

Supplementary Information

1 Extended Methods

1.1 Sociality and Interaction Envelope

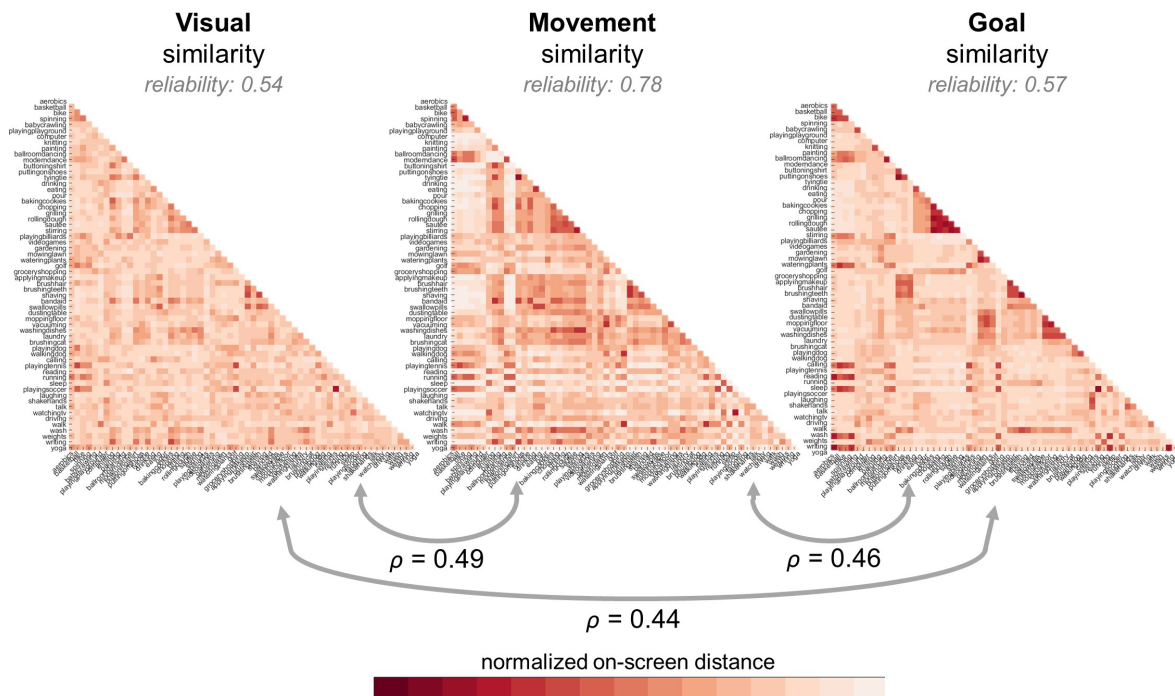
In previous work Tarhan and Konkle (2020b), we found evidence that the visual cortex’s responses to action videos are organized by two features: *sociality* (whether an action is directed at a person) and the size of the *interaction envelope* (the spatial extent of an action’s effect on the world). The fMRI experiment in that work used the same set of action videos as in the current study. Because these features predicted action processing well in a large swathe of the brain, we wondered whether they captured information that informs the downstream processing supporting intuitive action understanding. To investigate this, we asked how well sociality and interaction envelope size could predict the intuitive similarity judgments.

To calculate the sociality and interaction envelope features for each video, we returned to the analysis that revealed this pattern in our previous work. In that work, a clustering analysis revealed 5 networks of brain regions. We interpreted the tuning of each network by examining the networks’ tuning to the different body parts involved in the action videos and the actions’ targets (what they were directed at, such as an object or a person). One network was tuned to actions directed at people, such as talking (sociality; **Supplemental Figure 2a**), while the remaining four were tuned to different interaction envelope sizes (**Supplemental Figure 2b**). For example, one network was tuned to small interaction envelopes, as in actions that involve fine, object-directed hand movements like knitting. At the other extreme, another network was tuned to larger interaction envelopes, as in actions that involved large movements of the whole body, directed at distant locations like a soccer penalty shot.

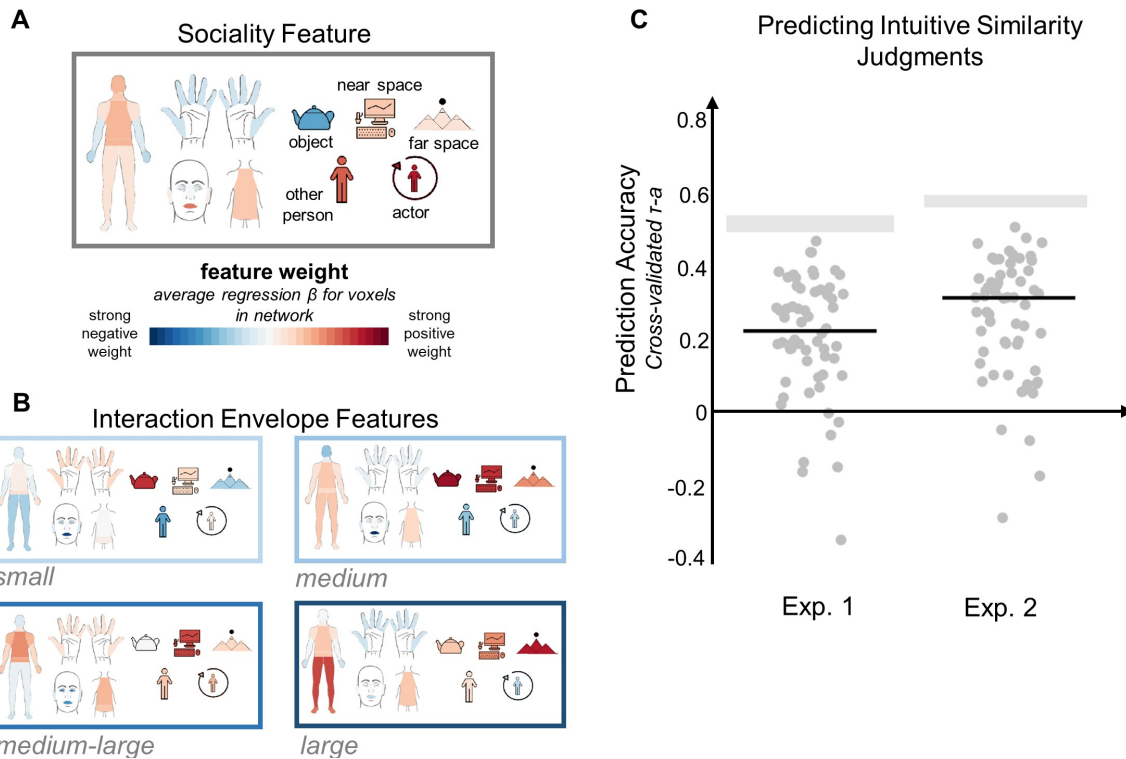
For each action video, we had measured which body parts were engaged by the action and what the action was directed at, using human ratings (see Tarhan and Konkle, 2020b for details). And for each network, we had calculated a tuning profile to these body parts and targets. So, to measure where each video fell along these five sociality and interaction envelope dimensions, we multiplied these body part and target ratings by each network’s tuning profile. This produced an estimate of how well each video aligned with each network’s preferred tuning. For example, a video of two people shaking hands was rated as being directed at other people and involved only the hands and arms. This video would then have a high value on the “sociality” and “small interaction envelope” dimensions, but lower values for the “medium interaction envelope,” “medium-large interaction envelope,” and “large interaction envelope” dimensions.

Finally, we assessed how well these features predict the intuitive similarity judgments in both experiments. In all three experiments, the sociality-interaction envelope features predicted the intuitive similarity judgments moderately well (**Supplemental Figure 2c**; mean cross-validated τ -a = 0.21, sd = 0.16 (Experiment 1); 0.26, s.d. = 0.16 (Experiment 2)). However, these features performed worse than the three model RDMs based on the actors’ goals, movements, and the videos’ visual appearance. This suggests that, while sociality and interaction envelope predict action responses well in the visual cortex, they do not add much when predicting intuitive judgments.

Supplemental Figures



Supplemental Figure 1: Comparing Model RDMs. Group-level Representational Dissimilarity Matrices (RDMs) are shown for the three types of action similarity judgments. Split-half reliabilities (after Spearman-Brown prophecy corrections) are listed for each type of judgment. Numbers beneath the grey arrows indicate the Spearman's correlations between these model RDMs.



Supplemental Figure 2: Sociality and Interaction Envelope. In prior work (Tarhan and Konkle, 2020b), we found that sociality and interaction envelope size predict action responses throughout the visual cortex. To assess whether these properties also influence downstream intuitive action processing, we asked how well similarity in actions' sociality and the size of their interaction envelopes could predict intuitive similarity judgments. (A) Illustration of the sociality feature dimension, showing how much voxels in a right-lateralized network of brain regions were tuned to the body parts and targets involved in an action (Tarhan and Konkle, 2020b). Each body part and target is colored according to the strength of the network's tuning – for example, this network was strongly tuned to actions directed at the actor or another person, but was not strongly tuned to actions directed at far space. (B) Illustrations of the four interaction envelope feature dimensions, which range from small envelopes around fine movements directed at objects (e.g., knitting) to large envelopes around coarser movements directed at distant locations (e.g., a soccer penalty shot). (C) We used these five dimensions – sociality and four sizes of interaction envelope – to predict intuitive similarity judgments about the actions. Prediction performance is plotted for the intuitive judgments measured in each experiment. Grey bars indicate the noise ceiling for each experiment. Horizontal black lines indicate the median prediction performance for each experiment, and grey dots plot performance on each iteration of the leave-1-condition-out procedure.