Neural development of social perception: Evidence from voxel-wise encoding model study in young children and adults

Elizabeth Jiwon Im (imelizabeth@jhu.edu)
Department of Cognitive Science, Johns Hopkins University

Angira Shirahatti (ashirahattia@gmail.com)
Department of Cognitive Science, Johns Hopkins University

Leyla Isik (lisik@jhu.edu)
Department of Cognitive Science, Johns Hopkins University

Abstract

From early in life, children exhibit sophisticated social perception and reasoning skills, but our understanding of the brain development behind these skills remains limited. To explore this, we analyzed fMRI data from children and adults as they viewed a socially engaging, animated film. The film was annotated with visual (motion energy) and social-affective features (faces, social interactions, theory of mind (ToM) events, valence, arousal). By training a voxel-wise encoding model on these features, we find that both visual and social features could predict brain activity in children as young as three, especially in areas along the lateral occipital cortex, including the motion-sensitive MT and superior temporal sulcus (STS). Additionally, we conducted individual social feature encoding model and find that faces and social interaction information are present in children as young as three years old. This study uniquely connects child brain activity to social visual information in a naturalistic setting, indicating that children's rich perceptions of social interactions are underpinned by neural mechanisms that develop early in life.

Keywords: fMRI, social interaction, development, encoding

Introduction

Very young children possess sophisticated social perceptual skills, such as the ability to recognize social behaviors and understand complex ideas about others and their relationships through observing these behaviors (Hamlin, Wynn, & Bloom, 2007; Powell & Spelke, 2018; Goupil, Papeo, & Hochmann, 2022; Thomas, Saxe, & Spelke, 2022).

In adult neuroimaging studies, specific brain regions within the superior temporal sulcus (STS) has been found to process social interactions, distinct from other visual or social information (Isik, Koldewyn, Beeler, & Kanwisher, 2017; Walbrin, Downing, & Koldewyn, 2018; Lee Masson & Isik, 2021). Recent studies also found social interaction selectivity in children's STS as young as six, but with reduced selectivity compared to adults (Walbrin, Mihai, Landsiedel, & Koldewyn, 2020), and work with functional near infrared spectroscopy (fNIRS) has shown stronger responses in infants' medial prefrontal cortex and STS to interacting dyad than non-interacting dyads(Farris et al., 2022).

Despite these advances, the developmental trajectory of social interaction perception and the relationship to other social and visual neural processing are still largely unknown due to challenges of scanning young children. However, in recent years, naturalistic stimuli have been gaining popularity in neuroimaging studies due to their more dynamic and engaging nature (Redcay & Moraczewski, 2020).These methods are particularly promising for child neuroimaging studies (Cantlon & Li, 2013). Using these approaches, Richardson et al. (2018) found that children's brain regions linked to theory of mind and pain are distinct by age three and become more specialized with age. However, identifying the specific movie content driving neural activity remains a challenge.

To better understand the developmental trajectory of brain responses to a specific visual and social features during natural viewing, we combined dense movie labeling and voxel-wise encoding, an approach that has been successful in previous adult studies(Lee Masson & Isik, 2021), to analyze children's brain responses to visual and social cues in fMRI data from Richardson et al. (2018). We found that encoding models based on visual and social-affective features could explain neural responses in children as young as three. We also find that individual social features are significantly predictive of children's and adult's brain responses, as well as evidence for social interaction information in the STS in children as young as three.

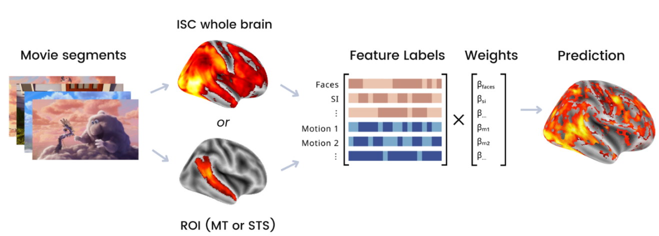

Figure 1: Schematic description. The voxel-wise encoding model predicts voxel-wise responses to held-out movie segments in both whole brain and ROI analyses.

Methods

fMRI data

We examined a publicly available fMRI dataset collected by Richardson et al. (2018), involving 155 participants—122 children ranging from 3 to 12 years and 33 adults (18-39 years). Participants viewed a 5.6-minute silent animation, "Partly Cloudy" (Pixar, 2009).

MRI images were aligned to the Montreal Neurological Institute template and smoothed using a 5mm Gaussian filter. Noise was managed using artifact timepoint and PCA-based regressors. Details on data collection and preprocessing are detailed in the original study.

For the whole-brain analysis, we employed inter-subject correlation (ISC) and identified voxels with stable neural activity, filtering out unique or noisy. Our analysis focused on two key regions of interest (ROIs): the middle temporal region (MT), linked to visual motion processing, and the superior temporal sulcus (STS), important for social processing. We utilized anatomical masks for MT(Wang, Mruczek, Arcaro, & Kastner, 2014) and STS(Deen, Koldewyn, Kanwisher, & Saxe, 2015), which also covered the temporoparietal junction (TPJ), a region associated with theory of mind and social perception.

Movie annotations

For each 2-second clip of the movie, we identified and labeled various social features such as the presence or absence of faces, social interactions, and characters' mentalization events, alongside emotional tone and arousal levels. These aspects cover both social perception and cognition and are known to influence neural activity in brain areas involved in social processing (Lee Masson & Isik, 2021; Richardson, Lisandrelli, Riobueno-Naylor, & Saxe, 2018).Additionally, to consider visual influences on neural activity, we analyzed motion energy features for each segment(A. Nunez-Elizalde, Deniz, Tour, Castello, & Gallant, 2021).

Encoding Model

To link brain activity with specific movie features, we trained a voxel-wise encoding model using banded ridge regression(A. O. Nunez-Elizalde, Huth, & Gallant, 2019). This approach effectively handles the complexity of rich visual features and simpler social-affective annotations. To account for temporal autocorrelations in the movie and fMRI responses, we divided the movie into 19 continuous segments (8 TRs each) and performed leave one out cross validation on these segments.

Results

In our whole-brain analysis, we found that the model significantly predicted brain activity in all age groups, even in the youngest ages (3-4 years old) (Figure 2, left). This was also true when only the visual motion energy (Figure 2, middle) or social-affective features (Figure 2, right) were used for prediction. Next we conducted ROI analysis in MT and STS (Figure 3).

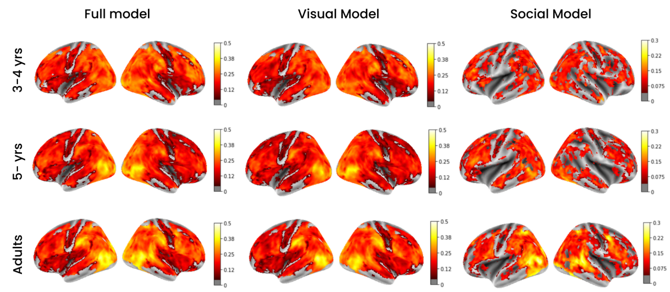

Figure 2: Lateral view of whole brain encoding model results for all voxels with significant inter-subject correlation in ages 3-4, 5, and adults. Results for child groups between 7 - 12 years old look similar to adults

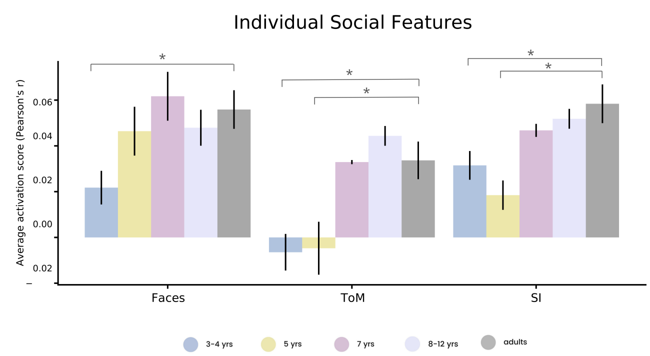

Figure 3: Average voxelwise predictivity of individual social features in the STS in different age groups.

The visual and social-affective models effectively predicted brain activity across all age groups for both ROIs. Visual responses were adult-like across all ages, indicating early maturation. However, the social-affective model showed significant developmental differences, particularly with younger children displaying distinct patterns compared to adults, suggesting ongoing development of social-affective neural representations throughout childhood.

In whole-brain analysis, no individual feature alone was able to predict the brain activity. To better understand the representations of individual social features throughout development, we looked at individual social features (faces, theory of mind, and social interaction) predictivity in STS. We find that face features predicted significant responses in the STS across all age groups, with notable differences between adults and the youngest children (3-4 years old).We also find significant social interaction information in the STS in all age groups, including children as young as three. For both ToM and social interaction predictivity, we see significant difference between adults and both 3-4 years old and 5-years old groups.

Discussion

In this study, we examined how visual and social-affective feature representations develop in children aged 3-12 years using a voxel-wise encoding model. This model, trained with a limited set of visual and social features, effectively predicted whole brain activity even in the youngest participants. Further, we explored how individual social features influenced brain activity in the superior temporal sulcus (STS). We found evidence of social interaction selectivity in children as young as three, with responses becoming adult-like at age seven. Together, this work provides some of the earliest evidence of social interaction processing in the STS, as well as a new framework for linking individual social features to developmental neuroimaging data during natural viewing.

Acknowledgments

This work was supported with funds from the Clare Boothe Luce Program for Women in STEM and NIMH R01MH132826 awarded to LI. We would like to thank Hillary Richardson and colleagues for data used in this study.

References

Cantlon, J. F., & Li, R. (2013). Neural activity during natural viewing of Sesame Street statistically predicts test scores in early childhood. *PLoS biology*, *11*(1), e1001462. doi: 10.1371/journal.pbio.1001462

Deen, B., Koldewyn, K., Kanwisher, N., & Saxe, R. (2015, November). Functional Organization of Social Perception and Cognition in the Superior Temporal Sulcus. *Cerebral Cortex (New York, N.Y.: 1991)*, *25*(11), 4596–4609. doi: 10.1093/cercor/bhv111

Farris, K., Kelsey, C., Krol, K., Thiele, M., Hepach, R., Haun, D., & Grossmann, T. (2022). Processing third-party social interactions in the human infant brain. infant behavior and development. *Infant Behavior and Development*, *68*. doi: 10.1016/j.infbeh.2022.101727

Goupil, N., Papeo, L., & Hochmann, J.-R. (2022). Visual perception grounding of social cognition in preverbal infants. *Infancy*, *27*(2), 210–231. (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/infa.12453) doi: 10.1111/infa.12453

Hamlin, J. K., Wynn, K., & Bloom, P. (2007, November). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557–559. (Number: 7169 Publisher: Nature Publishing Group) doi: 10.1038/nature06288

Isik, L., Koldewyn, K., Beeler, D., & Kanwisher, N. (2017, October). Perceiving social interactions in the posterior superior temporal sulcus. *Proceedings of the National Academy of Sciences*, *114*(43), E9145–E9152. (Publisher: Proceedings of the National Academy of Sciences) doi: 10.1073/pnas.1714471114

Lee Masson, H., & Isik, L. (2021, December). Functional selectivity for social interaction perception in the human superior temporal sulcus during natural viewing. *NeuroImage*, *245*, 118741. doi: 10.1016/j.neuroimage.2021.118741

Nunez-Elizalde, A., Deniz, F., Tour, T. D. l., Castello, M. V. d. O., & Gallant, J. L. (2021, January). *pymoten: motion energy features from video using a pyramid of spatio-temporal Gabor filters*. Zenodo. doi: 10.5281/zenodo.4437446

Nunez-Elizalde, A. O., Huth, A. G., & Gallant, J. L. (2019, August). Voxelwise encoding models with non-spherical multivariate normal priors.

Powell, L. J., & Spelke, E. S. (2018, December). Third-Party Preferences for Imitators in Preverbal Infants. *Open Mind*, *2*(2), 61–71. doi: 10.1162/opmi_a_00018

Redcay, E., & Moraczewski, D. (2020, August). Social cognition in context: A naturalistic imaging approach. *NeuroImage*, *216*, 116392. doi: 10.1016/j.neuroimage.2019.116392

Richardson, H., Lisandrelli, G., Riobueno-Naylor, A., & Saxe, R. (2018). Development of the social brain from age three to twelve years. *Nature Communications*, *9*(1), 1027. doi: 10.1038/s41467-018-03399-2

Thomas, A. J., Saxe, R., & Spelke, E. S. (2022, August). Infants infer potential social partners by observing the interactions of their parent with unknown others. *Proceedings of the National Academy of Sciences*, *119*(32), e2121390119. doi: 10.1073/pnas.2121390119

Walbrin, J., Downing, P., & Koldewyn, K. (2018, April). Neural responses to visually observed social interactions. *Neuropsychologia*, *112*, 31–39. doi: 10.1016/j.neuropsychologia.2018.02.023

Walbrin, J., Mihai, I., Landsiedel, J., & Koldewyn, K. (2020, April). Developmental changes in visual responses to social interactions. *Developmental Cognitive Neuroscience*, *42*, 100774. doi: 10.1016/j.dcn.2020.100774

Wang, L., Mruczek, R. E., Arcaro, M. J., & Kastner, S. (2014, December). Probabilistic Maps of Visual Topography in Human Cortex. *Cerebral Cortex*, *25*(10), 3911–3931. doi: 10.1093/cercor/bhu277