

# Cortical regions preferentially engaged during social and physical processing represent obstacles to agent action and object motion

Minjae Kim (mkim19@jhu.edu) Sam Maione (samaione1@jhu.edu)  
Assiya Drissi (adrissi1@jhu.edu) Shari Liu (sliu199@jhu.edu)

Dept Psychological and Brain Sciences, Johns Hopkins University  
3400 N. Charles Street, Baltimore, MD 21218 USA

## Abstract

In cognitive neuroscience, we routinely study the neural substrates of social cognition and physical cognition by directly contrasting neural responses in social and physical tasks. Yet, this approach can occlude computations that are shared across domains. Using an open fMRI dataset, we test the hypothesis that two cortical regions previously shown to be preferentially engaged for social and physical perception – the superior temporal sulcus (STS) and the supramarginal gyrus (SMG) – nonetheless contain representations relevant for both domains. Participants were scanned as they watched and made predictions about two dots interacting like social agents or inanimate objects. In exploratory functional region-of-interest analyses, we found that multivariate patterns in STS and SMG contained information about whether the scene included a physical barrier, despite showing strong and opposite preferences for social and physical videos. These findings suggest that cortical regions specialized for social and physical functions may share representational content.

**Keywords:** cognitive neuroscience; domain-specificity; physical reasoning; social cognition

Every day, we navigate a physical and social world, filled with objects, obstacles, bodies, and minds. Distinct cortical regions in the human brain support processing physical and social information. Regions in the premotor and parietal cortex are responsive during physical prediction and inference and contain representations of mass and stability (Pramod et al., 2022, Schwettmann, Tenenbaum, & Kanwisher, 2019, Fischer et al., 2016). Regions in the Theory of Mind (ToM) network, including superior temporal sulcus and temporoparietal junction, are preferentially engaged during social reasoning, and contain representations about social interactions (Dodell-Feder et al., 2011, Isik, Koldewyn, Beeler, & Kanwisher, 2017). These two systems have been hypothesized to serve distinct functions (Mitchell, Heatherton, & Macrae, 2002) and to compete with each other (Jack et al., 2013). Yet, the social and physical worlds are not independent (Liu, Outa, & Akbiyik, 2024); people’s mental states, actions, and interactions depend on their current physical states, their desired goal states, and what obstacles stand in their way (Baker et al., 2017, Jara-Ettinger et al., 2016, Gergely & Csibra, 2003). So how do our minds and brains combine these two domains of information? If physics regions and ToM regions are functionally segregated (Jack et al., 2013), then they should

only be sensitive to physical and social features, respectively, which are then combined downstream. In contrast, we hypothesize that these regions work together, in part by representing shared information (e.g., physical information relevant for interpreting social and physical interactions).

## Methods

To test these alternative hypotheses, we studied neural responses to simplified social and physical interaction events in a publicly available dataset. Analyses were conducted in four regions-of-interest (ROIs): left and right supramarginal gyrus (SMG), which are implicated in physical reasoning, and left and right superior temporal sulcus (STS), which are implicated in ToM. Do STS and SMG only represent social and physical information, respectively? Or are these regions also sensitive to information from the “opposite” domain? Here, we test whether the STS and SMG are sensitive to features that should be relevant for both domains: physical obstacles, and physical collisions.

## Dataset, task, and stimulus features

We analyzed an open fMRI dataset, including data from 45 participants (Liu, Lydic, et al., 2024). Here we present exploratory analyses in one subset of the data (N = 16 18- to 45-yo adults). In the scanner, participants saw two runs, each containing 32 8s videos of interactions between two 2D circles (Fig. 1B; (Fischer et al., 2016). In the physical interaction condition, the circles moved like solid objects (e.g., bouncing off the walls and each other); in the social interaction condition, the circles moved like social agents (e.g., chasing and imitating each other). Half of the videos in each condition included a physical barrier. Participants predicted the trajectory of one of the dots after it disappeared, and reported whether its reappearance at the end of the trial was plausible or implausible by button press. Each video was characterized in terms of the following three features: (1) domain (physical vs. social interaction); (2) physical constraint (present vs. absent); and (3) total number of collisions (always 0 for social videos, range 7 to 44 for physical videos).

## Overview of analyses

We used fMRIPrep (Esteban et al., 2019) to pre-process the data (see [here](#) for a full description), and a custom pipeline written using nipy (Gorgolewski et al., 2011) to extract run-level responses for all videos for all subjects. Scripts and fROI data required to reproduce the results are openly available on [OSF](#).

**fROI identification.** We used data and parcels from an independent task involving video stimuli of agents and objects (Liu, Lydic, et al., 2024) to functionally identify ROIs (fROIs) in each participant (Fig. 1A). LSMG and RSMG were identified using the contrast: objects > agents; LSTS and RSTS were identified using the opposite contrast. The top 100 voxels within each parcel were selected for further analysis.

**Univariate analyses.** For each fROI, we measured the average response to each video stimulus. We then used linear mixed effects models in R (“lme4”; Bates, 2010) to model the average response as a function of domain (physical vs. social interaction), constraint (present vs. absent), and collisions. Models included by-subject and by-video random intercepts. The resulting coefficients represent how much each feature drives the amplitude of response in an fROI, controlling for the other two features. The “afex” package (Singmann et al., 2015) was used to get  $p$  values for fixed effects.

**Representational similarity analyses.** For each of the three features, we computed a representational dissimilarity matrix (RDM) of feature distances across pairs of stimuli (Fig. 1D). The domain and constraint RDMs included 0s and 1s (same domain vs. not; both have/lack a barrier vs. not), and the collision RDM consisted of absolute differences in collisions. Next, for each fROI in each participant, we computed an RDM that represents neural pattern distances ( $1 - \text{Pearson's } r$ ) across the same pairs of stimuli. Then, for each feature, we computed the partial Spearman's  $\rho$  between the feature RDM and the neural RDM, controlling for the other two feature RDMs. We used permutation tests for significance testing (Stelzer, Chen, & Turner, 2013): We generated the null distribution of the  $t$  statistic by repeating the same partial correlation analysis over RDMs with scrambled labels ( $10^4$  iterations). The  $p$  value was the proportion of null  $t$  values that were greater than the empirical  $t$  value.

## Results

First, we replicated the finding that STS responds preferentially to social interactions (LSTS:  $p = 0.001$ ; RSTS:  $p < .0001$ ), and SMG responds preferentially to physical interactions (LSMG:  $p = 0.035$ ; RSMG did not show a domain preference,  $p = 0.478$ <sup>1</sup>). Second, none of these regions preferentially responded to events with or without a barrier ( $p > 0.109$ ), or parametrically increased their activity with the number of collisions ( $p > 0.308$ ). Third, multivariate patterns in these regions encoded the domain of the stimulus (LSMG:  $\rho = 0.023$ ,  $p < 0.0001$ ; RSMG:  $\rho = 0.045$ ,  $p = 0.004$ ; LSTS:  $\rho = 0.062$ ,  $p < 0.0001$ ; RSTS:  $\rho = 0.047$ ,  $p < 0.0001$ ), and the presence of a physical obstacle in that stimulus (LSMG:  $\rho = 0.013$ ,  $p = 0.006$ ; RSMG:  $\rho = 0.022$ ,  $p = 0.0001$ ; LSTS:  $\rho = 0.016$ ,  $p = 0.001$ ; RSTS:  $\rho = 0.017$ ,  $p = 0.006$ ). No fROI contained multivariate information about the number of collisions.

<sup>1</sup>The null univariate result in RSMG appears to hinge on our ROI definition: When separate runs of our focal task are used to choose fROIs, RSMG responds preferentially to physical interactions ( $p = 0.0049$ ).

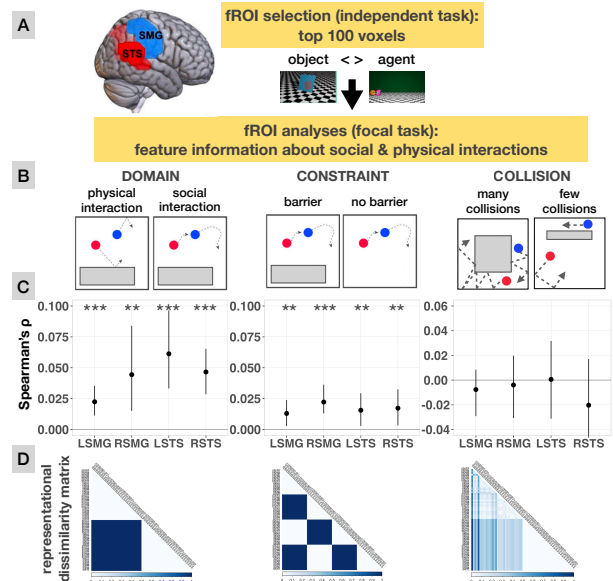


Figure 1: fROI selection, stimulus features, and multivariate results. (A-B) An independent task was used to identify SMG and STS; 3 features were tested via RSA. (C) All regions contained information about domain and physical constraints. (D) RDMs represent feature distances between pairs of videos.  $p$  \* < .05; \*\* < .01; \*\*\* < .001. See OSF for full-sized figure.

sions. See Fig. 1C. All RSA results still obtain when the other features are not partialled out.

## Discussion

Distinct cortical regions in human brains are engaged when we see and think about the social and physical world. But if our understanding of the social world depends on information about the physical world (Liu, Outa, & Akbiyik, 2024), then it is plausible that these regions share functions. For example, sensitivity to obstacles to agents' actions and object motion is one of the most early-emerging cognitive abilities in human infants (Gergely & Csibra, 2003; Spelke, Breinlinger, Macomber, & Jacobson, 1992), and is thus likely fundamental to adults' social and physical cognition. Here, we found that regions classically associated with social and physical functions, the STS and SMG, contain both univariate and multivariate information about domain (social vs. physical interaction). Yet, both regions also contained representations about the presence vs. absence of a physical constraint, despite a univariate preference for social information in STS. Our results are consistent with work showing that frontoparietal and superior temporal regions represent the physics of both actions and object events (Karakose-Akbiyik, Caramazza, & Wurm, 2023). So far, our work does not address how early visual information contributes to these representations. It is also unclear whether they originate in STS, SMG, or elsewhere, and what causal role they play. Nevertheless, the current results raise the possibility that ToM regions represent not only social

information, but also physical information relevant for social cognition.

## References

- Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017, March). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nat. Hum. Behav.*, *1*(4), 1–10.
- Bates, D. M. (2010). *lme4: Mixed-effects modeling with R*. Springer.
- Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fmri item analysis in a theory of mind task. *neuroimage*, *55*(2), 705–712.
- Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., ... others (2019). fmriprep: a robust preprocessing pipeline for functional mri. *Nature methods*, *16*(1), 111–116.
- Fischer, J., Mikhael, J. G., Tenenbaum, J. B., & Kanwisher, N. (2016, August). Functional neuroanatomy of intuitive physical inference. *Proc. Natl. Acad. Sci. U. S. A.*, *113*(34), E5072–81.
- Gergely, G., & Csibra, G. (2003, July). Teleological reasoning in infancy: the naive theory of rational action. *Trends Cogn. Sci.*, *7*(7), 287–292.
- Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in neuroinformatics*, *5*, 13.
- Isik, L., Koldewyn, K., Beeler, D., & Kanwisher, N. (2017). Perceiving social interactions in the posterior superior temporal sulcus. *Proceedings of the National Academy of Sciences*, *114*(43), E9145–E9152.
- Jack, A. I., Dawson, A. J., Begany, K. L., Leckie, R. L., Barry, K. P., Ciccio, A. H., & Snyder, A. Z. (2013, February). fMRI reveals reciprocal inhibition between social and physical cognitive domains. *Neuroimage*, *66*(C), 385–401.
- Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in cognitive sciences*, *20*(8), 589–604.
- Karakose-Akbiyik, S., Caramazza, A., & Wurm, M. F. (2023, June). A shared neural code for the physics of actions and object events. *Nat. Commun.*, *14*(1), 3316.
- Liu, S., Lydic, K., Mei, L., & Saxe, R. (2024). Violations of physical and psychological expectations in the human adult brain. *Imaging Neuroscience*, *2*, 1–25.
- Liu, S., Outa, J., & Akbiyik, S. (2024, Mar). *Naive psychology depends on naive physics*. PsyArXiv. Retrieved from [osf.io/preprints/psyarxiv/u6xdz](https://osf.io/preprints/psyarxiv/u6xdz) doi: 10.31234/osf.io/u6xdz
- Mitchell, J. P., Heatherton, T. F., & Macrae, C. N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences*, *99*(23), 15238–15243.
- Pramod, R. T., Cohen, M. A., Tenenbaum, J. B., & Kanwisher, N. (2022, May). Invariant representation of physical stability in the human brain. *Elife*, *11*.
- Schwettmann, S., Tenenbaum, J. B., & Kanwisher, N. (2019). Invariant representations of mass in the human brain. *Elife*, *8*, e46619.
- Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M., Højsgaard, S., ... others (2015). Package 'afex'. URL <http://afex.singmann.science/>, <https://github.com/singmann/afex>.
- Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological review*, *99*(4), 605.
- Stelzer, J., Chen, Y., & Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (mvpa): random permutations and cluster size control. *Neuroimage*, *65*, 69–82.