

## Journal Club

**Editor's Note:** These short, critical reviews of recent papers in the *Journal*, written exclusively by graduate students or postdoctoral fellows, are intended to summarize the important findings of the paper and provide additional insight and commentary. For more information on the format and purpose of the Journal Club, please see [http://www.jneurosci.org/misc/ifa\\_features.shtml](http://www.jneurosci.org/misc/ifa_features.shtml).

## Decoding Face Exemplars from fMRI Responses: What Works, What Doesn't?

Johan D. Carlin

MRC Cognition and Brain Sciences Unit, Cambridge CB2 7EF, United Kingdom  
Review of Dubois et al.

Understanding how the human brain encodes individual faces is a central challenge for face recognition research. Such within-category selectivity has mainly been reported in spike rate analyses of single-unit data from macaque inferotemporal cortex (IT) (Hasselmo et al., 1989; Tsao et al., 2006; Freiwald and Tsao, 2010). In contrast, human face processing studies are generally restricted to noninvasive, spatially imprecise neuroimaging methods such as functional MRI (fMRI) and have focused on brain responses that are organized on a coarser spatial scale, such as mapping areas with selective responses between object categories such as faces and scenes (Kanwisher et al., 1997). Thus, a key strength of single-unit recording has been its sensitivity to subtle within-category distinctions, while a strength of fMRI has been its applicability in humans and the ease with which large parts of cortex can be mapped.

One can leverage the relative strengths of these methods by using macaque fMRI to identify category-selective patches, and then target these patches with single-unit recording for within-category comparisons. Using this method, Freiwald and Tsao (2010) found functional dissociations between individual exemplar representations in adjacent face patches of the macaque temporal

lobe. Briefly, cells in the middle face patch ML/MF were strongly view-selective with limited identity sensitivity. Anterior to this, cells in patch AL showed mirror-symmetric view tuning and some sensitivity to identity, while cells in the anterior-most patch AM exhibited strikingly view-tolerant identity tuning. These results constitute a landmark achievement in the field and suggest the power of combining fMRI for large-scale mapping of selectivity at the between-category level with single-unit recording for measuring subtle within-category selectivity.

However, concurrent with the development of fMRI-guided macaque single-unit recording, several recent studies have used multivariate pattern analysis (MVPA) methods to discriminate human fMRI responses to individual faces, which suggests that fMRI may be more sensitive to within-category information than previously thought (Kriegeskorte et al., 2007; Natu et al., 2010; Nestor et al., 2011; Gao and Wilson, 2013; Goesaert and Op de Beeck, 2013; Nestor et al., 2013; Verosky et al., 2013; Anzellotti et al., 2014; Axelrod and Yovel, 2015; Anzellotti and Caramazza, 2015). This is an exciting development because the central goal of face perception research is arguably to understand human face recognition, and unlike single-unit recording, fMRI is widely applicable to human subjects. In contrast, macaque single-unit recording data can only be used to support inferences about human face recognition if humans and macaques are assumed to process faces similarly. In fact, some research suggests that there may be differences. For instance, stimulus inversion affects face recognition disproportionately

compared with other object categories in humans (Yin, 1969), but whether macaques exhibit a similar inversion effect is controversial (Bruce, 1982; Parr et al., 1999). Thus, there is a need to validate face effects from macaque single-unit recording against brain measurements in humans, and fMRI MVPA appears to provide the necessary sensitivity to enable this species comparison.

In a recent paper in *The Journal of Neuroscience*, Dubois et al. (2015) investigated the correspondence between fMRI and single-unit recording by applying similar MVPA classification methods to both fMRI responses and single-unit data recorded from the macaque face patch system. Their study has several important features. First, instead of comparing human fMRI MVPA to tuning analyses of macaque single-unit data, the authors used the same analysis method to compare data acquired with the two techniques in the same species, making it easier to attribute differences to modality effects. Second, the known functional dissociations between neighboring face patches provide a clear set of reference results against which the fMRI MVPA results can be compared to test whether fMRI captures the same effects as single-unit recording. Finally, these functional dissociations involve changes both in view and identity tuning across the face patches, which enables tests of how different within-category face effects are preserved across modalities.

Dubois et al. (2015) first trained classifiers to discriminate brain responses to

Received April 10, 2015; revised May 14, 2015; accepted May 15, 2015.

Correspondence should be addressed to Johan D. Carlin, MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, UK. E-mail: [johan.carlin@mrc-cbu.cam.ac.uk](mailto:johan.carlin@mrc-cbu.cam.ac.uk).

DOI:10.1523/JNEUROSCI.1385-15.2015

Copyright © 2015 the authors 0270-6474/15/359252-03\$15.00/0

different face viewpoints. These classifiers were trained on brain responses from all but one of the presented face identities and tested on the left-out identity to ensure that successful classification could not be attributed to low-level visual features in particular stimulus images. This analysis yielded above-chance classification performance in all face patches and in both fMRI and single-unit data. Furthermore, the pattern of classification errors across the viewpoints correlated between the two modalities. For instance, both the single-unit and fMRI data from patch AL/AF exhibited a similar pattern of mirror image confusions. Thus, fMRI MVPA appeared to capture all the main viewpoint effects in the single-unit data.

Unlike the viewpoint results, classifiers that were trained to discriminate face identity and tested on a left-out viewpoint performed quite differently on single-unit and fMRI data. Single-unit identity classification was above chance throughout the face patches, with increasingly accurate classification on a posterior–anterior axis ending in patch AM on the ventral temporal surface of anterior IT. In contrast, fMRI identity classification was reliably above chance across subjects in the posterior/lateral region ML/MF, and declined to nonsignificant levels anterior to this. The pattern of identity classification errors also did not correlate across modalities, although this result is less surprising since the sensitivity of such confusion analyses are limited by the absolute level of classification performance. Overall, peak identity classification was substantially less accurate than peak view classification in the fMRI data while the view and identity classifiers reached similar maxima in the single-unit data, suggesting that fMRI MVPA exhibited particularly low sensitivity to identity information compared with view information.

Dubois et al. (2015) suggest that fMRI MVPA may capture viewpoint effects to a greater extent than identity effects because the limited spatial selectivity of fMRI renders this method insensitive to representations that are not spatially clustered. In support of this interpretation, single units recorded within 1 mm along the same track had more strongly correlated view tunings than identity tunings (although this difference was modest in AM, which was the only patch to exhibit significant identity clustering). This analysis underlies the authors' concluding remark that representations that are not sufficiently clustered at the single-unit level are "unlikely ever to be successfully

decoded by fMRI multivoxel pattern analysis" (Dubois et al., 2015, p. 2802).

An alternative account of why the modalities differ in this manner is that the quality of fMRI data from the face patches that most strongly encode identity is particularly poor. Note that the strongest identity effects in the single-unit data were reported in the anterior temporal patches AL/AF and especially AM, where fMRI functional signal-to-noise ratios (fSNR) generally are lower than in the more posterior and lateral ML/MF. Dubois et al. (2015) confirm that fSNR declines on a posterior–anterior axis in their data, but dismiss this account on the grounds that fMRI view classification remained above-chance in AL/AF and AM despite low fSNR. Instead, the conclusions of the study are centered on issues of spatial clustering.

However, a strict clustering account appears problematic upon closer inspection of the authors' results. First, fMRI classification of identity in ML/MF was reliably greater than chance even though this patch exhibited the lowest identity clustering of the three patches, suggesting that an absence of clustering at the single-unit level does not preclude above-chance fMRI MVPA. Furthermore, ML/MF had the highest fSNR of the three patches and was the region where both view and identity classification effects were most similar to the single-unit data, which is consistent with the view that fSNR is critical for a close correspondence between recording modalities. Second, a strict clustering account predicts that fMRI classification of identity and viewpoint should be similarly accurate in AM, since this patch exhibited comparable clustering for identity and viewpoint. Contrary to this prediction, only viewpoint could be classified reliably across the three subjects. Instead, the results include a hint that fSNR may be important for fMRI classification in AM: although fSNR was generally lower in AM than in ML/MF, one of the three subjects did not exhibit this effect, and that subject was also the only one for which AM fMRI identity classification was reliably above-chance (Dubois et al., 2015, their Fig. 6*b*, white bars). The same subject also showed the strongest fMRI view classification in AM. Thus, under sufficiently high fSNR conditions, fMRI identity classification appears possible even with minimal spatial clustering, although these effects certainly were weaker than in the single-unit data.

The study by Dubois et al. (2015) highlights both the promise and challenge of

fMRI MVPA. In some cases, the correspondence with single-unit data is striking; while in other cases, high noise levels in the fMRI data reduce classification to chance-level performance. I disagree with the authors on whether fMRI data fails to match single-unit data because facial identity representations are insufficiently spatially clustered or because facial identity is coded primarily in regions with insufficient fSNR in macaques. This distinction matters because spatial clustering is a fixed property of the brain, while fSNR is a technical issue that can be improved with experimental design, neuroimaging protocols, analysis methods, or the use of a subject species with a more easily imaged ventral temporal lobe. Note that macaques exhibit particularly severe ventral temporal dropout with conventional fMRI methods (Ku et al., 2011), which might explain how several human fMRI MVPA studies obtained reliable identity classification in ventral temporal regions while Dubois et al. (2015) did not. Thus, I remain optimistic about the prospects of fMRI MVPA. As Dubois et al. (2015) themselves point out, null results should be interpreted with caution.

## References

- Anzellotti S, Caramazza A (2015) From parts to identity: invariance and sensitivity of face representations to different face halves. *Cereb Cortex*. Advance online publication. doi: 10.1093/cercor/bhu337. CrossRef Medline
- Anzellotti S, Fairhall SL, Caramazza A (2014) Decoding representations of face identity that are tolerant to rotation. *Cereb Cortex* 24: 1988–1995. CrossRef Medline
- Axelrod V, Vovel G (2015) Successful decoding of famous faces in the fusiform face area. *PLoS One* 10:e0117126. CrossRef Medline
- Bruce C (1982) Face recognition by monkeys: absence of an inversion effect. *Neuropsychologia* 20:515–521. CrossRef Medline
- Dubois J, de Berker AO, Tsao DY (2015) Single-unit recordings in the macaque face patch system reveal limitations of fMRI MVPA. *J Neurosci* 35:2791–2802. CrossRef Medline
- Freiwald WA, Tsao DY (2010) Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science* 330:845–851. CrossRef Medline
- Gao X, Wilson HR (2013) The neural representation of face space dimensions. *Neuropsychologia* 51:1787–1793. CrossRef Medline
- Goesaert E, Op de Beeck HP (2013) Representations of facial identity information in the ventral visual stream investigated with multivoxel pattern analyses. *J Neurosci* 33:8549–8558. CrossRef Medline
- Hasselmo ME, Rolls ET, Baylis GC (1989) The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behav Brain Res* 32:203–218. CrossRef Medline

- Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J Neurosci* 17:4302–4311. [Medline](#)
- Kriegeskorte N, Formisano E, Sorger B, Goebel R (2007) Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proc Natl Acad Sci U S A* 104:20600–20605. [CrossRef Medline](#)
- Ku SP, Tolia AS, Logothetis NK, Goense J (2011) fMRI of the face-processing network in the ventral temporal lobe of awake and anesthetized macaques. *Neuron* 70:352–362. [CrossRef Medline](#)
- Natu VS, Jiang F, Narvekar A, Keshvari S, Blanz V, O'Toole AJ (2010) Dissociable neural patterns of facial identity across changes in viewpoint. *J Cogn Neurosci* 22:1570–1582. [CrossRef Medline](#)
- Nestor A, Plaut DC, Behrmann M (2011) Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proc Natl Acad Sci U S A* 108:9998–10003. [CrossRef Medline](#)
- Nestor A, Behrmann M, Plaut DC (2013) The neural basis of visual word form processing: a multivariate investigation. *Cereb Cortex* 23:1673–1684. [CrossRef Medline](#)
- Parr LA, Winslow JT, Hopkins WD (1999) Is the inversion effect in rhesus monkeys face-specific? *Animal Cogn* 2:123–129. [CrossRef](#)
- Tsao DY, Freiwald WA, Tootell RB, Livingstone MS (2006) A cortical region consisting entirely of face-selective cells. *Science* 311:670–674. [CrossRef Medline](#)
- Verosky SC, Todorov A, Turk-Browne NB (2013) Representations of individuals in ventral temporal cortex defined by faces and biographies. *Neuropsychologia* 51:2100–2108. [CrossRef Medline](#)
- Yin RK (1969) Looking at upside-down faces. *J Exp Psychol* 81:141–145. [CrossRef](#)