

Response to the Journal Club by Pelekanos et al.

Yaoda Xu (Harvard University) and Sukeun Jeong (Princeton University)

We thank the authors for highlighting our article. Nevertheless, we disagree with some of the assertions made in the Journal Club article and would like to clarify a few important issues. We hope the authors and the readers of the Journal Club article will find our response informative.

In our study we used a block fMRI design and averaged response patterns from different face images sharing an identity but varied in viewpoints and expression. Pelekanos et al. argued that if there was joint coding of identity and other features, such as viewpoint, then identity representations might not survive image averaging and become undetectable. This could explain why we failed to find identity decoding in posterior ventral regions while others did using an event-related fMRI design (Nestor et al., 2011; Anzellotti et al., 2014). However, Pelekanos et al. failed to note that Nestor et al. (2011) combined responses from images of the same identity across expressions before pattern classification was performed. Likewise, Anzellotti et al. (2014) trained the classifier across multiple viewpoints to obtain a viewpoint invariant representation. Thus a similar averaging procedure was used in those studies. In Hong et al. (2016), although location and category were coded together, there was still a separation in the representational space such that category could be decoded across changes in location. A similar scheme could work for faces and an averaging procedure is critical in reducing the complexity of the representational space to isolate an identity representation across changes in identity-irrelevant visual features. Our failure to decode identity in posterior ventral region thus could not be due to the use of a block design, but rather to the removal of low-level visual features that likely contributed to identity decoding in the other studies.

Pelekanos et al. argued that differences in spatial frequency distribution among the image sets could have contributed to identity decoding in superior IPS. Although Vuilleumier et al. (2003) showed that spatial frequency could modulate parietal face responses, their parietal activation (Talairach coordinates: left -37/-68/23 and right 51/-54/6, converted from MNI space) was located differently than superior IPS (-21/-66/32 and 23/-52/45). Moreover, we found no name decoding in superior IPS despite the fact that the name sets also differed in spatial frequency distribution. Since word processing relies on high spatial frequency information just like faces and parietal cortex is not intrinsically face selective, it would be rather peculiar that spatial frequency only benefited face decoding but nothing else in superior IPS. Lastly, without any difference in spatial frequency distribution between the car image sets, we still found car identity decoding in superior IPS. Differences in spatial frequency distribution thus could not explain identity decoding in superior IPS.

Pelekanos et al. wondered if perhaps our identity decoding was contaminated by differences in eye-movements or spatial attention for the different face identities. In our study, the between-identity similarity was high as all target faces were from mid-aged white Caucasian males, and within-identity image similarity was low as face images from the same identity varied in viewpoint and expression. During the experiments, we never asked participants to directly compare the different target faces. But rather, they viewed target faces one at a time with different faces from the same identity shown sequentially within the same trial block. The main task was to detect the presence of one of eight (Experiment 1) or sixteen (Experiment 2) distractor actor faces that occurred in 20% of the blocks or less (blocks containing a distractor face were removed from further analysis). Because different face features were visible in different face images and different diagnostic features would be informative in distinguishing different target and distractor faces with the participants unable to anticipate the identity of the distractor actor, it is unclear how a strategy involving different eye-movement or spatial attention for each target face could be implemented to aid task performance. To the extent that face processing is holistic, this seems even less plausible. We actually monitored eye-movements in Experiments 2 and 3 but omitted to mention it in the paper. We did not find systematic bias in eye-movement patterns.