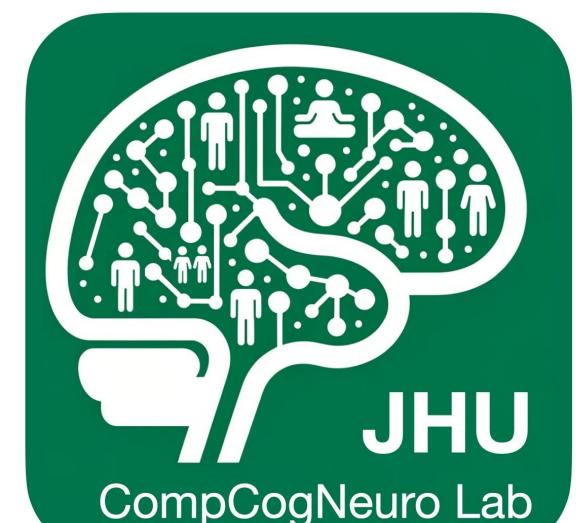




Relational Information Predicts Human Behavior and Neural Responses to Complex Social Scenes

Wenshuo Qin, Manasi Malik, Leyla Isik

Department of Cognitive Science at Johns Hopkins University

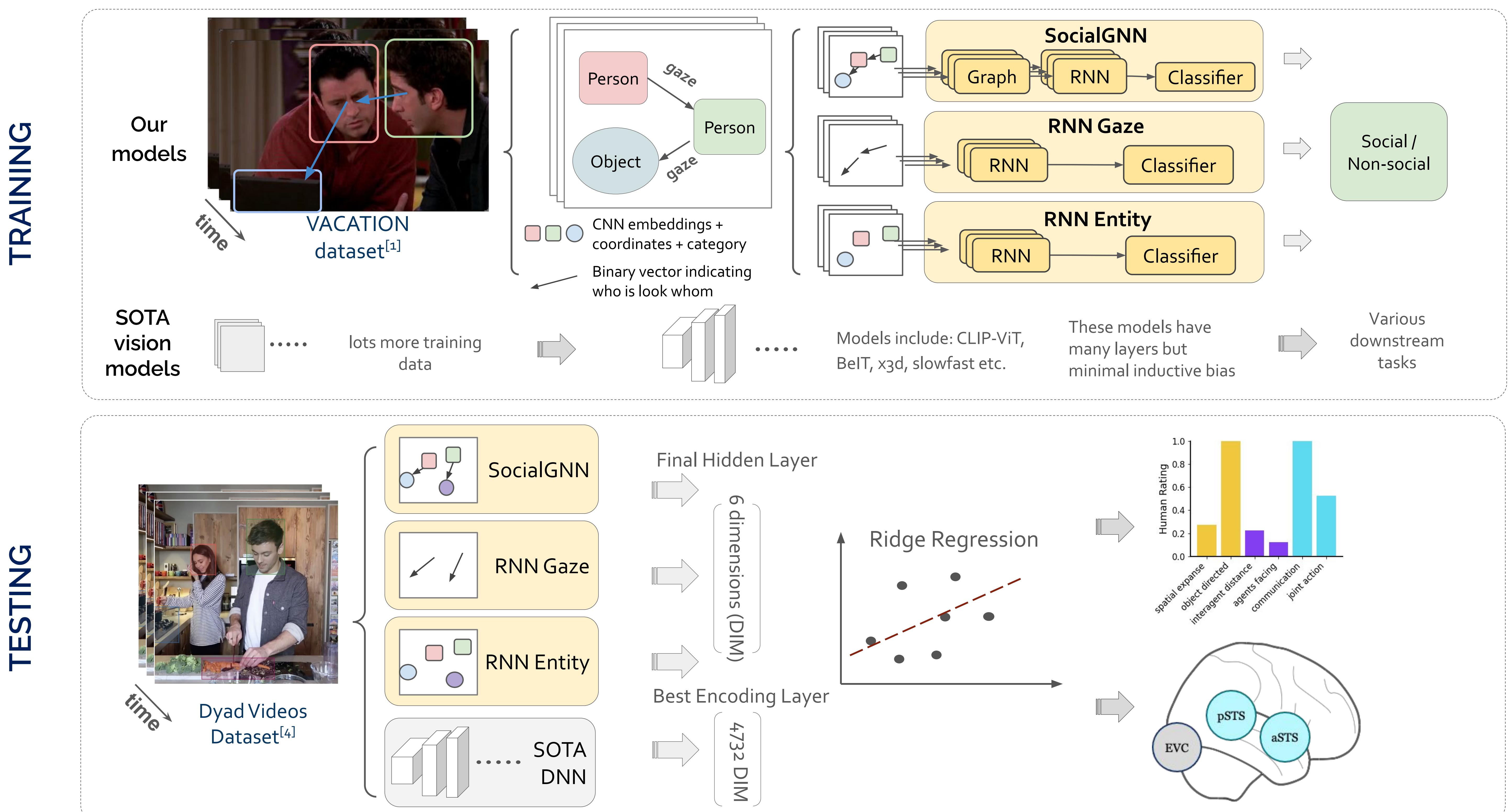


Introduction

Background: Relational information is critical for understanding social interactions and is represented in the human brain—particularly in the superior temporal sulcus (STS). However, most computational models overlook its importance.

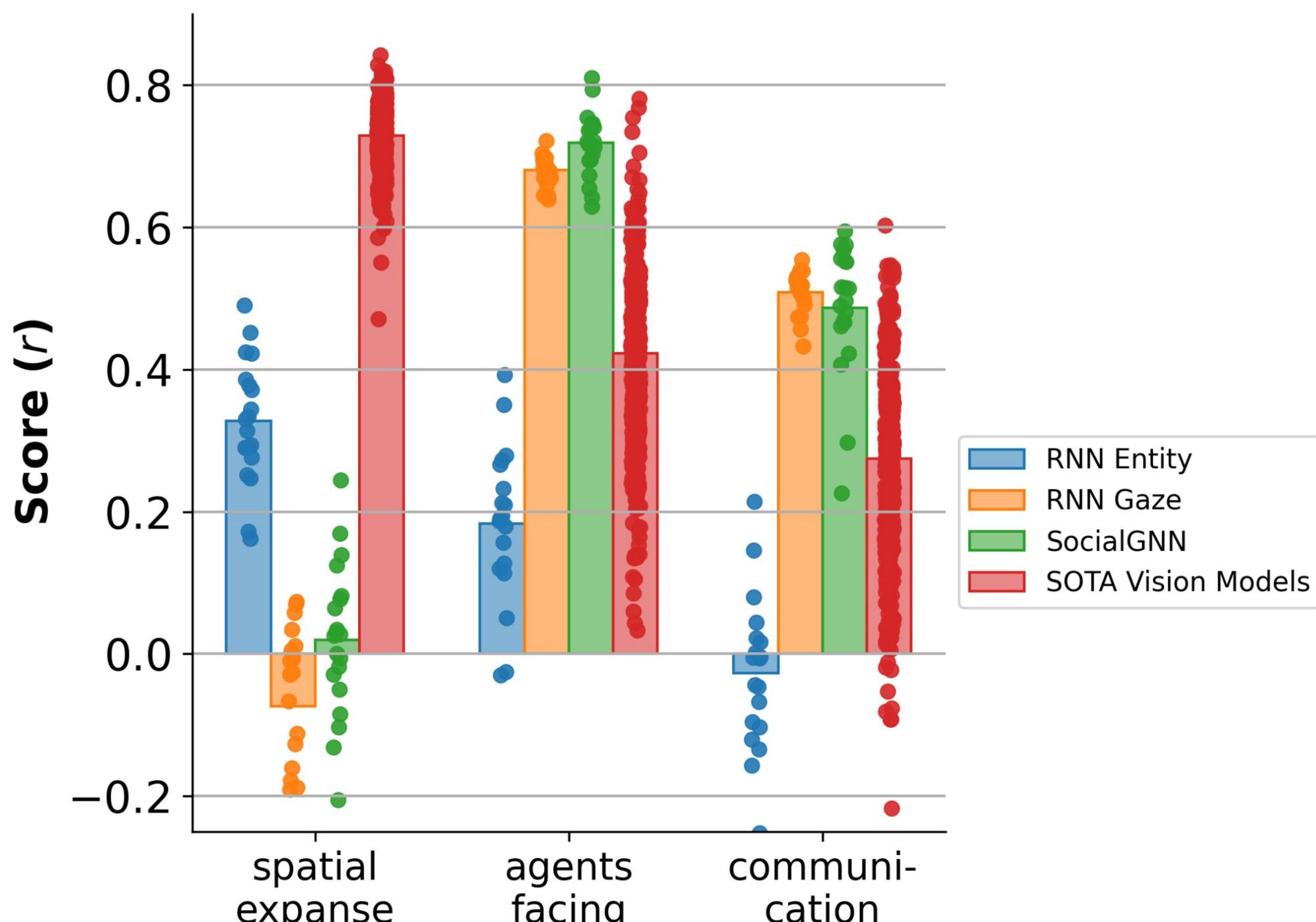
Goal: Compare **relational models (SocialGNN^[3] & RNN Gaze)** against a **non-relational model (RNN Entity)** and SOTA vision models^[2] in predicting human behavioral and neural responses to social scenes.

Methods



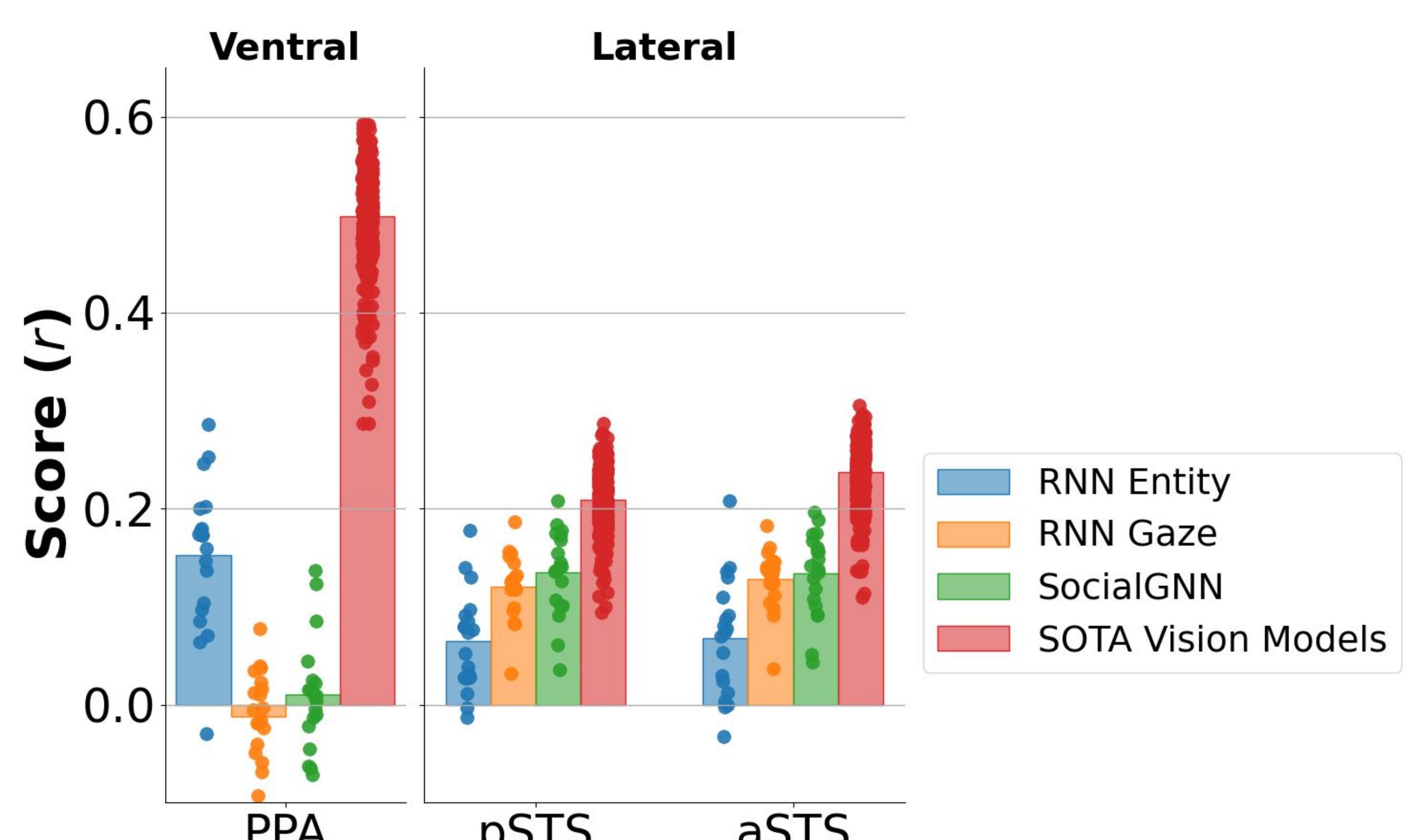
Results

Behavioral Encoding



- Relational models better predict facing and communication than the best SOTA vision model overall (DeiT)^{[2][6]}.
- Non-relational RNN Entity better predicts scene size.

Neural Encoding



- Relational models better predict the STS regions.
- RNN Entity better predicts the ventral region.

Discussion

Simple neural networks trained on **relational cues** (e.g., gaze)

- Better predict **human judgments of social interaction**.
- Better predict neural responses in the **superior temporal sulcus (STS)**.

These findings highlight the key role of relational cues. Future work should disentangle their contributions from non-relational cues in social perception.



- [1] Fan et al. (2019) ICCV
[2] Garcia et al. (2024) OSF
[3] Malik & Isik (2023) Nat Commun
[4] McMahon et al. (2023) Curr Biol
[5] Toviron et al. (2022) arXiv

