

## 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).

## Extracting Object Identity: Ventral or Dorsal Visual Stream?

Vassilis Pelekanos,<sup>1,2</sup>  Marieke Mur,<sup>1</sup> and  Katherine R. Storrs<sup>1</sup>

<sup>1</sup>Medical Research Council, Cognition and Brain Sciences Unit, Cambridge CB2 7EF, United Kingdom, and <sup>2</sup>Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, United Kingdom

Review of Jeong and Xu

The influential “two-streams hypothesis” of visual processing proposes that features related to object recognition are primarily encoded in the visual ventral–temporal stream (the “what” pathway), while spatial relationships among objects are primarily encoded in the dorsal–parietal stream (the “where” pathway; Ungerleider and Mishkin, 1982; Goodale and Milner, 1992). Consistent with this proposal is evidence that, in both humans and monkeys, neural activity in the ventral stream contains information about visual object identity and category (Kriegeskorte et al., 2008; Bell et al., 2009) without being modulated by image properties such as object position, viewing angle, size, or context (Li et al., 2009; Rust and DiCarlo, 2010; Anzellotti et al., 2014). However, other evidence suggests that simple visual features (e.g., motion direction) required to categorize stimuli are also represented in dorsal regions (Toth and Assad, 2002; Freedman and Assad, 2006). The hypothesis that two functionally independent streams process visual information has been challenged and revisited (Milner and Goodale, 2008), and we still do not know with confidence whether category and identity information are present exclusively in one visual stream and not in the

other. In a recent article in *The Journal of Neuroscience*, Jeong and Xu (2016) raised further questions about the two-streams hypothesis by reporting that abstract identity information for complex stimuli is represented in the dorsal stream.

Jeong and Xu (2016) used functional magnetic resonance imaging (fMRI) to track response patterns elicited in ventral and dorsal regions during object identification. Human participants viewed the names and faces of two famous actors (experiment 1), and the names and images of two well known car models (experiment 2). Experiment 3 extended experiment 1 by presenting the faces of eight famous actors. Stimuli were presented in blocks, and each block contained a set of 10 unique images of a single actor or car model. Two stimulus sets per identity were used in total. Faces differed in viewpoint, hairstyle, expression, and age, while car images differed in viewpoint, size, and background scene. Participants were asked to detect the presence of an oddball stimulus that appeared infrequently and was an actor/car with an identity other than the frequently presented identity of the block. The oddball-detection task, thus, required recognition of the identity of the presented face/car in images that varied in appearance. The authors used multivoxel pattern analysis to compare the response patterns elicited by two sets of the same identity (“within identity”) with that elicited by two sets of different identities (“between identity”). A region showing higher correlation for objects varying in appearance but matching in identity, compared with objects varying in both appearance and identity,

could be interpreted as representing abstract object identity.

In all three experiments, superior intraparietal sulcus (IPS), in the dorsal stream, was the only region to show significantly higher within-identity than between-identity correlation, for both faces and cars [Jeong and Xu (2016), their Figs. 2, 3, 4*B*]. Intriguingly, this was not true of any of the ventral regions examined [lateral occipital (LO), fusiform face area (FFA), parahippocampal place area (PPA), visual word-form area (VWFA)]. The authors used representational similarity analysis (Kriegeskorte et al., 2008), a mathematical technique used to relate neural representational structure to behavior and/or computational models, to compare the similarity among neural representations of faces to the similarity among perceptual judgments of face identity. The former was measured as the correlation between response patterns to each identity pair in each region, and the latter as the reaction time to detect one identity among distractors of another identity (such that the greater the similarity, the slower the reaction time). A correlation between the neural and behavioral representations was again found solely in IPS [Jeong and Xu (2016), their Fig. 4*D*]. The authors interpret their results as collectively showing that task-relevant object identity information is represented dorsally, in IPS.

These results are particularly interesting in light of another recently published study (Hong et al., 2016). Those authors showed that spatial properties of an object, such as the position, size, and pose, can be decoded from neural activity in

Received April 2, 2016; revised May 10, 2016; accepted May 13, 2016.

We thank Nikolaus Kriegeskorte and Andrew Bell for useful discussions regarding the article by Jeong and Xu (2016).

The authors declare no competing financial interests.

Correspondence should be addressed to Vassilis Pelekanos, Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK. E-mail: vassilis.pelekanos@mrc-cbu.ox.ac.uk.

DOI:10.1523/JNEUROSCI.1102-16.2016

Copyright © 2016 the authors 0270-6474/16/366368-03\$15.00/0

macaque inferotemporal cortex (a part of the ventral stream), where representations are often thought to be invariant to spatial particulars. Similar findings have been reported by Schwarzlose et al. (2008) and Kravitz et al. (2010). Combined with the study by Jeong and Xu (2016), these studies present results opposite to the predictions of the two-stream model and render the what versus where dichotomy increasingly questionable.

The dorsal identity representations discovered by Jeong and Xu (2016) provide a richer picture of neural visual object processing, although the division of labor between dorsal and ventral pathways remains unclear. Differences in anatomical coverage, experimental design, stimuli, and task render their findings difficult to integrate with previous studies showing that facial identity information is present in ventral regions including FFA (Nestor et al., 2011; Anzellotti et al., 2014) and more anterior temporal regions (Kriegeskorte et al., 2007). Jeong and Xu (2016) argue that the ventral face-identity information that these previous studies found may be due to low-level stimulus confounds, or that face-identity information may be present only in anterior temporal cortex, from which their MRI slice coverage precluded measurement. This will likely remain contentious until a study uses comprehensive anatomical coverage (including early visual cortex and the entire ventral and dorsal streams) combined with comprehensive exploration of the recorded areas (e.g., by searchlight analysis; Kriegeskorte et al., 2006).

Beyond anatomical coverage, an important difference between the experimental design in the study by Jeong and Xu (2016) and those in previous studies, which can potentially explain the negative findings in ventral cortex by Jeong and Xu (2016), is the use of an fMRI protocol with a block design. In a block design, images within a single-identity condition are presented sequentially, and a single multivoxel pattern is estimated for that identity, effectively averaging across individual images. Successive presentation has the benefit of increasing signal strength. However, the extraction of a single pattern within each block works only if the identity-specific component of the activity pattern survives after averaging across the variant components (e.g., position, background). That is, in block designs one assumes that pattern information related to identity combines additively with pattern information relating to

nonidentity attributes. If the neural code has, instead, a more complex format, where individual neurons encode conjunctions of identity and other attributes [as Hong et al. (2016) suggest], averaging across images would destroy much identity-specific information. An event-related design with single-image level analyses would be sensitive to both types of identity information.

To avoid low-level confounds (e.g., viewpoint, size, or shape), Jeong and Xu (2016) used widely varying photographs of each face and car identity. However, even in this stimulus set, spatial frequency distributions were somewhat predictive of identity as assessed by a linear classifier trained by the authors to classify images based on their spatial frequency profiles. The authors contend that the larger spatial frequency differences found for between-identity sets, compared with within-identity sets, cannot explain their main finding (lower fMRI pattern correlation for the former). They argue that those spatial frequency differences are unlikely to have driven the effects in IPS, since they had no measurable effect on responses in the “sensory regions” LO and FFA (these regions showed similar within-identity and between-identity correlations). This argument implicitly assumes that ventral visual regions are sensitive to a wide range of low-level image features, and that if they fail to encode differences in a particular low-level feature (spatial frequency), these differences will not be encoded by other regions either. The authors provide support for the first part of this argument. They compared the correlations of responses across odd and even runs within the same set of images with those between different image sets depicting the same identity. Regions sensitive to low-level image features should show higher within-set than between-set correlation. This was indeed found for ventral regions LO, FFA, PPA, and VWFA, but not for IPS [Jeong and Xu (2016), their Fig. 6]. However, the second part of the argument rests on an assumption. The fact that ventral regions were not sensitive to identity-related spatial frequency differences does not exclude the possibility that identity effects in IPS were driven by these differences. Besides, it has been previously shown that fMRI signals in response to faces in parietal cortex can be modulated by spatial frequency content (Vuilleumier et al., 2003).

Jeong and Xu (2016) highlight the role of IPS in visual short-term memory (VSTM). Given the diversity of tasks used in previous studies [image-anomaly-detection (Kriegeskorte et al., 2007); identification (Nestor et al., 2011); target vs nontarget categorization (Anzellotti et al., 2014)], it will be interesting for future studies to clarify the extent to which IPS results reflect VSTM content. Specifically, the information decoded in IPS may be a visual search template of the frequently presented identity, which is mentally formed by participants during each block as they searched for the oddball stimulus.

Finally, given that human fMRI studies have found increased IPS activity during spatial attention and eye movements (Müri et al., 1996; Corbetta et al., 1998), future studies should also carefully rule out the possibility that these could have partially driven the IPS results. Systematic differences in saccade distribution or spatial attention might occur if, for example, different individuals carry identity-diagnostic information in different facial features. The results of the study by Jeong and Xu (2016) highlight a need for future studies of visual object representation to extensively explore both dorsal and ventral regions, use tight stimulus controls, and explore the effects of different cognitive tasks within a stimulus set. Comprehensive surveys of this kind are crucial to understand how the primate brain accomplishes object recognition.

## References

- Anzellotti S, Fairhall SL, Caramazza A (2014) Decoding representations of face identity that are tolerant to rotation. *Cereb Cortex* 24:1988–1995. [CrossRef Medline](#)
- Bell AH, Hadj-Bouziane F, Frihauf JB, Tootell RB, Ungerleider LG (2009) Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. *J Neurophysiol* 101:688–700. [CrossRef Medline](#)
- Corbetta M, Akbudak E, Conturo TE, Snyder AZ, Ollinger JM, Drury HA, Linenweber MR, Petersen SE, Raichle ME, Van Essen DC, Shulman GL (1998) A common network of functional areas for attention and eye movements. *Neuron* 21:761–773. [CrossRef Medline](#)
- Freedman DJ, Assad JA (2006) Experience-dependent representation of visual categories in parietal cortex. *Nature* 443:85–88. [CrossRef Medline](#)
- Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. *Trends Neurosci* 15:20–25. [CrossRef Medline](#)
- Hong H, Yamins DL, Majaj NJ, DiCarlo JJ (2016) Explicit information for category-orthogonal object properties increases along the ventral stream. *Nat Neurosci* 19:613–622. [CrossRef Medline](#)

- Jeong SK, Xu Y (2016) Behaviorally relevant abstract object identity representation in the human parietal cortex. *J Neurosci* 36:1607–1619. [CrossRef Medline](#)
- Kravitz DJ, Kriegeskorte N, Baker CI (2010) High-level visual object representations are constrained by position. *Cereb Cortex* 20:2916–2925. [CrossRef Medline](#)
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103:3863–3868. [CrossRef 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](#)
- Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, Bandettini PA (2008) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60:1126–1141. [CrossRef Medline](#)
- Li N, Cox DD, Zoccolan D, DiCarlo JJ (2009) What response properties do individual neurons need to underlie position and clutter “invariant” object recognition? *J Neurophysiol* 102:360–376. [CrossRef Medline](#)
- Milner AD, Goodale MA (2008) Two visual systems re-viewed. *Neuropsychologia* 46:774–785. [CrossRef Medline](#)
- Müri RM, Iba-Zizen MT, Derosier C, Cabanis EA, Pierrot-Deseilligny C (1996) Location of the human posterior eye field with functional magnetic resonance imaging. *J Neurol Neurosurg Psychiatry* 60:445–448. [CrossRef Medline](#)
- Nestor A, Plaut DC, Behrmann M (2011) Unravelling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proc Natl Acad Sci U S A* 108:9998–10003. [CrossRef Medline](#)
- Rust NC, Dicarlo JJ (2010) Selectivity and tolerance (“invariance”) both increase as visual information propagates from cortical area V4 to IT. *J Neurosci* 30:12978–12995. [CrossRef Medline](#)
- Schwarzlose RF, Swisher JD, Dang S, Kanwisher N (2008) The distribution of category and location information across object-selective regions in human visual cortex. *Proc Natl Acad Sci U S A* 105:4447–4452. [CrossRef Medline](#)
- Toth LJ, Assad JA (2002) Dynamic coding of behaviourally relevant stimuli in parietal cortex. *Nature* 415:165–168. [CrossRef Medline](#)
- Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In: *Analysis of visual behavior* (Ingle DJ, Goodale MA, Mansfield RJW, eds), pp 549–586. Cambridge, MA: MIT.
- Vuilleumier P, Armony JL, Driver J, Dolan RJ (2003) Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nat Neurosci* 6:624–631. [CrossRef Medline](#)