exemplars progressively more accurately and more quickly with training, reaching an asymptote a few sessions before the end of the training program. This naming task was particularly important because it required fast and accurate individual-level recognition. More generally, participants in both groups performed equally well across all training tasks. The results from the 4AFC delayed matching task (Laguesse et al., 2012) during the testing sessions also showed little training-related differences between participants in the facelike and nonfacelike groups, suggestive of general learning effects. Participants in both groups responded more efficiently after training, and all participants showed an overall advantage for stimuli in the facelike stimulus orientation (Vuong et al., 2016).

In this study, we focused on training-specific neural changes related to rapid and automatic perceptual processes that are important for visual expertise. We found evidence for these changes for the facelike group but not for the nonfacelike group. Participants in both groups showed strong general learning; that is, their overall performance on the different behavioral tasks improved after training. However, we did not find any neural correlates of this general learning in either group. Our whole-scalp analyses further showed no neural changes related to general learning outside visual cortex. Thus, we favor the interpretation that visual expertise is best characterized by changes to rapid and automatic perceptual processes that require little to no attention and are likely localized in visual cortex. Harel and colleagues (Harel, 2016; Harel et al., 2010, 2013), on the other hand, have proposed an interactive framework of visual expertise highlighting the importance of expertise-related attentional modulation of behavior by a much more distributed network of brain regions across the entire brain (including frontal and parietal regions). They focused on other domains of expertise (e.g., birds and cars) and on fMRI results. Future work is needed to determine whether visual expertise is better characterized by perceptual or interactive changes. Here, we note a few factors that need to be considered. First, in other natural domains of visual expertise (e.g., birds, dogs, or cars), experts discriminate exemplars at the subordinate level and rarely at the individual level. Second, whether perceptual or interactive changes occur may depend on the extent to which different types of features are used to discriminate the to-be-learned object exemplars (e.g., spatial configuration of parts vs. features of intermediate complexity; see Harel et al., 2013).

Overall, our results provide two important interrelated constraints for the acquisition of visual expertise by adults and the neural mechanisms underlying this expertise. First, they support the idea that face resemblance is an important perceptual property to facilitate the emergence of neural changes related to visual expertise (Brants et al., 2011; Biederman & Kalocsai, 1997). Previous studies have shown that adults can learn to name individual Greebles as quickly and accurately as they can categorize them and that Greebles can elicit responses in face-sensitive areas after training (Gauthier & Tarr, 2002; Gauthier et al., 1999). However, it was not clear whether these results were due to the Greebles' resemblance to faces, to the fact that they may have diagnostic parts, or to any physical or perceptual (dis)similarity between them (Brants et al., 2011; McKone et al., 2007; Biederman & Kalocsai, 1997). Our results disentangle this issue. The very same stimuli, differing only by a rotation in the picture plane, induced neural changes after extensive training only in the orientation that gave rise to a strong facelike percept.

Second and more importantly, our results suggest that preexisting face templates shape early neural responses to objects that resemble faces, which can facilitate the acquisition of visual expertise in adulthood. That is, the adult perceptual system may automatically use face templates when learning facelike exemplars from a new object category to tune neural mechanisms for the new category leading to rapid and automatic object individuation. This conclusion is consistent with the evidence that newborn infants selectively look at nonface stimuli that have shapes arranged in a facelike spatial configuration more so than at faces with spatially scrambled facial parts (Turati et al., 2002; Morton & Johnson, 1991). Studies with adults show that different types of facelike stimuli (e.g., Arcimboldo images and face pareidolia) can elicit BOLD activity in face-sensitive brain regions (e.g., Liu et al., 2014; Davidenko et al., 2012; Churches et al., 2009; Hadjikhani et al., 2009). Critically, the neural responses to these stimuli depend on the perception of the stimulus as a face irrespective of training. It remains possible that other preexisting nonface templates (e.g., developed by dog judges or car hobbyists) may similarly facilitate the acquisition of expertise for novel objects that are physically and perceptually similar to objects in the domain of expertise. This possibility requires future research, which is beyond the scope of this study.

Conclusion

From birth (or even before; see Reid et al., 2017), human adults have had a lifetime experiencing the typical face template: "two eyes above a nose above a mouth." We show that this highly familiar spatial configuration of

parts can help adults develop perceptual processes that rapidly and automatically individuate exemplars from new object categories, which is an important step in the acquisition of visual expertise for new categories. We used the most extensive training program to date with stimuli, which differed only in orientation, with one orientation being judged as more facelike than the other. We also applied the FPVS protocol to our program, which is novel with respect to studies on visual expertise. This protocol is a powerful means to objectively quantify neural responses to rapid periodic changes occurring to different aspects of faces, such as face categories (Rossion, Torfs, Jacques, & Liu-Shuang, 2015), face identities (Liu-Shuang et al., 2014), and facial expressions (Dzhelyova, Jacques, & Rossion, 2017). The current study thus presents new insights into the acquisition of visual expertise in adulthood. The reliance on familiar templates highlights both the efficiency of the adult perceptual system to exploit familiar spatial configurations when learning new object categories and the limitations the system faces if new categories do not have a familiar configuration.

Acknowledgments

REFERENCES

Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex, 9,* 415–430.

Benton, A. L., Sivan, A. B., Hamsher, K. d. S., Varney, N. R., & Spreen, O. (1983). *Contributions to neuropsychological assessment*. New York: Oxford University Press.

Biederman, I., & Kalocsai, P. (1997). Neurocomputational bases of object and face recognition. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 352,* 1203–1219.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Brants, M., Wagemans, J., & Op de Beeck, H. P. (2011). Activation of fusiform face area by Greebles is related to face similarity but not expertise. *Journal of Cognitive Neuroscience, 23,* 3949–3958.

Cao, X., Jiang, B., Li, C., Xia, N., & Floyd, R. J. (2015). The commonality between the perceptual adaptation mechanisms involved in processing faces and nonface objects of expertise. *Neuropsychology, 29,* 715–725.

Carey, S. (1992). Becoming a face expert. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 335,* 95–103.

Churches, O., Baron-Cohen, S., & Ring, H. (2009). Seeing face-like objects: An event-related potential study. *NeuroReport, 20,* 1290–1294.

Davidenko, N., Remus, D. A., & Grill-Spector, K. (2012). Face likeness and image variability drive responses in human face selective ventral regions. *Human Brain Mapping, 33,* 2334–2349.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology, 115,* 107–117.

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia, 44,* 576–585.

Dzhelyova, M., Jacques, C., & Rossion, B. (2017). At a single glance: Fast periodic visual stimulation uncovers the spatio-temporal dynamics of brief facial expression changes in the human brain. *Cerebral Cortex, 27,* 4106–4123.

Dzhelyova, M., & Rossion, B. (2014). The effect of parametric stimulus size variation on individual face discrimination indexed by fast periodic visual stimulation. *BMC Neuroscience, 15,* 1–12.

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference evidence for a non-modular account of face processing. *Nature Neuroscience, 6,* 428–432.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3,* 191–197.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring the mechanisms for face recognition. *Vision Research, 37,* 1673–1682.

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behaviour. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 431–444.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience, 2,* 568–573.

Hadjikhani, N., Kveraga, K., Naik, P., & Ahlfors, S. P. (2009). Early (N170) activation of face-specific cortex by face-like objects. *NeuroReport, 20,* 403–407.

Harel, A. (2016). What is special about expertise? Visual expertise reveals the interactive nature of real-world object recognition. *Neuropsychologia, 83,* 88–99.

Harel, A., Gilaie-Dotan, S., Malach, R., & Shlomo, B. (2010). Top–down engagement modulates the neural expressions of visual expertise. *Cerebral Cortex, 20,* 2304–2318.

Harel, A., Kravitz, D., & Baker, C. I. (2013). Beyond perceptual expertise: Revisiting the neural substrates of expert object recognition. *Frontiers in Human Neuroscience, 7,* 885.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4,* 223–233.

Heisz, J. J., Watter, S., & Shedden, J. M. (2006). Progressive N170 habituation to unattended repeated faces. *Vision Research, 46,* 47–56.

Jacques, C., d'Arripe, O., & Rossion, B. (2007). The time course of the inversion effect during individual face discrimination. *Journal of Vision, 7,* 1–9.

Jacques, C., & Rossion, B. (2004). Concurrent processing reveals competition between visual representations of faces. *NeuroReport, 15,* 2417–2421.

Jolicoeur, P., & Milliken, B. (1989). Identification of disoriented objects: Effects of context of prior presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 200–210.

Jonas, J., Jacques, C., Liu-Shuang, J., Brissart, H., Colnat-Coulbois, S., Maillard, L., et al. (2016). A face-selective ventral occipito-temporal map of the human brain with intracerebral potentials. *Proceedings of the National Academy of Sciences, U.S.A., 113,* E4088–E4097.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17,* 4302–4311.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception, 36,* 1.

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.

Liu, J., Li, J., Feng, L., Li, L., Tian, J., & Lee, K. (2014). Seeing Jesus in toast: Neural and behavioral correlates of face pareidolia. *Cortex, 53,* 60–77.

Liu-Shuang, J., Norcia, A. M., & Rossion, B. (2014). An objective index of individual face discrimination in the right occipito-temporal cortex by means of fast periodic oddball stimulation. *Neuropsychologia, 52,* 57–72.

Liu-Shuang, J., Torfs, K., & Rossion, B. (2016). An objective electrophysiological marker of face individualisation impairment in acquired prosopagnosia with fast periodic visual stimulation. *Neuropsychologia, 83,* 100–113.

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences, 11,* 8–15.

Mondloch, C. J., Maurer, D., & Ahola, S. (2006). Becoming a face expert. *Psychological Science, 17,* 930–934.

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review, 98,* 164–181.

Op de Beeck, H. P., James, B., DiCarlo, J. J., & Kanwisher, N. G. (2006). Discrimination training alters object representations in human extrastriate cortex. *Journal of Neuroscience, 26,* 13025–13036.

Pelli, D. G. (1997). The Video Toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Regan, D. (1989). *Human brain electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine.* New York: Elsevier.

Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., & Reissland, N. (2017). The human fetus preferentially engages with facelike visual stimuli. *Current Biology, 27,* 1825–1828.

Retter, T. L., & Rossion, B. (2016). Visual adaptation provides objective electrophysiological evidence of facial identity discrimination. *Cortex, 80,* 35–50.

Rezlescu, C., Barton, J. J., Pitcher, D., & Duchaine, B. (2014). Normal acquisition of expertise with Greebles in two cases of acquired prosopagnosia. *Proceedings of the National Academy of Sciences, U.S.A., 111,* 5123–5128.

Rhodes, G., Byatt, G., Michie, P. T., & Puce, A. (2004). Is the fusiform face area specialized for faces, individuation, or expert individuation? *Journal of Cognitive Neuroscience, 16,* 189–203.

Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition, 103,* 34–79.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382–439.

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition, 21,* 139–253.

Rossion, B. (2014). Understanding individual face discrimination by means of fast periodic visual stimulation. *Experimental Brain Research, 232,* 1599–1621.

Rossion, B., & Boremanse, A. (2011). Robust sensitivity to facial identity in the right human occipito-temporal cortex as revealed by steady-state visual-evoked potentials. *Journal of Vision, 11,* 1–21.

Rossion, B., Collins, D., Goffaux, V., & Curran, T. (2007). Long-term expertise with artificial objects increases visual competition with early face categorization processes. *Journal of Cognitive Neuroscience, 19,* 543–555.

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., & Crommelinck, M. (2002). Expertise training with novel objects leads to left-lateralized facelike electrophysiological responses. *Psychological Science, 13,* 250–257.

Rossion, B., Hanseeuw, B., & Dricot, L. (2012). Defining face perception areas in the human brain: A large-scale factorial fMRI face localizer analysis. *Brain and Cognition, 79,* 138–157.

Rossion, B., Kung, C. C., & Tarr, M. J. (2004). Visual expertise with nonface objects leads to competition with the early perceptual processing of faces in the human occipitotemporal cortex. *Proceedings of the National Academy of Sciences, U.S.A., 101,* 14521–14526.

Rossion, B., Schiltz, C., Robaye, R., Pirenne, D., & Crommelinck, M. (2001). How does the brain discriminate familiar and unfamiliar faces?: A PET study of face categorical perception. *Journal of Cognitive Neuroscience, 13,* 1019–1034.

Rossion, B., Torfs, K., Jacques, C., & Liu-Shuang, J. (2015). Fast periodic presentation of natural images reveals a robust face-selective electrophysiological response in the human brain. *Journal of Vision, 15,* 1–18.

Sadr, J., & Sinha, P. (2004). Object recognition and random image structure evolution. *Cognitive Science, 28,* 259–287.

Schneider, W., & Chein, J. M. (2003). Controlled & automatic processing: Behavior, theory, and biological mechanisms. *Cognitive Science, 27,* 525–559.

Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology, 75,* 221–242.

Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain, 115,* 15–36.

Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science, 12,* 43–47.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, 46A,* 225–245.

Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level processing automatized by expertise. *Nature Neuroscience, 3,* 764–769.

Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modelling of elementary psychological processes.* Cambridge: Cambridge University Press.

Turati, C., Simion, F., Milani, I., & Umiltà, C. (2002). Newborns' preference for faces: What is crucial? *Developmental Psychology, 38,* 875–881.

Vuong, Q. C., Willenbockel, V., Zimmermann, F. G. S., Lochy, A., Laguesse, R., Dryden, A., et al. (2016). Facelikeness matters: A parametric multipart object set to understand the role of spatial configuration in visual recognition. *Visual Cognition, 24,* 406–421.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Attention, Perception, & Psychophysics, 33,* 113–120.

Wojciulik, E., Kanwisher, N., & Driver, J. (1998). Covert visual attention modulates face-specific activity in the human fusiform gyrus: fMRI study. *Journal of Neurophysiology, 79,* 1574–1578.

Wong, A. C. N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for face-like expertise with objects: Becoming a Ziggerin expert—But which type? *Psychological Science, 20,* 1108–1117.

Wong, A. C. N., Palmeri, T. J., Rogers, B. P., Gore, J. C., & Gauthier, I. (2009). Beyond shape: How you learn about objects affects how they are represented in visual cortex. *PLoS One, 2,* e8405.

Xu, Y. (2005). Revisiting the fusiform face area in visual expertise. *Cerebral Cortex, 15,* 1234–1242.