



Low-frequency oscillations employ a general coding of the spatio-temporal similarity of dynamic faces



Nicholas Furl^{a,b,*}, Michael Lohse^{b,c}, Francesca Pizzorni-Ferrarese^a

^a Department of Psychology, Royal Holloway, University of London, Surrey TW20 0EX, United Kingdom

^b Cognition and Brain Sciences Unit, Medical Research Council, Cambridge CB2 7EF, United Kingdom

^c Department of Physiology, Anatomy, and Genetics, University of Oxford, Oxford OX1 3QX, United Kingdom

ARTICLE INFO

Keywords:

Action perception
Emotional expressions
Face perception
Neural oscillations
Magnetoencephalography
Representational similarity analysis

ABSTRACT

Brain networks use neural oscillations as information transfer mechanisms. Although the face perception network in occipitotemporal cortex is well-studied, contributions of oscillations to face representation remain an open question. We tested for links between oscillatory responses that encode facial dimensions and the theoretical proposal that faces are encoded in similarity-based “face spaces”. We quantified similarity-based encoding of dynamic faces in magnetoencephalographic sensor-level oscillatory power for identity, expression, physical and perceptual similarity of facial form and motion. Our data show that evoked responses manifest physical and perceptual form similarity that distinguishes facial identities. Low-frequency induced oscillations (< 20 Hz) manifested more general similarity structure, which was not limited to identity, and spanned physical and perceived form and motion. A supplementary fMRI-constrained source reconstruction implicated fusiform gyrus and V5 in this similarity-based representation. These findings introduce a potential link between “face space” encoding and oscillatory network communication, which generates new hypotheses about the potential oscillation-mediated mechanisms that might encode facial dimensions.

Introduction

Neural oscillations (rhythmic neural firing) are ubiquitous features of the brain and furnish mechanisms contributing to network communication (Engel and Singer, 2001; Salinas and Sejnowski). Synchronization of membrane potentials enhances coupling between brain regions, allowing them to control information flow and organize specific functional networks (Fries, 2005, 2009). Hierarchical processing among visual areas may be mediated by oscillatory mechanisms, with forward (bottom-up) and backward (top-down) communication between higher- and lower-level visual areas carried respectively by high- (gamma) and low- (beta) frequency oscillations (Michalareas et al., 2016). These connectivity mechanisms could enable “binding” of visual dimension representations into unitary object percepts (Engel and Singer, 2001). Although these mechanisms have perhaps been best-studied for visual processes in non-human animals, neural oscillations are also a hallmark of visual processing in humans. Low-frequency power modulation is a ubiquitous feature of visual responses measured by electroencephalography (EEG) and magnetoencephalography (MEG). A negatively-deflected alpha/beta (10–30 Hz) response, in particular, putatively indexes visual object encoding (Hanslmayr et al., 2012). Nevertheless, more could be learned about how this low-frequency power

deflection gives rise to visual encoding, what information is encoded, and in what format.

An example of a brain network in the human whose communication may be mediated by oscillatory mechanisms is the well-studied network of discrete functional areas in ventral occipital cortex, fusiform gyrus, V5 and superior temporal sulcus (Haxby et al., 2001; Furl et al., 2015) associated with perception of dynamic faces and localized using functional magnetic resonance imaging (fMRI). These face-selective and motion-sensitive areas encode the form and motion information used to recognize faces and their emotional expressions and presumably give rise to oscillatory signals that reflect this encoding and that would be detectable using MEG. For example, spatial locations in static photographs of facial forms useful for expression categorization is reflected in both power and phase of oscillations below 25 Hz (Schyns et al., 2011). Several studies have now also examined dynamic facial movements and illustrated a role for low-frequency oscillations. This frequency range is modulated by motion and form information present in facial video (Muthukumaraswamy et al., 2006; Virji-Babul et al., 2007; Popov et al., 2013; Furl et al., 2014; Güntekin and Başar, 2014; Jabbi et al., 2015; Fox et al., 2016; Symons et al., 2016).

* Corresponding author at: Department of Psychology Royal Holloway, University of London, Surrey TW20 0EX, United Kingdom.
E-mail address: nicholas.furl@rhul.ac.uk (N. Furl).

These findings suggest that oscillations, especially in low frequencies, may be transmitting the information about form and motion processed in the aforementioned face perception network.

We propose to go beyond these existing studies by investigating the role of neural oscillations in face perception from the standpoint of similarity-based representations. A longstanding theory (Valentine, 1991) of face recognition posits a similarity-based “face space”, where faces are encoded relative to a set of constituent attributes in a multidimensional feature space and evaluated based on their similarity with learned representations. This formulation motivated us to test whether oscillatory power might also reflect representational distances between faces based on their physical and perceptual similarity. Such similarity-based object representations have been discovered using time-domain data from EEG (Kaneshiro et al., 2015), MEG (Cichy et al., 2014), intra-cranial recording (Op de Beeck et al., 2001; Kiani et al., 2007) and fMRI (Haushofer et al., 2008; Drucker and Aguirre, 2009; Proklova et al., 2016) and for static facial attributes such as identities (Vida et al., 2017), configurations (Goesaert and Op de Beeck, 2013) and gaze directions (Carlin et al., 2011). However, these results are limited to time-domain data, and they cannot link stimulus information content with potential neural mechanisms manifested by oscillatory power. Much, therefore, remains to be learned about how oscillations might (or might not) reflect similarities among faces, relative to constituent features in a multidimensional similarity space.

Here, we tested for similarity-based oscillatory responses using representational similarity analysis (RSA) to compare similarity distances between MEG response patterns with similarity values derived from physical and perceptual measures of high-level facial dimensions and categories (Su et al., 2012). To this end, we developed “physical similarity spaces” by extracting configurations of facial form and patterns of facial motion from videos of dynamic facial expressions. We also developed “perceptual similarity spaces”, based on participants’ similarity judgments of facial form and motion. Lastly, “categorical similarity spaces” were based on the between- versus within-category structure for identity and emotional expression. Using these spaces, we were able to behaviorally test for inter-relationships between physical and perceptual measures of facial similarity and whether they contain information about facial identities and emotional expressions. Our main aim, however, was to establish whether any of these similarity spaces was manifested by induced oscillatory MEG responses, as measured at the sensor-level. As a basis for further comparison, we also tested whether time-domain evoked response similarity corresponded to physical perceptual or categorical face spaces. We therefore could determine whether any facial encoding we found for induced responses was also present in evoked signals. Lastly, as a supplemental analysis, we optimized a source reconstruction to localize our sensor space RSA effects within the aforementioned, well-studied face perception network. We acquired fMRI functional localizer data in the same participants as those who underwent behavioral and MEG testing and exploited the superior spatial resolution of fMRI to constrain our source solution. This multimodal dataset of physical data extracted from video, behavioral data, evoked and induced sensor-level MEG responses and fMRI-guided source localizations provided us with a rich set of measures to fully explore several novel tests about representations of facial similarity spaces.

Methods and materials

Participants

Twenty participants (> 18 years) were scanned using fMRI. Of these, two did not return for the behavioral experiment, one additional participant did not return for MEG, and behavioral data for one more participant were lost due to technical issues. Analyses proceeded with the sixteen participants who possessed the full complement of data. All participants were right-handed, had normal or corrected-to-normal

vision and reported no history of psychiatric or neurological disorder. The local Cambridge, UK ethics committee granted approval.

fMRI procedures and analysis

Structural scans were obtained to facilitate data registration during MEG source reconstruction. The results of fMRI localizer scans were also used to constrain source solutions to fMRI-defined functional regions of interest (ROIs). fMRI scans were collected using a 3 T Siemens Tim Trio MRI scanner with 32 channel head coil. Functional scans included whole-brain T2*-weighted echo-planar volumes with 64 × 64 matrix and 3 mm² resolution in-plane and 3.75 mm thick axial slices, TR 2 s, TE 30 ms, flip angle 78°. Structural scans were T1-weighted MPRAGE with 1 mm³ voxels. The two localizer runs (175 volumes) were separated by runs related to a different experiment on faces, not reported here. The localizer procedures were adapted from Furl et al. (2013, 2015). The experiment was controlled using E-Prime (Psychology Software Tools, Pittsburgh, PA). In each run, participants viewed four types of block, each containing grayscale presentations of a stimulus category: dynamic faces, dynamic objects or static versions of the same faces or objects (taken from the last frame of each video). There were six blocks of each block type per run and block order was pseudo-random. Each block comprised eight presentations of 1375 ms stimuli and a 1 s inter-block interval. Each participant fixated on a white dot overlaid on the center of each presentation and pressed a button-box key with the right index finger when the dot turned red on a pseudo-random one-third of stimulus presentations. Four male and four female facial identities, exhibiting transitions from neutral to disgust, fearful, happy and sad expressions were taken from the Amsterdam Dynamic Face Expression Set (ADFES) (Van der Schalk et al., 2011). Face blocks comprised eight identities and four randomly-selected expressions, with each expression appearing twice. Object blocks included eight objects, previous used in functional localizers (Fox et al., 2009; Furl et al., 2013, 2015). Dynamic object videos included various plants blowing in the wind, a spinning globe, a spinning ceiling fan, a burning flame, operating machinery and a running tap.

fMRI data were preprocessed and analyzed using SPM12 (Wellcome Trust Center for Neuroimaging, London <http://www.fil.ion.ucl.ac.uk/spm/>) and MATLAB (The Mathworks, Natick, MA, USA). Data were motion-corrected, spatially-normalized to an EPI template in MNI space, and smoothed to 8 mm FWHM. At a first level of analysis, we estimated within-participant effects using an AR(1) corrected general linear model with a 128 ms high pass filter. Four regressors were added by convolving onset times and durations for dynamic faces, static faces, dynamic objects and static objects with a canonical hemodynamic response function. Regressors were also added for head motion parameters. We tested contrasts of the block types at a second level, where a group analysis was conducted to identify locations in MNI space of occipitotemporal areas associated with form and motion representations of dynamic faces (Haxby et al., 2001; Furl et al., 2015). We localized face-selective areas: bilateral occipital face area (OFA), bilateral fusiform face area (FFA) and right superior temporal sulcus (STS) (defined by contrasting face blocks > object blocks) and motion-sensitive areas: right and left V5 (defined by contrasting dynamic blocks > static blocks). For ROI definition, we identified the coordinates of the peaks of clusters observed at $P < 0.001$ uncorrected that achieved family-wise error correction at the voxel level using random-field theory (Brett et al., 2003).

Behavioral procedures

The behavioral experiment was conducted using PsychoPy (Peirce, 2009) in a separate testing session either immediately following fMRI or within two weeks. Participants viewed the 630 possible unique pairings of 36 dynamic faces. The 36 faces were taken from the BU-

Download English Version:

<https://daneshyari.com/en/article/5630933>

Download Persian Version:

<https://daneshyari.com/article/5630933>

[Daneshyari.com](https://daneshyari.com)