Neural representation of occluded objects in visual cortex

Fraser W Smith, School of Psychology, University of East Anglia
Fraser.Smith@uea.ac.uk

Courtney Mansfield, Department of Psychology, University of Cambridge

Tim Kietzmann, Institute of Cognitive Science, University of Osnabruck, Germany

Ian Charest, Department of Psychology, University de Montreal

Marieke Mur, Brain & Mind Institute, Western University

Niko Kriegeskorte, Zuckerman Institute, Columbia University

Abstract

The ability of the human visual system to recognize occluded objects is striking, yet current models of vision struggle to account for this successfully. Our goal was to understand what best explains neural representations of occluded objects under more realistic occlusion i.e., when objects occlude other objects. In an event-related fMRI design, participants performed a one-back task while being presented with unoccluded objects, objects occluded by another object, or cut out by a corresponding object silhouette. Decoding analyses showed that EVC responses to occluded objects were better determined by the visible features whereas in IT inferred features better predicted the response. Projection analyses showed that the weights assigned to occluded objects in IT were predicted by independent categorization judgements. Our results demonstrate that IT better decouples responses to real-world occluded objects with robust representations evident across multiple competing objects.

Keywords: High-level vision; occlusion; fMRI; decoding; MVPA;

Introduction

The ability of the human visual system to recognize occluded objects is striking, yet current models of vision struggle to account for this successfully. Previous studies investigating occlusion at both the behavioural and neural levels typically used simple shapes or cut outs as occluders, rather than other objects (e.g. Smith & Muckli 2010; Tang et al 2014, 2017; Johnson & Olhausen 2005; Fyall et al 2017; but see Spoerer et al 2017). However in our visual world, it is typically meaningful objects that occlude other meaningful objects. The goal of the present study was to understand what best explains neural representations of occluded objects under more realistic occlusion i.e., when objects occlude other objects. We approached this question by developing a novel stimulus set that enabled us to explicitly relate activity patterns of occluded objects (e.g. a cup occluding a face) with those generated when viewing the same objects in isolation (the cup or the face). This provides a way to determine how the representation of an object when occluded relates to the same object representation in the absence of occlusion.

Methods

Participants

The sample included N=21 participants, pooled across two testing sites (N = 12 Cambridge; N = 9 Norwich UK; both using Siemens Prisma 3T). The experiments were reviewed and approved by relevant ethics committees in each site.

Stimuli

We created a novel stimulus set where 8 original objects were combined in a systematic manner to create all possible versions of occluded object pairs (Occluded Object Trials N=56). We also created a series of control images where the same features of the occluded object were present with the occluding object deleted (Deleted Trials N=56). The original unoccluded object images were also presented (single object trials N=8).

Procedure

Each visual stimulus was presented for 1s with a 3.5s ITI (25% null trials). Participants fixated and performed a one-back task, where they had to respond if any object in the current trial was present in the immediately preceding trial. Each run comprised all trial types presented in a pseudo-random order (128 trials) and participants completed 4-6 runs each.

fMRI Acquisition & Processing

Blood-oxygen level dependent (BOLD) signals were recorded using a multiband echo-planar imaging (EPI) sequence: (TR = 1240ms; TE = 30ms; voxel size = 2X2X2; multiband factor 2). The fMRI data was pre-processd with Brain Voyager (v 20.4). The Glasser parcellation was used in FreeSurfer to define three ROIs (EVC, MID, IT - see Kietzmann et al 2019). GLM analysis was conducted per run per participant, with a different predictor for each stimulus presentation.

Decoding & Projection Analyses

Linear support vector machine (SVM) cross-decoding of object identity (N=8) with leave one run out cross-validation (e.g. Bailey et al. 2023) was used to determine how occluded objects related to unoccluded single object presentations. Projection analyses (Reddy et al 2009) were carried out to estimate the weight assigned to the occluded (back) object in terms of the response to the unoccluded version of that object (single).

Results

Cross-decoding analyses, where the classifiers were trained on responses to either unoccluded objects (single object trial) or the corresponding visible features (deleted features trial) and tested to predict identity of the occluded object (back), reveal higher accuracy in EVC for the visible features whereas in IT the model trained on unoccluded objects resulted in higher accuracy (Figure 1; 2 way interaction, p = .0006; permutation ANOVA).

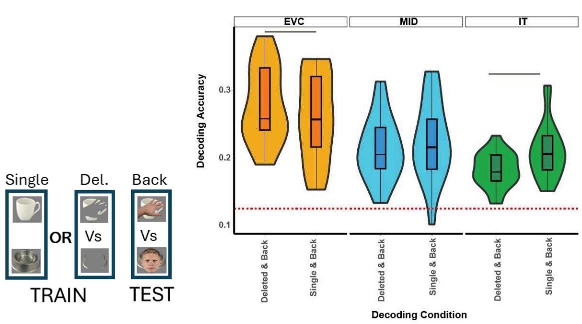

Figure 1: Cross-decoding which features best predict responses to occluded objects

Additional cross-decoding between responses to unoccluded single objects and other conditions revealed that EVC was affected more strongly by the presence of multiple objects compared to IT (Figure 2, 2 way interaction, p = .0004; Permutation ANOVA).

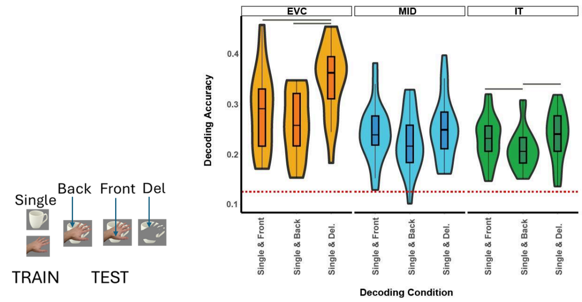

Figure 2: Cross-decoding reveals the cost of multiple objects in a display

Finally projection analyses revealed that the weights assigned to occluded objects was predicted by independent human behavioural reaction times (Figure 3, rS=0.425, p=.0012, by items; behavioural recognition data acquired from independent group of participants N=29).

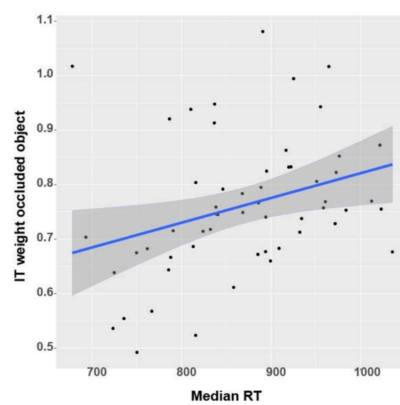

Figure 3: Correlation between weights assigned to occluded objects in IT and behaviour (RT)

Discussion

In sum the results demonstrate that IT better decouples responses to real-world occluded objects with robust representations evident across multiple competing objects. Responses in IT to occluded objects are also directly related to human categorization behaviour. Thus, our data support the importance of investigating neural mechanisms underlying object recognition under more naturalistic occlusion scenarios.

References

Bailey KM, Giordano BL, Kaas A & Smith FW (2023). Decoding sounds depicting hand-object interactions in primary somatosensory cortex. *Cerebral Cortex*, 33, 7, p. 3621–3635.

Fyall, A.M., El-Shamayleh J., Choi H, Shea-Brown, E. & Pasupathy A. (2017) Dynamic representation of partially occluded objects in primate prefrontal and visual cortex *eLife* 6:e25784.

Johnson, J. S., & Olshausen, B. A. (2005). The recognition of partially visible natural objects in the presence and absence of their occluders. *Vis. Res*. 45, 3262–3276.

Kietzmann,T.C., Spoerer,C.J., Sörensen,L.K.A., Cichy, R.M., Hauk, O., & Kriegeskorte,N. (2019). Recurrence is required to capture the representational dynamics of the human visual system, *PNAS*, 116 (43) 21854-21863,

Reddy, L, Kanwisher., N., & Van Rullen R. (2009). Attention and biased competition in multi-voxel object representations. *PNAS*, 106 (50) 21447-21452.

Smith, F.W. & Muckli, L. (2010). Non-stimulated early visual areas carry information about surrounding context. *PNAS*, 107 (46), 20099-20103.

Spoerer, CJ., McClure, P., & Kriegeskorte, N. (2017). Recurrent convolutional neural networks: a better model of biological object recognition. *Frontiers in Psychology*, 8,10.3389/fpsyg

Tang, H., Buia, C., Madhavan, R., Crone, N. E., Madsen, J. R., Anderson, W. S., et al. (2014). Spatiotemporal dynamics underlying object completion in human ventral visual cortex. *Neuron* 83, 736–748.

Tang, H., Schrimpf, M., Lotter, W., Moerman, C., Paredes, A., Ortega Caro, J., W. Hardesty, W., Cox, D., & Kreiman, G. (2018). Recurrent computations for visual pattern completion, *PNAS*,. 115 (35) 8835-8840.