Neural Representation of Social Relationship Graphs through Multidimensional Modeling of Dynamic Social Interactions

Dasom Kwon (dasom.kwon@g.skku.edu)

Center for Neuroscience Imaging Research, Institute for Basic Science,
Suwon, 16419, Republic of Korea
Department of Intelligent Precision Healthcare Convergence, Sungkyunkwan University,
Suwon, 16419, Republic of Korea

Eshin Jolly (e3jolly@ucsd.edu)

Department of Psychology, University of California San Diego,
La Jolla, CA, 92093, USA

Luke J. Chang (luke.j.chang@dartmouth.edu)

Department of Psychological & Brain Sciences, Dartmouth College,
Hanover, NH, 03755, USA

Won Mok Shim (wonmokshim@skku.edu)

Center for Neuroscience Imaging Research, Institute for Basic Science,
Suwon, 16419, Republic of Korea
Department of Intelligent Precision Healthcare Convergence, Sungkyunkwan University,
Suwon, 16419, Republic of Korea
Department of Biomedical Engineering, Sungkyunkwan University,
Suwon, 16419, Republic of Korea

Abstract

Social interactions continuously evolve, shaping our understanding of interpersonal relationships. Yet, how does the human brain construct relational knowledge from such dynamics? Prior research has primarily relied on unidimensional co-occurrence metrics, failing to capture the complexity of real-world social dynamics. Here, we introduce a multidimensional modeling framework characterizing dynamic social interactions as valence-weighted graphs. Using fMRI data collected during movie-viewing and subsequent relationship rating tasks, we show that distributed brain regions track dynamic interactions and represent social relationship graphs. These representations were preserved across tasks, with higher dimensionality observed in the medial prefrontal cortex (mPFC) and lower dimensionality in the posterior superior temporal sulcus (pSTS). These findings bridge online social perception and structured relational knowledge, elucidating how the brain organizes dynamic social interactions into multi-layered interpersonal relationship graphs.

Keywords: social interaction; multidimensional network modeling; naturalistic fMRI

Introduction

One of the critical challenges for human social inference is navigating the *dynamic* nature of social interactions and relationships (Meijerink-Bosman et al., 2023). Friends may become enemies, or romantic partners, highlighting how evolving interactions serve as a strong social signal that facilitates inferences about different types of relationships. Even infants incorporate the valence of interactions (e.g., helping vs. hindering) to make inferences about others (Hamlin et al., 2013).

Recent naturalistic neuroscience approaches have explored how humans accomplish this by using rich social stimuli (e.g. TV dramas) combined with models based on the visual co-occurrence of individuals. Social graphs based upon co-occurrence can explain how individuals represent and remember others (Jolly et al, 2023) and graph neural networks using only co-occurrence can predict social evaluations without complex computations like inverse planning (Malik and Isik, 2023). Neural encoding models based on co-occurrence predict voxel-wise activity in the lateral-temporal cortices during narrative viewing and recall (Masson & Isik, 2021; Masson et al., 2024) and population response patterns in these regions reflect generalized metrics such as social distance (Parkinson et al., 2017; Peer et al., 2021).

However, co-occurrence alone is not sufficient for capturing the dynamic and nuanced nature of interactions: characters may co-occur as friends or enemies. To capture these, we used directed, multidimensional graphs that reflect distinct relational features for each edge between individuals. We examined whether this valence-based graph model provides a comprehensive account of relationship evaluations and their underlying neural representations.

Methods

Task. During fMRI (3T, 3 mm iso, TR = 1 sec) scans, participants (N = 24) viewed a movie depicting evolving relationships among six characters. After the movie, they performed a relationship-rating task, reporting their beliefs about each character’s feelings toward others across six dimensions: closeness, love, trust, listening, relationship duration, and perceived co-occurrence.

Dynamic social interaction model. We modeled pairwise interactions in the movie using valence tensors (positive, neutral, negative) derived from moment-by-moment annotations of characters’ actions and emotions, scored with a Korean sentiment lexicon (Park et al., 2018). These tensors were then aggregated over time into multidimensional valence matrices that captured the evolving structure of social relationships.

Neural encoding and reconstructing social relationship graphs. We trained a voxel-wise encoding model (Naselaris et al., 2011) to predict BOLD responses during movie-viewing based on moment-by-moment co-occurrence and valence-based graphs. These models were used to reconstruct multi-layered social graphs (e.g., friendship-based, romantic, and emotionally salient relationships) from brain activity.

Post-movie relationship representations. During the relationship-rating task, neural responses were modeled with a GLM using separate regressors for each character. We then performed whole-brain searchlight representational similarity analysis (RSA) (Kriegeskorte et al., 2008) to predict neural RSMs using multi-layered social relationships.

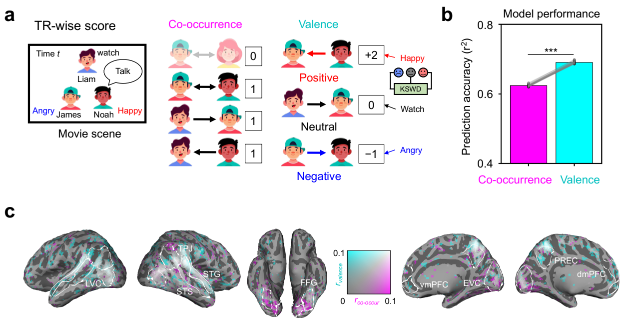

Figure 1: a) Multidimensional information in social interactions. b) Model comparison for predicting participants' relationship ratings. c) Neural encoding of co-occurrence vs. valence information (qs < .05).

Results

We used partial least squares regression (PLS) to predict participants' ratings from either unidimensional co-occurrence or multidimensional valence tensors (Figure 1a). Valence model significantly outperformed the co-occurrence model (Figure 1b; $t(23) = 147.77$, $p < .001$). When fit to brain activity, neural encoding model robustly predicted voxel activity across distributed regions, with valence tensors encoded in the medial-prefrontal and lateral-temporal cortices, and co-occurrence encoded in the visual cortex (Figure 1c).

To explore whether the brain represents the multi-layered social graph in the movie, we used the voxel-wise encoding weights and reconstructed graphs learned by participants' ratings using PLS regression (Figure 2a, 2b). These reconstructions revealed regional variation in representational complexity: graphs reconstructed from sensory areas reflected a single relationship dimension, while posterior-medial and prefrontal reconstructions reflected additional relationship complexity, such as romantic and emotionally rich ties between characters (Figure 2c). These results suggest that the brain forms layered relational representations during online social perception that vary in complexity across the cortex.

To assess whether these neural representations are specific to online social perception or reflect learned stable relationships, we analyzed neural responses during the post-movie rating task. Linear regression predicting neural RSMs from relationship graphs revealed that the similar brain regions represent the same structured social graphs even after movie-viewing. Dimensionality analysis demonstrated that while the mPFC maintained high-dimensional graph representations across tasks, temporal regions showed task-dependent compression, indicating distinct cortical roles in organizing multi-layered social knowledge.

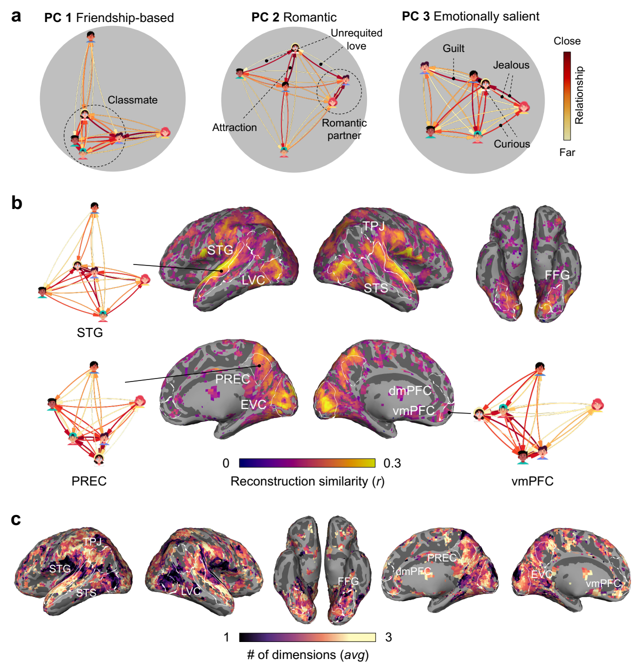

Figure 2: a) Data-driven relationship dimensions derived from the movie. b) Voxel-wise reconstruction of social relationship graphs (qs < .05). c) Dimensionality of relationship representations during movie-viewing.

Discussion

Here, we demonstrate that the human brain encodes dynamic social interactions into structured, multi-dimensional relationship graphs. Higher-order regions, such as the mPFC, support rich, multi-layered social knowledge, while sensory areas encode more compressed, low-dimensional representations. These representations persist across encoding and retrieval, yet their dimensionality and localization vary, suggesting differential encoding of social relationship aspects not fully captured by previous approaches using co-occurrence. Our findings suggest that understanding the richness of real-world social relationships crucially depends on tracking how interactions unfold. Interactions can vary along multiple dimensions and models that reflect these multidimensional social graphs provide a possible explanation for how the human brain navigates and represents complex social structures.

Acknowledgements

This work was supported by RS-2024-00348130, RS-2025-02304581, and the Fourth Stage of Brain Korea 21 Project in Department of Intelligent Precision Healthcare, Sungkyunkwan University (S-2023-0794-000).

References

Hamlin, J. K. (2013). Failed attempts to help and harm: intention versus outcome in preverbal infants' social evaluations. *Cognition, 128*(3), 451–474.

Jolly, E., Sadhukha, S., Iqbal, M., Molani, Z., Walsh, T. M., Manning, J. R., & Chang, L. J. (2023). People are represented and remembered through their relationships with others. *PsyArxiv.*

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience, 2*, 4.

Malik, M., & Isik, L. (2023). Relational visual representations underlie human social interaction recognition. *Nature Communications, 14*(1), 7317.

Lee Masson, H., & Isik, L. (2021). Functional selectivity for social interaction perception in the human superior temporal sulcus during natural viewing. *NeuroImage, 245*, 118741.

Lee Masson, H., Chen, J., & Isik, L. (2024). A shared neural code for perceiving and remembering social interactions in the human superior temporal sulcus. *Neuropsychologia, 196*, 108823.

Meijerink-Bosman, M., Back, M., Geukes, K., Leenders, R., & Mulder, J. (2023). Discovering trends of social interaction behavior over time: An introduction to relational event modeling : Trends of social interaction. *Behavior Research Methods, 55*(3), 997–1023.

Naselaris, T., Kay, K. N., Nishimoto, S., & Gallant, J. L. (2011). Encoding and decoding in fMRI. *NeuroImage, 56*(2), 400–410.

Park, S.-M., Na, C.-W., Choi, M.-S., Lee, D.-H., & On, B.-W. (2018). KNU Korean Sentiment Lexicon: Bi-LSTM-based Method for Building a Korean Sentiment Lexicon. *Journal of Intelligence and Information Systems, 24*(4), 219–240.

Parkinson, C., Kleinbaum, A. M., & Wheatley, T. (2017). Spontaneous neural encoding of social network position. *Nature Human Behaviour, 1*(5), 1–7.

Peer, M., Hayman, M., Tamir, B., & Arzy, S. (2021). Brain Coding of Social Network Structure. *Journal of Neuroscience, 41*(22), 4897–4909.