bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The neural representation of personally familiar and unfamiliar faces in the distributed system for face perception

Matteo Visconti di Oleggio Castello[1,+,*], Yaroslav O. Halchenko[1,+,*], J. Swaroop Guntupalli[1], Jason D. Gors[1], M. Ida Gobbini[1, 2, *]

1: Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755, USA

2: Dipartimento di Medicina Specialistica, Diagnostica e Sperimentale (DIMES), Medical School, University of Bologna, Bologna, Italy

+: these authors contributed equally to the work

*: corresponding authors

M. Ida Gobbini: mariaida.gobbini@unibo.it

Matteo Visconti di Oleggio Castello: mvdoc.gr@dartmouth.edu

Yaroslav O. Halchenko: yoh@dartmouth.edu

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

Personally familiar faces are processed more robustly and efficiently than unfamiliar faces. The human face processing system comprises a core system that analyzes the visual appearance of faces and an extended system for the retrieval of person-knowledge and other nonvisual information. We applied multivariate pattern analysis to fMRI data to investigate aspects of familiarity that are shared by all familiar identities and information that distinguishes specific face identities from each other. Both identity-independent familiarity information and face identity could be decoded in an overlapping set of areas in the core and extended systems. Representational similarity analysis revealed a clear distinction between the two systems and a subdivision of the core system into ventral, dorsal and anterior components. This study provides evidence that activity in the extended system carries information about both individual identities and personal familiarity, while clarifying and extending the organization of the core system for face perception.

Keywords: personally familiar faces; mvpa; decoding; brain networks; core and extended systems; representational similarity analysis.

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

A wide and distributed network of brain areas underlies face processing. The model by Haxby and colleagues (Gobbini & Haxby, 2007; Haxby & Gobbini, 2011; Haxby, Hoffman, & Gobbini, 2000) posited a division between a core system involved in the processing the visual appearance of faces—comprising the Occipital Face Area (OFA), the Fusiform Face Area (FFA), and the posterior Superior Temporal Sulcus (pSTS)—and an extended system, comprising parietal, frontal, and subcortical areas, involved in inferring socially relevant information from faces, such as direction of attention, intentions, emotions, and retrieval of person knowledge (Gobbini, 2010; Gobbini & Haxby, 2007; Haxby & Gobbini, 2011; Haxby et al., 2000).

The definition of the core system has been extended to include areas in the anterior fusiform gyrus (the anterior temporal face area, ATFA; Collins & Olson, 2014; Rajimehr, Young, & Tootell, 2009), the anterior superior temporal sulcus (aSTS-FA; Carlin, Calder, Kriegeskorte, Nili, & Rowe, 2011; Duchaine & Yovel, 2015; Pitcher, Dilks, Saxe, Triantafyllou, & Kanwisher, 2011), and the inferior frontal gyrus (IFG-FA; Duchaine & Yovel, 2015; Guntupalli, Wheeler, & Gobbini, 2017; J. V. Haxby et al., 1994). For example, in a recent fMRI neural decoding study with visually familiar faces (Guntupalli et al., 2017), we showed that the representation of face identity is progressively disentangled from image-specific features along the ventral visual pathway. While early visual cortex and the OFA represented head view independently of the identity of the face, we recorded an intermediate level of representation in the FFA in which identity was emerging but was still entangled with head view. The human face processing pathway culminated in the right ATFA and IFG-FA where we recorded a view-invariant representation of face identity.

While both unfamiliar and familiar faces effectively activate the core system (Duchaine & Yovel, 2015; & Haxby, 2006; Guntupalli et al., 2017; Natu & O'Toole, 2011; Pitcher et al., 2011), familiar faces activate the extended system more strongly than unfamiliar faces (Bobes, Lage Castellanos, Quiñones, García, & Valdes-Sosa, 2013; Cloutier, Kelley, & Heatherton, 2011; Gobbini & Haxby, 2007; Natu & O'Toole, 2011; Taylor et al., 2009). Personally familiar faces recruit Theory of Mind (ToM) areas such as the

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

medial prefrontal cortex (MPFC) and the temporo-parietal junction (TPJ), because they are more strongly associated with person knowledge (Cloutier et al., 2011; Gobbini & Haxby, 2007; Gobbini, Leibenluft, Santiago, & Haxby, 2004); they activate the precuneus and the anterior temporal cortices, suggesting retrieval of long-term episodic memories; they modulate the activity in the amygdala and insula, suggesting an increased emotion processing (Gobbini & Haxby, 2007; Gobbini et al., 2004; Natu & O'Toole, 2011). Because the core and extended systems have been mostly studied separately, we lack a clear understanding of how personal familiarity, consolidated through repeated interactions, affects the representations in the core system, and how core and extended systems interact to create the known behavioral advantages for personally familiar faces.

The behavioral literature on face processing (Bruce, Henderson, Newman, & Burton, 2001; A. M. Burton, Wilson, Cowan, & Bruce, 1999; Gobbini et al., 2013; Ramon, Vizioli, Liu-Shuang, & Rossion, 2015; Visconti di Oleggio Castello, di Oleggio Castello, Wheeler, Cipolli, & Gobbini, 2016; Visconti di Oleggio Castello, Guntupalli, Yang, & Gobbini, 2014; Visconti di Oleggio Castello & Gobbini, 2015) suggests that, despite the subjective impression of efficient or "expert" perception of natural faces (Diamond & Carey, 1986), only familiar faces are detected and recognized more robustly and efficiently, in stark contrast with the surprisingly inefficient identification of unfamiliar faces. Recognition of personally familiar faces is highly accurate even when images are severely degraded, while recognition of unfamiliar faces is markedly impaired by variation in head position or lighting, even with good image quality (Bruce et al., 2001; A. Mike Burton, Jenkins, & Schweinberger, 2011; A. M. Burton et al., 1999; Hancock, Bruce, & Burton, 2000; Jenkins & Burton, 2011). Detection of personally familiar faces is facilitated even in conditions of reduced attentional resources and without awareness (Gobbini et al., 2013).

The representations of familiar and unfamiliar faces may differ in multiple ways. Familiar identities could have more robust, individually-specific representations, which are learned and consolidated over the course of personal interactions. Alternatively, familiar face representations could be enhanced with attributes that are similar across many personally familiar faces. For example, personally familiar faces (especially those

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

used in the present and our previous experiments that are faces of close relatives of personal friends) are associated with person-knowledge and emotional attachment that lead to social interactions that are different from the interactions with strangers, and these attributes may be shared across many familiar—one may be more open and unguarded with family and personal friends (Gobbini et al., 2004).

Here we applied multivariate pattern analyses (MVPA; Haxby et al., 2001; Haxby, Connolly, & Guntupalli, 2014), including MVP classification (MVPC) and representational similarity analysis (RSA; Kriegeskorte & Kievit, 2013) with two goals in mind. First, we wanted to dissociate familiarity information from identity information in the core and extended systems. Second, we wanted to investigate the relationships among core and extended face processing areas by examining the similarities of their representational spaces using second-order representational geometry (Guntupalli et al., 2016; Kriegeskorte & Kievit, 2013; Kriegeskorte, Mur, & Bandettini, 2008).

We first derived independent neural measures of identification and familiarity. To prevent any effect of familiarity information in identity decoding, we performed identity classification separately for familiar and unfamiliar faces. To control for the effect of identity-specific visual information in familiarity decoding, we trained classifiers to distinguish familiar from unfamiliar faces, and tested them on left-out identities. The results replicated the distinction between the representations of personally familiar and unfamiliar faces in the extended system that was previously revealed only with univariate analysis (Gobbini & Haxby, 2007), showing that this effect captured factors that were common across familiar faces and invariant across identities.

To unravel the representational structure of the face processing network, we investigated the relationships among the areas of the core and extended system uncovered by the classification analyses. Using the approach used by Guntupalli et al. (2016) (see also Kriegeskorte et al., 2008), we studied the similarities between representational geometries (Kriegeskorte & Kievit, 2013) in different face-processing areas (second-order representational geometry). This analysis revealed clear distinctions between the core system and the extended system, supporting the model by Gobbini & Haxby (2007), Haxby & Gobbini (2011), Haxby et al. (2000). In addition,

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the results support the extension of the core system to more anterior areas, such as the ATFA, the aSTS-FA and IFG-FA (Collins & Olson, 2014; Duchaine & Yovel, 2015; Fairhall & Ishai, 2007; Guntupalli et al., 2017; Rajimehr et al., 2009), and reveal a finer subdivision of this system into ventral, dorsal, and anterior components.

Results

In this experiment, we investigated the face processing network while participants performed an oddball-detection task with faces of friends and strangers (see Figure 1). We first investigated which areas responded more strongly to familiar faces than unfamiliar ones with a standard GLM analysis. Because familiarity information (whether a face is a familiar one) is necessarily confounded with identity information (who that person is), we next used MVPC to dissociate which areas of the core and extended system encode identity-independent familiarity information (familiar vs. unfamiliar classification across identities), and which parts of the network encode identity information. We performed two classification analyses using different cross-validation schemes to control for the effect of identity on the representation of general familiarity and to control for the effect of familiarity on the representation of identity. For the familiarity classification, we employed a leave-two-identities-out cross-validation scheme, where the classifier was trained on six faces (three familiar, three unfamiliar) to distinguish between familiar and unfamiliar faces, and tested on two left-out identities. This cross-validation scheme reduced the effect of identity information (see Supplementary Figures 1 and 2). For the identity classification, we decoded the four familiar faces and the four unfamiliar faces separately to eliminate the effect of familiarity information in the classification of identity information. Finally, we investigated the network structure derived from the similarities of representations to investigate relationships among areas in the core and extended system.

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

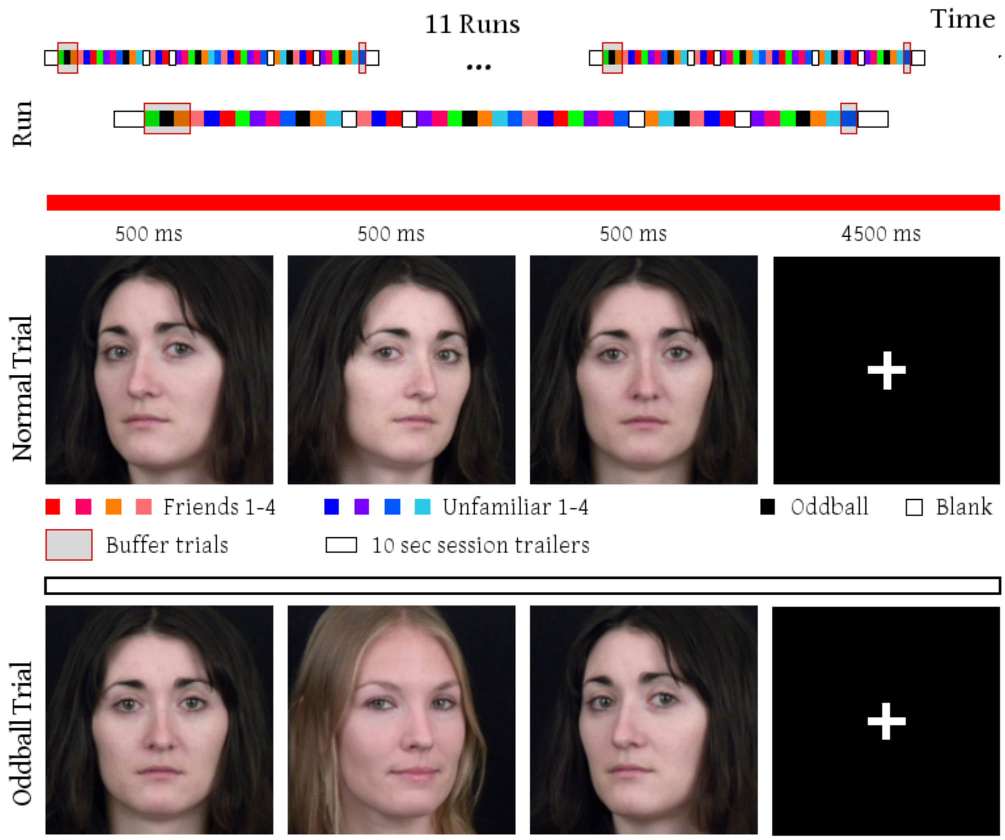


Figure 1: Slow event-related fMRI design. During each trial, images were presented in sequences of three pictures of the same identity (normal trial) or two different identities (oddball trials) in front-view or 30-degree profile views. Subjects engaged in an oddball-detection task to ensure that they paid attention to each stimulus.

GLM

In the univariate analysis contrasting Familiar > Unfamiliar we found significant activation in bilateral MTG/STS extending along the full length of the right STS. Additionally, we found significant clusters in the bilateral precuneus and bilateral MPFC, as well as in the right IFG. Familiar faces also evoked stronger responses in the left mid fusiform gyrus and the right anterior fusiform gyrus near the locations of the FFA (Grill-Spector & Weiner, 2014; Weiner et al., 2013) and ATFA (Collins & Olson, 2014). For the contrast Unfamiliar > Familiar we found only one significant cluster in

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the right inferior parietal lobule encroaching on the TPJ. Figure 2 shows the resulting statistical maps projected on the surface.

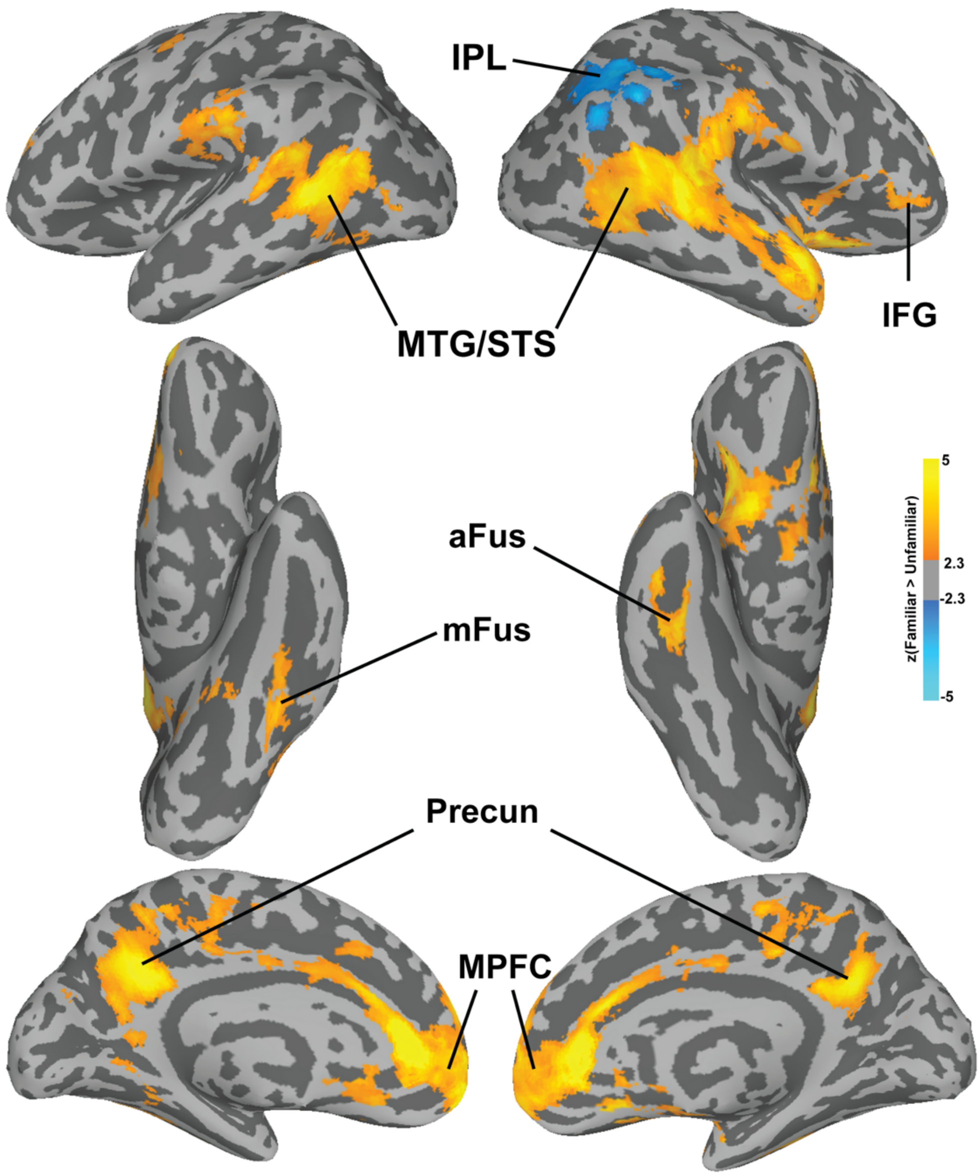


Figure 2. Cluster-corrected ($p < .05$) z-values for the univariate contrast Familiar > Unfamiliar. Abbreviations: IPL: inferior parietal lobule; mFus: middle fusiform gyrus; aFus: anterior fusiform gyrus; TPJ: temporo-parietal junction; MTG/STS: middle temporal gyrus/superior temporal sulcus; Precun: precuneus; MPFC: medial prefrontal cortex; IFG: inferior frontal gyrus.

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

MVPA

Familiarity Classification

The results of searchlight MVPC of identity-independent familiarity largely overlapped with the univariate maps, showing significant classification in the bilateral MTG/STS, mid and anterior right fusiform gyrus, right IFG, TPJ, precuneus, and MPFC (Figure 3). Surprisingly, small patches of cortex in early visual cortex also showed significant MVPC of identity-independent familiarity. We further investigated MVPC in early visual cortex with additional analyses on probabilistic ROI masks from Wang, Mruczek, Arcaro, & Kastner (2015), and found statistically significant decoding performance in V2 and V3 (see Supplementary Methods and Supplementary Figure 7). Since testing was performed on left-out familiar and unfamiliar identities, and all pictures were taken with the same equipment and settings, it is unlikely that this result was due simply to low-level features that distinguished familiar from unfamiliar faces. To test this further, we extracted features from the layers C1 and C2 of the HMAX model (Riesenhuber & Poggio, 1999; Serre, Wolf, Bileschi, Riesenhuber, & Poggio, 2007) and performed the same classification analysis, and found that decoding performance was not statistically significant (accuracy with C1 features 52%, $p = 0.66$; accuracy with C2 features 49%, $p = 0.95$; see Supplementary Methods and Supplementary Figure 8).

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

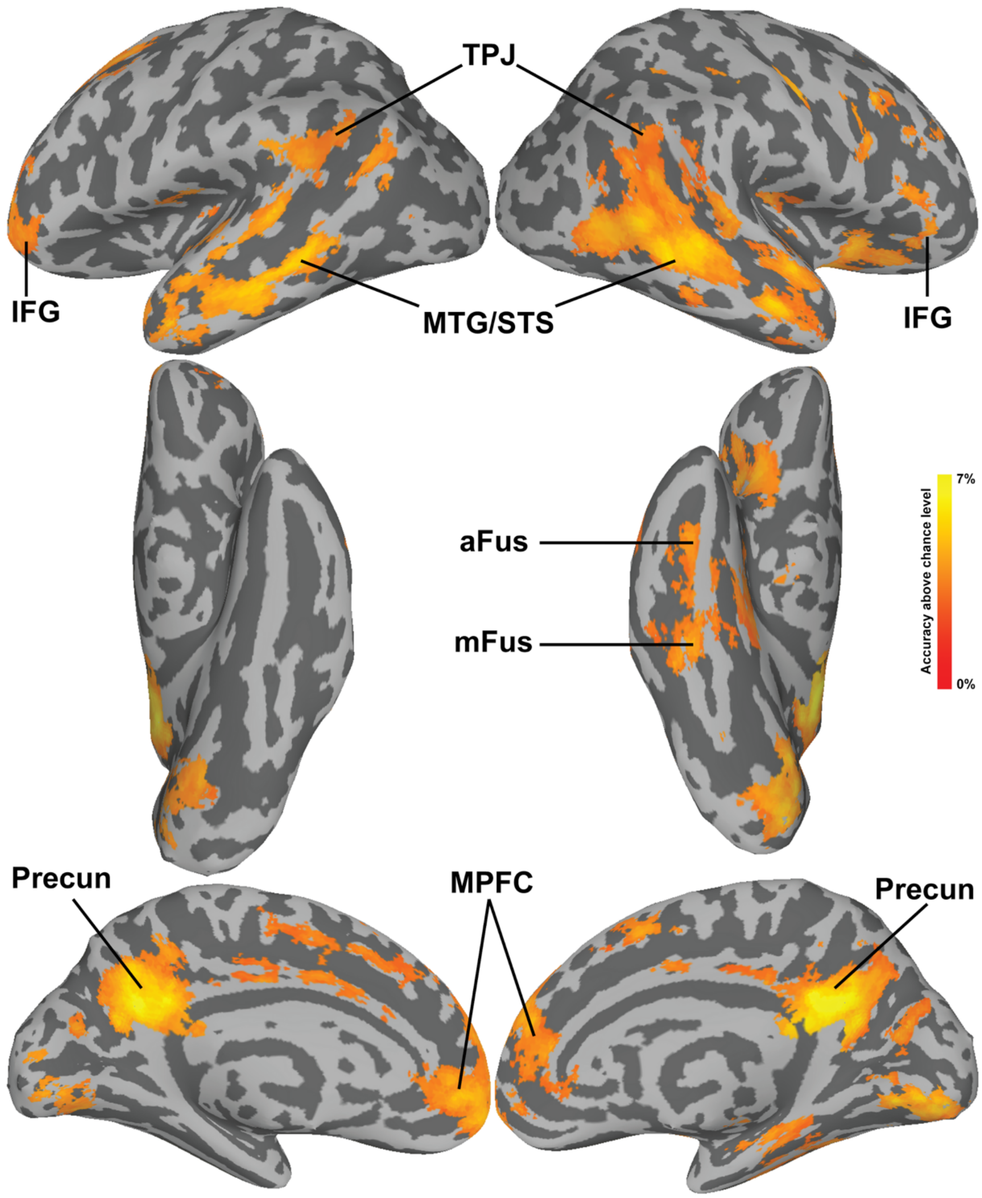


Figure 3. Searchlight maps for the Familiarity classification projected onto the surface. Maps were thresholded at a z-TFCE score of 1.65, corresponding to $p < 0.05$ one-tailed (corrected for multiple comparisons). Abbreviations: mFus: middle fusiform gyrus; aFus: anterior fusiform gyrus; TPJ: temporo-parietal junction; MTG/STS: middle temporal gyrus/superior temporal sulcus; Precun: precuneus; MPFC: medial prefrontal cortex; IFG: inferior frontal gyrus.

180

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Identity Classification

The identity classification analysis showed that identity could be decoded in many of the same areas as identity-independent familiarity (Figure 4). Significant classification was found in the MPFC and precuneus, and in the bilateral MTG/STS, TPJ, and IFG. The area in the precuneus with significant identity classification, however, was quite dorsal, whereas that for significant familiarity classification was ventral and included the posterior cingulate. Identity classification was significant in bilateral visual cortex starting in EV and extending to occipital, posterior, and mid fusiform cortices. Although MVPC of familiar identities showed a weak trend towards higher accuracies than for unfamiliar identities in the IFG and MTG/STS (Supplementary Figures 4, 5, and 6), these differences were not significant despite the large number of subjects.

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

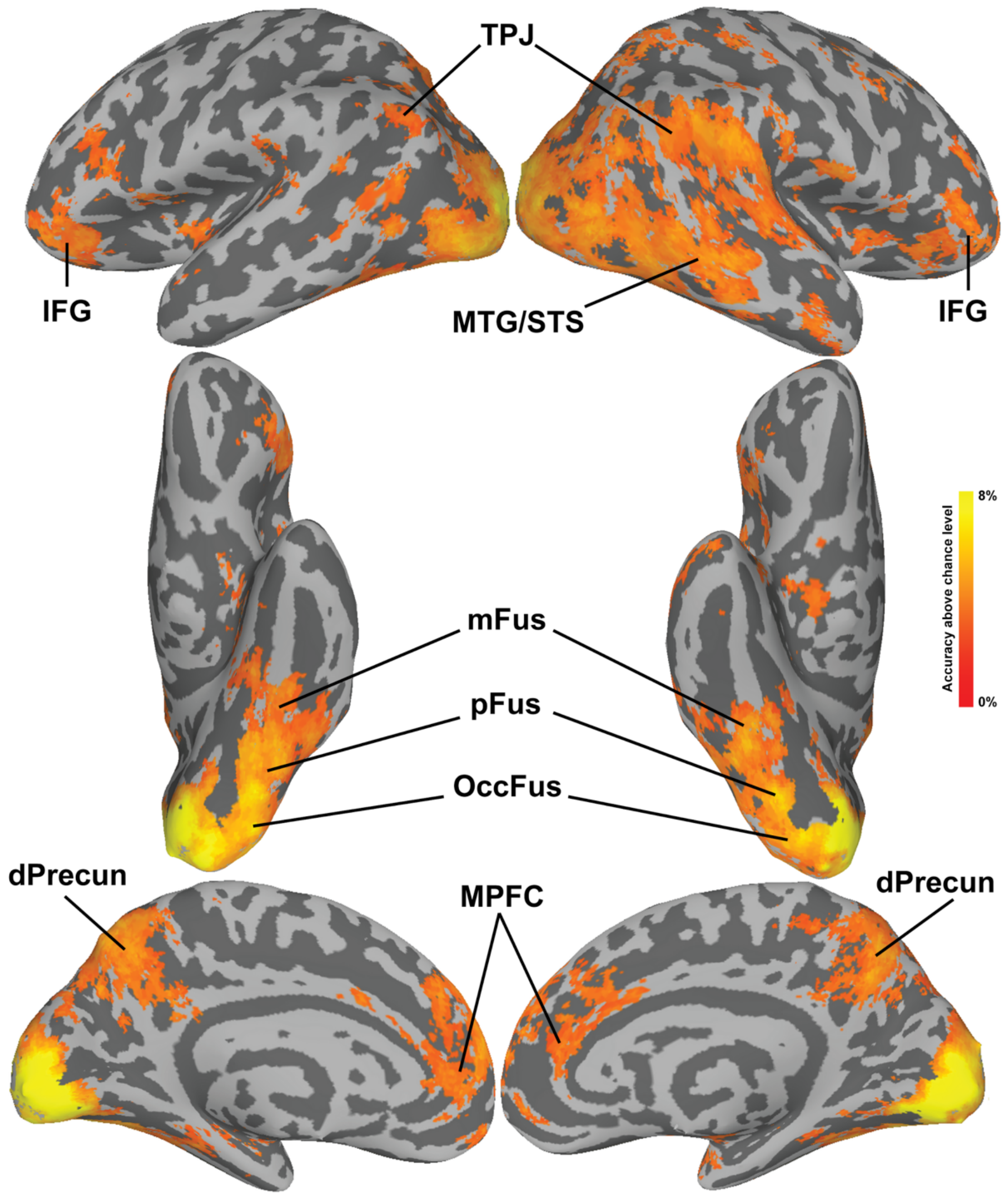


Figure 4. Searchlight maps for the Identity classification. The classification was run separately for familiar and unfamiliar identities (4-way), and the resulting maps were averaged. Maps were thresholded at a z-TFCE score of 1.65, corresponding to $p < 0.05$ one-tailed (corrected for multiple comparisons). Abbreviations: OccFus: occipital fusiform gyrus; pFus: posterior fusiform gyrus; mFus: middle fusiform gyrus; TPJ: temporo-parietal Junction; MTG/STS: middle temporal gyrus/superior temporal sulcus; dPrecun: dorsal precuneus; MPFC: medial prefrontal cortex; IFG: inferior frontal gyrus.

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

ROI Analysis and Second-order Representational Geometry

We investigated the relationships among the areas uncovered by the classification analysis as a second-order, inter-areal representational geometry. We selected 30 spherical ROIs (see Methods for how they were selected, Figure 5 for their location, and Supplementary Table 1 for their MNI coordinates) and computed a cross-validated representational dissimilarity matrix (Henriksson et al., 2015) in each ROI. We then constructed a distance matrix quantifying the similarity of these RDMs between all pairs of ROIs. Then, we computed an MDS solution to visualize the geometry of this inter-ROI matrix. Figure 6 shows the results of a 2D MDS. Supplementary Figure 10 shows the distance matrix, and Supplementary Figures 11 and 12 show the full MDS solution.

The 2D solution captured relationships among areas in the ventral portion of the core system in the first dimension, and relationships among areas in the dorsal and anterior parts of the core system and areas in the extended system in the second dimension. The first dimension showed a progression from EV areas to the posterior, mid, and anterior fusiform areas. Extended system areas were all at the distant end of the first dimension, as were the areas in the dorsal part of the core system (MTG/STS) and the IFG. The second dimension captured distinctions among these extended and core system areas, with the precuneus areas clustered together at one end, the MPFC and TPJ in the middle, and the dorsal and anterior core system areas at the other end.

We replicated this second-order RSA on an independent fMRI dataset collected while different subjects watched a full-length audiovisual movie, *Raiders of the Lost Ark* (Haxby et al. 2011; Guntupalli et al. 2016). This naturalistic stimulus contained a rich variety of dynamic faces that rapidly became familiar while the plot unfolded. The inter-ROI similarity matrix and MDS plot replicated the results based on representational geometry for the eight faces in the experiment (Figure 6). The results tend to be more clearly defined for the movie data, probably due to the dynamic videos, the larger data set, and hyperalignment of the data. Contributions from scene context, language, music, and narrative structure might also play a role (Huth, de Heer, Griffiths, Theunissen, & Gallant, 2016; Simony et al., 2016). The 2D solution

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

cleanly captured distinctions in the ventral core system in the first dimension and in the extended, dorsal core, and anterior core systems in the second dimension, with remarkably similar placement of ROIs on each of these dimensions between task data and movie data.

We quantified the similarity of the within-system RDMs by running a linear mixed-effect model on the correlation values and contrasting within-systems correlations with between-systems correlations. We found a clear distinction between the core and extended systems in terms of similarity of representational geometries. For the task data, the correlations within the extended system were significantly higher than the between-system correlations (estimate of the contrast "Within Extended > Between" 0.0993 [0.0875, 0.1111] 95% confidence interval, t-value = 16.36), while the correlations within the core system were not significantly different from the between-system correlations (estimate of the contrast "Within Core > Between" 0.0044 [-0.0043, 0.0130], t-value = 1.00). For the movie data, both contrasts were significant: within-core vs. between 0.0678 [0.0619, 0.0738], t-value = 22.47; and within-extended vs. between 0.1479 [0.1398, 0.1565], t-value = 35.07. Supplementary Tables 2 and 3 show the full parameter estimates for both models, while Supplementary Tables 4 and 5 report additional statistics on the subsystems.

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

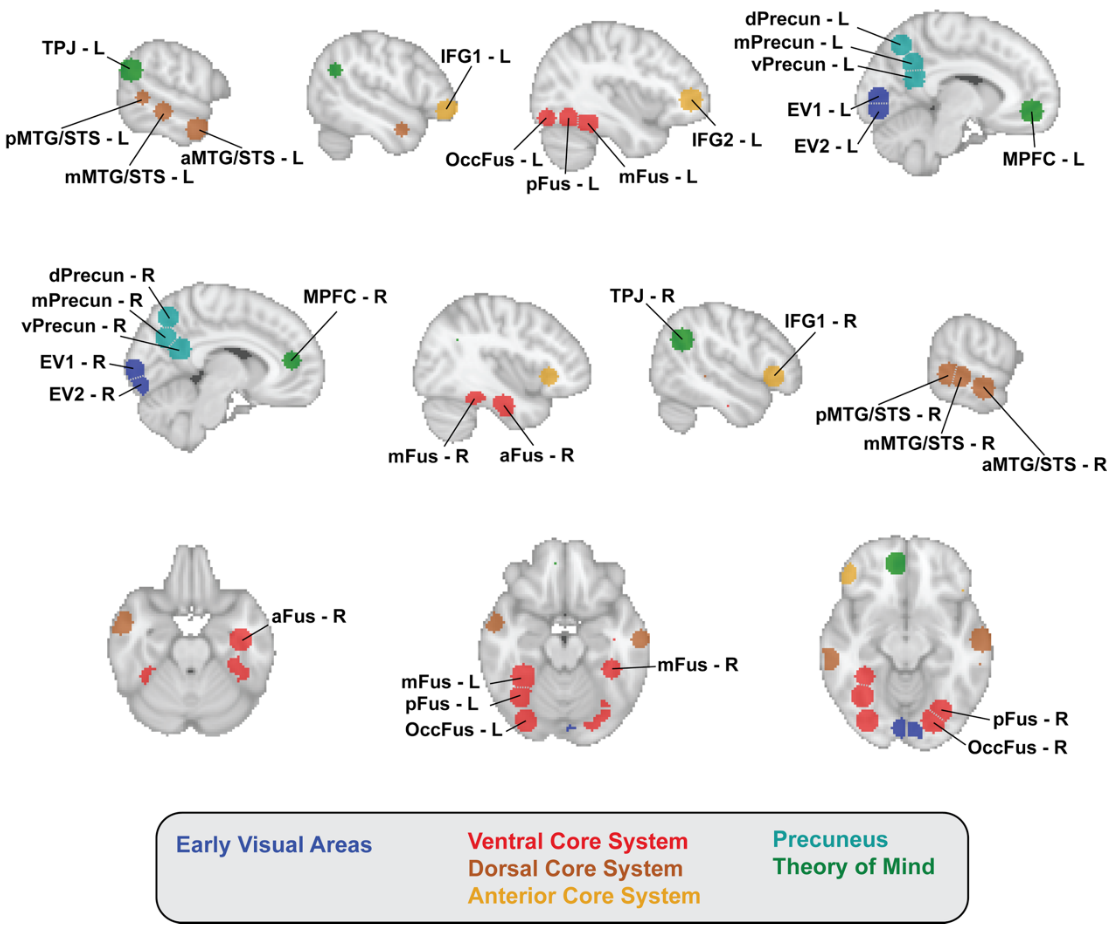
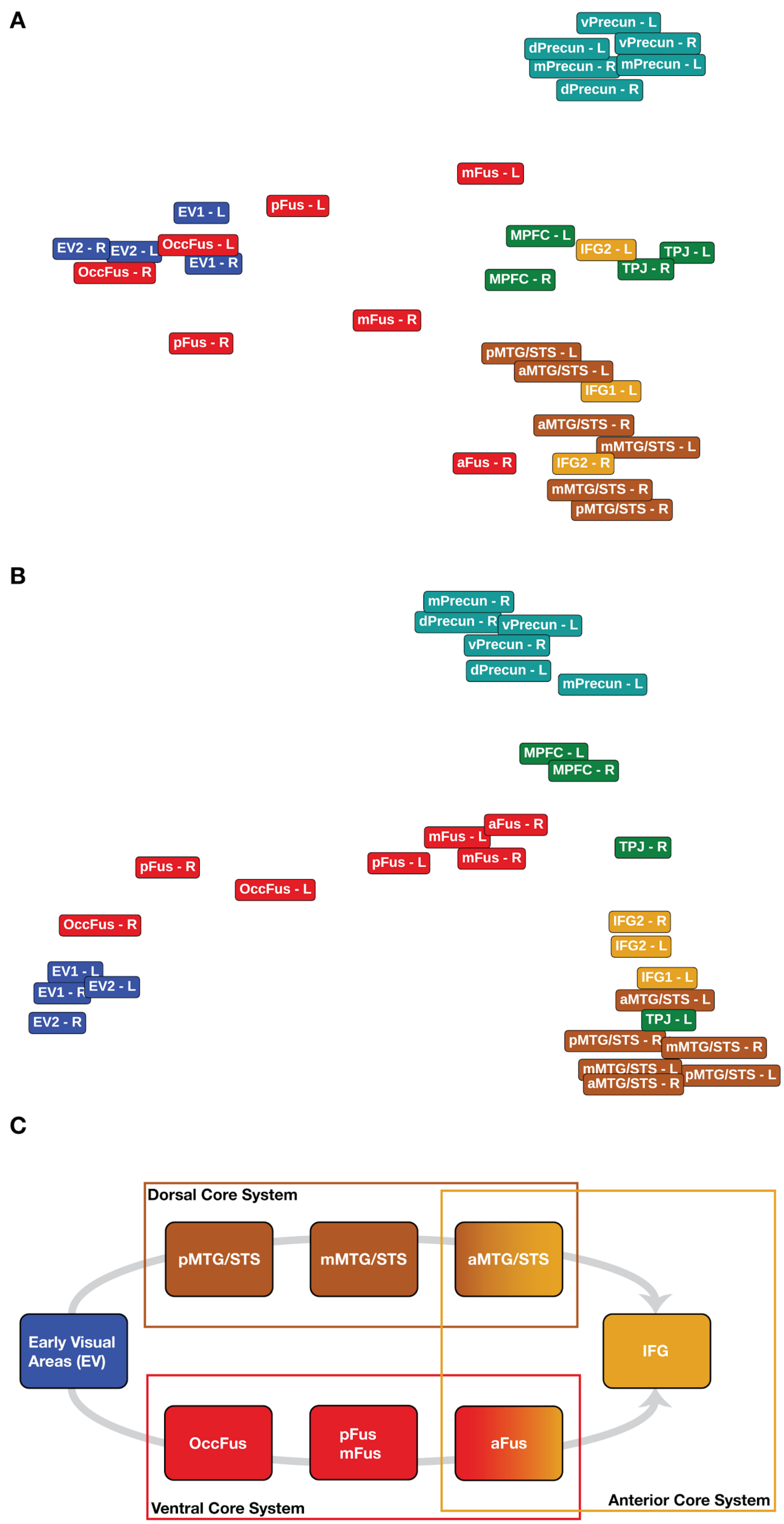


Figure 5. Spherical ROIs used to analyze the similarity of representational geometries. Top row shows left sagittal slices; middle row shows right sagittal slices; bottom row shows axial slices. Regions are color coded according to the system they belong to. Grey dotted lines between ROIs indicates that they were contiguous but not overlapping (see Methods for details).

240

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 6. Similarity of neural representations in ROIs derived from familiarity and identity decoding. Top panel and middle panel show MDS solutions based on the task data (A) and the hyperaligned movie data (B) (Guntupalli et al., 2016; Haxby et al., 2011) (see Methods section for more details). The color of the labels indicates the system to which the ROI belongs to (see Figure 5 for their location and Supplementary Table 1 for the MNI coordinates). With both datasets the MDS solution shows the hierarchy from early visual cortex to ventral core system (first dimension, x-axis), as well as a segregation between the precuneus, theory of mind areas, and areas of the anterior and dorsal core system (second dimension, y-axis). Panel (C) shows the proposed division of the core system into dorsal, ventral, and anterior portions. Representation of identity and gaze in the anterior core areas are disentangled from variations in head view (Carlin, Rowe, Kriegeskorte, Thompson, & Calder, 2012; Guntupalli et al., 2017).

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Discussion

In this experiment we investigated how familiar and unfamiliar faces are represented in the distributed neural system for face perception. We distinguished between familiarity information, abstracted from the visual appearance of the faces, and the identification of individual faces, controlling for the added information of personal familiarity. These analyses revealed an extensive network of areas that carry information about face familiarity and identity, replicating previous studies that used univariate analyses, but providing more details about the type of information present in those areas. We then analyzed the second-order representational geometry of this extensive network, revealing a clear distinction between the core and the extended systems for face perception and a new subdivision of the areas in the core system.

The results suggest that the core system for face perception can be separated into ventral, dorsal, and anterior subsystems. The ventral core system consists of fusiform areas extending from the occipital lobe to the anterior ventral temporal lobe. The dorsal system extends from the posterior MTG/STS to anterior lateral temporal cortex. The representations in the dorsal core system did not appear to have strong similarities with those in the ventral core system, consistent with the functional distinction between dorsal and ventral areas suggested by O'Toole, Roark, & Abdi (2002) and Pitcher et al. (2011). The anterior areas in the fusiform gyrus, the anterior MTG/STS, and the IFG may be the convergence of the ventral and dorsal pathways in which representations of faces become invariant to facial attributes such as head position (Carlin et al., 2011; Guntupalli et al., 2017) and perhaps other social attributes. For example, the right anterior STS plays a role in the representation of the dangerousness of animals (Connolly et al., 2016) and may play a role in the representation of social impressions, such as trustworthiness and aggressiveness (Todorov, Gobbini, Evans, & Haxby, 2007).

We teased apart neural responses due to factors that are shared by familiar faces from factors that are specific to familiar and unfamiliar identities. To separate identity-independent familiarity information from identity-specific visual information, we employed a cross-validation scheme in MVPC of face familiarity in which we tested

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the classifier on identities that were not included in the training data. To investigate identity-specific information that was independent of familiarity, we tested MVPC of familiar and unfamiliar identities separately.

We found reliable decoding of identity-independent familiarity in extended system areas that showed stronger responses to familiar faces in univariate analyses, such as theory of mind areas (precuneus, TPJ, and MPFC), consistent with previous reports (Gobbini & Haxby, 2007; Natu & O'Toole, 2011). Importantly, MVPC of familiarity was designed to test for a familiarity effect that was not specific to familiar individuals, revealing that this network does carry such identity-independent information about the familiarity of faces. Both the univariate and MVPC results expand the areas reported previously to include additional areas that are components of the dorsal and anterior core system for face perception in the MTG/STS, anterior fusiform cortex, and IFG. We suspect that our relatively large sample size made it possible to identify this more extensive network.

Unexpectedly, we found significant decoding of familiarity information in early visual cortex while controlling for identity information. Additional ROI decoding analyses in early visual areas (Wang et al., 2015) revealed that familiarity information could be decoded in V2 and V3 (see Supplementary Material). Low-level image differences did not seem to explain this finding: familiar and unfamiliar faces were indistinguishable using features extracted from the HMAX model (Riesenhuber & Poggio, 1999; Serre et al., 2007). Recent studies have shown that feedback information from higher-order visual areas to early visual cortex carries fine-grained information about the category of the stimuli being observed (Morgan, Petro, & Muckli, 2016; Muckli et al., 2015), suggesting that feedback processes might have contributed to the significant familiarity decoding in early visual areas. However, future studies with paradigms designed to address the nature of these feedback processes are needed to further test this possibility.

In addition to identity-independent familiarity, the same network carries information about specific identities. We tested for this type of information with separate MVPC analyses of four familiar identities and four unfamiliar identities. By not including

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

familiar and unfamiliar identities in the same analysis, we could test for identity-specific neural patterns that were not dependent on familiarity. Again, this network was more extensive than that reported in previous studies (e.g. Anzellotti, Fairhall, & Caramazza, 2013; Guntupalli et al., 2017; Kriegeskorte, Formisano, Sorger, & Goebel, 2007; V. S. Natu et al., 2010; Nestor, Plaut, & Behrmann, 2011), most probably due to the larger number of subjects and, perhaps, the inclusion of personally familiar faces. Importantly, this network included the IFG, consistent with Guntupalli et al. (2017), and extended into the MTG/STS, TPJ, precuneus, and MPFC.

Identity decoding was also found in early visual cortex and the posterior ventral core system, likely reflecting to some extent image-specific information. In Guntupalli et al. (2017) we showed that view-dependent representation of faces was the dominant factor in early visual cortex and the OFA. We did not find a significant difference in MVPC of familiar identities as compared to MVPC of unfamiliar identities, despite the large number of subjects in this study. There was a nonsignificant trend towards higher MVPC accuracies for familiar identities in the IFG and MTG/STS, but more work is needed to establish whether these trends are real.

Conclusions

Our results revealed new structure in the distributed system for face perception, suggesting that the core system can be subdivided into ventral, dorsal, and anterior components based on differences of representations. The anterior portion of the core system may be the point at which the ventral and dorsal pathways converge to generate view-independent representations of identity and of socially-relevant visual information, such as direction of attention. Identity-independent information about familiarity could be decoded in extended system areas such as the TPJ, precuneus, and MPFC, as well as in dorsal and anterior core system areas such as the MTG/STS, anterior fusiform cortex, and IFG. In sum, these results reveal new information about how face perception, one of the most highly developed and socially relevant visual functions, is realized in an extensive distributed system involving cortical fields in occipital, temporal, parietal, and prefrontal cortices.

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Materials and Methods

Participants

Thirty-three young adults participated in the experiment (mean age 23 y.o. +/- 3.33 SD, 13 males). They were recruited from the Dartmouth College community and all had normal or corrected-to-normal vision. Prior to the imaging study we took pictures of four friends for each participant to use as familiar stimuli. Some of these friends also were study participants (pictures of 76 individuals were taken as familiar stimuli). Photos of unfamiliar individuals were collected at the University of Vermont (Burlington) using the same camera and lighting conditions. Prior to participation in the fMRI study, subjects were screened for MRI compliance and provided informed consent in accordance with the Committee for the Protection of Human Subjects at Dartmouth College.

Stimuli

The stimuli for the fMRI experiment were pictures portraying different familiar and unfamiliar identities: four friends' faces, four unknown faces, and the subject's own face. For each identity we used three images with different head orientations: frontal view and 30-degree profiles to the left and right with gaze towards the camera. All photos on both sites (Dartmouth College and University of Vermont) were taken using the same consumer-grade digital camera in a dedicated photo-studio room with black background and uniform lighting.

Each familiar face was matched with an unfamiliar individual face, similar in age, gender and ethnicity. Twenty-seven images (9 individuals, 3 head positions) were used in the experimental design per each subject. Stimuli were presented to the subjects in the MRI scanner using a projection screen positioned at the rear of the scanner and viewed through a mirror mounted on the head coil.

The original high-resolution digital images were cropped to include the face from the top of the head to the neck visible under the chin, centered on the face. Images were

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

scaled to 400x400 pixels. Images subtended approximately 10x10 degrees of visual angle.

Procedure

The stimuli were presented using a slow event-related design while subjects were engaged in a simple oddball task (Figure 1). A typical trial consisted of three different images of the same individual, each presented for 500 ms with no gap. On catch trials, one of the three images was of a different individual. The order of head orientations within trials was randomized. The task was included to make sure that subjects paid attention to the identity of the faces. Before entering the scanner, subjects had a short practice session with each condition (one trial for each of 9 identities, one blank trial, and one catch trial) to be familiarized with the design and the stimuli.

The order of the events was pseudo-randomized to approximate a first-order counterbalancing of conditions (Aguirre, 2007). A functional run comprised 48 trials: four trials for each of the nine individuals (four familiar, four unfamiliar and self), four blank trials, four oddball and four buffer trials (three at the beginning and one at the end). The buffer trials were added to optimize the trial order and were discarded from the analysis. Each run had 10 seconds of fixation at the beginning (to stabilize the hemodynamic response) and at the end (to collect the response to the last trials). Each session consisted of 11 functional runs, resulting in 396 non-oddball trials (44 for each of the nine identities).

Image acquisition

Brain images were acquired using a 3T Philips Achieva Intera scanner with a 32-channel head coil. Functional imaging used gradient-echo echo-planar-imaging with SENSE reduction factor of 2. The MR parameters were TE/TR = 35/2000 ms, Flip angle = 90°, in-plane resolution = 3×3 mm, matrix size of 80×80 and FOV = 240×240 mm. 35 axial slices were acquired with no gap covering the entire brain except the most dorsal portion (Supplementary Figure 9). Slices were acquired in the Philips-specific interleaved order (slice step of 6, i.e., ceiled square root of total number of

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

slices). Each of the 11 functional runs included 154 dynamic scans with 4 dummy scans for a total time of 316 seconds per run. After the functional runs a single high-resolution T1-weighted (TE/TR = 3.7/8.2 ms) anatomical scan was acquired with a 3D-TFE sequence. The voxel resolution was 0.938×0.938×1.0 mm with a bounding box matrix of 256×256×160 (FOV = 240×240×160 mm).

Image preprocessing

All preprocessing steps were run using a Nipype workflow (version 0.11.0; FSL version 5.0.9) (K. Gorgolewski, Burns, Madison, & Clark, 2011; Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012), which also used functions from SciPy (Jones, Oliphant, & Peterson, 2001) and NumPy (van der Walt, Colbert, & Varoquaux, 2011). We modified the preprocessing pipeline *fmri_ants_openfmri.py* and adapted it for our analyses. The modified version is available at https://www.github.com/mvdoc/famface. All the preprocessing analyses were run on a computing cluster running Debian Jessie with tools provided by the NeuroDebian repository (Halchenko & Hanke, 2012).

Preprocessing Steps

We used a standard FSL preprocessing pipeline (FEAT) as implemented in Nipype (*nipype.preprocess.create_featreg_preproc*), using a FWHM smoothing of 6 mm, a highpass filter at 60 s cutoff, and the first volume of the first run as a reference for EPI alignment. After motion correction, the BOLD time-series were masked with a dilated gray-matter mask, smoothed, and then high-pass filtered. The preprocessed data were then used for a GLM and MVPA analysis, with additional preprocessing steps as described in the following sections.

Template Registration

Each subject's data (functional or second-level betas) were resliced into the MNI template with 2 mm isotropic voxel size. First, a reference volume was created by cmputing a median temporal SNR volume across functional runs. Then, we computed an affine transformation registering this median tSNR volume to the subject's aatomical scan using FSL's FLIRT tool (Jenkinson, Bannister, Brady, & Smith, 2002),

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

and the transformation was improved using the BBR cost function. A second non-linear transformation registering the subject's anatomical image to the MNI template was computed using ANTs (Avants, Tustison, & Song, 2009) with default parameters. The affine and nonlinear transformations were then combined to reslice the reference volume and all the functional volumes and second-level betas into the MNI template. Results from this registration pipeline were visually inspected for each subject.

MVPA Preprocessing

First, we resliced the bold time-series into the MNI template using a combination of linear and nonlinear transformations (see Template Registration section). Then, we extracted beta parameters associated with each condition for each run using PyMVPA's *fit_event_hrf_model* (Hanke et al., 2009) function based on NiPy's functionality (Millman & Brett, 2007). Additional nuisance regressors comprised motion estimates, artifacts (volumes were marked as artifact if their intensity exceeded three standard deviations of the normalized intensity), and noise estimates. To obtain noise estimates we used the CompCor method (Behzadi, Restom, Liau, & Liu, 2007). In brief, we performed a GLM on the BOLD timeseries in the voxels belonging to each subject's white-matter mask projected in MNI space. The regressors of this GLM were the motion estimates and volumes marked as artifacts. We then performed PCA on the residuals, and took the first 5 components as noise estimates.

GLM analyses

The first-level and second-level analyses (fixed effect) for each subject were performed in the subject's individual space, and the results were then projected into a standard template (FSL's MNI152, 2 mm isotropic, see details in the Template Registration section). These analyses followed a standard FSL pipeline as implemented in Nipype (*nipype.estimate.create_modelfit_workflow* and *nipype.estimate.create_fixed_effects_flow*). A standard GLM analysis was performed separately for each run to extract beta values associated with each condition and the planned contrasts. Additional nuisance regressors comprised motion estimates, artifacts (volumes were marked as artifact if their intensity exceeded three standard

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

deviations of the normalized intensity), and first-order derivatives. A second-level analysis was performed to obtain per-subject statistical maps associated with each condition and contrast using FSL's *FLAMEO* (fixed-effect model). The statistical maps were then resliced into the MNI152 template (see details above), and a third-level analysis was performed across subjects using FSL's *FLAMEO* (mixed-effect model). The resulting z-stat maps were then corrected for multiple comparisons using FSL's *cluster* routine, with a voxel z-threshold set at 2.3, and cluster p-value of $p = .05$. The Nipype pipeline we used for third-level analysis can be found at https://www.github.com/mvdoc/famface[1].

MVPA analyses

Classification methods

MVPC was implemented in Python using PyMVPA (Hanke et al., 2009) http://www.pymvpa.org). GLM betas were estimated within each run for each condition (see MVPA Preprocessing section). For all analyses we kept only the betas for the four familiar and the four unfamiliar identities, discarding trials where subjects saw their own face, or responded to an oddball presentation. The betas were then z-scored within each run (separately for each voxel) and used as features for classification. We used Linear C-SVM as a classifier, as implemented in LIBSVM (Chang & Lin, 2011). The C parameter was set to the PyMVPA default, which scales it according to the mean norm of the training data.

Cross-validation

We used a leave-one-out (LOO) scheme for cross-validation. The splitting unit was dependent on the type of classification (familiarity or identity). For familiarity classification, we cross-validated across pairs of identities. We trained the classifier on three familiar and three unfamiliar identities, and tested on the left-out identities. This resulted in 16 cross-validation splits that allowed us to control for identity information (see Supplementary Figures 1 and 2 for a comparison of leave-one-run-

[1] We thank Satrajit Ghosh and Anne Park for sharing the original pipeline.

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

out and leave-two-identities-out cross-validation schemes). For identity classification, we cross-validated across runs, resulting in a leave-one-run-out scheme (11 splits). To remove the effect of familiarity on classification of face identity, we performed identity classification independently for familiar and unfamiliar identities, and averaged the resulting accuracy maps.

Searchlight

We used sphere searchlights (Kriegeskorte, Goebel, & Bandettini, 2006) to extract local features for classification. We selected a 5-voxel radius (10 mm), and moved the searchlight sphere across the voxels belonging to a union mask in which at least 26 subjects (~80%, arbitrarily chosen) had fMRI coverage (see Supplementary Figure 9), as well as selecting only gray- and white-matter voxels in the cerebrum. For each center voxel in this mask, we selected nearby voxels contained in a sphere, and used them as features for classification. The classifier's accuracy was stored in the central voxel, and the process was repeated for every voxel.

Statistical assessment

To determine statistical significance for the MVPC analyses, we performed permutation testing (Stelzer, Chen, & Turner, 2013) coupled with Threshold-Free Cluster Enhancement (TFCE, (Smith & Nichols, 2009), as implemented in CoSMoMVPA (Oosterhof, Connolly, & Haxby, 2016). For each subject and each classification analysis, we computed a null distribution by randomly permuting the labels and performing classification. For identity classification analysis, we randomly shuffled the identity labels within each run, and performed classification. This procedure was repeated 20 times for each subject. For familiarity analysis, we randomly permuted the familiarity labels across the entire experiment. This was repeated exhaustively, resulting in 35 permutations (see Supplementary Materials for a short proof that only 35 unique permutations are possible in this case). To create a null distribution of TFCE values for each voxel, permutation maps were randomly sampled and averaged across subjects, and this process was repeated 10,000 times. Note that we selected a smaller number of permutations than suggested by (Stelzer et al., 2013) (100 per subject) because of the large number of subjects we had: with 33

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

subjects, the number of possible average maps for identity classification was 20^{33} and for familiarity classification was 35^{33}.

Similarity of neural representations within ROIs

Second-order Representational Similarity Analysis

We defined ROIs based on the searchlight results for both the familiarity and identity classification. Thirty spherical ROIs were centered on voxels selected manually at or near peak values, with a 10 mm radius (five voxels). Voxels belonging to more than one ROI were assigned to the ROI with the closest center (Euclidean distance), resulting in some contiguous but not overlapping ROIs (see Figure 5). On average, ROIs contained 412 voxels at a 2 mm isotropic resolution (SD: 73 voxels).

For each ROI we computed a cross-validated representational dissimilarity matrix (RDM) (Henriksson, Khaligh-Razavi, Kay, & Kriegeskorte, 2015) between the eight identities (four familiar faces, four unfamiliar faces). First, we z-scored the beta estimates within each run, which were computed as described in the MVPA Preprocessing section. Then, we divided all runs into two partitions of six and five runs, and averaged the beta values within each partition. The data between these two partitions were correlated (Pearson correlation) to obtain an 8x8 matrix of dissimilarities between pairs of identities. Note that because correlations were computed between data from two different partitions, the diagonal could be different from one. This process was repeated for every possible combination of runs, yielding 462 RDMs that were averaged to obtain a final RDM for each ROI and each subject. The final RDMs were made symmetrical by averaging them with their transpose. All averaging operations were performed on Fisher-transformed (r-to-z) correlation values, then mapped back to correlation using the inverse transformation.

We used these final RDMs to compute pairwise distances between ROIs for each subject individually using correlation distance. The resulting 33 distance matrices (one for each subject) were averaged to obtain a group-level distance matrix. This distance matrix was used to compute a three-dimensional MDS solution, using classical MDS as implemented in R (*cmdscale*) interfaced in Python using *rpy2 (Gautier, 2008)*.

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Comparison with movie data

To investigate the reproducibility of the network formed by the ROIs defined above, we computed between-subject correlation distances across these ROIs using hyperaligned data from a different study, in which eleven participants watched "Raiders of the Lost Ark" (Guntupalli et al., 2016; Haxby et al., 2011). Since data were functionally aligned with hyperalignment (Guntupalli et al., 2016; Haxby et al., 2011), we performed a between-subject analysis instead of a within-subject analysis, where distances between pairwise ROIs were computed across subjects, replicating the approach in (Guntupalli et al., 2016). Additional details on the experimental paradigm and scanning parameters can be found in the Supplementary Material.

Because data were in two different resolutions of the same template (task: MNI 2 mm; movie: MNI 3 mm), center coordinates of the spherical ROIs were recalculated assigning the closest voxel in MNI 3 mm using Euclidean distance. The median displacement was 1.41 mm (min: 1 mm, max: 1.73 mm). As described above, spherical ROIs were drawn around these center voxels using a radius of 9 mm (3 voxels) to account for the different voxel size. Overlapping voxels were assigned to the ROI with the closest center, resulting in possibly contiguous but not overlapping ROIs. On average ROIs contained 100 voxels (SD: 20 voxels).

The movie data were masked selecting only white- and gray-matter voxels, and divided into two parts for cross-validation. For each of the two parts, whole-brain searchlight hyperalignment parameters were derived from one part of the movie, and the second part was projected into the common model space in functional alignment (Guntupalli et al., 2016; Haxby et al., 2011). The aligned data were z-scored, and timepoint-by-timepoint RDMs were computed in each ROI for each subject individually, yielding a 1322 x 1322 RDM within each ROI (1336 x 1336 for the second fold of hyperalignment). Following the analysis in (Guntupalli et al., 2016) we estimated a distance matrix between ROIs while cross-validating across subjects. For each pair of ROIs, the correlation between their RDMs was computed for all 55 pairs of subjects, and averaged to compute the cross-validated correlation between those ROIs. This process resulted in two 30x30 cross-validated distance matrices (one for

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

each hyperalignment fold), which were made symmetrical by averaging them with their transpose, and finally averaged together to obtain one final 30x30 matrix. All averaging operations were computed on Fisher-transformed (r-to-z) correlation values, then mapped back to correlation using the inverse transformation. Finally, a dissimilarity index (D) was computed for each pair of ROIs to normalize the correlation according to the maximum possible correlation within each ROI (Guntupalli et al., 2016):

$$D_{ROI1 \cdot ROI2} = 1 - \frac{r_{ROI1 \cdot ROI2}}{\sqrt{r_{ROI1} \cdot r_{ROI2}}}$$

The final matrix containing dissimilarity indices was then used to compute an MDS solution as described previously.

Differences between core and extended system representational geometries

In order to quantify differences in representational geometries between areas of the core and extended systems, we divided the pairwise distances between ROIs in the upper triangular RDM into within-system and between-system cells, and converted them back to correlations (by subtracting them from 1). Then, we ran a Linear Mixed-Effect Model on the correlations using *lme4* (Bates, Maechler, Bolker, & Walker, 2014), fitting a linear model of the form

$$r_{i,j} = \beta_0 + \beta_1 C_{i,j} + \beta_2 E_{i,j} + z_i \,,$$

where $i = 1 \ldots N$ indicates either the subjects for task data ($N = 33$) or the pairwise subjects for hyperaligned movie data ($N = 55$); $j = 1 \ldots 465$ indicates the index of the pairwise correlations between ROIs, $C_{i,j}$ and $E_{i,j}$ indicate whether $r_{i,j}$ is a within-system correlation for the core or extended system respectively, $\beta_0, \beta_1, \beta_2$ are fixed-effects parameters, and z_i are the subject-level random effects. Using this model, β_1 corresponds to the contrast “Within Core > Between”, and β_2 to the contrast “Within Extended > Between”. After fitting, we performed parametric bootstrapping to obtain 95% bootstrapped confidence intervals on the model parameters.

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Visualization

Volumetric results were visualized using Nilearn (Abraham et al., 2014), and projected on template surfaces using AFNI and SUMA (Cox, 1996; Saad, Reynolds, Argall, Japee, & Cox, 2004).

Data and code availability

Non-thresholded statistical maps can be found on neurovault.org (K. J. Gorgolewski et al., 2015) at the following URL: http://neurovault.org/collections/NEUNABLT. All data can be found at http://datasets.datalad.org/?dir=/labs/gobbini/famface/data[2]. The code used for the analyses is available at the following GitHub repository: https://www.github.com/mvdoc/famface.

Acknowledgments

The authors would like to thank Jim Haxby, Brad Duchaine, and the members of the GobbiniLab and HaxbyLab for helpful comments and discussions on this work.

References

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., … Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*, 14.

Aguirre, G. K. (2007). Continuous carry-over designs for fMRI. *NeuroImage*, *35*(4), 1480–1494.

Anzellotti, S., Fairhall, S. L., & Caramazza, A. (2013). Decoding Representations of Face Identity That are Tolerant to Rotation. *Cerebral Cortex* . Retrieved from http://www.cercor.oxfordjournals.org/cgi/doi/10.1093/cercor/bht046

Avants, B. B., Tustison, N., & Song, G. (2009). Advanced normalization tools (ANTS). *The Insight Journal*, *2*, 1–35.

[2] All data will be released after publication. Currently, data from the first subject are available from the DataLad repository.

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). Package “lme4.” *R Foundation for …*. Retrieved from http://cran.r-mirror.de/web/packages/lme4/lme4.pdf

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, *37*(1), 90–101.

Bobes, M. A., Lage Castellanos, A., Quiñones, I., García, L., & Valdes-Sosa, M. (2013). Timing and tuning for familiarity of cortical responses to faces. *PloS One*, *8*(10), e76100.

Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology. Applied*, *7*(3), 207.

Burton, A. M., Jenkins, R., & Schweinberger, S. R. (2011). Mental representations of familiar faces. *British Journal of Psychology* , *102*(4), 943–958.

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.

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: CB*, *21*(21), 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*(4), 735–744.

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*(3), 1–27.

Cloutier, J., Kelley, W. M., & Heatherton, T. F. (2011). The influence of perceptual and knowledge-based familiarity on the neural substrates of face perception. *Social Neuroscience*, *6*(1), 63–75.

Collins, J. A., & Olson, I. R. (2014). Beyond the FFA_ The role of the ventral anterior temporal lobes in face processing. *Neuropsychologia*, *61*(C), 65–79.

Connolly, A. C., Sha, L., Guntupalli, J. S., Oosterhof, N., Halchenko, Y. O., Nastase, S.

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

A., … Others. (2016). How the Human Brain Represents Perceived Dangerousness or “Predacity” of Animals. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(19), 5373–5384.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, an International Journal*, *29*(3), 162–173.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology. General*, *115*(2), 107.

Duchaine, B., & Yovel, G. (2015). A Revised Neural Framework for Face Processing. *Annual Review of Vision Science*, *1*(1), 393–416.

Fairhall, S. L., & Ishai, A. (2007). Effective Connectivity within the Distributed Cortical Network for Face Perception. *Cerebral Cortex* , *17*(10), 2400–2406.

Gautier, L. (2008). rpy2: A Simple and Efficient Access to R from Python. *URL Http://rpy. Sourceforge. net/rpy2. Html*.

Gobbini, M. I. (2010). Distributed process for retrieval of person knowledge. *Social Neuroscience: Toward Understanding the Underpinnings of the Social Mind*, 40–53.

Gobbini, M. I., Gors, J. D., Halchenko, Y. O., Rogers, C., Guntupalli, J. S., Hughes, H., & Cipolli, C. (2013). Prioritized Detection of Personally Familiar Faces. *PloS One*, *8*(6), e66620.

Gobbini, M. I., & Haxby, J. V. (2006). Neural response to the visual familiarity of faces. *Brain Research Bulletin*, *71*(1-3), 76–82.

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, *45*(1), 32–41.

Gobbini, M., Leibenluft, E., Santiago, N., & Haxby, J. V. (2004). Social and emotional attachment in the neural representation of faces. *NeuroImage*, *22*(4), 1628–1635.

Gorgolewski, K., Burns, C. D., Madison, C., & Clark, D. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Front …*. Retrieved from http://www.research.ed.ac.uk/portal/files/18241445/fninf_05_00013.pdf

Gorgolewski, K. J., Varoquaux, G., Rivera, G., Schwarz, Y., Ghosh, S. S., Maumet, C.,

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

… Margulies, D. S. (2015). NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Frontiers in Neuroinformatics*, *9*, 8.

Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews. Neuroscience*, *15*(8), 536–548.

Guntupalli, J. S., Hanke, M., Halchenko, Y. O., Connolly, A. C., Ramadge, P. J., & Haxby, J. V. (2016). A Model of Representational Spaces in Human Cortex. *Cerebral Cortex* , bhw068.

Guntupalli, J. S., Wheeler, K. G., & Gobbini, M. I. (2017). Disentangling the Representation of Identity from Head View Along the Human Face Processing Pathway. *Cerebral Cortex* . https://doi.org/10.1093/cercor/bhw344

Halchenko, Y. O., & Hanke, M. (2012). Open is Not Enough. Let's Take the Next Step: An Integrated, Community-Driven Computing Platform for Neuroscience. *Frontiers in Neuroinformatics*, *6*, 22.

Hancock, P., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences*. Retrieved from http://www.sciencedirect.com/science/article/pii/S1364661300015199

Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009). PyMVPA: a Python Toolbox for Multivariate Pattern Analysis of fMRI Data. *Neuroinformatics*, *7*(1), 37–53.

Haxby, J. V., Connolly, A. C., & Guntupalli, J. S. (2014). Decoding Neural Representational Spaces Using Multivariate Pattern Analysis. *Annual Review of Neuroscience*, *37*(1), 435–456.

Haxby, J. V., & Gobbini, M. I. (2011). Distributed neural systems for face perception. In A. Calder, G. Rhodes, M. Johnson, & J. Haxby (Eds.), *Oxford Handbook of Face Perception*. OUP Oxford.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425–2430.

Haxby, J. V., Guntupalli, J. S., Connolly, A. C., Halchenko, Y. O., Conroy, B. R.,

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Gobbini, M. I., … Ramadge, P. J. (2011). A Common, High-Dimensional Model of the Representational Space in Human Ventral Temporal Cortex. *Neuron*, *72*(2), 404–416.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233.

Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., & Grady, C. L. (1994). The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *14*(11 Pt 1), 6336–6353.

Henriksson, L., Khaligh-Razavi, S.-M., Kay, K., & Kriegeskorte, N. (2015). Visual representations are dominated by intrinsic fluctuations correlated between areas. *NeuroImage*, 1–48.

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*(7600), 453–458.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*(2), 825–841.

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790.

Jenkins, R., & Burton, A. M. (2011). Stable face representations. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1571), 1671–1683.

Jones, E., Oliphant, T., & Peterson, P. (2001). SciPy: Open source scientific tools for Python. Retrieved from http://www.scipy.org/

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 of the United States of America*, *104*(51), 20600–20605.

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences*, *103*(10), 3863–

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

3868.

Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, *17*(8), 401–412.

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 4.

Millman, K. J., & Brett, M. (2007). Analysis of Functional Magnetic Resonance Imaging in Python. *Computing in Science Engineering*, *9*(3), 52–55.

Morgan, A. T., Petro, L. S., & Muckli, L. (2016, January 1). *Cortical feedback to V1 and V2 contains unique information about high-level scene structure*. *bioRxiv*. https://doi.org/10.1101/041186

Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., … Yacoub, E. (2015). Contextual Feedback to Superficial Layers of V1. *Current Biology: CB*, *25*(20), 2690–2695.

Natu, V., & O’Toole, A. J. (2011). The neural processing of familiar and unfamiliar faces: A review and synopsis. *British Journal of Psychology* , *102*(4), 726–747.

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*(7), 1570–1582.

Nestor, A., Plaut, D. C., & Behrmann, M. (2011). Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(24), 9998–10003.

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. *Frontiers in Neuroinformatics*, *10*, 27.

O’Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: a psychological and neural synthesis. *Trends in Cognitive Sciences*, *6*(6), 261–266.

Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

cortical regions. *NeuroImage*, *56*(4), 2356–2363.

Rajimehr, R., Young, J. C., & Tootell, R. B. H. (2009). An anterior temporal face patch in human cortex, predicted by macaque maps. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(6), 1995–2000.

Ramon, M., Vizioli, L., Liu-Shuang, J., & Rossion, B. (2015). Neural microgenesis of personally familiar face recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(35), E4835–E4844.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*(11), 1019–1025.

Saad, Z. S., Reynolds, R. C., Argall, B., Japee, S., & Cox, R. W. (2004). SUMA: an interface for surface-based intra- and inter-subject analysis with AFNI. In *2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821)* (pp. 1510–1513 Vol. 2).

Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M., & Poggio, T. (2007). Robust object recognition with cortex-like mechanisms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(3), 411–426.

Simony, E., Honey, C. J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., & Hasson, U. (2016). Dynamic reconfiguration of the default mode network during narrative comprehension. *Nature Communications*, *7*, 12141.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, *44*(1), 83–98.

Stelzer, J., Chen, Y., & Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): Random permutations and cluster size control. *NeuroImage*, *65*(C), 69–82.

Taylor, M. J., Arsalidou, M., Bayless, S. J., Morris, D., Evans, J. W., & Barbeau, E. J. (2009). Neural correlates of personally familiar faces: parents, partner and own faces. *Human Brain Mapping*, *30*(7), 2008–2020.

Todorov, A., Gobbini, M. I., Evans, K. K., & Haxby, J. V. (2007). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia*, *45*(1), 163–173.

bioRxiv preprint doi: https://doi.org/10.1101/138297; this version posted May 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

van der Walt, S., Colbert, S. C., & Varoquaux, G. (2011). The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science & Engineering*, *13*(2), 22–30.

Visconti di Oleggio Castello, M., di Oleggio Castello, M. V., Wheeler, K. G., Cipolli, C., & Gobbini, M. I. (2016). Familiarity Facilitates Feature-based Face Processing. https://doi.org/10.1101/058537

Visconti di Oleggio Castello, M., & Gobbini, M. I. (2015). Familiar Face Detection in 180ms. *PloS One*, *10*(8), e0136548.

Visconti di Oleggio Castello, M., Guntupalli, J. S., Yang, H., & Gobbini, M. I. (2014). Facilitated detection of social cues conveyed by familiar faces, 1–11.

Wang, L., Mruczek, R. E. B., Arcaro, M. J., & Kastner, S. (2015). Probabilistic Maps of Visual Topography in Human Cortex. *Cerebral Cortex* , *25*(10), 3911–3931.

Weiner, K. S., Golarai, G., Caspers, J., Chuapoco, M. R., Mohlberg, H., Zilles, K., … Grill-Spector, K. (2013). The mid-fusiform sulcus: A landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *NeuroImage*, 1–13.