

# Representational Dynamics of Facial Viewpoint Encoding

Tim C. Kietzmann<sup>1,2</sup>, Anna L. Gert<sup>2</sup>, Frank Tong<sup>3</sup>, and Peter König<sup>2,4</sup>

## Abstract

■ Faces provide a wealth of information, including the identity of the seen person and social cues, such as the direction of gaze. Crucially, different aspects of face processing require distinct forms of information encoding. Another person's attentional focus can be derived based on a view-dependent code. In contrast, identification benefits from invariance across all viewpoints. Different cortical areas have been suggested to subserve these distinct functions. However, little is known about the temporal aspects of differential viewpoint encoding in the human brain. Here, we combine EEG with multivariate data analyses to resolve the dynamics of face processing with high temporal resolution. This revealed a distinct sequence of view-

point encoding. Head orientations were encoded first, starting after around 60 msec of processing. Shortly afterward, peaking around 115 msec after stimulus onset, a different encoding scheme emerged. At this latency, mirror-symmetric viewing angles elicited highly similar cortical responses. Finally, about 280 msec after visual onset, EEG response patterns demonstrated a considerable degree of viewpoint invariance across all viewpoints tested, with the noteworthy exception of the front-facing view. Taken together, our results indicate that the processing of facial viewpoints follows a temporal sequence of encoding schemes, potentially mirroring different levels of computational complexity. ■

## INTRODUCTION

Faces are among the most important categories of visual stimuli, providing a rich set of information that is essential to our everyday behavior. One aspect of facial processing that has attracted considerable interest in the past is the cortical encoding of 3-D viewpoints resulting from rotations in depth. The variety of possible viewpoints presents a considerable challenge to the visual system, as different computational goals may rely on specific viewpoints and different levels of viewpoint invariance. For example, invariant face identification mechanisms must generalize across a large range of viewpoints, whereas shared attention must distinguish different gaze directions. Head orientation provides a strong cue for the recognition of another person's attentional focus, and therefore, different head orientations, too, need to be distinguishable from each other (Haxby, Hoffman, & Gobbini, 2000; Perrett, Hietanen, Oram, & Benson, 1992). Finally, the frontal viewpoint has important social and emotional relevance because it frequently co-occurs with direct eye contact (Carlin, Calder, Kriegeskorte, Nili, & Rowe, 2011; Senju & Johnson, 2009), whereas seeing a face from all other viewpoints usually signals an averted gaze.

Befitting the complexity and variety of the underlying computational demands, the primate brain contains a large network of face-selective regions (Pitcher, Walsh,

& Duchaine, 2011; Freiwald & Tsao, 2010; Yovel & Kanwisher, 2005; Gauthier et al., 2000; Haxby et al., 2000; Kanwisher, McDermott, & Chun, 1997), and distinct cortical subsystems may exist that support different aspects of face processing (Freiwald & Tsao, 2010; Hoffman & Haxby, 2000). Despite our increasing understanding of the selectivity in the various nodes of this network (Anzellotti, Fairhall, & Caramazza, 2013; Axelrod & Yovel, 2012; Kietzmann, Swisher, König, & Tong, 2012; Carlin et al., 2011; Freiwald & Tsao, 2010; Natu et al., 2010; Pourtois, Schwartz, Seghier, Lazeyras, & Vuilleumier, 2005; Tong, Nakayama, Moscovitch, Weinrib, & Kanwisher, 2000) and a large body of research focusing on electrophysiological signatures of face processing (Caharel, Collet, & Rossion, 2015; Rossion, Prieto, Boremanse, Kuefner, & Van Belle, 2012; Eimer, 2000, 2011; Rossion & Jacques, 2008; Schweinberger, Kaufmann, Moratti, Keil, & Burton, 2007; Itier, Herdman, George, Cheyne, & Taylor, 2006; Joyce, Schyns, Gosselin, Cottrell, & Rossion, 2006; Liu, Harris, & Kanwisher, 2002; McCarthy, Puce, Belger, & Allison, 1999; Bentin, Allison, & Puce, 1996), temporal aspects of differential viewpoint encoding schemes have not been explored in detail. The latencies at which different types of information emerge from the cortical network can ultimately provide valuable insights into the underlying cortical mechanisms and could mirror different levels of computational complexity and task importance. That is, later effects might either be due to more complex visual inference that requires longer computation time or result from a less preferential treatment of the visual system that focuses on other aspects first.

<sup>1</sup>Medical Research Council, Cambridge, UK, <sup>2</sup>University of Osnabrück, Germany, <sup>3</sup>Vanderbilt University, Nashville, TN, <sup>4</sup>University Medical Center Hamburg-Eppendorf, Germany

To investigate the temporal sequence of viewpoint processing in the human brain, we used EEG to record cortical responses while participants viewed faces shown from 37 different viewpoints spanning a total of 180° (Figure 1A). We applied multivariate pattern analyses to the visually evoked responses to extract fine-grained information about the representation of facial viewpoints at different temporal latencies. Crucially, the high temporal resolution of the EEG measurements allowed us to investigate fast changes in the underlying representational structures, allowing us to determine the point in time at which the visual system exhibits different viewpoint encoding schemes and invariances.

To analyze the visually evoked potentials, we employed representational similarity analysis (RSA; Kriegeskorte, Mur, & Bandettini, 2008), a multivariate approach to study similarities between cortical activation patterns and their relation to the structure of experimental conditions. The resulting representational similarity structures were analyzed using data-driven visualization techniques, followed by a model-driven approach that explicitly tested three different encoding schemes, each based on a different computational demand for viewpoint encoding. The first model tested how well different viewing directions could be distinguished, as would be required for mechanisms of face perception and shared attention. This model predicted that similar viewpoints lead to similar neural responses, whereas increasing angular differences between viewpoints lead to larger differences in the visually evoked response. It should be noted that cortical activity sensitive to viewpoint similarity could either reflect face-specific processes or processes sensitive to low-level differences, because these increase as a function of viewpoint difference. Contrasting this separation of different viewpoints, our second model focused on partial viewpoint invariance by investigating the effects of viewpoint symmetry. This effect describes the neural preference for select viewpoints as well as their horizontally mirror-symmetric counterparts (e.g., 30° and -30° rotated away from the front view are processed similarly, compared with 30° and the intermediate angles). Such joint viewpoint selectivity would be particularly efficient for bilaterally symmetric objects, such as faces, for which a computationally simple mirror operation would allow the system to either effectively reduce the number of viewpoints required to achieve full invariance or increase the signal-to-noise ratio. Consequently, the effects of mirror-symmetric response tuning were suggested to constitute an important intermediate computational step in building a fully viewpoint-invariant cortical representation (Vetter, Poggio, & Bühlhoff, 1994). Neural effects of mirror symmetry were first observed in the macaque temporal lobe (Dubois, de Berker, & Tsao, 2015; Freiwald & Tsao, 2010; Logothetis, Pauls, & Poggio, 1995; Perrett et al., 1991). In humans, the effects of viewpoint symmetry have been reported across a wide range of higher-level visual areas, in studies using multivariate analyses of fMRI data (Axelrod & Yovel, 2012; Kietzmann et al., 2012), and

TMS (Kietzmann et al., 2015), highlighting the importance of this computational step. As the third model of interest, we investigated whether the frontal viewpoint is processed differently compared with neighboring, slightly oblique views, as the former frequently co-occurs with direct eye contact and might consequently posit a special status.

## METHODS

### Participants

Nineteen healthy participants (aged 19–29 years, seven women) took part in the experiment. All participants had normal or corrected-to-normal visual acuity. They were informed about their right to withdraw from the experiment at any time and gave written informed consent to participate. Because of poor task performance or technical error, 3 of the 19 participants had to be excluded from the analyses. The study was approved by the institutional review board of the Osnabrück University and conformed to the Declaration of Helsinki.

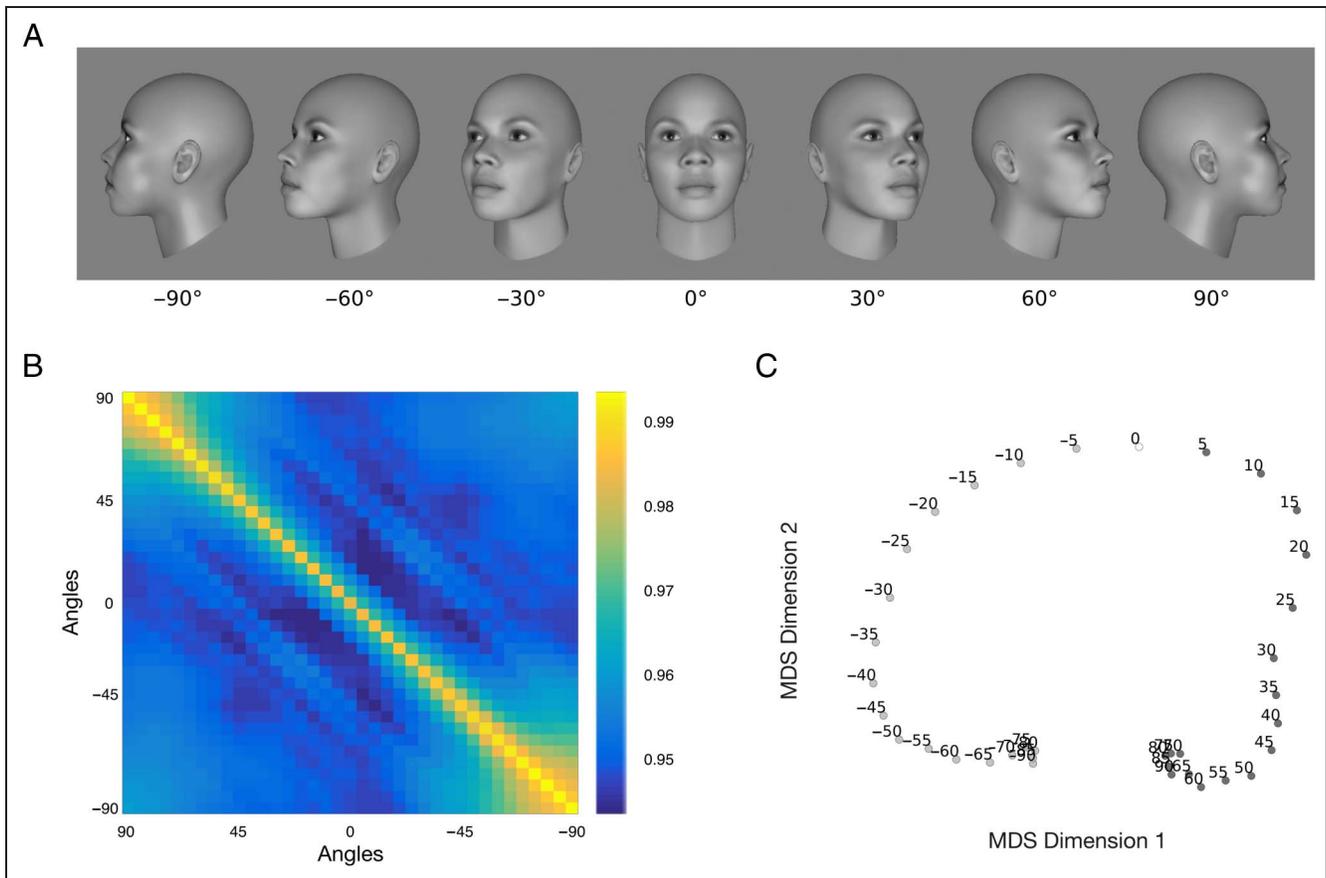
### Stimuli

The stimulus set was created using FaceGen (Singular Inversions, Inc., Toronto, Canada). It included four individuals (two women), shown from 37 different angles ranging from -90° to 90° in steps of 5°, resulting in 37 experimental conditions (Figure 1A). The stimuli were presented in grayscale on a gray background. The luminance histograms of all stimuli were matched using the SHINE toolbox (Willenbockel et al., 2010). To match previous experimental work from the literature, the face models did not include hair.

To better understand the similarity structure of the stimulus space, we followed our previous approach (Kietzmann et al., 2012) and estimated a low-level stimulus similarity matrix based on a computational model of V1 simple cells. The model consists of a set of 2-D Gabor functions with 17 spatial scales and four orientations (Serre & Riesenhuber, 2004; Lee, 1996). Once applied to the experimental stimuli, the model yields a high-dimensional response vector for each stimulus, which can be used to compute the corresponding similarity matrix using a Pearson product-moment correlation (Figure 1B). To visualize the similarity structure across all viewpoints to be expected from low-level stimulus features alone, the resulting V1 similarity matrix was projected into a 2-D space using multidimensional scaling (MDS; Figure 1C).

### Experimental Setup and Design

Each experimental session consisted of 16 blocks, with 312 trials each (4992 trials in total). Within each block, each identity and viewpoint combination was shown twice (296 trials) in pseudorandomized order, while preventing the direct repetition of identical viewpoints. Each trial included 400 msec of stimulus presentation, followed



**Figure 1.** Exemplary stimuli. (A) A subset of all stimuli used. In the experiment, participants viewed 37 different head orientations, ranging from  $-90^\circ$  to  $90^\circ$  in steps of  $5^\circ$ . (B) The results of a V1 model, displayed as similarity matrix. (C) Same data as in B, displayed in 2-D after MDS was applied.

by a random ISI of 300–500 msec. The presentation of the experimental blocks was self-paced by the participants.

To keep the participants' attention on the stimulus display, they were instructed to maintain fixation on a central target and to perform a color change detection task at fixation by reporting these rare target events (16 trials in each block) via button press. This distractor task was orthogonal to the experimental question. Target trials and trials including button presses by the participants were excluded from the analyses.

During the EEG measurements, the participants were seated in a dark room. Stimuli were presented using a 24-in. BenQ screen (Model 2420T, Taipei, Taiwan) running at a resolution of  $1920 \times 1080$  pixels and a refresh rate of 120 Hz. The average latency between EEG trigger and stimulus onset was 7.5 msec. The data shown were corrected for this delay. The distance to the screen was 80 cm, yielding a stimulus display size of approximately  $7.9^\circ \times 9.7^\circ$  of visual angle.

In addition to the EEG recordings, we measured the eye movements of our participants using an EyeLink 1000 remote eye tracker (SR Research Ltd., Mississauga, Ontario, Canada), running at a sampling frequency of 500 Hz. The eye tracker was calibrated before the experiment, and recalibration was conducted after every

fourth experimental block. Throughout the experiment, the eye-tracking validation error was kept below an average of  $0.5^\circ$  visual angle.

### Data Acquisition and Preprocessing

Electrophysiological data were recorded using a 128-Ag/AgCl electrode system (ANT Neuro, Enschede, The Netherlands), with electrodes placed on a Waveguard cap according to the 5% electrode system (Oostenveld & Praamstra, 2001). The data were recorded at a sampling rate of 1024 Hz, ensuring premeasurement that the impedances of all electrodes were below 10 k $\Omega$ .

Data were preprocessed using the EEGLAB toolbox (Delorme & Makeig, 2004) in the following order: First, the data were down-sampled to 500 Hz and subsequently filtered using a 1-Hz high-pass filter and a 120-Hz low-pass filter as well as a notch filter around 50 and 100 Hz. Channels exhibiting either excessive noise or strong drifts were removed (four channels on average). After this, the continuous data were manually cleaned, rejecting data sequences including jumps, muscle artifacts, and other sources of noise (on average, 17% of data were rejected for each participant). To remove eye movement-related artifacts, an independent component analysis (based on

the AMICA algorithm; Palmer, Makeig, Kreutz-Delgado, & Rao, 2008) was computed on the cleaned data. The independent components corresponding to eye, heart, or muscle activity were manually selected and removed, based on criteria matching previous work on component selection (Plöchl, Ossandón, & König, 2012), before transforming the data back into the original sensor space. The initially removed channels were interpolated based on the activity of their neighboring channels, selected via channel triangulation. Subsequently, the continuous data were divided into epochs for each trial by including data from 100 msec prestimulus to 700 msec poststimulus, using the time window between  $-100$  msec and stimulus onset for baseline correction. The resulting temporal resolution of the multivariate pattern analysis is 2 msec. ERP topographies were plotted using Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011) and custom code. To aid visibility, ERP topographies (Figure 3) and overview similarity matrices (Figure 4) are based on down-sampled data (100 Hz).

## Data Analysis

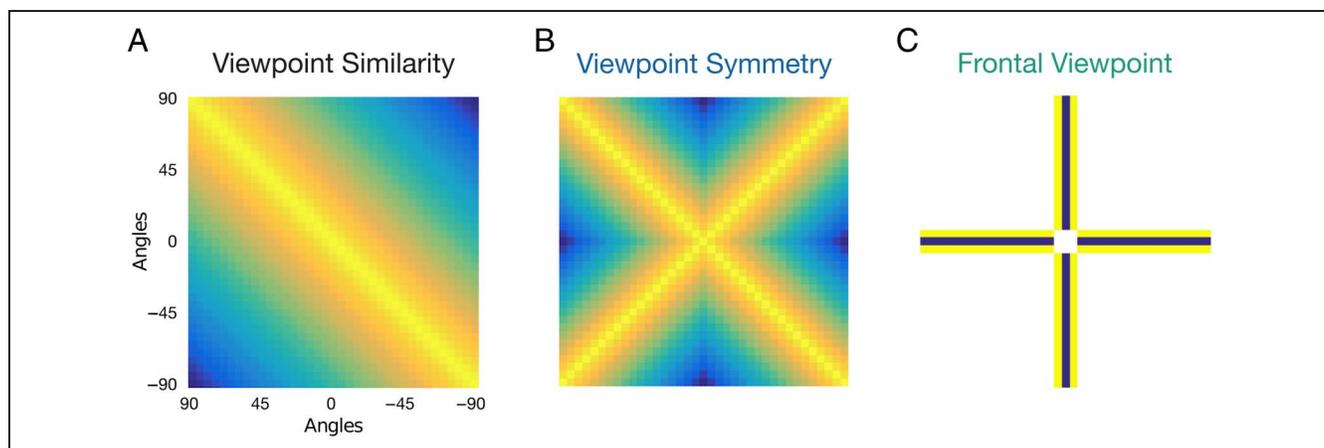
### Multivariate Analyses of EEG Data

To test for the temporal development of facial viewpoint encoding, we performed a multivariate pattern analysis on the visually evoked potentials, computing representational similarity matrices across all experimental conditions. Each entry in a given similarity matrix was computed using a Pearson product-moment correlation (Khaligh-Razavi & Kriegeskorte, 2014; Kriegeskorte, Mur, Ruff, et al., 2008) on the respective ERP activation patterns across all EEG channels. The resulting matrices were ordered based on the rotational angle of the faces, ranging from  $90^\circ$  to  $-90^\circ$ . Neighboring cells in each similarity matrix therefore represent neighboring viewing angles, and the resulting matrix depicts the overall simi-

ilarity structure across all conditions. Elements along the main diagonal were estimated by comparing the same viewpoint across different identities (comparisons of the same identity and viewpoint yield a correlation of 1 by definition and were therefore excluded). Computed for every point in time and participant, the resulting 3-D structure (Condition  $\times$  Condition  $\times$  Time) reflects the temporal development of representational similarity across all experimental conditions (Cichy, Pantazis, & Oliva, 2014; Carlson, Tovar, Alink, & Kriegeskorte, 2013).

Representational similarity matrices allow for data-driven analyses as well as comparisons with model predictions, such as low-level similarity, control conditions, and other behavioral measures (Kietzmann et al., 2012; Kriegeskorte, Mur, & Bandettini, 2008). To investigate the dynamics of viewpoint encoding across time, we performed data- and model-driven analyses on the spatiotemporal similarity matrices. For the data-driven approach, we first visualized the average empirical similarity at each point in time by projecting the data into a 2-D space using nonmetric MDS. The resulting 2-D arrangement of conditions closely resembles the similarity structure in the original data but is visually more accessible. To account for the temporal smoothness of the underlying data, each MDS optimization was seeded with the MDS result computed for the previous time point (Carlson et al., 2013). The initial MDS solution was based on a random position seed.

As noted earlier, different computational goals likely rely on different viewpoint encoding schemes. On the basis of the respective requirements, distinct predictions about the optimal representational similarity structure can be derived and used as model predictions that are then matched against the empirical data. In addition to the data-driven analyses, three encoding models were defined and tested: The viewpoint similarity model predicts decreasing similarity values with increasing angular difference, the viewpoint symmetry model additionally



**Figure 2.** Model predictions. (A) The viewpoint similarity model predicted decreasing neural response similarity with increasing angular difference between viewpoints. (B) The viewpoint symmetry model focused on partial viewpoint invariance by explicitly testing whether mirror-symmetric viewpoints are processed more similarly than nonsymmetric angles. (C) The frontal viewpoint model compared  $0^\circ$  (frontal) against  $\pm 5^\circ$  viewpoints by testing their respective similarities against all other viewpoints in the data set. Blue = negative model values; yellow = positive; white = not a number.

predicts high similarity values for mirror-symmetric viewing angles, and the frontal viewpoint model predicts distinct EEG patterns for the front-facing view as compared with adjacent views (Figure 2).

- Viewpoint similarity: Communication and shared attention require the recognition of another person's direction of gaze. For this, the overall head orientation is a reliable indicator, raising the question where and when different facial viewpoints are neurally separated. Potentially based on low-level stimulus features, early visual areas in macaque and human cortex have been shown to allow for reliable discrimination between different face viewpoints (Carlin, Rowe, Kriegeskorte, Thompson, & Calder, 2012; Kietzmann et al., 2012; Freiwald & Tsao, 2010). In contrast, higher-level visual areas are increasingly invariant to changes in viewpoint (Axelrod & Yovel, 2012, 2015; Anzellotti et al., 2013; Kietzmann et al., 2012; Freiwald & Tsao, 2010; Natu et al., 2010; Kriegeskorte, Formisano, Sorger, & Goebel, 2007). Nevertheless, it is possible that different viewpoints remain separable, even in higher visual areas, if different readout strategies are used (DiCarlo & Cox, 2007). To assess at what latencies head orientations can be discriminated, we designed a model that predicts that neighboring viewpoints elicit similar neural responses, whereas decreasing similarities are expected with increasing viewpoint difference. Maximal similarity is expected for identical viewpoints (tested across facial identities), that is, across the main diagonal of the similarity matrix. The model implicitly tests the extent to which disparate viewing angles can be distinguished by the activity patterns they evoke (Figure 2A). Because rank-based methods are employed to estimate the fit between model and empirical data (detailed below), the viewpoint similarity model does not make specific predictions about the exact rate of similarity falloff across viewpoints and matches any strictly decreasing similarity function. It therefore presents a more general approach, compared with the predictions of the Gabor model used to visualize the stimulus space, which is potentially better applicable to smooth EEG data.
- Viewpoint symmetry: Whereas some cortical computations on face stimuli require a separation of differently oriented viewpoints, others require information that is viewpoint invariant. Achieving invariance is a computationally challenging task, as faces seen from different perspectives differ largely in their retinal projections. To cope with this complexity, the visual system relies on multiple view-specific representations that span the whole space of possible viewpoints. Together, these units support invariant recognition (Perrett, Oram, & Ashbridge, 1998; Ullman, 1998; Logothetis et al., 1995; Bühlhoff & Edelman, 1992). In addition to this view-based code, however, it has been noted that

bilaterally symmetric objects, such as faces, allow for a computational shortcut, because mirror-symmetric viewpoints are left/right mirror images of each other (Vetter et al., 1994). Indeed, corresponding effects were reliably shown using single-cell recordings (Freiwald & Tsao, 2010; Perrett et al., 1991) and fMRI (Dubois et al., 2015) in the macaque. In humans, the effects of viewpoint symmetry have also been reliably observed in behavioral (Vetter et al., 1994), fMRI (Axelrod & Yovel, 2012; Kietzmann et al., 2012), and TMS (Kietzmann et al., 2015) experiments. To investigate the temporal latency of the effect, we included a viewpoint symmetry model in our RSA predictors. This model predicts that directly neighboring viewpoints as well as mirror-symmetric views will elicit similar cortical response patterns. Because the experimentally tested viewpoints are sorted from  $+90^\circ$  to  $-90^\circ$  in the corresponding representational similarity matrices, values around the principal diagonal (top left to bottom right) indicate similarity across directly neighboring views. The secondary diagonal (top right to bottom left) defines mirror-symmetric viewpoints, that is, the similarity between selected viewpoints and their mirror-symmetric counterparts. Because of this layout, the effects of viewpoint symmetry predict increased similarity values along both diagonals (X shape, shown in Figure 2B). To derive the model prediction, we used the viewpoint similarity model and added a horizontally mirrored version to define the secondary diagonal.

- Frontal viewpoint: Compared with oblique viewpoints, a direct frontal view of a face has distinct implications for communication and shared attention. It is horizontally symmetric and frequently co-occurs with direct eye contact. In the current experiment, the directly frontal viewpoint was the only condition that exhibited direct eye contact, whereas neighboring viewpoints, differing by only a  $5^\circ$  rotation, signaled an averted gaze. In line with this observation, the frontal viewpoint model predicts that the  $0^\circ$  (frontal) viewpoint elicits a neural response that is distinct from the response to all other viewpoints. The most conservative approach to testing this prediction is to contrast the frontal with its directly neighboring, slightly oblique viewpoints ( $+5^\circ$  and  $-5^\circ$ ), which function as controls. Specifically, a "special status" of the frontal view predicts that the average representational similarity between the  $0^\circ$  and all other viewpoints is smaller than the average similarity between  $\pm 5^\circ$  face views and all other viewpoints. Accordingly, the model matrix contains negative weights for all similarity estimates involving the frontal viewpoint (shown in blue in Figure 2C) and positive weights for correlations of slightly leftward- and rightward-facing viewpoints (shown in yellow). Please note that, although the viewpoints differ in the direction of gaze (direct vs. averted), the overall visual differences between the tested conditions are small.

To statistically evaluate the different models, we used a Spearman's rank correlation (Khaligh-Razavi & Kriegeskorte, 2014; Kriegeskorte, Mur, & Bandettini, 2008) between the model prediction and the empirical similarity matrices of each participant. This approach assumes that the relationship between model and data should maintain ordinal relationships and to avoid assuming linearity of the relationship between model and data. We then calculated a  $t$  test across participants to test, for every point in time, and to determine whether the correlation values (Fisher  $z$  corrected) were significantly different from zero.

To control for the multiple statistical comparisons performed, one for each time point, we applied a cluster-based permutation test (Maris & Oostenveld, 2007). All connected time points exhibiting a  $p$  value of  $<.01$  were considered as empirical candidate clusters. That is, only time points that were individually significant at  $p < .01$  were included for further analysis. The cluster null distribution was computed by randomly flipping the sign of the correlation values of each participant (in line with the prediction of a zero correlation; Good, 2013). In each iteration, the corresponding sign flip for a participant was applied to the whole time series, that is, to all time points, to preserve the temporal smoothness of the original data, as clusters are defined in the temporal domain. For every permutation (100,000 in total), we followed the same analysis steps as before and estimated positive and negative clusters to be expected under the null hypothesis. In every iteration, only the strongest positive and negative clusters were kept for the null distribution (max sum  $t$  statistic). The originally observed, empirical clusters were then compared with this cluster null distribution, computing the probability of observing a cluster equal or larger than the empirical one by chance alone. Only clusters with a  $p$  value of  $<.05$  are reported in the following (because we separately tested for positive and negative clusters, a  $p < .025$  was applied to either side, similar to a two-sided  $t$  test).

### *Noise Ceiling Estimates*

To estimate an upper bound for model performance, we computed the noise ceiling of the RSA data (Nili et al., 2014) across time. Given an (unknown) true similarity structure that underlies data generation for all participants, this approach uses the empirically observed similarity matrices to estimate the extent to which noise present in the data might limit the maximal level of performance that can be achieved by any model prediction. The noise ceiling consists of a lower bound and an upper bound. The lower bound is computed by a leave-one-participant-out approach, in which the average similarity matrix of all but one participant is correlated with the data of the left-out participant. The upper bound is overfitted to the individual participant, as it is based on

the overall average similarity matrix. Because the noise ceiling is estimated from the empirical similarity matrices, it depends on the cells that are included in the analyses. Thus, if models predict different parts of the similarity matrix, they require different noise ceiling estimates. Viewpoint similarity and viewpoint symmetry both rely on the whole matrix and therefore share a noise ceiling estimate. In contrast to this, the frontal viewpoint model provides a prediction for the subset of cells in the similarity matrix and therefore yields its own noise ceiling.

### *Sensor Searchlight Analysis*

To estimate the spatial origin of the effects observed, we implemented a sensor-based searchlight analysis. For each EEG channel, we defined a set of neighboring sensors (i.e., searchlight window within a distance of 2 cm in the 2-D spatial layout; 19.8 neighbors on average per channel). All channels in the local searchlight were then used together to estimate the local representational similarity matrices for each point in time. The resulting similarity structure has size  $n_{\text{chan}} \times n_{\text{time}} \times n_{\text{cond}} \times n_{\text{cond}}$ . Once computed, the different model predictions (viewpoint similarity, viewpoint symmetry, and frontal viewpoint) were tested for each searchlight location, storing the corresponding correlation values in the respective searchlight center. Please note that neural dipoles can also project to distal sensor groups, which are not necessarily spatial neighbors. The results of the current sensor searchlight analysis are therefore limited to spatially connected regions exhibiting a selected effect, neglecting effects that could be observed across more distant sensors.

Because the effect latencies were known for the individual models from our previous analyses including all sensors, the searchlight analysis was applied specifically for these predefined temporal windows of interest.

### *Eye-tracking Control*

Our participants were instructed to maintain fixation on a central target while performing a color change detection task at fixation. We nevertheless tested whether eye movements might have contributed to any of the observed effects. This was accomplished by computing a similarity matrix from the recorded eye-tracking data of the respective experimental conditions. We first computed a smoothed, 2-D probability distribution (fixation density map) for every participant, condition, and time point. Similar to the EEG activity and the V1 model, we then correlated the eye-tracking data of each condition with every other condition. This resulted in a 3-D correlation matrix (Condition  $\times$  Condition  $\times$  Time), which describes the similarity structure of the eye-tracking data for each individual participant and time point. As a next step, we again focused on previously established temporal clusters of significant model fits and used each participant's

eye-tracking similarity matrices as control data for a partial Spearman’s rank correlation analysis. Specifically, for a given temporal window of interest, we tested the fit between model and EEG similarity matrix, while using the eye-tracking similarity matrix as control. The (Fisher  $z$  transformed) partial correlation values were then subject to a  $t$  test across participants, while controlling for multiple statistical comparisons via Bonferroni correction.

## RESULTS

The basis of all analyses performed was the transformation of the high-dimensional EEG data into representational similarity matrices, which depict the pattern similarity of the visually evoked responses across all conditions. Similarity matrices were computed and statistically evaluated for each time point individually at a temporal resolution of 2 msec, allowing us to investigate fast changes in the representational similarity structure, indicative of changes in the underlying viewpoint encoding. Exemplary EEG topographies, underlying the multivariate analyses, are shown in Figure 3.

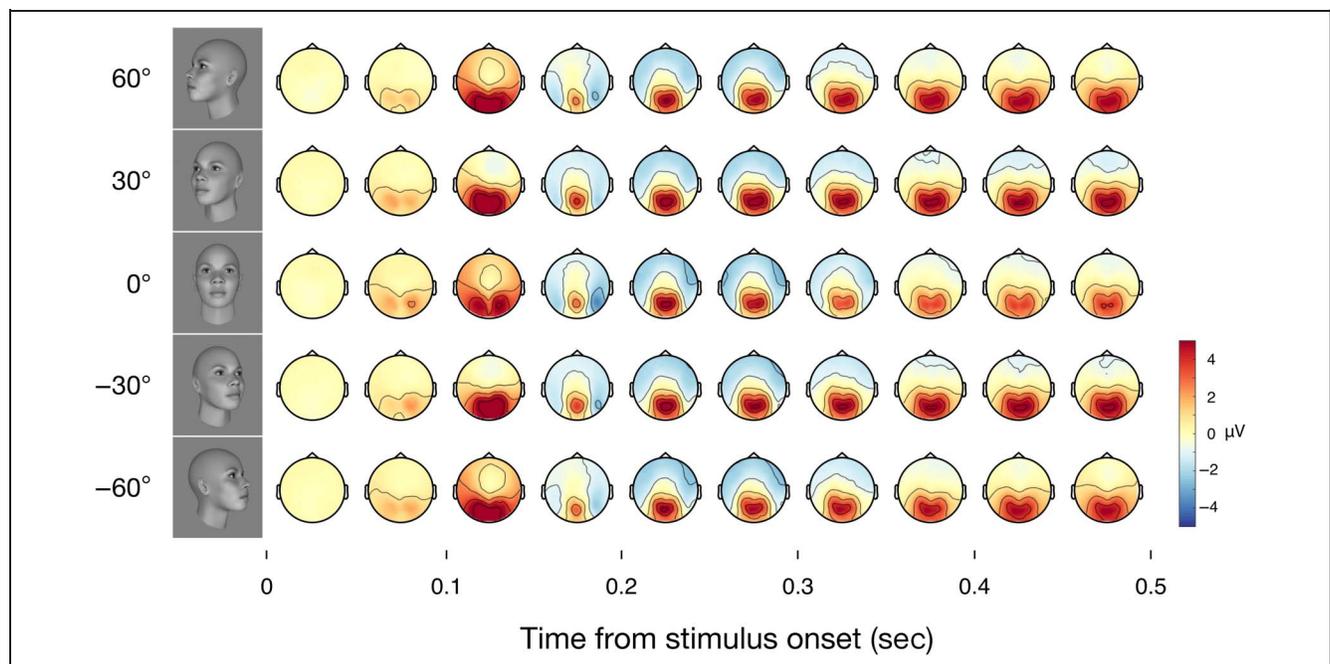
### Data-driven Analyses

To explore the temporal development of viewpoint processing, we visualized the average similarity matrix at each point in time and performed an MDS analysis on the data to project it into 2-D space.

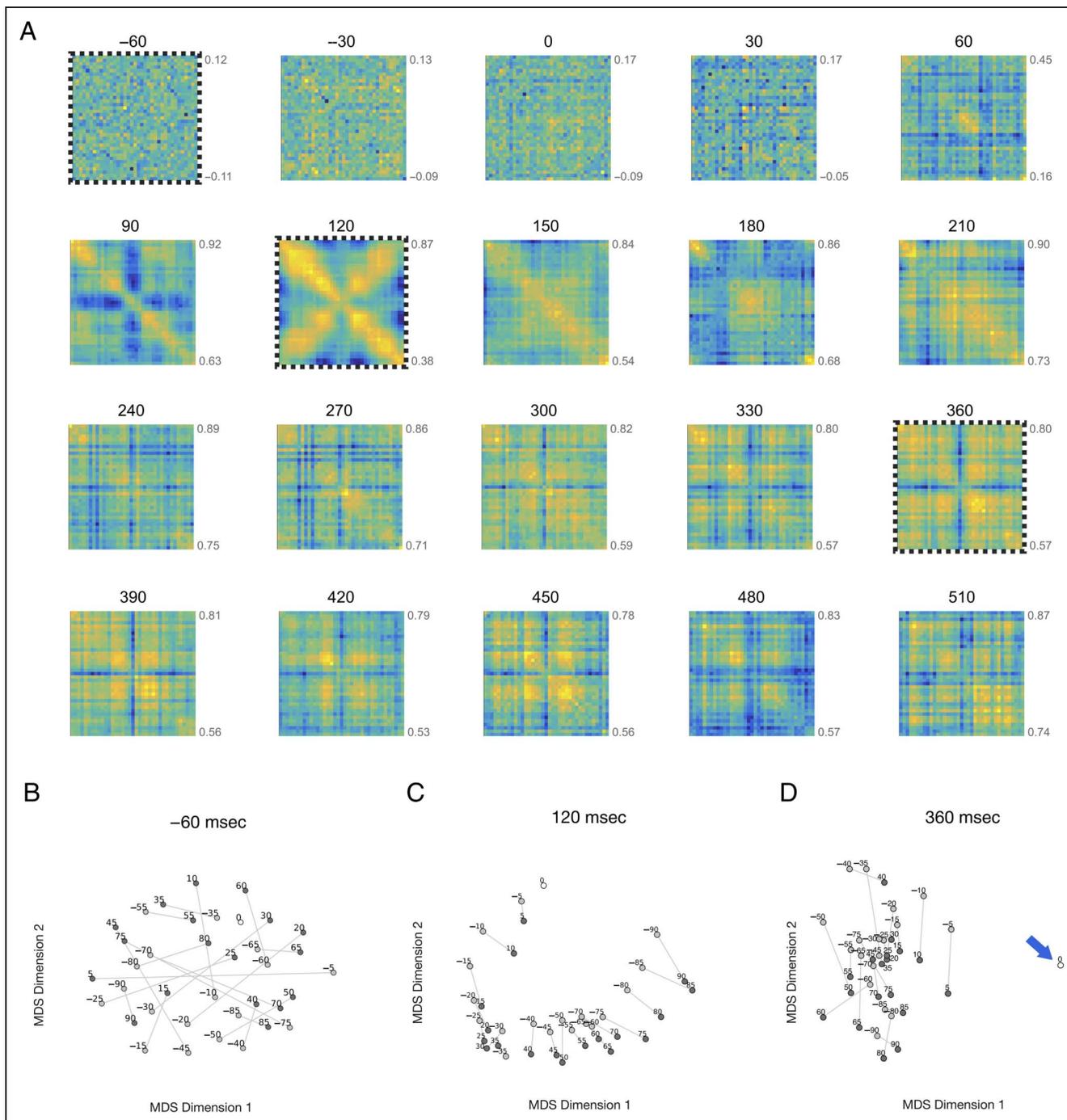
The similarity matrices indicated distinct and temporally separated stages of face processing (Figure 4). Initially, the similarity matrices appear unstructured and random

in their patterns. Then, starting around 60 msec after stimulus onset, the principal diagonal exhibits enhanced correlation values. At this point in time, neighboring viewing angles exhibit similar evoked activity patterns, in line with the response properties of lower-level retinotopic visual areas (Figure 4A and B). After this, at around 120 msec, we observed strong effects of viewpoint symmetry: Mirror-symmetric viewpoints exhibited increased correlations compared with intermediate viewpoints (Figure 4C). At a later point in time, around 300 msec, an effect specific to frontal viewpoints was observed. Whereas all other viewpoints exhibited comparable similarity in the evoked response patterns, the cortical responses to frontal views were decidedly different. The same pattern of results can be observed in the MDS projection. The alignment of conditions during baseline is largely random, whereas the arrangement after about 120 msec of processing is almost perfectly ordered with respect to overall rotational angle, with the addition of close proximity for symmetric viewpoints in this multidimensional space. At a considerably later point in time, the order of almost all viewpoints seems mostly random with one marked exception—frontal viewpoints. They appear at a large distance to all other conditions (Figure 4D), including viewpoints that differ by only small rotation angles, such as  $\pm 5^\circ$ .

The two later effects observed in the empirical similarity structures, viewpoint symmetry and the special status of the frontal view, are distinctly different from the representational similarity obtained from a model of V1 responses to our face stimuli, which was analyzed using the same processing stream as the EEG data. For the



**Figure 3.** EEG topographies. Scalp topographies of the visually evoked potentials across time (shown for a subset of viewpoints and latencies). These data were used to estimate representational similarity matrices across time to reveal changes in the underlying encoding schemes.



**Figure 4.** Empirical similarity matrices and MDS results. The representational similarity matrices obtained from the high-dimensional EEG data were visualized in their original form and projected into two dimensions using MDS. (A) The development of representational similarity across time. Please note that the color code changes in each panel to best visualize the underlying structure. B, C, and D show data at time points representative of the overall development of the similarity structure based on their MDS projection. Symmetric viewpoints are connected via thin gray lines.

V1 model, neighboring viewpoints exhibited strong similarity, which decreased with increasing angular difference (Figure 1B and C). For more extreme viewpoints, a slightly increased similarity for mirror-symmetric viewpoints can be observed. This effect has been reported previously for a similar stimulus set of faces without hair (Kietzmann et al., 2012). However, the effect is small, compared with the strong effects of viewpoint symmetry

observed in the EEG data. A more detailed statistical analysis of the contribution of low-level similarity to the effects observed is provided below in the model-based analyses.

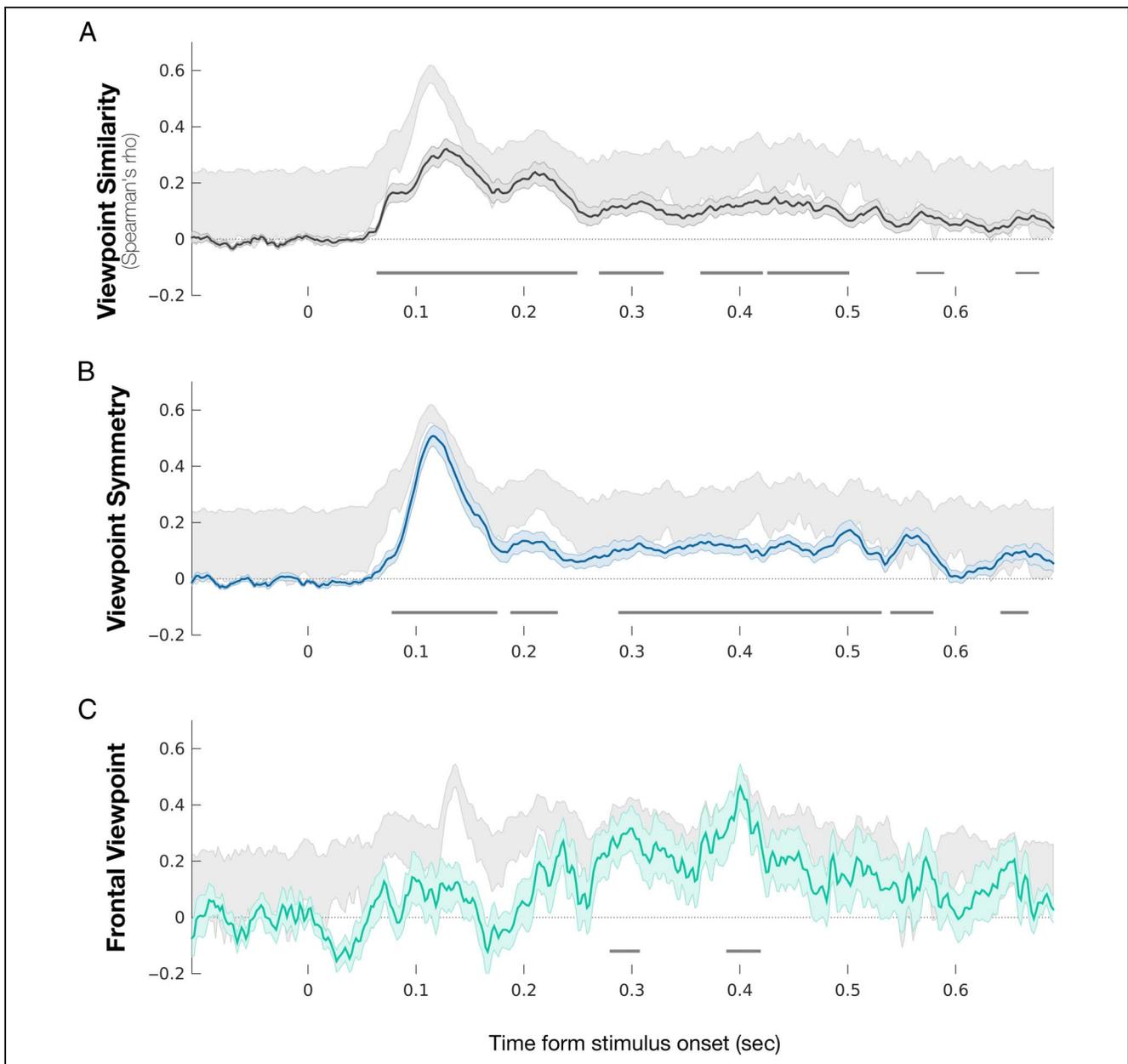
### Model-based Analyses

To better understand the dynamics of viewpoint encoding, we extended our data-driven analyses with a

model-based approach. Three models were defined in close agreement with the predictions of different task requirements: viewpoint similarity, an encoding scheme contributing to mechanisms of shared attention; viewpoint symmetry, supporting viewpoint-invariant mechanisms of face identification; and frontal viewpoint, focusing on the special status of frontal views, presumably differentiating direct from averted gaze.

The results of the three model fits are presented in Figure 5 and Table 1. All statistical analyses were cluster-corrected at a threshold of  $p < .05$  (cluster inclusion

subject to  $p < .01$  for each individual time point). The analysis of the viewpoint similarity model (Figure 5A) revealed multiple significant clusters. The earliest cluster emerged at around 60 msec after stimulus onset and remained statistically significant throughout the whole period analyzed (see Table 1 for exact latencies and cluster  $p$  values). These results indicate that viewpoints remain distinguishable from early to late periods of visual processing. Effects of viewpoint symmetry (Figure 5B) are expressed slightly later, starting after around 80 msec of visual processing, peaking around 115 msec. Finally,



**Figure 5.** Model results and noise ceiling (shown as gray backdrop). (A) Results of the viewpoint similarity model. Gray bars below the figure mark significant time windows, corrected for multiple comparisons using a cluster-based correction method. (B) The viewpoint symmetry model estimated whether mirror-symmetric viewpoints lead to larger similarity estimates as compared with nonsymmetric ones. (C) The frontal viewpoint model tested whether the frontal viewpoint leads to different similarity structures as compared with slightly oblique viewpoints, rotated 5° away from the frontal view.

**Table 1.** Cluster Statistics

| <i>Time</i>                 | <i>Cluster Statistic</i> | <i>Cluster <math>p</math></i> |
|-----------------------------|--------------------------|-------------------------------|
| <i>Viewpoint similarity</i> |                          |                               |
| 0.065–0.249                 | 550.5                    | <.001                         |
| 0.271–0.329                 | 111.8                    | .003                          |
| 0.365–0.421                 | 105.1                    | .003                          |
| 0.427–0.501                 | 141.2                    | .002                          |
| 0.565–0.589                 | 41.8                     | .017                          |
| 0.657–0.677                 | 36.3                     | .020                          |
| <i>Viewpoint symmetry</i>   |                          |                               |
| 0.079–0.175                 | 328.2                    | <.001                         |
| 0.189–0.231                 | 77.0                     | .006                          |
| 0.289–0.531                 | 522.3                    | <.001                         |
| 0.541–0.579                 | 91.9                     | .005                          |
| 0.643–0.667                 | 45.8                     | .016                          |
| <i>Frontal viewpoint</i>    |                          |                               |
| 0.281–0.307                 | 48.5                     | .007                          |
| 0.389–0.419                 | 71.0                     | .002                          |

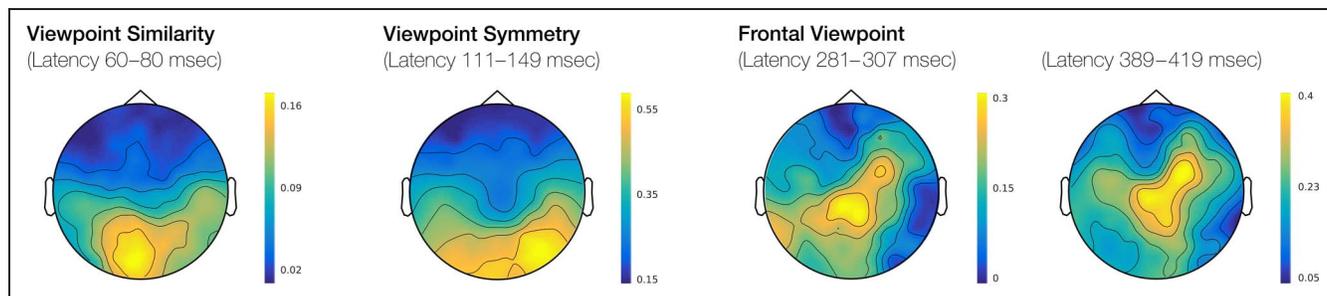
This table shows the latencies and statistics for all significant temporal clusters. Clusters with a  $p$  value smaller than .025 were considered significant (applied to both negative and positive clusters, this yields an overall cluster-corrected value of  $p < .05$ ). Cluster inclusion was contingent on an individual sample statistic of  $p < .01$ .

differences in similarity structure between frontal and directly neighboring views ( $\pm 5^\circ$ ) were observed in two later clusters (Figure 5C), starting at around 280 msec and ending around 420 msec. Taken together, our model-based analysis revealed a temporally distinct sequence of viewpoint encoding stages in the human visual system. Effects of viewpoint similarity occurred first, followed by effects of viewpoint symmetry, which are again followed by

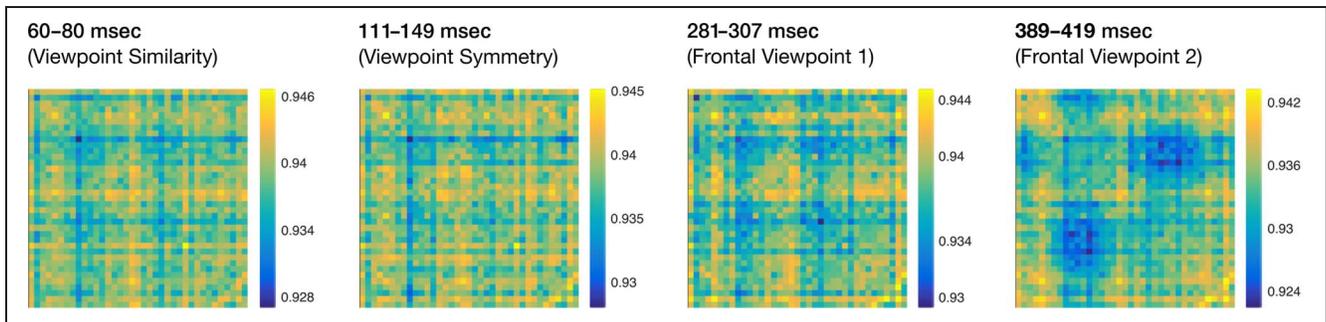
differences between frontal and oblique viewpoints that occurred at comparably late stages of processing. Interestingly, whereas the viewpoint symmetry model approaches the noise ceiling (i.e., maximal expected correlation for true model given the noise in the data) during the early phase of processing, the frontal viewpoint model reaches the noise ceiling at later stages. This may indicate a true change in the underlying neural encoding scheme from early to late face processing.

By definition, the viewpoint symmetry model contains aspects related to viewpoint similarity. It is therefore possible that the effects of viewpoint similarity contribute to the observed effects of viewpoint symmetry. To differentiate these two representational aspects, we split the viewpoint symmetry model into its two constituent diagonals: One part corresponds to viewpoint similarity ( $\backslash$ ); the other corresponds to the horizontally flipped counterpart ( $/$ ), called symmetry-only in the following. We then used a partial Spearman's rank correlation to estimate the effect of symmetry-only while controlling for effects of viewpoint similarity. This revealed a later effect of viewpoint symmetry, ranging from 111 to 149 msec after stimulus onset (cluster  $p = .003$ , cluster statistic = 88.2). This more conservative measure of viewpoint symmetry reaches statistical significance much later than the onset of viewpoint similarity, starting around 50 msec earlier, and therefore indicates two separate representational stages. The effects of frontal viewpoint remained unchanged after using viewpoint similarity as a control in a corresponding partial correlation analysis.

Having demonstrated a distinct representational sequence of viewpoint encoding schemes based on the whole set of EEG sensors, we performed a more fine-grained, sensor-based searchlight analysis to spatially localize the effects observed (Figure 6). The effects of viewpoint similarity were localized in the earliest time window (60–80 msec) to ensure that the effects of viewpoint symmetry are not (yet) present. In line with the early latency, the resulting effect topography suggests an occipital generator. Next, we localized the effects of viewpoint symmetry in a more conservative time window (111–149 msec). This revealed a more lateralized topography, in line with



**Figure 6.** Searchlight results. To better localize the observed effects, a sensor-based searchlight analysis was performed. Focusing on previously established temporal windows of interest for the given models, the results depict a progression from occipital to occipitotemporal to central sensor locations. Please note, however, that the analysis can only detect effects in spatially connected sensor clusters, omitting effects across spatially distinct regions.



**Figure 7.** Eye-tracking control. To test and control for the effects of residual eye movements, we computed similarity matrices based on the average fixation distribution of each participant in each condition for each time window of interest. The resulting similarity matrices revealed only minor effects, which run counter to the observed EEG effects.

stronger neural responses to faces in the right hemisphere. Finally, the special status of the frontal viewpoint (281–307 and 389–419 msec) was observed most strongly in central electrode locations.

We also performed a control analysis to ensure that any residual effects of eye position, which might have occurred despite the color change detection task at fixation, could not account for the results. Using the eye-tracking data recorded during the experimental trials, we estimated 2-D fixation density maps, one for every participant, condition, and time point. These were then used to compute eye-tracking similarity matrices for each participant (Figure 7). These individually defined matrices were then used as controls in a partial Spearman's rank correlation between model prediction and EEG data. That is, we asked whether the explanatory power of our models could be explained away by differential eye movements during the respective temporal window of interest. Focusing on previously established windows of interest and testing partial model correlations against zero, all model predictions remained statistically significant (viewpoint similarity:  $p < .0001$ ,  $t(15) = 5.8$ ; viewpoint symmetry:  $p < .0001$ ,  $t(15) = 9.1$ ; frontal viewpoint:  $p < .005$ ,  $t(15) = 3.9$ ;  $p < .0005$ ,  $t(15) = 5.4$ ). Residual effects of eye position are therefore unlikely to explain the effects observed.

Finally, to rule out explanations based on differences in attentional load introduced by the distractor task, we tested the behavioral performance of our participants across all viewpoint conditions using a repeated-measures ANOVA. We found no significant differences in  $d'$  ( $p = .34$ ,  $F(36, 15) = 1.09$ ), hit rate ( $p = .30$ ,  $F(36, 15) = 1.12$ ), or false alarm rate ( $p = .53$ ,  $F(36, 15) = 0.96$ ) across experimental conditions, indicating that differences in attentional load cannot account for the observed EEG effects.

## DISCUSSION

We investigated the representational dynamics of face viewpoint encoding in the human brain by applying multivariate analyses to high-dimensional EEG data, recorded at high temporal resolution. Our data-driven and model-

based analyses indicate a temporal sequence of distinct viewpoint encoding stages in the EEG activation patterns. Neural responses that reflected viewpoint similarity occurred first, followed by strong effects of viewpoint symmetry. At a later processing stage, the frontal view led to significantly different activation patterns compared with all other viewpoints. Ruling out alternative explanations for the effects observed, we showed that viewpoint symmetry and specialized processing of front-facing views cannot be explained in terms of viewpoint similarity and that none of the effects observed can be explained in terms of residual eye movements or differences in distractor task performance.

The distinction between different head orientations was present early and lasted throughout the entire analyzed time window of 700 msec. The early onset of this viewpoint similarity effect and the results of the sensor-based searchlight analysis suggest that the initial discrimination of head orientation was mainly driven by low-level stimulus properties. A possible explanation of the fact that the effects lasted throughout the whole trial is that low-level stimulus properties remain an important factor in visually evoked responses, even at later temporal stages of visual processing. These results are in line with the results of recent fMRI studies demonstrating residual low-level effects in higher-level visual areas (Rice, Watson, Hartley, & Andrews, 2014; Wardle & Ritchie, 2014; Kietzmann et al., 2012; Yue, Cassidy, Devaney, Holt, & Tootell, 2011). A second, nonexclusive explanation is that the initial sensory coding of head orientation with a specific similarity structure was transformed to action-oriented activity patterns that obey identical relations. Indeed, the mapping of head orientation to the location of joint attention is continuous, that is, similar visual stimuli indicate similar locations for joint attention. Thus, this interpretation is also compatible with an action-oriented interpretation of cortical representations (Engel, Maye, Kurthen, & König, 2013).

Using multivariate analyses of fMRI data, the effects of viewpoint symmetry were previously shown to be prevalent across a large range of higher-order visual areas (Kietzmann et al., 2012). In addition to replicating the general pattern of these results with a different method for measuring neural activity, our current findings extend

these results by providing an estimate of the latency of these face processing effects. In general agreement with electrophysiological data from macaque monkeys (Freiwald & Tsao, 2010), the effects of viewpoint symmetry were found to occur at around 110 msec after stimulus onset. This finding constrains possible models of viewpoint symmetry and renders explanations based on extended, recurrent processing less likely. Viewpoint symmetry was observed later than the effects of head orientation but significantly earlier than one might expect for fully viewpoint-invariant effects of face identification, as has been estimated in previous experimental work. For instance, behavioral data from a rapid go/no-go face identification task suggest that people need about 260 msec to identify a face, when accounting for the time required to perform the corresponding motor response (Barragan-Jason, Besson, Ceccaldi, & Barbeau, 2013). Electrophysiological experiments suggest a similar latency. Although the N170 component, with an onset of around 130 msec, was previously shown to exhibit effects of face identification (Jacques & Rossion, 2006), the N170 does not represent a fully viewpoint-invariant code (Caharel et al., 2015; Ewbank, Smith, Hancock, & Andrews, 2008; Miyakoshi, Kanayama, Nomura, Iidaka, & Ohira, 2008). Studies of a later face-selective N250 component, however, do show evidence of viewpoint invariance processing (Caharel et al., 2015; Schweinberger, Pickering, Jentsch, Burton, & Kaufmann, 2002). Finally, the estimated latency of viewpoint symmetry directly matches the typically observed latencies of the higher-level visual areas (Ghuman et al., 2014; Parvizi et al., 2012; Liu, Agam, Madsen, & Kreiman, 2009), which were previously found to exhibit viewpoint symmetry using fMRI (Kietzmann et al., 2012). Taken together, these results are in line with the proposal that viewpoint symmetry might act as an intermediate step in achieving full viewpoint invariance, after the effects of head orientation and preceding a fully invariant code (Freiwald & Tsao, 2010).

In line with the vast majority of face processing studies, we decided to present the face stimuli foveally to study cortical responses under more natural viewing conditions and to obtain an improved signal-to-noise ratio. Recently, studies have evaluated the extent to which viewpoint encoding effects can generalize across changes in retinal position in the fusiform face area (FFA). Testing across peripheral stimulus positions along the vertical meridian, no position-invariant effects of viewpoint symmetry were observed (Ramírez, Cichy, Allefeld, & Haynes, 2014). Although further experiments are required to fully understand the origins of this symmetry null effect, an explanation can be given based on the fact that higher-level visual representations are not position-invariant (Hong, Yamins, Majaj, & DiCarlo, 2016; Golomb & Kanwisher, 2012; Kravitz, Vinson, & Baker, 2008; Hung, Kreiman, Poggio, & DiCarlo, 2005). Furthermore, FFA exhibits a much reduced signal-to-noise ratio in response to stimuli presented in the periphery (Hasson, Levy, Behrmann, Hendler, & Malach, 2002; Lerner,

Hendler, & Malach, 2002; Levy, Hasson, Avidan, Hendler, & Malach, 2001) and is better tuned for foveally presented stimuli (Kay, Weiner, Kay, & Weiner, 2015). Both factors strongly affect the ability to find effects driven by distinct activation patterns.

In addition to these effects of viewpoint symmetry, we observed distinct cortical activation patterns for the front-facing viewpoint at a much later processing stage. As a possible explanation for this effect, we noted that the front view differed from the others in terms of direct versus averted gaze. Such effects, which are also known as the “eye contact effect,” were previously studied in the context of social attention (Nummenmaa & Calder, 2009) and social neuroscience (Itier & Batty, 2009; Senju & Johnson, 2009). Interestingly, we observed such effects while our participants performed a color change detection task at fixation, consistent with behavioral studies suggesting that automatic processing of direct gaze can occur, even when attention is drawn away from the face stimulus (Yokoyama, Sakai, Noguchi, & Kita, 2014). Despite the importance of direct eye contact for social cognitive inferences (Nummenmaa & Calder, 2009; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), the results of previous electrophysiological studies on the topic are mixed. Whereas one study suggested that the effects of eye contact occur after only 160 msec of processing (Conty, N’Diaye, Tijus, & George, 2007), others have not found differences between direct and averted gaze at similar latencies (Taylor, Itier, Allison, & Edmonds, 2001). The latter result is in line with univariate analyses performed in developmental studies that indicate diminishing effects of eye contact with adulthood (Grice, Halit, Farroni, & Baron-Cohen, 2005). Finally, data from intracranial recordings suggest that the effects of direct versus averted gaze occur during later stages of processing (>200 msec; Pourtois & Spinelli, 2010), in agreement with the current observations. Given the strength of the effects observed here, it would be interesting to revisit studies that produced negative results and to reanalyze the data using multivariate pattern analyses. As indicated above, the front-facing view is a reliable predictor for direct eye contact. It is not yet clear, however, which mechanisms could aid the visual system in detecting this particular viewpoint. A potential factor could be the reflectional symmetry of front-view faces. Studies of visuospatial regularity report a sustained posterior negativity for reflectionally symmetric dot patterns (Makin, Wilton, Pecchinenda, & Bertamini, 2012). Although previous effects of eye contact were shown based on nonsymmetric stimuli (Pourtois & Spinelli, 2010), demonstrating that the effects of eye contact exist independently of reflectional symmetry, it is conceivable that the reflectional symmetry of frontal face views contributes to their detection in a visual scene.

In search for a cortical origin of gaze direction effects, converging evidence from neurophysiology (Perrett et al., 1992; Perrett & Smith, 1985) and fMRI experiments has emphasized the role of the STS. In addition to the work by Carlin et al. (2011), experiments by Hoffman and Haxby

(2000) suggest that the processing of facial information is separated between the STS and FFA. Whereas the FFA was suggested to process information related to facial identity (Axelrod & Yovel, 2015), the STS was found to be involved in the processing of changeable aspects of faces, including the perception of gaze direction (Haxby et al., 2000). In line with this, Calder et al. (2007) used an fMRI adaptation paradigm to study the encoding of gaze direction and found that the anterior STS and inferior parietal cortex contain information that allows for a separation of different gaze directions.

It should be noted that head orientation and gaze direction were congruent in the currently used stimulus set. It was therefore not possible to separate cortical signals related to gaze and head orientation in this study. Evidence in favor of such a separation was provided by Carlin et al. (2011), who demonstrated that the STS contains finely graded information about the direction of gaze, independently of head orientation and physical image features. Such evidence of a high-level encoding of gaze direction is in line with data from a behavioral adaptation paradigm, in which effects of gaze direction were demonstrated despite changes in size and head direction (Jenkins, Beaver, & Calder, 2006). Moreover, electrophysiological studies using the same paradigm found late effects, starting around 250 msec after stimulus onset (Kloth & Schweinberger, 2010; Schweinberger, Kloth, & Jenkins, 2007). In context of the current analyses and results, it would be interesting to combine both methodologies, gaze adaptation and spatiotemporal pattern similarity, to investigate how gaze adaptation affects the dynamics of viewpoint encoding. Future work might further benefit from the currently used analysis approach to investigate the temporal development and separability of head and gaze direction signals.

In summary, our multivariate analyses of visually evoked potentials revealed that the cortical representations of facial viewpoints traverse a distinct sequence, expressing different encoding schemes at different latencies. Such representational stages may reflect the complexity of the underlying task and the priority that the brain devotes to the respective computation.

## Acknowledgments

The work was supported by the Research and Innovation Programs of the European Union (FP7-ICT-270212, H2020-FETPROACT-2014, and SEP-210141273), the European Research Council (ERC-2010-AdG #269716), and an DFG Postdoctoral Fellowship for Tim C. Kietzmann.

Reprint requests should be sent to Tim C. Kietzmann, Cognition and Brain Sciences Unit, Medical Research Council, 15 Chaucer Rd., Cambridge, United Kingdom CB2 7EF, or via e-mail: Tim.Kietzmann@mrc-cbu.cam.ac.uk.

## REFERENCES

Anzellotti, S., Fairhall, S. L., & Caramazza, A. (2013). Decoding representations of face identity that are tolerant to rotation. *Cerebral Cortex*, *24*, 1988–1995.

- Axelrod, V., & Yovel, G. (2012). Hierarchical processing of face viewpoint in human visual cortex. *Journal of Neuroscience*, *32*, 2442–2452.
- Axelrod, V., & Yovel, G. (2015). Successful decoding of famous faces in the fusiform face area. *PLoS One*, *10*, e0117126.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*, 241–251.
- Barragan-Jason, G., Besson, G., Ceccaldi, M., & Barbeau, E. J. (2013). Fast and famous: Looking for the fastest speed at which a face can be recognized. *Frontiers in Psychology*, *4*, 100.
- Bentin, S., Allison, T., & Puce, A. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*, 551–565.
- Bülthoff, H. H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences, U.S.A.*, *89*, 60–64.
- Caharel, S., Collet, K., & Rossion, B. (2015). The early visual encoding of a face (N170) is viewpoint-dependent: A parametric ERP-adaptation study. *Biological Psychology*, *106*, 18–27.
- Calder, A. J., Beaver, J. D., Winston, J. S., Dolan, R. J., Jenkins, R., Eger, E., et al. (2007). Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. *Current Biology*, *17*, 20–25.
- Carlin, J. D., Calder, A. J., Kriegeskorte, N., Nili, H., & Rowe, J. B. (2011). A head view-invariant representation of gaze direction in anterior superior temporal sulcus. *Current Biology*, *17*, 1817–1821.
- Carlin, J. D., Rowe, J. B., Kriegeskorte, N., Thompson, R., & Calder, A. J. (2012). Direction-sensitive codes for observed head turns in human superior temporal sulcus. *Cerebral Cortex*, *22*, 735–744.
- Carlson, T., Tovar, D., Alink, A., & Kriegeskorte, N. (2013). Representational dynamics of object vision: The first 1000 ms. *Journal of Vision*, *13*, 1–19.
- Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space and time. *Nature Neuroscience*, *17*, 455–462.
- Conty, L., N'Diaye, K., Tijus, C., & George, N. (2007). When eye creates the contact! ERP evidence for early dissociation between direct and averted gaze motion processing. *Neuropsychologia*, *45*, 3024–3037.
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21.
- DiCarlo, J. J., & Cox, D. D. (2007). Untangling invariant object recognition. *Trends in Cognitive Sciences*, *11*, 333–341.
- Dubois, J., de Berker, A. O., & Tsao, D. Y. (2015). Single-unit recordings in the macaque face patch system reveal limitations of fMRI MVPA. *Journal of Neuroscience*, *35*, 2791–2802.
- Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *NeuroReport*, *11*, 2319–2324.
- Eimer, M. (2011). The face-sensitivity of the n170 component. *Frontiers in Human Neuroscience*, *5*, 119.
- Engel, A. K., Maye, A., Kurthen, M., & König, P. (2013). Where's the action? The pragmatic turn in cognitive science. *Trends in Cognitive Sciences*, *17*, 202–209.
- Ewbank, M. P., Smith, W. A. P., Hancock, E. R., & Andrews, T. J. (2008). The M170 reflects a viewpoint-dependent

- representation for both familiar and unfamiliar faces. *Cerebral Cortex*, *18*, 364–370.
- Freiwald, W. A., & Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science*, *330*, 845–851.
- Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform “face area” is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, *12*, 495–504.
- Ghuman, A. S., Brunet, N. M., Li, Y., Konecky, R. O., Pyles, J. A., Walls, S. A., et al. (2014). Dynamic encoding of face information in the human fusiform gyrus. *Nature Communications*, *5*, 5672.
- Golomb, J. D., & Kanwisher, N. (2012). Higher level visual cortex represents retinotopic, not spatiotopic, object location. *Cerebral Cortex*, *22*, 2794–2810.
- Good, P. (2013). *Permutation tests: A practical guide to resampling methods for testing hypotheses*. New York: Springer.
- Grice, S., Halit, H., Farroni, T., & Baron-Cohen, S. (2005). Neural correlates of eye-gaze direction in young children. *Cortex*, *141*, 342–353.
- Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron*, *34*, 479–490.
- 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.
- Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, *3*, 80–84.
- Hong, H., Yamins, D. L. K., Majaj, N. J., & DiCarlo, J. J. (2016). Explicit information for category-orthogonal object properties increases along the ventral stream. *Nature Neuroscience*, *19*, 613–622.
- Hung, C. P., Kreiman, G., Poggio, T., & DiCarlo, J. J. (2005). Fast readout of object identity from macaque inferior temporal cortex. *Science*, *310*, 863–866.
- Itier, R. J., & Batty, M. (2009). Neural bases of eye and gaze processing: The core of social cognition. *Neuroscience and Biobehavioral Reviews*, *33*, 843–863.
- Itier, R. J., Herdman, A. T., George, N., Cheyne, D., & Taylor, M. J. (2006). Inversion and contrast-reversal effects on face processing assessed by MEG. *Brain Research*, *1115*, 108–120.
- Jacques, C., & Rossion, B. (2006). The speed of individual face categorization. *Psychological Science*, *17*, 485–492.
- Jenkins, R., Beaver, J. D., & Calder, A. J. (2006). I thought you were looking at me: Direction-specific aftereffects in gaze perception. *Psychological Science*, *17*, 506–513.
- Joyce, C. A., Schyns, P. G., Gosselin, F., Cottrell, G. W., & Rossion, B. (2006). Early selection of diagnostic facial information in the human visual cortex. *Vision Research*, *46*, 800–813.
- 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.
- Kay, K. N., Weiner, K. S., Kay, K. N., & Weiner, K. S. (2015). Attention reduces spatial uncertainty in human ventral temporal cortex. *Current Biology*, *25*, 595–600.
- Khaligh-Razavi, S.-M., & Kriegeskorte, N. (2014). Deep supervised, but not unsupervised, models may explain IT cortical representation. *PLoS Computational Biology*, *10*, 1–29.
- Kietzmann, T. C., Poltoratski, S., König, P., Blake, R., Tong, F., & Ling, S. (2015). The occipital face area is causally involved in viewpoint symmetry judgments of faces. *Journal of Neuroscience*, *35*, 16398–16403.
- Kietzmann, T. C., Swisher, J. D., König, P., & Tong, F. (2012). Prevalence of selectivity for mirror-symmetric views of faces in the ventral and dorsal visual pathways. *Journal of Neuroscience*, *32*, 11763–11772.
- Kloth, N., & Schweinberger, S. (2010). Electrophysiological correlates of eye gaze adaptation. *Journal of Vision*, *10*, 1–13.
- Kravitz, D. J., Vinson, L. D., & Baker, C. I. (2008). How position dependent is visual object recognition? *Trends in Cognitive Sciences*, *12*, 114–122.
- Kriegeskorte, N., Formisano, E., Sorger, B., & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *104*, 20600–20605.
- Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 4.
- Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., et al. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, *60*, 1–16.
- Lee, T. (1996). Image representation using 2D Gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *18*, 1–13.
- Lerner, Y., Hendler, T., & Malach, R. (2002). Object-completion effects in the human lateral occipital complex. *Cerebral Cortex*, *12*, 163–177.
- Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience*, *4*, 533–539.
- Liu, H., Agam, Y., Madsen, J. R., & Kreiman, G. (2009). Timing, timing, timing: Fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron*, *62*, 281–290.
- Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience*, *5*, 910–916.
- Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology*, *5*, 552–563.
- Makin, A., Wilton, M., Pecchinenda, A., & Bertamini, M. (2012). Symmetry perception and affective responses: A combined EEG/EMG study. *Neuropsychologia*, *50*, 3250–3261.
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*, 177–190.
- McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex*, *9*, 431–444.
- Miyakoshi, M., Kanayama, N., Nomura, M., Iidaka, T., & Ohira, H. (2008). ERP study of viewpoint-independence in familiar-face recognition. *International Journal of Psychophysiology*, *69*, 119–126.
- Natu, V. S., Jiang, F., Narvekar, A., Keshvari, S., Blanz, V., & O’Toole, A. J. (2010). Dissociable neural patterns of facial identity across changes in viewpoint. *Journal of Cognitive Neuroscience*, *22*, 1570–1582.
- Nili, H., Wingfield, C., Walthers, A., Su, L., Marslen-Wilson, W., & Kriegeskorte, N. (2014). A toolbox for representational similarity analysis. *PLoS Computational Biology*, *10*, e1003553.
- Nummenmaa, L., & Calder, A. J. (2009). Neural mechanisms of social attention. *Trends in Cognitive Sciences*, *13*, 135–143.
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 1–10.

- Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, *112*, 713–719.
- Palmer, J. A., Makeig, S., Kreutz-Delgado, K., & Rao, B. D. (2008). Newton method for the ICA picture model. In *Proceedings of the 33rd IEEE International Conference on Acoustics and Signal Processing* (pp. 1805–1808). Las Vegas, NV.
- Parvizi, J., Jacques, C., Foster, B. L., Witthoft, N., Withoft, N., Rangarajan, V., et al. (2012). Electrical stimulation of human fusiform face-selective regions distorts face perception. *Journal of Neuroscience*, *32*, 14915–14920.
- Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *335*, 23–30.
- Perrett, D. I., Oram, M. W., & Ashbridge, E. (1998). Evidence accumulation in cell populations responsive to faces: An account of generalisation of recognition without mental transformations. *Cognition*, *67*, 111–145.
- Perrett, D. I., Oram, M. W., Harries, M. H., Bevan, R., Hietanen, J. K., Benson, P. J., et al. (1991). Viewer-centered and object-centered coding of heads in the macaque temporal cortex. *Experimental Brain Research*, *86*, 159–173.
- Perrett, D. I., & Smith, P. A. J. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, *223*, 293–317.
- Pitcher, D., Walsh, V., & Duchaine, B. (2011). The role of the occipital face area in the cortical face perception network. *Experimental Brain Research*, *209*, 481–493.
- Plöchl, M., Ossandón, J. P., & König, P. (2012). Combining EEG and eye tracking: Identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Frontiers in Human Neuroscience*, *6*, 278.
- Pourtois, G., Schwartz, S., Seghier, M. L., Lazeyras, F., & Vuilleumier, P. (2005). Portraits or people? Distinct representations of face identity in the human visual cortex. *Journal of Cognitive Neuroscience*, *17*, 1043–1057.
- Pourtois, G., & Spinelli, L. (2010). Modulation of face processing by emotional expression and gaze direction during intracranial recordings in right fusiform cortex. *Journal of Cognitive Neuroscience*, *22*, 2086–2107.
- Ramírez, F. M., Cichy, R. M., Allefeld, C., & Haynes, J.-D. (2014). The neural code for face orientation in the human fusiform face area. *Journal of Neuroscience*, *34*, 12155–12167.
- Rice, G. E., Watson, D. M., Hartley, T., & Andrews, T. J. (2014). Low-level image properties of visual objects predict patterns of neural response across category-selective regions of the ventral visual pathway. *Journal of Neuroscience*, *34*, 8837–8844.
- Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage*, *39*, 1959–1979.
- Rossion, B., Prieto, E. A., Boremanse, A., Kuefner, D., & Van Belle, G. (2012). A steady-state visual evoked potential approach to individual face perception: Effect of inversion, contrast-reversal and temporal dynamics. *Neuroimage*, *63*, 1585–1600.
- Schweinberger, S. R., Kaufmann, J. M., Moratti, S., Keil, A., & Burton, A. M. (2007). Brain responses to repetitions of human and animal faces, inverted faces, and objects: An MEG study. *Brain Research*, *1184*, 226–233.
- Schweinberger, S. R., Kloth, N., & Jenkins, R. (2007). Are you looking at me? Neural correlates of gaze adaptation. *NeuroReport*, *18*, 693–696.
- Schweinberger, S. R., Pickering, E. C., Jentsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Cognitive Brain Research*, *14*, 398–409.
- Senju, A., & Johnson, M. H. (2009). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences*, *13*, 127–134.
- Serre, T., & Riesenhuber, M. (2004). Realistic modeling of simple and complex cell tuning in the HMAX model, and implications for invariant object recognition in cortex. Technical Report CBCL Paper 239/AI Memo 2004–2017. Retrieved from <http://18.7.29.232/handle/1721.1/30491>.
- Taylor, M. J., Itier, R. J., Allison, T., & Edmonds, G. E. (2001). Direction of gaze effects on early face processing: Eyes-only versus full faces. *Cognitive Brain Research*, *10*, 333–340.
- Tong, F., Nakayama, K., Moscovitch, M., Weinrib, O., & Kanwisher, N. (2000). Response properties of the human fusiform face area. *Cognitive Neuropsychology*, *17*, 257–280.
- Ullman, S. (1998). Three-dimensional object recognition based on the combination of views. *Cognition*, *67*, 21–44.
- Vetter, T., Poggio, T., & Bühlhoff, H. H. (1994). The importance of symmetry and virtual views in three-dimensional object recognition. *Current Biology*, *4*, 18–23.
- Wardle, S. G., & Ritchie, J. B. (2014). Can object category-selectivity in the ventral visual pathway be explained by sensitivity to low-level image properties? *Journal of Neuroscience*, *34*, 14817–14819.
- Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, *42*, 671–684.
- Yokoyama, T., Sakai, H., Noguchi, Y., & Kita, S. (2014). Perception of direct gaze does not require focus of attention. *Scientific Reports*, *4*, 3858.
- Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology*, *15*, 2256–2262.
- Yue, X., Cassidy, B. S., Devaney, K. J., Holt, D. J., & Tootell, R. B. H. (2011). Lower-level stimulus features strongly influence responses in the fusiform face area. *Cerebral Cortex*, *21*, 35–47.