MANUSCRIPT ACCEPTED 10/14/20 AT JOURNAL OF EXPERIMENTAL PSYCHOLOGY: HUMAN PERCEPTION AND PERFORMANCE

Single-session label training alters neural competition between objects and faces

Gabriella Silva, Harold A. Rocha, Ethan Kutlu, Maeve R. Boylan, Lisa S. Scott, and Andreas Keil

University of Florida

Author Note

Gabriella Silva, Harold A. Rocha, Maeve R. Boylan, Lisa S. Scott, and Andreas Keil, Department of Psychology, University of Florida. Funding for this research was provided in part to A. Keil from the National Institute of Mental Health (R01MH112558) and to L. Scott and A. Keil from by the National Science Foundation (BCS: #1728133). Funders of this work had no role in study design; in the collection, analysis and interpretation of data; in the writing of this report; or in the decision to submit this article for publication. We thank members of UF's Center for the Study of Emotion and Attention and The UF Brain, Cognition and Development Lab and R. Barry-Anwar for relevant discussion; research assistance and technical and programming assistance.

Correspondence should be addressed to Andreas Keil, akeil@ufl.edu**;** Department of Psychology, University of Florida, 945 Center Drive, P.O. Box 112250, Gainesville, FL 32608.

Declaration of interest: none

Abstract

The extent to which visuocortical processing is altered when observers learn to categorize novel visual stimuli via labeling is not well understood. The present investigation used steady state visual evoked potential (ssVEP) frequency tagging to test the hypothesis that learning to categorize novel objects via labeling prompts a competitive advantage over concurrently presented stimuli. In the learning (label training) phase, participants (*n*=24) categorized objects according to two different species labels and faces according to gender. A control group (n=26) viewed the same stimuli without label learning. Before and after learning, faces and objects were superimposed and viewed concurrently while periodically turned on and off at unique temporal rates (5/sec or 6/sec). The spectral power of the ssVEP at each frequency was projected to an L2 (minimum) norm estimated source space, and competition between faces and objects was compared using permutation-controlled mass univariate t-tests. Results showed that, only in the training group, learning to label novel objects led to a competitive advantage over faces across a network of occipito-temporal and fronto-parietal cortical regions. These changes were more pronounced in participants showing more improvement across the label learning phase. Together, the findings support the notion that learning to label novel object categories affects neural competition though recurrent neural interactions in regions commonly associated with visual perception and selective attention.

Keywords: Attention; Visual Selective Attention; Visual Perception, Categorization, Visuocortical Processing; Electroencephalography (EEG); Steady State Visual Evoked Potentials (ssVEPs).

Significance Statement: The present study demonstrates that learning to categorize novel visual stimuli via labeling amplifies the visual representation of the novel stimuli. This neural amplification competes with neural processing of existing categories (here: male and female faces) in visual cortex, and in frontoparietal networks that are often linked to selective attention. These changes were greater in participants who showed more evidence of learning. The findings suggest that visuocortical selection is affected by learning labels, supporting theoretical notions that emphasize an active role of sensory cortex in higher-order cognition. They also highlight the potential usefulness of indices of neural competition in studies of education, training, and in interventions.

Introduction

Humans recognize and categorize familiar faces and objects quickly and efficiently (e.g., Barragan-Jason, Cauchoixa & Barbeaub, 2015; Miyakoshi, Nomura & Ohira, 2007; Nemrodov et al., 2016; Pierce et al., 2011; Thorpe et al. 1996), even when obscured or embedded in noisy, complex scenes (e.g., Burton et al., 1999). Fast and accurate visual discrimination is facilitated after objects are paired with a specific verbal label or name (e.g., Scott et al., 2006; 2008; Jones et al., 2018). However, it is unclear how the brain supports these attentional, perceptual and conceptual processes. Here, we test the hypothesis that learning to categorize novel visual stimuli via label training alters visuocortical processing. The findings reported here can be used to inform computational models of vision and have potential practical applications. For example, quantifying the locus and nature of neural changes as observers learn to discriminate objects, locations, or features can be used to monitor and potentially enhance, through neuromodulation or feedback, performance in cognitive domains including memory, spatial navigation, and recognition (e.g., Dowsett et al. 2016).

A body of work examining the acquisition of perceptual expertise used laboratory training interventions to train novices to recognize objects from natural (birds, Devillez et al., 2019; Tanaka, Curran & Sheinberg, 2005; Scott et al., 2006), human-made (cars: Scott et al., 2008) and artificial categories (Gauthier & Tarr, 1997; Gauthier et al., 1998; Jones et al., 2018). Findings from these training studies consistently demonstrated the efficiency of laboratory training for increasing visual discrimination of object exemplars across stimulus categories (Wong, Palmeri, & Gauthier, 2009; Tanaka, Curran & Sheinberg, 2005; Scott et al., 2006; 2008; Jones et al., 2018). These behavioral changes co-varied with both immediate and lasting neural changes (Scott et al., 2006; 2008). In addition, training that included labeling at specific levels of abstraction increased perceptual discrimination of within-group exemplars and led to differential occipital-temporal neural responses (Scott et al., 2006; 2008; Jones et al., 2018). For example, learning a specific subordinate-level label for each species led to increased discrimination and differential neural processing relative to learning basic-level labels (e.g., all the species within a family are “other”; Scott et al., 2006).

Based on these previous investigations, verbal or written labels have been hypothesized to differentially direct attention and perceptual processing of visual objects. This hypothesis is supported by a previously proposed framework of the development of face processing that predicts mechanistic interactions between attention and perceptual learning processes in the service of acquiring conceptual knowledge (Markant & Scott, 2018). However, it is unclear whether label learning impacts selection of a labeled stimulus among other visual stimuli, i.e., how competition is resolved at the level of visuocortical processing.

One approach for quantifying the extent to which learning to associate a label with a novel object category impacts visuocortical selection is to measure whether learning to label novel objects results in altered visuocortical competition between these newly acquired categories and existing categories. This approach has been used in perceptual expertise research and shows interference and heightened competition between faces and objects of expertise (Curby & Gauthier, 2014; Gauthier & Curby, 2005; Gauthier et al., 2003; Rossion et al., 2004; 2007; McGugin et al., 2012; 2014). For example, face selective neurons in the fusiform face area (FFA) respond to cars in car experts more so than novices (McGugin, et al., 2012; 2014). Additionally, McGugin and colleagues (2015) reported that diverted attention reduces, but does not eliminate, car expertise effects in car experts. By contrast, a task that induces competition between cars with faces, another domain of expertise, eliminates car expertise effects across much of the occipital temporal cortex. In another investigation, concurrent presentations of faces and cars reduced the N170 event-related potential (ERP) amplitude to faces in car experts, but not in novices (Rossion et al., 2007). These competition and interference effects are consistent with shared, but limited capacity, neural mechanisms. If acquiring perceptual expertise through label learning contributes to heightened neural competition between faces and objects seen in expert observers, we expect a competitive advantage for novel objects to emerge after a label learning intervention in non-experts. To address this question, we measured neural responses to concurrently presented faces and novel objects before and after a brief (i.e., 5 minutes), experimentally controlled, label training intervention. Changes in competition between the concurrently presented stimuli were quantified using frequency-tagged steady state visual evoked potentials (ssVEPs; Wieser et al., 2016; Norcia et al., 2015). To leverage the spatial information in the dense-array EEG signal, scalp potentials were projected to a source space using a minimum norm estimation algorithm (Hauk et al., 2002; 2004).

Steady-state visual evoked potentials (ssVEPs) are large-scale oscillatory brain responses to stimuli that are rapidly and periodically modulated in luminance (i.e., flickered) or in contrast (Norcia, Appelbaum, Ales, Cottereau, & Rossion, 2015; Vialatte, Maurice, Dauwels, & Cichocki, 2010). The ssVEP is extracted from scalp-recorded EEG signals in the frequency domain at the exact stimulus presentation frequency (Wieser, Miskovic, & Keil, 2016). In ssVEP frequency tagging, multiple stimuli are flickered at different frequencies (Wang et al., 2007). Frequency tagging allows for separate quantification of the degree of large-scale visuocortical engagement for each stimulus. In the present study, objects and faces were turned on and off at a rate of 5Hz and 6Hz (frequency counterbalanced) against a Brownian noise background. Frequency tagging has been previously used for quantifying competitive or synergistic interactions between visuocortical representations evoked by fully overlapping stimuli differing in task-relevance or intrinsic saliency (e.g., Wieser & Keil, 2011; Moratti & Keil, 2009; Thigpen et al., 2019). For example, using a change detection task with grating stimuli and overlapping faces or scenes, Wieser and Keil (2011) reported early increased ssVEPs to detected target gratings, followed by a reduction, whereas no temporal change was found for non-target or non-detected target stimuli. These results suggest that the population activity in visual cortex is sensitive to task-based competition dynamics. They also demonstrate that frequency tagging of concurrently presented stimuli is useful for quantifying neural competition, especially in lower-tier visual areas near the occipital pole.

In addition to being a useful measure of visual attention and neural competition, the ssVEP signal has excellent signal-to-noise ratio (SNR), making it well suited for source space analyses, which require high-SNR data. The present study leverages this favorable property of ssVEPs to analyze sources using L2 (minimum) norm estimation (MNE). By comparing source configurations across frequency tagging conditions, we can identify cortical regions in which competitive interactions between faces and objects exist, and how competition differs before and after object label training. Because ssVEP frequency tagging yields independent estimates of neural mass activity linked to multiple concurrent stimuli, it allows the precise characterization of the nature of competition between them. If a region is characterized by a fixed limited capacity, then increase of one stimulus representation will be at the cost of the competing stimulus. Other competitive interactions would be expected under a flexible capacity assumption, where competition may be resolved by selectively facilitating one stimulus while leaving the representation of the competing stimulus unaffected (Thigpen et al., 2019). Here, we use complementary analytical approaches to characterize regions in which these different types of competitive interactions are observed.

In the present investigation a label training intervention was used to promote categorization of novel objects and examine learning related visuocortical changes. It should be noted that other forms of categorization learning exist, which do not involve label learning (e.g., Smith et al., 2015), and although not tested here may or may not prompt similar changes. In the present study, it was predicted that a brief label training intervention would prompt competitive interactions in visuocortical brain networks between concurrently displayed faces and novel objects. No predictions were made regarding the laterality of these effects. These competitive interactions were expected to change with training and be primarily driven by changes in object, but not face representations. Competitive interactions were not expected for untrained faces and objects in a comparison group in which observes viewed faces and objects twice but did not complete label training for these stimuli. If supported, the presence of competitive interactions for faces and novel objects after label training would suggest that learning labels for novel object categories heightens their visuocortical representation. This heighted representation taps into limited capacity systems that also represent faces, and thus prompting increased interference and competition between the two stimuli.

Materials and Methods

Participants

Two different groups of participants were tested. One group received label training for objects and faces and the other group did not receive label training. All participants in both groups reported normal or corrected-to-normal vision, no personal or familial history of seizures, and (3 in the training group) of 50 participants included in the study reported being left-handed. All participants provided informed consent and received course credit for participation. All procedures and methods used in this study were approved by the Institutional Review Board of the University of Florida.

Label training group: Twenty-six (14 women, 12 men) undergraduate students participated in the current study. Two participants were excluded from the analysis due to equipment failure (1 participant) and participant fatigue (1 participant). The remaining 24 participants (13 female, 11 male; aged 18 to 29 years; M = 20.08, Median age = 19.5, SD = 2.36; 71% White, 21% Asian, 4.2% Native Hawaiian or other Pacific Islander, and 25% Hispanic) were included in analyses.

No-training group: In a separate sample, twenty-six (13 women, 13 men) undergraduate students participated in a similar study using the same pre- and post-training task. The same objects were used, however they were shown in full color. In this experiment, participants did learn to label the objects that were presented at pre- and post-training. No participants were excluded from this group (13 female, 13, male; aged 18 to 26 years; M=18.9, Median age =18, SD = 1.67; 65% White, 15% Black, 15% Asian and 23% Hispanic).

Stimuli

The visual stimuli used in this experiment included 70 grayscale pictures of human faces (35 male, 35 female) and 70 grayscale pictures of novel objects (Figure 1). Pictures of human faces were chosen from the Karolinska Directed Emotional Faces (KDEF) database and were forward facing portraits with neutral expressions. The novel objects, ‘Sheinbugs’ were previously used by Jones et al. (2018) and were originally created and edited using Modo© (Luxology, LLC). To minimize effects of low-level physical stimulus features, each stimulus was normalized to have the same mean luminance (64 cd/m^2) and contrast (both assessed by means of a Gossen MavoSpot luminance meter), as well as the same mean and standard deviation of grayscale values, using functions from the MATLAB image processing toolbox. All stimuli spanned a visual angle of 6.82º vertically and 7.57º horizontally, with participants' eyes at a 100 cm distance from the presentation monitor. The novel objects were divided into two categories based on their distinctive physical characteristics (including body shape and the shape/size/number/position of appendages, see Jones et al., 2018) and were labeled as Species 1 and Species 2. Thirty-five images from each category (Species 1 and Species 2) were used in the experiment.

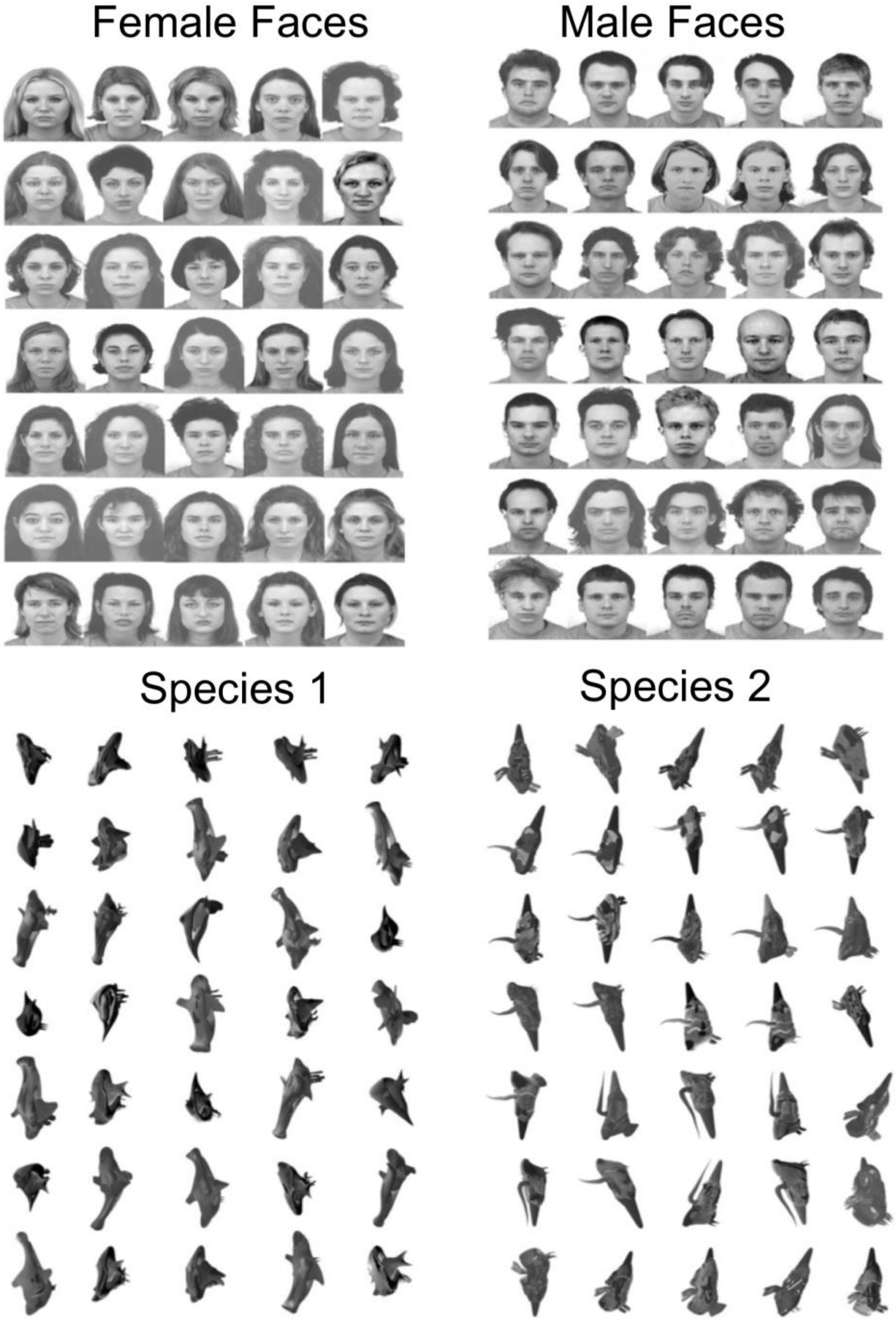

Figure 1. Stimuli included 70 neutral male and female faces from the Karolinska Directed Emotional Faces database and 70 novel object images called Sheinbugs (Jones et al., 2018) divided into two subset “Species.”

Experimental Procedure

Label training group. The experimental task lasted approximately 35 minutes and consisted of a pre-training (baseline) phase, training phase, and post-training (test) phase. Each participant was seated in an electrically shielded and dimly lit room and a 129 channel geodesic EEG recording net (Philips EGI, OR, USA) was placed on the head, used to record neural activity. Participants viewed all stimuli on a 23” 3D-LED (Samsung LS23A950, refresh rate of 120 Hz) monitor located 100 cm away from the participant’s eyes. Stimulus presentation and response registration was controlled by MATLAB code using functions from the Psychtoolbox suite of programs (Brainard, 1997). The baseline phase and test phase were identical in design, using the same 140 images, each shown once, and with equal numbers of male/female faces superimposed with concurrently presented Sheinbugs of each species. The order of stimuli presented, as well as the assignment of faces to be paired with Sheinbugs was randomized, with the constraint that equal proportions of male and female faces were assigned to be shown concurrently with species 1 and species 2 exemplars. Thus, participants passively viewed different superpositions of concurrent stimuli.

During the baseline and test phases, the faces and novel objects (Sheinbugs) were concurrently presented, fully spatially overlapping with each other, and rapidly contrast-modulated. To implement ssVEP frequency-tagging, we used two different temporal rates, 5 Hz and 6 Hz, one used for faces and one for objects, with the tagging frequency corresponding to faces versus objects counterbalanced across participants. This type of stimulation primarily drives luminance-sensitive neurons in occipital cortex, but also engages higher-order cortices sensitive to stimulus content (Norcia et al., 2015). To minimize effects of stimulus edges or systematic shape differences between faces and Sheinbugs on the ssVEP, both stimuli periodically emerged at their respective tagging frequency from a same-size Brownian noise (spatial noise with a 1/f^2 characteristic) patch with the same mean luminance and contrast as the experimental stimuli. The noise patch was randomly generated in each trial, to ensure the absence of edges and spatial concentration of high contrast in specific locations of the visual field. Thus, this approach reduces the presence of visuocortical signals in the ssVEP that are not specific to the stimulus of interest. To achieve this, each stimulus was demeaned and the residual gray values (with their sign) added to the grayscale values of the Brownian patch. Phenomenologically, this creates an impression of faces and objects "appearing within the noise".

Each trial (see Figure 2 for a timeline) began with the presentation of a gray screen for an inter-trial interval randomly (drawn from a rectangular distribution) varying between 3 and 5 seconds. Then the Brownian noise patch generated for this trial was presented alone, for the duration of 1 second, to minimize any transient brain responses to a compound stimulus onset (noise, face, and object) which would contaminate the ssVEP segment. Subsequently, one face and one object were added periodically to the Brownian noise, each at its specific tagging frequency, for a duration of 6000 ms. The overall trial structure and trial number was identical in for the baseline and test phase of the experiment, but both phases contained different, randomized sequences of random face-object pairs, to minimize any memory or repetition effects.

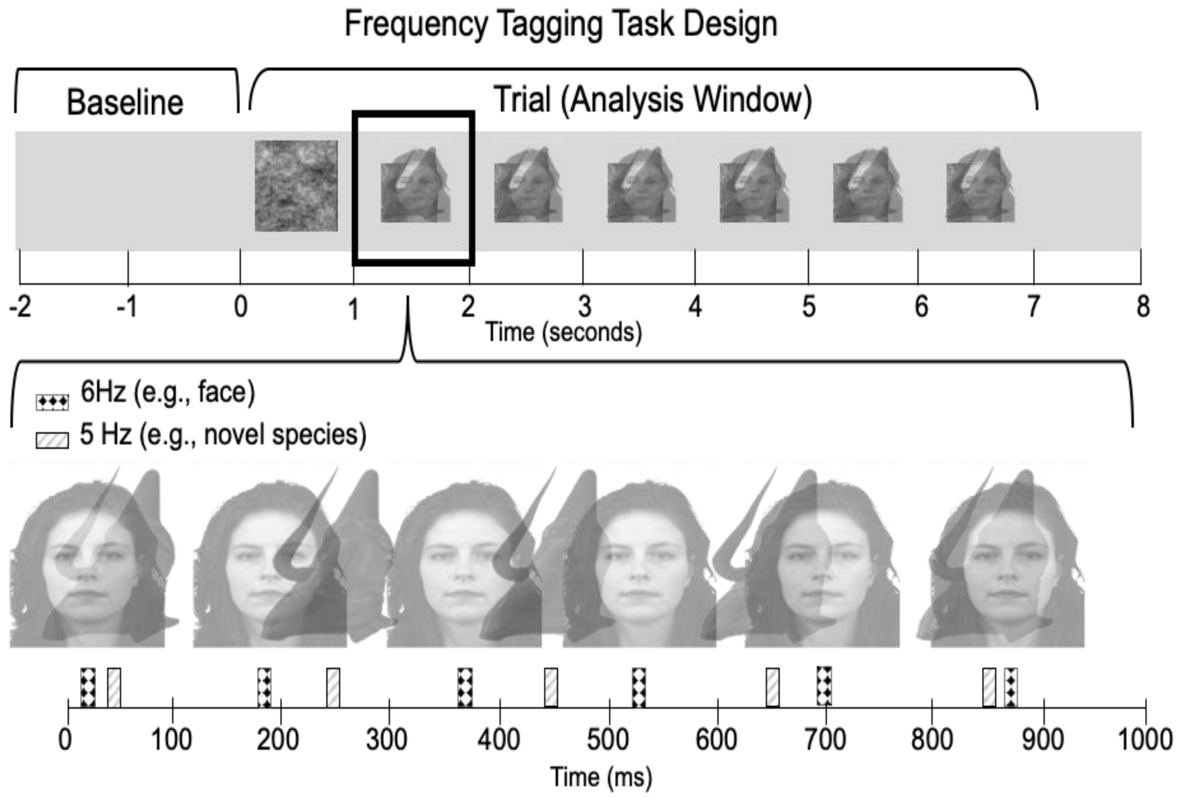

Figure 2. Stimuli concurrently emerged from a Brownian noise (spatial noise with a $1/f^2$ characteristic) background with the same mean luminance and contrast as the experimental stimuli. Each trial was 6 seconds long. Presentation Frequency was counterbalanced between subjects, who passively viewed the stimuli, while maintaining fixation.

Throughout, participants were encouraged to remain focused on the center of the screen, to restrict eye blinking to times when visual stimuli were not being presented, and to keep their head and extremity movements to a minimum, while viewing the stimuli. No other task was required during the baseline and test phases.

At the beginning of the training phase, participants were serially presented with five examples (freely randomized order) from each species that were not part of the experimental stimulus set (total of 10 exemplars) and given explanations of how they differed in terms of key features such as overall shape, location and size of appendages, pattern, etc. Then, six practice trials were completed in which the participant discriminated Sheinbugs of both species. Explanation and additional training trials were given as needed. In the subsequent main portion of the training, participants were asked to complete a discrimination task that included all 70 human faces (35 male; 35 female) and all 70 Sheinbugs (35 species 1; 35 species 2). During the task, each image appeared in the center of the screen with two written answer choices on each side of the screen. If a face image was presented, the response options were “male” and “female.” If a Sheinbug image was presented, the response options were “species 1” and “species 2.” Participants selected a response option by clicking on it with a standard computer mouse. This setup was used instead of a two-key choice response to allow future studies to use more than two species. Participants were instructed to emphasize accuracy over speed. If the question was answered incorrectly, a non-aversive sound (i.e., “beep”) was played for incorrect responses. For each stimulus, the selection latency (from picture onset to mouse click) and accuracy were recorded.

No-training group. The no-training group completed the same task as the label training group, with two differences. First, there was no label training intervention between the first and second session; second, the objects and faces were shown in color instead of grayscale. We include this non-training group as a reference point for effects related to repeated viewing of the same stimulus material in two sessions, without label training, and provide a non-parametric comparison of the effects seen in both groups.

EEG Recording. EEG was continuously recorded using an Electrical Geodesic (EGI) HCGSN 129 channel high-density sensor net. Electrode impedances were kept below 60 kΩ and the vertex sensor (Cz) was used as the recording reference. All channels were digitized at a rate of 500 Hz and were filtered online using Butterworth filters with 3dB points set at 0.1 Hz (high-pass) and 50 Hz (low-pass). All additional data processing was completed offline.

EEG Data Processing and Analyses. Preprocessing and artifact handling were conducted using the ElectroMagnetoEncephalography (EMEGS) toolbox for MATLAB (Peyk, De Cesarei, & Junghofer, 2011), version 2.8. In these analyses, the EEG was first digitally filtered to eliminate non-ssVEP signals, by means of a 20-Hz low-pass (12th order Butterworth) and a 1-Hz high-pass (4th order Butterworth) filters. Stimulus-locked segments were extracted from continuous EEG data, containing 400 ms pre- and 7400 ms post-stimulus onset for both the pre and post training phases. Channels and trials for data that contained artifacts were then identified using the distribution of three statistical parameters (Junghöfer et al., 2000): the absolute value of the voltage, the maximum standard deviation of the voltage, and the maximum temporal differential. These were calculated across channels as well as across trials and resulted in data quality histograms, the tails of which (defined as 3 standard deviations above the median) were discarded (trials), or interpolated from the full channel set (channels). The original recording reference (Cz) was first used to detect channels that were contaminated, followed by detecting bad channels per trial. The maximum of interpolated channels was 12, with the mean across participants being 8.0 interpolated channels. A trial was rejected if interpolated channels clustered in the same scalp region. To this end, the emegs algorithm compares a forward model with the full channel set to the channel set without bad channels and discards a trial if the discrepancy between the two forward solutions exceeds a critical value. Next, the data were arithmetically re-referenced to the common average, and trial-specific artifacts were detected using the same three statistical metrics, calculated for each trial-channel combination, and bad trials rejected. Data were then referenced to the average reference. After artifact rejection, an average of 76% (range: 59% to 91%) of the trials in the baseline, and 70% (range: 55% to 88%) of the trials in the test phase were retained. Time domain averages across artifact-free trials were submitted to source space estimation. Data and code for this study is freely available on the Open Science Framework (https://osf.io/a53s9/?view_only=1966f70fac954bac886381f908c7a275).

Source Space Projection. To address the central hypothesis of the study, namely that competition between objects and faces changes in specific brain regions after a brief labeling intervention, we conducted source estimation using the L2 (minimum) norm method. The minimum norm estimation (MNE) approach leverages the dense sensor montage, using a simple and conservative inverse projection onto a source space that in the present implementation consisted of 4 concentric spherical shells approximating the brain volume, at radii corresponding to the brain surface, the cortex, and two deep shells modeling any contributions from deep sources. Using the algorithm described in Hauk et al., (2002) and Hauk (2004), we first calculated a leadfield matrix, mapping the weights for each of 655 model dipoles equidistantly located on the 4 spheres relative to each sensor of the 129-channel montage, in three directions, 1 radial (perpendicular) direction, and 2 tangential directions relative to the surface of each spherical shell. Then the inverse of the leadfield was estimated using Tikhonov-Philips regularization and multiplied to the data to yield MNE time series reflecting source densities for each model dipole and orientation/direction, on each shell. The regularization parameter lambda used for Tikhonov-Philips regularization (0.021) was determined for the group-level data by plotting spatial variance and goodness of fit (residual variance) across all trials and participants, and finding the turning point, i.e., the L-curve method. Because the scalp EEG primarily reflects radial source activity, the tangential information was discarded, and the MNE time series representing the radial orientation from the outer shell (350 dipoles) were submitted to discrete Fourier transform (see below). For visualization, power values, differences, and statistical parameters were mapped onto a standard brain, which was normalized to the same coordinates, using the fiducial positions (nasion, inion, left and right periauricular points) co-registered with the standard sensor locations of the 129-channel net, by means of MATLAB functions implemented in fieldtrip software.

Frequency Domain Analyses. Only the ssVEP period of each epoch was used for the frequency-domain analyses, and only fundamental frequencies (the driving frequencies of 5 Hz and 6 Hz) were analyzed. An exploration of other frequencies of potential interest such as higher harmonics or so-called intermodulation frequencies (see e.g., Boremanse, Norcia, and Rossion, 2013) showed that these signals were relatively small in the present study, showed a different topography than the fundamental frequencies, and were therefore not included in the present analysis. A cosine-square window with a 50 ms taper was applied to the beginning and end of this time segment. After multiplication with the window function, the data were submitted to Discrete Fourier Transform, and power obtained as the norm of the real and imaginary Fourier coefficients, normalized by the number of sample points entering the Fourier transform. The resulting normalized spectrum had a resolution of 0.166 Hz, reflecting the length of the ssVEP segment, i.e., 6000 ms. The power at the tagging frequency for each stimulus was then extracted for each dipole, and the respective power map stored according to the frequency tagging condition (i.e., the frequency for the face and object; counterbalanced across participants), separately for both phases of the experimental session. Thus, for pre- and for post-training, two topographical maps were obtained, one representing the response evoked by the face and one the response evoked by the object, for each participant and dipole location. These maps entered the permutation controlled statistical comparisons described in the next paragraph. In addition, to assess data quality, we computed the signal-to-noise ratio (SNR) for each participant and tagging frequency, by dividing the spectral power at each tagging frequency by the mean power measured at 6 adjacent frequencies (3 higher, 3 lower), leaving out the two immediately neighboring frequencies respectively.

Sample Size Estimation and Statistical Analyses. Traditional power analysis is limited in repeated-measures designs for which the inter-variable correlations are not known a-priori. We therefore used a simulation based approach to determine trial numbers and sample sizes. Using trials from an existing frequency tagging data set (Wieser & Keil, 2011), we determined the trial count needed to obtain a signal to noise ratio > 6 dB in each participant, considered satisfactory. We also calculated the sample size necessary to observe a phase (before vs after training) by stimulus interaction in a repeated measures ANOVA design by randomly selecting trials and participants, while changing the effect size of the interaction effect. Using this method, we determined that 46 trials were needed per condition. To accommodate artifacts, 70 trials were run per condition in the present experiment. Assuming 46 good trials per subjects, the simulations showed that to detect a difference between stimuli across blocks corresponding to a partial omega square of 0.1 (a medium sized effect), 16 participants are needed. When a permutation controlled t-test model applied to each dipole was used instead of ANOVA, 20 participants were needed to robustly detect the same effect size. Following extant recommendations for powering studies, we ran 26 participants in each group, to accommodate any data loss.

The overall analytical strategy was to first conduct the same analyses in the two groups (label training group; no-training group). We then compared the effects seen in the no-training group with the effects seen in the label training group by means of non-parametric Mann-Whitney U tests, comparing the competition maps for each dipole location as described below.

This approach accounts for the fact that both groups viewed slightly different stimuli but also allows us to assess the specificity of any effects, anywhere in the source space, for the label training group. Throughout this report, mass univariate permutation control is used to address alpha error accumulation in map comparisons, which involve many dipole locations.

In order to determine the effectiveness of the brief labeling intervention, changes in accuracy (percent correct) and selection response time (RT) were examined during the training phase for face and object trials. Here, RT was defined as the duration between the onset of each image and the choice selection made by each participant during the training phase. As discussed above, selection responses were executed using a mouse operation, thus adding variability to the selection response latency. RT measures are therefore reported in this study, but cannot be interpreted as choice RTs with standard response key setups. Averages comparing the first and second half of training were computed for each participant. The first half included the first 70 trials of training and the second half included the last 70 trials of training. 2 (object, face) x 2 (1st half, 2nd half) ANOVAs were computed separately for percent correct values and selection time. Follow-up analyses were corrected for multiple comparisons using the Bonferroni method.

To directly assess the hypothesis that a brief training intervention changes the competitive interactions between concurrent visuo-cortical representations of novel objects (Sheinbugs) and faces, we conducted permutation-controlled mass univariate tests at each model dipole, separately for the label training group and the non-training group. To this end, we quantified the difference between objects and faces before and after training, by subtracting the spectral power at the face tag from the power measured at the object tag, separately for pre-training and post-training. The resulting difference maps were then evaluated by t-tests comparing pre-intervention and post-intervention. The critical *t*-values (corresponding to a significance level of .05) were obtained by calculating distributions of t-values on data shuffled between the conditions, within each participant (8000 permutations): Specifically, based on the procedure proposed by Blair and Karniski (1993), we randomly permuted the pre-intervention and post-intervention condition within each participant 1,000 times and computed *t*-values for each of the 350 model dipoles. Then, the tails of each permutation's *t*-value distribution were determined and stored in a $t_{\min}$ and $t_{\max}$ distribution, respectively, each having 1,000 values corresponding to the 1,000 permutations. Finally, 2.5th and 97.5th percentiles from the resulting distribution of $t_{\min}$ and $t_{\max}$ values were used as critical *t*-values, and only empirical *t*-values crossing that threshold were considered statistically significant.

To further characterize the nature of the differences that led to the permutation-controlled effects in the difference-of-difference maps, we used two follow-up strategies. First, we conducted post-hoc t-tests comparing pre- and post-training amplitudes for objects and faces separately, for each location identified, to explicitly test the hypothesis that changes in the difference-of-differences map reflected pre-to-post changes in object representations, but not face representations. Second, we calculated a simple index that reflects the trade-off aspect of competition. Trade-off competition effects in the present study would be present if a given brain region shows post-training ssVEP power <u>increase</u> for objects that co-occurs with post-training power <u>decrease</u> for faces. To assess this specific type of competition, which implies limited capacity, we calculated a competition index (Boylan et al., 2019), based on the product of the amplitude changes across two blocks for each stimulus type:

$$(\text{Object ssVEP}_{\text{post-training}} - \text{Object ssVEP}_{\text{pre-training}}) * (\text{Face ssVEP}_{\text{post-training}} - \text{Face ssVEP}_{\text{pre-training}})$$

This index is negative if object-evoked power increases from pre- to post-training and there is concurrent power reduction in the face-evoked ssVEP, or vice versa (trade-off). The index will be positive if both stimuli display amplitude reduction or amplitude enhancement from pre-training to post-training. For both follow-up strategies, the analysis was conducted at each dipole, and the resulting maps were thresholded using the same permutation-controlled t-tests as discussed above. We also used the competition index maps to conduct an exploratory analysis, comparing competition between the two groups (label training versus no training), calculating Wilcoxon Rank-Sum tests for each dipole location in the MNE source volume. Here, a large positive Wilcoxon z-score would suggest that competition indices were consistently higher for the training group in a given region, whereas larger negative z-scores would indicate greater competition indices for the no-training (control) group.

In order to examine the extent to which neural competition was associated with discrimination accuracy, rank correlations (Spearman's Rho) were computed across participants, for each dipole location. To this end, the relation between the change in neural competition directed by the novel object from before to after training (i.e., the difference-of-difference maps as described above) and the change in object species discrimination accuracy from the first half to the second half of training was examined. The resulting rank correlations maps were again thresholded with the same permutation approach (i.e., $\mathrm{Rho}_{\mathrm{max}}$ distribution based on 1000 random permutations of condition labels within subject).

Results

Response time (RT) and accuracy

For analyses of accuracy, there were significant main effects of stimulus type, $F(1, 23) = 13.486$, $p = .001$, $\eta_p^2 = .370$, and half of training $F(1, 23) = 7.709$, $p = .011$, $\eta_p^2 = .251$, as well as a significant interaction between stimulus type and half of training $F(1, 23) = 5.398$, $p = .029$, $\eta_p^2 = .190$. The main effect of stimulus type was due to greater accuracy for faces than objects (see

Table 1). The main effect of half was due to increased accuracy in the second half of the intervention compared to the first half for species. The interaction between stimulus type and half of training was driven by increased accuracy identifying objects in the second half of trials compared to the first half, $t(23) = -2.923$, $p = .008$, (Bonferroni corrected $p<.05$; see Table 1). There was no significant accuracy difference between the first half and second half of trials for faces. Accuracy also was significantly greater for faces in the first half when compared with objects $t(23) = 4.190$, $p < .001$ (Bonferroni corrected $p<.005$), but no significant difference was found between faces and objects in the second half of trials.

Analyses of RT showed main effects of stimulus type, $F(1, 23) = 31.009$, $p < .000$, $\eta_p^2 = .574$, and half of training, $F(1, 23) = 37.778$, $p = .000$, $\eta_p^2 = .622$. The main effect of stimulus type was due to faster RT to faces than objects (see Table 1 for Means and SE). The main effect of half of training was due to faster RTs during the second half relative to the first half of trials. There were no significant interactions for RT. RT was significantly less (participants responded faster) for faces than objects in the first half $t(23) = -4.392$, $p < .001$(Bonferroni corrected $p<.005$), and in the second half of trials $t(23) = -2.600$, $p = .016$ (Bonferroni corrected $p<.05$).

Table 1
Percent correct and average response times during training (Mean +/-SEM).

	Percent Correct		*Average Response Time (ms)*	
Stimulus	1st Half	2nd Half	1st Half	2nd Half
Faces	98.452 ± 0.45	98.571 ± 0.42	1603.490 ± 92.33	1326.074 ± 95.93
Objects	93.691 ± 1.17	96.786 ± 0.97	2153.256 ± 167.78	1568.344 ± 110.18

Steady-state visual evoked potentials: Label training group

The frequency-tagged stimulus array prompted robust ssVEP signals in all participants, with SNRs for each participant and experimental condition exceeding 3.2 dB. Figure 3A shows the mean spectrum for the pre-intervention session in the label training group, averaged across 12 participants in the same tagging group, to illustrate the satisfactory data quality attained with this paradigm, and with the present source estimation method. Power for both stimulus types, pre- and post-intervention, displayed a strong occipital maximum, as illustrated in Figure 3B. The raw difference maps (objects minus faces) are shown in Figure 3C, evincing overall greater responses to faces compared to objects at mid-occipital locations, but greater responses to novel objects at anterior sites.

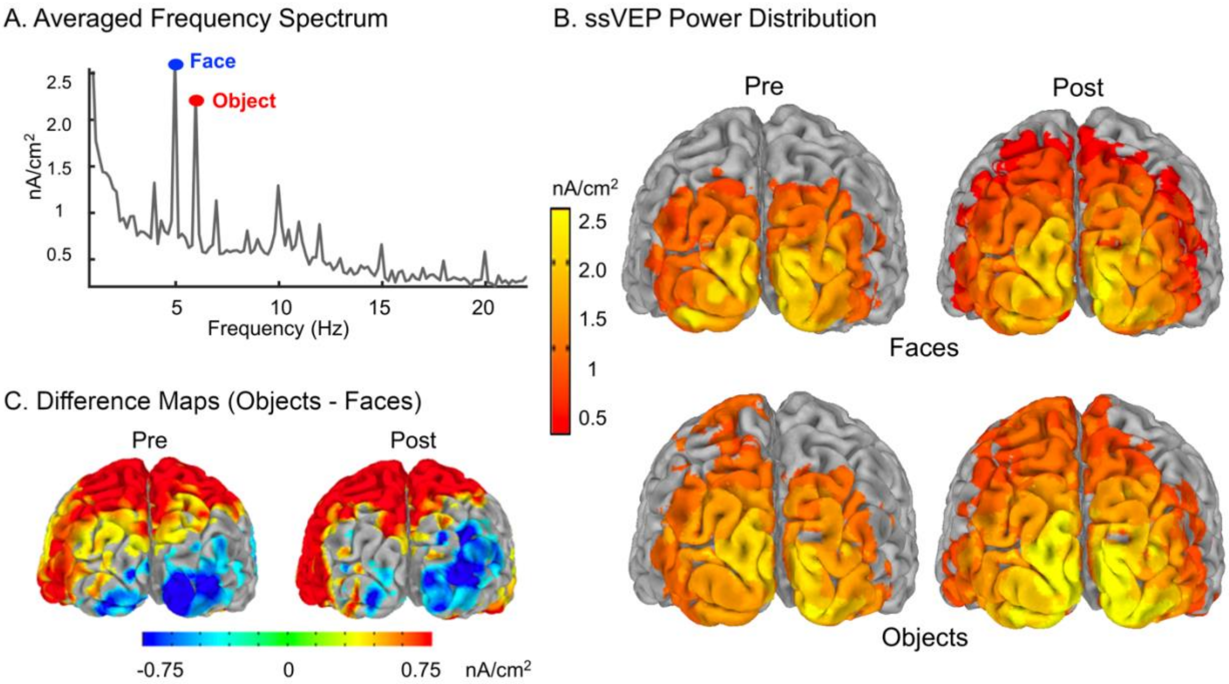

Figure 3. A. Averaged frequency spectrum from a model dipole located at the occipital pole, during the pre-intervention session in one of two tagging groups (n=12), where faces were tagged at 5 Hz, and objects tagged at 6 Hz. B. Grand mean (n=24) ssVEP power in response to both faces and objects during the pre- and post- intervention sessions. C. Grand mean (n=24) difference maps illustrating object minus face power differences in source space.

When comparing post- and pre-intervention competition maps using paired t-tests, we found that a total of 19 dipole locations out of 350 exceeded the conservative permutation-based threshold of $t_{crit} = 3.17$ (corresponding to $p < .05$) for the post > pre comparison of the competition map defined as the difference of $ssVEP_{objects}$ minus $ssVEP_{faces}$ (Figure 4). A control analysis with three left-handed individuals excluded showed the same pattern of results.

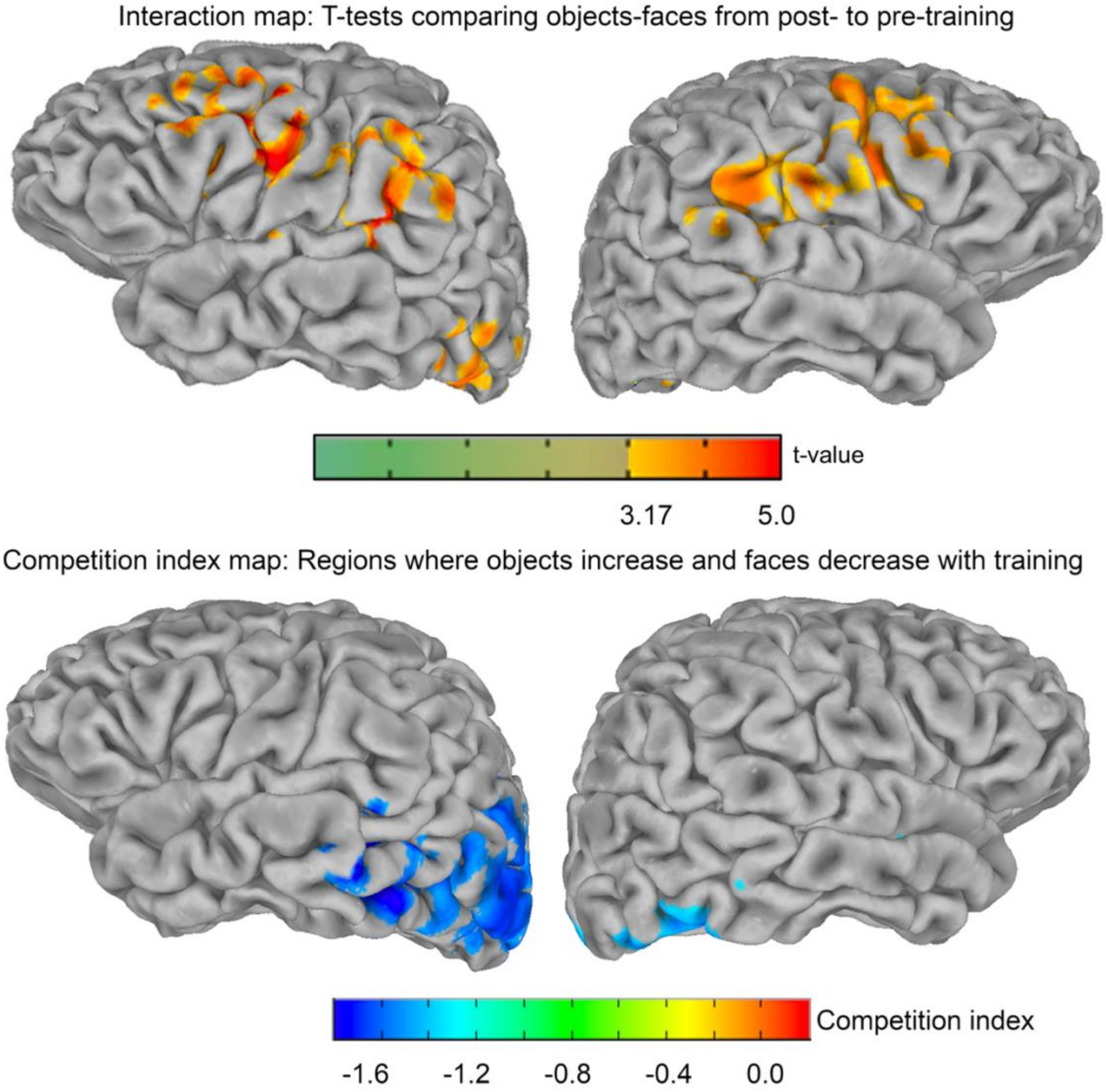

Figure 4. Top panel: Dipole locations exceeding the permutation-based threshold when comparing post- and pre-intervention competition maps. Bottom panel: Dipole locations in

which the unitless competition indices were different from zero (permutation controlled t-test against 0). The competition index is the product of the post-training minus pre-training differences for each stimulus, thus rendering negative numbers when one response (i.e. object) increases, and the other decreases. It is positive when both decrease or increase with training. Only negative values were observed in the present data set, and all were related to increase in object-evoked ssVEPs.

Table 2 shows all significant regions and the distributions of ssVEP power values in source space. No dipole location exceeded the threshold for the difference in the opposite direction, i.e. greater competitive advantage for the faces compared to the objects after training compared to before training. No dipole location showed a significant ssVEP amplitude increase from pre- to post-training for faces (all ts<2.41). Out of 19 dipole location showing significant changes in the object minus face difference from pre- to post-training, 16 showed significant amplitude enhancement (from pre- to post-training) in object-evoked ssVEPs, with 3 supramarginal gyrus dipole locations not reaching the permutation-controlled threshold.

Table 2

Competitive Dipole Locations

		Hemisphere	*Focus point MNI coordinates*			*Statistic (dipoles)*
Cortical region	Brodmann's area		x	y	z	T
Inferior occipital gyrus	18	L	-22	-82	-5	3.46 (2)
Middle occipital gyrus	19, 39	L	-43	-76	14	3.23 (2)
Inferior temporal gyrus	37, 19	L	-51	-66	-8	3.55 (1)
Supramarginal gyrus	40, 1, 3, 2	R	67	-18	26	3.21 (2)
	40, 39	L	-52	-37	29	3.66 (2)
Postcentral gyrus	4, 3b, 1	R	64	-6	19	3.33 (1)
Precentral gyrus	6, 44	R	58	6	23	3.40 (1)
	6	L	-59	5	35	3.47 (2)
Middle frontal gyrus	10, 9	R	52	40	22	3.33 (2)
	9, 8	L	-47	26	36	4.01 (4)

Note. Dipole locations with significant changes in competition indices (ssVEP differences sheinbug minus faces), from pre- to post-intervention. Focus points (where the largest t-values was observed) within each region are given in MNI coordinates, along with the number of dipoles crossing the threshold of t=3.17 in that region.

Dipoles from the left inferior occipital gyrus and left middle frontal gyrus dipoles are shown in Figure 5. Consistent with the competition maps, and the results reported above, they illustrate a selective increase in ssVEP power for objects, compared to faces, after the intervention, with 18 participants showing interaction effects consistent with selective facilitation of object representations after training, compared to pre-training. When comparing the two tagging groups (face tagged at 5 Hz, face tagged at 6 Hz), neither significant main effects of group nor group by condition interactions emerged (all Fs < 3).

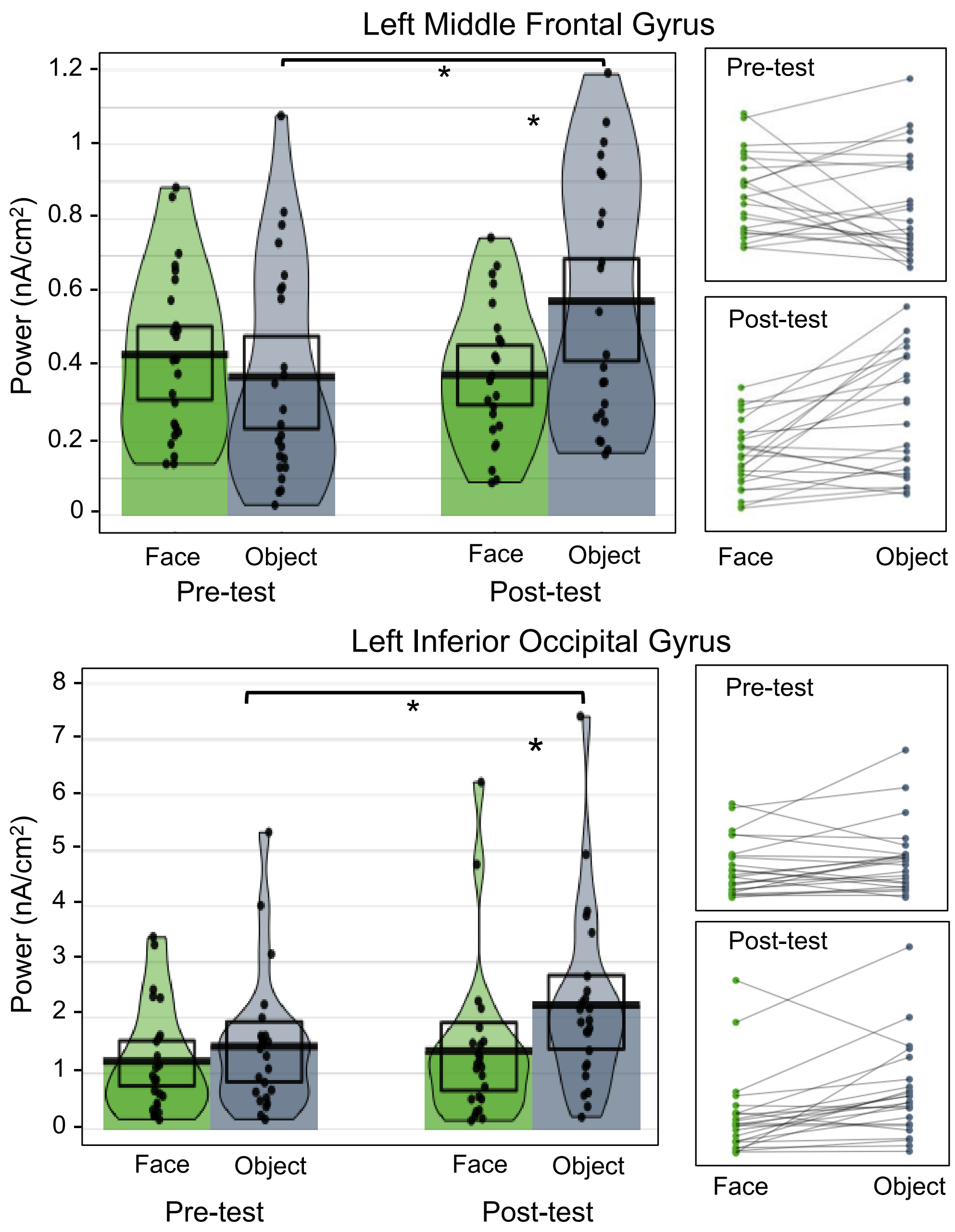

Figure 5. Illustration of data from two regions (Left Middle Frontal Gyrus (top); Left Inferior Occipital Gyrus (bottom) showing a selective increase in ssVEP power (in source space) for objects, compared to faces from pre-intervention to post-intervention. The bar plots include the mean (dark line) and 95% CI (box) as well as individual data points. Individual power differences between faces and objects are shown in the boxes on the right at pre-test and post-test

for each region. Note that these plots are given to further illustrate the results of the topographical analysis in Figure 4, but do not represent a separate or additional statistical test.

Steady-state visual evoked potentials: No-training group

The competition analyses described above were repeated in the non-training group, which also showed satisfactory SNRs and clear occipital maxima in the MNE projection across all experimental conditions (See Figure 6A). Likewise, the overall pattern visible in the competition maps (Figure 6B) paralleled those from the label training group in that they showed an occipital dominance of face-evoked responses over object-evoked responses (blue hues), and greater object-evoked responses (red hues) at lateral sites. In contrast to the label training group however, when comparing post- and pre-intervention difference maps using paired t-tests, none of the t-tests crossed the permutation controlled critical value of 3.13. The distribution of all t-values is shown in Figure 6C. Session did not affect competition between objects and faces in the non-training group. An exploratory comparison of competition index maps for the label training group and the no-training group (Figure 6D) showed that the competition in the training group was greater than in the no-training group, at 5 left occipital dipole sites, where competition was most pronounced for the training group. At these 5 contiguous locations, z-scores exceeded the threshold of 2.90, determined by permuting group labels, as described above.

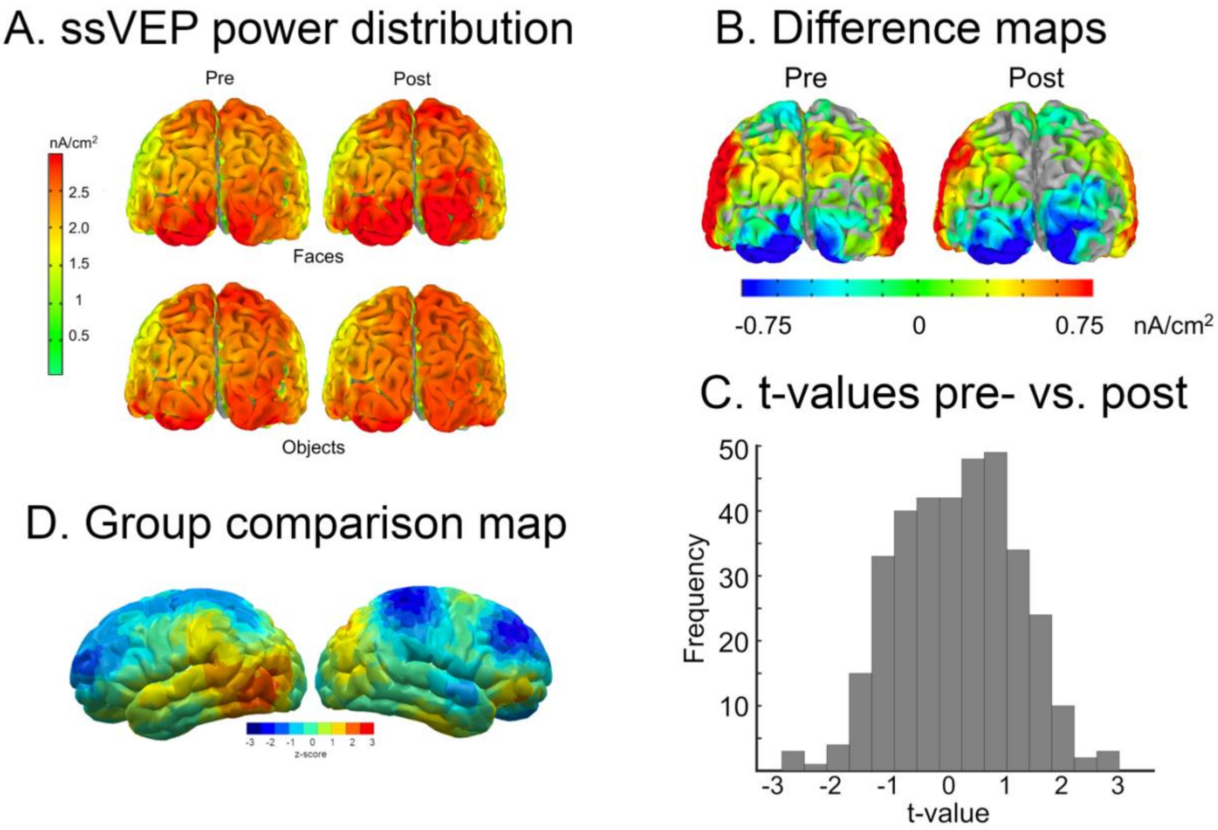

Figure 6. Data from the no-training group. A. Grand mean (n=26) ssVEP power projected to MNE space in response to both faces and objects during the pre-break and post-break sessions. B. Grand mean (n=26) competition map illustrating object minus face power difference in source space, again shown for the sessions before and after the break. C. Histogram of t-values from the competition analysis, which examined changes in competition (objects minus faces) from pre- to post-break sessions. No location in the 350-source space exceeded the threshold determined by permutation (3.13) D. Exploratory comparison of competition indices in the label-training group and in the no-training group, using Wilcoxon Rank-Sum tests at each dipole location.

Neural Competition and Behavior: Label training group

To quantify the extent to which altered competition from before to after training was related to differences in discrimination accuracy, we related the competition change maps to the change in accuracy for object species discrimination, recorded during the training phase. Using Spearman's Rho, a strong positive correlation would indicate that participants with high ranks in object discrimination improvement also had high ranks in terms of increased neural competition towards objects, and vice versa. The permutation-controlled .975 quantile of the $\mathrm{Rho}_{\mathrm{max}}$ distribution was .40, used as a threshold for the resulting correlation map, shown in Figure 8. Three clusters of dipoles crossed this threshold, 6 dipoles in left occipital cortex, 5 dipoles in right parietal cortex, and 3 dipoles in left precentral gyrus. No dipoles showed a negative relationship between change in accuracy and change in competition.

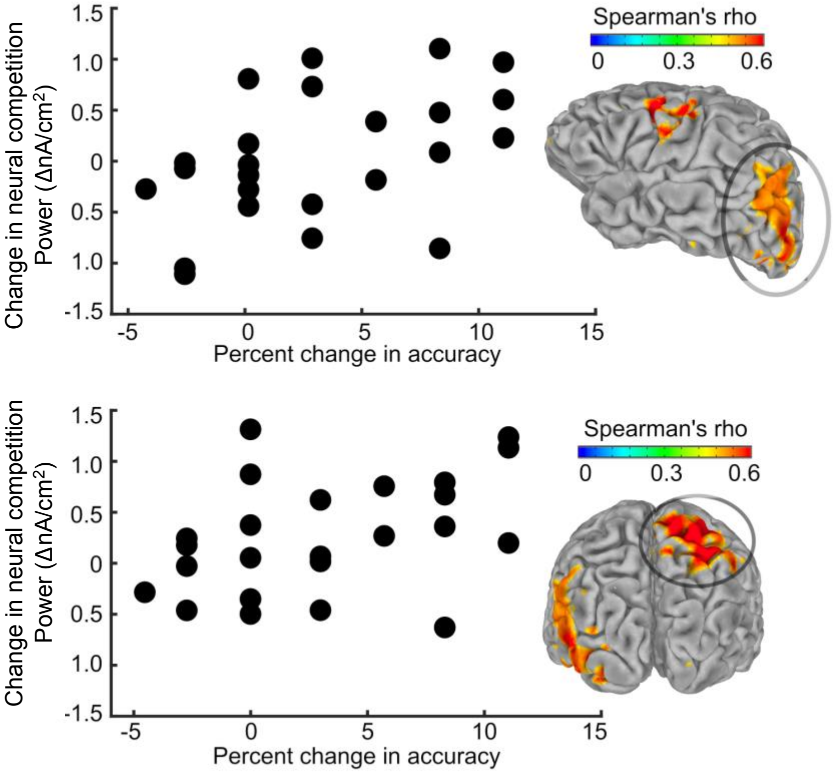

Figure 7. Rank correlation map between change in neural competition towards objects from before to after training and discrimination accuracy in the label training group. Top panel: Left view of the correlation map in the estimated source space, projected to a standard brain. A scatter plot of the change in competition in left occipital dipoles (circled) and change in detection accuracy during training is shown. Bottom panel: Back view of the correlation map in the estimated source space, projected to a standard brain. A scatter plot of the change in competition in right parietal (circled) and change in detection accuracy during training is shown.

Discussion

The primary aim of the current investigation was to determine the extent to which neural competition between objects and faces changes after an object labeling intervention in which adults learned to categorize exemplars of novel objects into two different species. Results suggest that object label training prompted a competitive advantage for exemplars from newly learned object categories over faces. Competition, assessed as the difference of the ssVEP power values evoked by concurrently viewed objects and faces, was altered selectively in left occipital, as well as left and right fronto-parietal cortical areas. When operationalizing competition more narrowly as occupation of a limited capacity, we found strong trade-off effects between novel objects and faces, only in occipital cortices. In these areas, most pronounced in the left hemisphere, labeling-related increases in object-evoked responses were associated with decreases in face-evoked responses. A comparison group in which observers viewed the same type of display without intervening label training did not show these changes in large-scale neural competition. Thus, learning to categorize novel objects via labeling prompts changes in large-scale competition. Future research is necessary to determine the extent to which categorization in the absence of labeling is sufficient to prompt changes in neural competition.

Behavioral performance (accuracy and selection time) from the first half of training was compared to the second half of training in order to determine whether the brief label training intervention impacted behavior (see Table 1 for results). As expected, accuracy as measured by percent correct responses was greater for categorizing gender compared to the novel objects. Accuracy also improved over the course of the training session, reflected in an increase in mean accuracy from the first to the second half of trials. No such increase was found for faces, for which performance was at ceiling. Overall, discrimination accuracy results suggest that participants in this study showed rapid learning from pre-intervention to post-intervention for categorizing the novel objects, with many observers being near-ceiling within the first half of training. We also related these behavioral changes to neural data: Left occipital and right parietal areas showed greater competitive advantages for objects after training in those with greater discrimination performance improvement, further supporting the notion that competitive advantage at the neural level accompanies successful label learning.

Frequency-tagging of overlapping faces and novel objects presented against Brownian noise evoked robust and reliable ssVEP signals associated with each specific stimulus category (face, object). The purpose of the Brownian noise used in this study was to minimize visuocortical engagement that is not specific to the faces and objects respectively. The presence of Brownian noise minimizes the effects of low level features such as edges, luminance, and non-specific contrast and emphasizes processing object and face-specific information. The acquisition of robust ssVEP signals allowed us to quantify the competition between the fully overlapping faces and objects. After a short instruction and label training intervention, cortical representations of novel objects, but not faces, significantly increased across a distributed network of neural regions including the left inferior and middle occipital gyri, and inferior temporal gyrus as well as bilateral supramarginal, postcentral, and middle frontal gyri, and right postcentral gyrus. In occipital cortex, these interactions took the form of competition for a fixed limited capacity: Pre- to post-training increase in ssVEP power for objects was associated with a decrease in ssVEP power evoked by faces. By contrast, frontal and parietal regions showed a selective increase of object-evoked ssVEP signals that was not at the expense of the concurrent face stimulus. This pattern of findings suggests that label learning facilitates the representation of labeled stimulus in lower tier visual cortex, where receptive fields are small and overlapping stimuli compete for limited capacity. By contrast, higher-order, anterior cortices do not show this bottleneck for spatially overlapping stimuli. Previous work has identified heightened competitive advantages of a stimulus in visual regions as a strong correlate of attentive selection, based on conceptual (Reynolds & Heeger, 2009) and empirical grounds (Andersen, Müller, & Hillyard, 2009; Wieser & Keil, 2011). Anterior brain regions showing selective changes in the present study have been linked with working memory, object categorization, motor and language processing, and selective attention (Gazzaley & Nobre, 2015). Specifically, frontal networks, including the prefrontal cortex, may be biasing downstream sensory regions in a top-down manner (e.g., Miller & Cohen, 2001; D'Esposito & Postle, 2015). Top-down modulation of face specific neural activity is consistent with previous FFA activity and N170 ERP responses (Gazzaley et al., 2005). Furthermore, the results presented here converge with work in other areas of associative learning, most notably Pavlovian conditioning (Gruss, Langee, & Keil, 2016). Specifically, conditioning-induced changes in competition have been shown to prompt re-entrant fronto-parietal signals that enhance or suppress information in occipital-temporal regions, which are capacity limited (Petro et al., 2017). Similar re-entrant signals may bias visual attention to the novel objects in the context of the present labeling intervention, and increase their competitive strength relative to faces. Future work may use multimodal imaging and directional analyses to examine this hypothesis.

The unpredicted left lateralized selective increase in ssVEP power for objects after label training (see Table 2 and Figures 5) is consistent with findings from a range of investigations reporting a distributed left lateralized network of regions involved in language processing (Friederici, 2012). The network of regions reported in the present investigation have been linked visual language perception (McCandliss et al., 2003) and semantic processing (Bookheimer, 2002; Pulvermüller, 2018; Tomasello, Garagnani, Wennekers & Pulvermüller, 2017). Left lateralized responses and a broad network of regions are consistent with one recent neurocomputational model (Tomasello et al., 2017) which implements Hebbian learning across frontal and temporal cortical regions as well as a previous proposal that includes a combination of top-down and bottom-up predictive processes involved in language comprehension (Friederici, 2012). The idea that associating a verbal label with a class of object aids in representation of the object by enriching the feature space overlaps conceptually with the label-as-feature hypothesis (Gliozzi et al., 2009; Sloutsky, 2010) discussed in developmental studies.

The present findings are also consistent with previous reports of competition and interference effects in real-world perceptual experts (e.g., Curby & Gauthier, 2014; Gauthier & Curby, 2005; Gauthier et al., 2003; Rossion et al., 2004; 2007; McGugin et al., 2012; 2014). However, these results extend the network of regions involved in cortical competition to include distributed regions in frontal as well as occipital temporal regions. In addition, the present investigation shows competition between face and object representations within minutes of learning. Since ssVEPs were measured during an free viewing period instead of in the context of a task it is difficult to determine whether the changes we report from pre-test to post-test reflect momentary changes in allocation of visual selective attention or the initial stages of expert-like perceptual indices of visual perceptual expertise. Future work, following adults across several training sessions with difficult expert-level discriminations and experimental manipulations of attention would allow for an unfolding of the time course of cortical competition underlying the role of selective attention in the acquisition of expertise. Finally, based on the present data it is unclear whether the specificity of the label plays an important role in competitive cortical processing. For example, if another family of artificial objects was trained but the two species were simply labeled “other” would competition between faces and objects decrease? It is also possible that labels are just one of many ways to promote categorization, individuation and neural competition (see Bukach et al., 2012). In a similar vein, objects other than the biologicalappearing “Sheinbug” species used here may prompt different patterns of competition. Previous work, for example. has found neural competition between faces and objects to be stronger when competing objects were more “facelike”, i.e., when they resembled a human face (Lochy et al., 2017). The ssVEP frequency tagging method coupled with source estimation holds significant promise for addressing these issues. Future work, replicating and extending the present findings, is needed to establish the extent to which label-learning-induced changes in neural competition follow the pattern observed here.

The present study is consistent with broader conceptual models suggesting that learning to label exemplars from novel categories involves interactions between attention, perception, and conceptual processing (Markant & Scott, 2018). The results reported here highlight the potential role of a broad network of frontoparietal cortical regions during the course of label learning. The selective facilitation of novel objects in frontoparietal and visuocortical regions is consistent with selective attention weighting and biasing of sensory processing towards the novel and recently learned category at the cost of an existing category (gender of faces). In addition, crosstalk between visuocortical areas and frontoparietal regions may result in optimized visuocortical tuning to the object features that are critical for the accurate recognition of exemplars from newly acquired categories (Mcteague, Gruss, & Keil, 2015). Visual features including color and spatial frequency have been previously associated with recognition and discrimination advantages in real world experts (Hagen et al., 2014, 2016) and after extensive laboratory training (Devillez et al, 2019; Jones et al, 2018). Thus, manipulation of these and other visual features using frequency tagging before and after label learning may allow for a more detailed understanding of how attention to these perceptual features is modulated by label learning as well as the extent to which features, like color and spatial frequency, compete with each other across this distributed network.

Figure Captions

Figure 1. Stimuli included 70 neutral male and female faces from the Karolinska Directed Emotional Faces database and 70 novel object images called Sheinbugs (Jones et al., 2018) divided into two subset "Species."

Figure 2. Stimuli concurrently emerged from a Brownian noise (spatial noise with a $1/f^2$ characteristic) background with the same mean luminance and contrast as the experimental stimuli. Each trial was 6 seconds long. Presentation Frequency was counterbalanced between subjects, who passively viewed the stimuli, while maintaining fixation.

Figure 3. A. Averaged frequency spectrum from a model dipole located at the occipital pole, during the pre-intervention session in one of two tagging groups (n=12), where faces were tagged at 5 Hz, and objects tagged at 6 Hz. B. Grand mean (n=24) ssVEP power in response to both faces and objects during the pre- and post- intervention sessions. C. Grand mean (n=24) difference maps illustrating object minus face power differences in source space.

Figure 4. Top panel: Dipole locations exceeding the permutation-based threshold when comparing post- and pre-intervention competition maps. Bottom panel: Dipole locations in which competition indices were different from zero (permutation controlled t-test against 0). The competition index is the product of the post-training minus pre-training differences for each stimulus, this rendering negative numbers when one response (i.e. object) increases, and the other decreases. It is positive when both decrease or increase with training. Only negative values were observed in the present data set, and all were related to increase in object-evoked ssVEPs.

Figure 5. Illustration of data from two regions (Left Middle Frontal Gyrus (top); Left Inferior Occipital Gyrus (bottom) showing a selective increase in ssVEP power (in source space) for

objects, compared to faces from pre-intervention to post-intervention. The bar plots include the mean (dark line) and 95% CI (box) as well as individual data points. Individual power differences between faces and objects are shown in the boxes on the right at pre-test and post-test for each region. Note that these plots are given to further illustrate the results of the topographical analysis in Figure 4, but do not represent a separate or additional statistical test.

Figure 6. Data from the no-training group. A. Grand mean (n=26) ssVEP power projected to MNE space in response to both faces and objects during the pre-break and post-break sessions. B. Grand mean (n=26) competition map illustrating object minus face power difference in source space, again shown for the sessions before and after the break. C. Histogram of t-values from the competition analysis, which examined changes in competition (objects minus faces) from pre- to post-break sessions. No location in the 350-source space exceeded the threshold determined by permutation (3.13) D. Exploratory comparison of competition indices in the label-training group and in the no-training group, using Wilcoxon Rank-Sum tests at each dipole location.

Figure 7. Rank correlation map between change in neural competition towards objects from before to after training and discrimination accuracy. Top panel: Left view of the correlation map in the estimated source space, projected to a standard brain. A scatter plot of the change in competition in left occipital dipoles (circled) and change in detection accuracy during training is shown. Bottom panel: Back view of the correlation map in the estimated source space, projected to a standard brain. A scatter plot of the change in competition in right parietal (circled) and change in detection accuracy during training is shown.

References

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*(10), 457-464. https://doi.org/10.1016/j.tics.2004.08.011

Andersen, S. K., Muller, M. M., & Hillyard, S. A. (2009). Color-selective attention need not be mediated by spatial attention. J Vis, 9(6), 2 1-7. https://doi.org/10.1167/9.6.2

Barragan-Jason, G., Cauchoix, M., & Barbeau, E. (2015). The neural speed of familiar face recognition. *Neuropsychologia*, *75*, 390–401. https://doi.org/10.1016/j.neuropsychologia.2015.06.017

Blair, R. C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology, 30*(5), 518-524.

Boremanse, A., Norcia, A. M., & Rossion, B. (2013). An objective signature for visual binding of face parts in the human brain. Journal of Vision, 13(11), 6–6. https://doi.org/10.1167/13.11.6

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. (9176952).

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience, 151(88).* DOI: 10.1146/annurev.neuro.25.112701.142946

Bukach, C. M., Vickery, T.J., Kinka, D., & Gauthier, I. (2012). Training Experts: Individuation without naming is worth it. Journal of Experimental Psychology: Human Perception and Performance, 38(1), 14-7.

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face Recognition in Poor-Quality Video: Evidence From Security Surveillance. *Psychological Science*, *10*(3), 243–248. https://doi.org/10.1111/1467-9280.00144

Curby, K. M., & Gauthier, I. (2014). Interference between face and non-face domains of perceptual expertise: A replication and extension. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00955

Devillez, H., Mollison, M.V., Hagen, S., Tanaka, J.W., Scott, L.S., & Curran, T. (2019). Color and spatial frequency differentially impact early stages of perceptual expertise training. Neuropsychologia, https://doi.org/10.1016/j.neuropsychologia.2018.11.011

D'Esposito, M & Postle, B.R. (2015). The Cognitive Neuroscience of Working Memory. *Annual Review of Psychology, 66*:115-142. https://doi.org/10.1146/annurev-psych-010814-015031.

Dowsett, A. J., Sandford, A., & Burton, A. M. (2016). Face learning with multiple images leads to fast acquisition of familiarity for specific individuals. *The Quarterly Journal of Experimental Psychology*, *69*(1), 1–10. https://doi.org/10.1080/17470218.2015.1017513

Friederici, A.D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cognitive Sciences, 16(5),* 262-268. https://doi.org/10.1016/j.tics.2012.04.001

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training 'greeble'experts: a framework for studying expert object recognition processes. *Vision Research*, *38*(15-16), 2401-2428. https://doi.org/10.1016/s0042-6989(97)00442-2

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience, 6*(4), 428-432. https://doi.org/10.1038/nn1029

Gauthier, I., & Curby, K. M. (2005). A Perceptual Traffic Jam on Highway N170. *Current Directions in Psychological Science, 14*(1), 30-33. https://doi.org/10.1111/j.0963-7214.2005.00329.x

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring mechanisms for face recognition. *Vision Research*, *37*(12), 1673-1682. https://doi.org/10.1016/s0042-6989(96)00286-6

Gazzaley, A., Cooney, J.W., McEvoy, K., Knight, R.T., D'Esposito M. (2005). Top-down enhancement and suppression of the magnitude and speed of neural activity. *Journal of Cognitive Neuroscience, 17*:507–17

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences*, *16*(2), 129–135. https://doi.org/10.1016/j.tics.2011.11.014

Gliozzi, V., Mayor, J., Hu, J.-F., & Plunkett, K. (2009). Labels as Features (Not Names) for Infant Categorization: A Neurocomputational Approach. Cognitive Science, 33(4), 709–738. https://doi.org/10.1111/j.1551-6709.2009.01026.x

Hagen, S. Vuong, Q., Scott, L.S., Curran, T. & Tanaka, J.W. (2014) The role of color in expert object recognition. *Journal of Vision, 14(9)*, 1-13, https://doi.org/10.1167/14.9.9

Hagen, S. Vuong, Q., Scott, L.S., Curran, T. & Tanaka, J.W. (2016) The role of spatial frequency in expert object recognition. *Journal of Experimental Psychology: Human Perception and Performance. 42(3)*, 413–422, https://doi.org/10.1037/xhp0000139

Hauk, O., Keil, A., Elbert, T., & Müller, M. (2002). Comparison of data transformation procedures to enhance topographical accuracy in time-series analysis of the human EEG. *Journal of Neuroscience Methods, 113*(2), 111-122. https://doi.org/10.1016/s0165-0270(01)00484-8

Hauk, O. (2004). Keep it simple: A case for using classical minimum norm estimation in the analysis of EEG and MEG data. *NeuroImage, 21*(4), 1612-1621. https://doi.org/10.1016/j.neuroimage.2003.12.018

Jones, T., Hadley, H., Cataldo, A. M., Arnold, E., Curran, T., Tanaka, J. W., & Scott, L. S. (2018). Neural and behavioral effects of subordinate-level training of novel objects across manipulations of color and spatial frequency. *European Journal of Neuroscience*. https://doi.org/10.1111/ejn.13889

Junghöfer M., Elbert T., Tucker D. M., & Rockstroh B. (2000) Statistical control of artifacts in dense array EEG/MEG studies. *Psychophysiology 37*, 523–532. https://doi.org/10.1111/1469-8986.3740523

Lochy, A., Zimmermann, F. G. S., Laguesse, R., Willenbockel, V., Rossion, B., & Vuong, Q. C. (2017). Does Extensive Training at Individuating Novel Objects in Adulthood Lead to Visual Expertise? The Role of Facelikeness. Journal of Cognitive Neuroscience, 30(4), 449–467. https://doi.org/10.1162/jocn_a_01212

Luxology, L. L. C. (2005). Luxology: modo-Advanced Ergonomics.

Markant, J. & Scott, L. S. (2018). Attention and Perceptual Learning Interact in the Development of the Other-Race Effect. *Current Directions in Psychological Science*, *27*(3), 163–169. https://doi.org/10.1177/0963721418769884

McCandliss BD, Cohen L, Dehaene S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Scienc*e, 7(7), 293-299. doi:10.1016/s1364-6613(03)00134-7

McGugin, R. W., Gatenby, J. C., Gore, J. C., & Gauthier, I. (2012). High-resolution imaging of expertise reveals reliable object selectivity in the fusiform face area related to perceptual performance. *Proceedings of the National Academy of Sciences, 109*(42), 17063-17068. https://doi.org/10.1073/pnas.1116333109

McGugin, R. W., Newton, A. T., Gore, J. C., & Gauthier, I. (2014). Robust expertise effects in right FFA. *Neuropsychologia, 63*, 135-144. https://doi.org/10.1016/j.neuropsychologia.2014.08.029

McTeague, L. M., Gruss, L. F., & Keil, A. (2015). Aversive learning shapes neuronal orientation tuning in human visual cortex. Nature Communications, 6, 7823. https://doi.org/10.1038/ncomms8823

Miller, E.K., & Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annual Reviews of Neuroscience, 24,* 167–202

Miyakoshi, M., Nomura, M., & Ohira, H. (2007). An ERP study on self-relevant object recognition. *Brain and Cognition*, *63*(2), 182–189. https://doi.org/10.1016/j.bandc.2006.12.001

Moratti, S., & Keil, A. (2009). Not what you expect: experience but not expectancy predicts conditioned responses in human visual and supplementary cortex. *Cerebral Cortex*, *19*(12), 2803-2809. https://doi.org/10.1093/cercor/bhp052

Nemrodov, D., Niemeier, M., Mok, J. N. Y., & Nestor, A. (2016). The time course of individual face recognition: A pattern analysis of ERP signals. *NeuroImage*, *132*, 469–476. https://doi.org/10.1016/j.neuroimage.2016.03.006

Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: a review. *Journal of Vision*, *15*(6), 4-4. https://doi.org/10.1167/15.6.4

Peyk, P., De Cesarei, A., & Junghöfer, M. (2011). ElectroMagnetoEncephalography software: overview and integration with other EEG/MEG toolboxes. *Computational Intelligence and Neuroscience*, *2011*. https://doi.org/10.1155/2011/861705

Pierce, L. J., Scott, L. S., Boddington, S., Droucker, D., Curran, T., & Tanaka, J. W. (2011). The N250 Brain Potential to Personally Familiar and Newly Learned Faces and Objects. *Frontiers in Human Neuroscience, 5(111), 1-13*. https://doi.org/10.3389/fnhum.2011.00111

Pulvermüller F. (2018). Neurobiological Mechanisms for Semantic Feature Extraction and Conceptual Flexibility. *Topics Cognitive Science*, 10(3), 590-620. doi:10.1111/tops.12367

Reynolds, J. H., & Heeger, D. J. (2009). The Normalization Model of Attention. Neuron, 61(2), 168–185. https://doi.org/10.1016/J.Neuron.2009.01.002

Rossion, B., Kung, 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, 101*(40), 14521-14526. https://doi.org/10.1073/pnas.0405613101

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*(3), 543-555. https://doi.org/10.1162/jocn.2007.19.3.543

Scott, L., Tanaka, J., Sheinberg, D., & Curran, T. (2006). The contributions of category experience and learning to perceptual expertise: A behavioral and neurophysiological study. *Journal of Vision,6*(6), 616-616. https://doi.org/10.1167/6.6.616

Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2008). The role of category learning in the acquisition and retention of perceptual expertise: A behavioral and neurophysiological study. *Brain Research, 1210*, 204-215. https://doi.org/10.1016/j.brainres.2008.02.054

Sloutsky, V. M. (2010). From Perceptual Categories to Concepts: What Develops? Cognitive Science, 34(7), 1244–1286. https://doi.org/10.1111/j.1551-6709.2010.01129.x

Smith, J. D., Zakrzewski, A. C., Herberger, E., Boomer, J., Roeder, J., Ashby, F. G., & Church, B. A. (2015). The Time Course of Explicit and Implicit Categorization. Attention, Perception & Psychophysics, 77(7), 2476–2490. https://doi.org/10.3758/s13414-015-0933-2

Tanaka, J. W., Curran, T., & Sheinberg, D. L. (2005). The training and transfer of real-world perceptual expertise. *Psychological Science*, *16*(2), 145-151. https://doi.org/10.1111/j.0956-7976.2005.00795.x

Thigpen, N., Petro, N. M., Oschwald, J., Oberauer, K., & Keil, A. (2019). Selection of Visual Objects in Perception and Working Memory One at a Time. Psychological Science, 30(9), 1259–1272. https://doi.org/10.1177/0956797619854067

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. https://doi.org/10.1038/381520a0

Tomasello R, Garagnani M, Wennekers T, Pulvermüller F. (2017). Brain connections of words, perceptions and actions: A neurobiological model of spatio-temporal semantic activation in the human cortex. *Neuropsychologia*, 98, 111-129. doi:10.1016/j.neuropsychologia.2016.07.004

Vialatte, F. B., Maurice, M., Dauwels, J., & Cichocki, A. (2010). Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Progress in Neurobiology*, *90*(4), 418-438. https://doi.org/10.1016/j.pneurobio.2009.11.005

Wang, J., Clementz, B. A., & Keil, A. (2007). The neural correlates of feature-based selective attention when viewing spatially and temporally overlapping images. *Neuropsychologia*, *45*(7), 1393-1399. https://doi.org/10.1016/j.neuropsychologia.2006.10.019

Wieser, M. J., & Keil, A. (2011). Temporal trade-off effects in sustained attention: dynamics in visual cortex predict the target detection performance during distraction. *Journal of Neuroscience*, *31*(21), 7784-7790. https://doi.org/10.1523/JNEUROSCI.5632-10.2011

Wieser, M. J., Miskovic, V., & Keil, A. (2016). Steady-state visual evoked potentials as a research tool in social affective neuroscience. *Psychophysiology*, *53*(12), 1763-1775. https://doi.org/10.1111/psyp.12768

Wong, A. C. N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for facelike expertise with objects: Becoming a Ziggerin expert—but which type?. *Psychological Science*, *20*(9), 1108-1117. https://doi.org/10.1111/j.1467-9280.2009.02430.x.