When Seating Matters: Modeling Graded Social Attitudes as Bayesian Inference

Zihan Wang (zihan.wang.zw446@yale.edu)
Department of Psychology, Yale University
New Haven, CT 06511 USA

Julian Jara-Ettinger (julian.jara-ettinger@yale.edu)
Department of Psychology, Yale University
New Haven, CT 06511 USA

Abstract

Humans can quickly infer social relationships from minimal cues, such as where people choose to sit in a meeting room. We investigated how people make graded, context-sensitive judgments about social attitudes beyond simple proximity-based heuristics. Using controlled seating scenarios, we compared participants' judgments to the predictions of Bayesian models: the interaction-probability model, which captures how one person's seat choice affects the probability that another person will initiate the conversation, and the interaction-cost model, which accounts for the effort required based on how far apart they sit from each other. Results showed that participants' inferences aligned best with the interaction-cost model, indicating sensitivity to effort and moving trajectory, rather than relying solely on proximity. Our findings suggest that higher-order cognition refines perceptual cues, enabling nuanced, graded social reasoning essential for complex social interactions.

Keywords: social cognition; computational modeling; social relationship inference; Bayesian inference

Introduction

Imagine arriving early to a meeting and taking the closest seat next to the door. Soon after, your colleague walks in, looks at where you are sitting, and goes all around the room to sit in the opposite corner. From this simple action, you might start to wonder how your presence influenced their choice – Perhaps they are busy and want to avoid conversation, and you would be less likely to start one. If instead, your colleague sat close to you, it might suggest that they are open to interaction, and you might be more inclined to say hi; however, this choice could also reflect that they simply took one of the most convenient seats. On the other hand, if you were sitting on the opposite side of the door and your colleague walked all the way across the room to sit next to you, you might feel confident to interpret this as a positive desire to interact. These small, everyday actions carry rich information about interpersonal attitudes, prompting intuitive social inferences.

These inferences are central to social behavior, helping us identify who likes us, who wants to interact, and with whom to engage to build relations and bonds. In line with their importance, research shows that people detect social inferences remarkably fast, often within 300 milliseconds (Isik, Mynick, Pantazis, & Kanwisher, 2020; Papeo, Stein, & Soto-Faraco, 2017; Papeo & Abassi, 2019). We can even extract social meaning from minimal or abstract stimuli, such as point-light displays (Flavell, Over, Vestner, Cook, & Tipper, 2022).

Not only are these abilities rapid, they also emerge early in development. Before the first year of life, infants have concepts of social relationships (Thomas, Woo, Nettle, Spelke, & Saxe, 2022; Kudrnova, Spelke, & Thomas, 2024) and are drawn to social interactions (Thiele, Hepach, Michel, & Haun, 2021).

But how do humans achieve this? What computational and inferential processes give rise to this capacity? One important recent account argues that social inferences emerge as part of the visual system (Pitcher & Ungerleider, 2021; McMahon & Isik, 2023). Under this view, the visual system goes beyond social primitives like proximity, orientation, and motion contingency, producing categorical labels of social relationships like 'friendly', 'neutral', or 'adversarial'(Malik & Isik, 2023).

However, this does not necessarily mean high-level cognition does not support this inference at all. While perception might provide a first-pass approximation for the inference procedure, higher-level cognition can contribute to these rich social inferences in two important ways. First, it might allow us to move beyond categorical judgments to support graded, continuous judgments about relationships, distinguishing between strong and weak preferences rather than making binary categorizations. For instance, we can discern whether someone really wants to interact with someone or is only mildly inclined to do so. Second, higher-level cognition can override perceptual cues when context suggests alternative explanations. For example, someone could sit right next to someone else because that is the closest seat to the door. In this way, their action may not reflect a social preference. Both accounts highlight the role of higher-level cognitive reasoning in refining and contextualizing social judgments and inferences.

Recent work validates seating choice as a rich paradigm for inferring social relationships. For example, Jern, Derrow-Pinion, and Piergiovanni (2021) showed that people make categorical judgments – whether someone likes or dislikes another – based on their relative seating in a row, guided by perceptual features and preference for simpler explanations. While supporting perception's role in social inference, this account remains limited to coarse categorical distinctions.

Here, we test whether people can make graded social inferences that go beyond categorical judgments and simple perceptual cues like proximity. We examined how participants infer social attitudes in ambiguous situations where low-level cues alone were insufficient. We then compared their responses to predictions from computational models that formalize high-level reasoning strategies people may be using, including a Bayesian framework that incorporates social intuitions about interaction likelihood and conversational effort.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

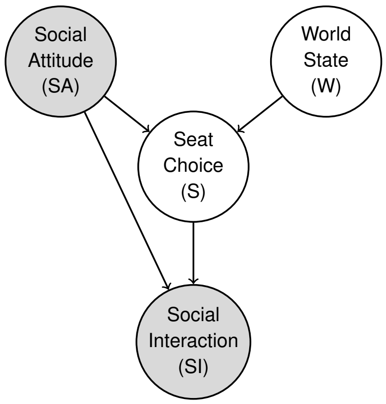

Figure 1: Causal structure behind modeling: Purple's seat choice (*S*) is affected by their feeling towards Yellow (Social Attitude; *SA*) and the world states (where the available seats are and where Yellow sits at; *W*). Based on the two intuitions we are testing, Purple's seat choice (*S*) and social attitude (*SA*) also affect the likelihood of social interaction (*SI*). Inaccessible information to observers is marked in gray.

Computational Framework

To understand how people infer social attitudes based on observable seating actions, we built a simplified version of our opening example: an agent (Purple) decides where to sit based on where another agent (Yellow) is sitting at. Figure 2 shows four example trials. Each image shows a conference room with a table and multiple seats (brown circles). Yellow is already seated at the table when Purple arrives, and Purple has to decide where to sit. For example, Purple could take the closest seat by the entrance (Fig. 2a&b), or could walk all the way to the other side of the room to sit next to Yellow (Fig. 2c) or even to avoid them (Fig. 2d). For each trial, the task is to infer how Purple feels about Yellow (Purple's Social Attitude toward Yellow).

Figure 1 shows the causal model behind how we conceptualize this task. When Purple comes in, the world state is fully observable, including the set of available seats and Yellow's seat position. We assume that Purple has a social attitude towards Yellow, represented as a value on a continuous scale from -7 (strongly dislikes) to $+7$ (strongly likes). This social attitude (*SA*), combined with the world state (*W*), determines where Purple chooses to sit. Together, Purple's seat choice and their social attitude towards Yellow influence the likelihood of a social interaction. Social attitudes can directly lead to a social interaction if Purple initiates a conversation – if Purple likes Yellow, they may initiate a conversation directly. Or indirectly, Purple's seat choice can also affect the probability of interaction: Purple sitting further away may decrease the likelihood of Yellow initiating a conversation, while Purple sitting closer may prompt Yellow to interact. When a social interaction occurs, we assume that Purple's potential reward gained from this social interaction is equivalent to their social attitude towards Yellow. If Purple has a positive *SA*, they get a positive reward from interacting with Yellow, a negative reward if they have a negative *SA* towards Yellow, and zero if they are neutral.

From the observer's perspective, only Purple's final seat choice (*S*) and the initial world state (*W*) are known (nodes marked in white in Fig. 1). Neither participants nor models have access to whether a social interaction occurred or to Purple's social attitude (*SA*) toward Yellow (nodes marked in gray in Fig. 1). The goal is to infer Purple's social attitude. Following much prior work suggesting that social inferences are Bayesian (e.g., Baker, Jara-Ettinger, Saxe, and Tenenbaum (2017)), we formalize this process using a Bayesian approach. Formally, the probability that Purple has social attitude (*SA*), given their seat choice (*S*) and the initial world state (*W*), is given by:

$$p(SA|S,W) \propto p(S|SA,W)p(SA) \quad (1)$$

Here, $p(SA)$ is the prior distribution of different social attitudes, which in this experiment we assume is uniform. $p(S|SA,W)$ is the likelihood that Purple would choose seat S given attitude SA and the initial world state W. To capture this likelihood function, we use a Markov Decision Processes (MDP), modeling how Purple builds a plan for navigating the space to choose a seat.

In our MDP, Purple could move through `MovingStates` that consist of all possible configurations of their position, orientation, and whether they are seated. Purple's actions include movement (`Move`, `TurnLEFT`, `TurnRIGHT`) and interaction-related actions (`SIT`, `Talk`), each with associated costs. Transition functions determine how actions lead to new states, with key decisions occurring in the `DecidingState`, where Purple chooses whether or not to initiate a conversation. Rewards depend on Purple's social attitude (*SA*) toward Yellow (and the distance between them, as shown in Table 1), if they entered `InteractingState` (see Kaelbling, Littman, and Cassandra (1998) for a technical introduction to MDPs).

In this formulation, we capture the intuition that Purple will sit close to Yellow in two ways: either to initiate an interaction with Yellow, or when they want Yellow to initiate it.

Cost Model If Purple chooses the action `Talk` in the `DecidingState`, they transition to an `InteractingState`. This model captures the intuition that as the distance between two agents increases, the effort and cost of interaction also increase, making it more difficult to engage in a conversation. This negative cost to Purple's reward is modeled as $R(SI) = SA - d$, where greater distances d reduce the benefit of interacting. This is called the Cost model.

Probability Model Purple may also choose to remain silent in the `DecidingState`, in which case sitting closer increases the likelihood of Yellow initiating a conversation. Formally, the probability of transitioning to an `InteractingState` increases as the distance d between the agents decreases, modeled as $P(SI) = 1 - \frac{d}{d_{max}}$, where d_{max} is the maximum possible seating distance. This is called the Probability model.

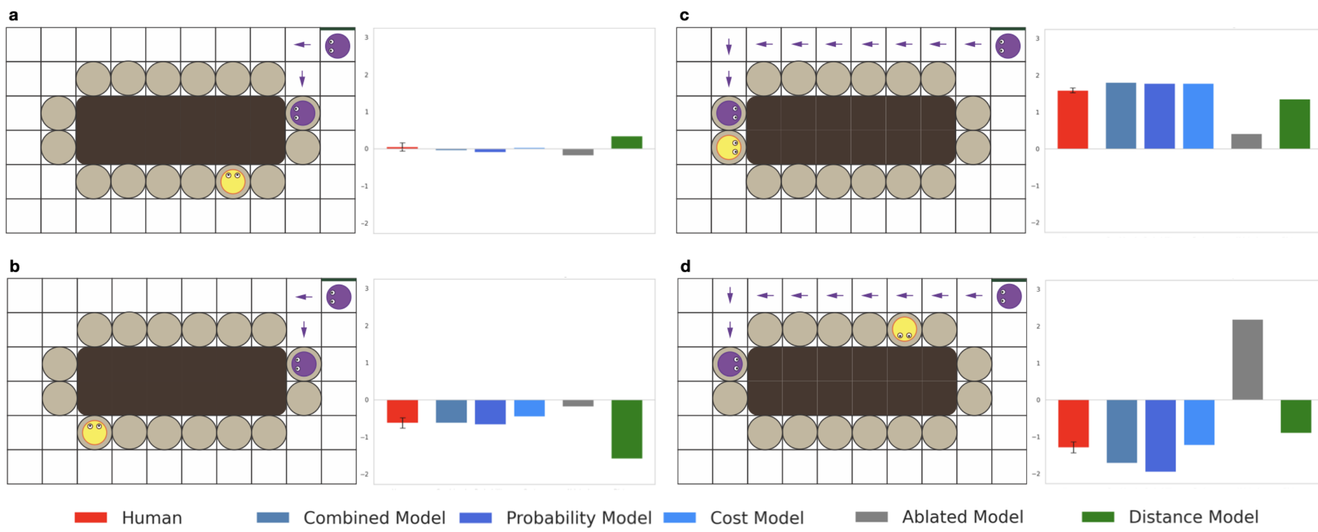

Figure 2: Four example trials, showing Purple's different seating choices relative to Yellow's seat. Each figure shows participant judgments and the model predictions on the right side.

Combined Model The combined model integrates both intuitions, accounting for both the increased likelihood of interaction and the reduced effort required when agents are interacting at a closer distance.

These are our three main models. To assess whether these intuitions are critical for explaining human judgments, we also tested two alternative models:

Ablated Model The ablation model removes both assumptions about social interactions, aiming to demonstrate that the effectiveness of any of our main models depends on these key assumptions. Without them, the ablated model is expected to fail. Table 1 summarizes the interaction probabilities and the rewards assumptions for each model.

Distance Heuristic Model The distance model tests for a low-level perceptual account. It uses a simple heuristic based solely on spatial proximity, assuming that Purple's closer seating always implies greater liking, regardless of contexts.

Model	**Interaction Probability**	**Interaction Reward**
Combined Model	$P(SI) = 1 - \frac{d}{d_{\max}}$	$R(SI) = SA - d$
Probability Model	$P(SI) = 1 - \frac{d}{d_{\max}}$	$R(SI) = SA$
Cost Model	$P(SI) = 0.5$	$R(SI) = SA - d$
Ablation Model	$P(SI) = 0.5$	$R(SI) = SA$

Table 1: Interaction Probabilities and Rewards in each model.

Behavioral Experiment

Method

All data, stimuli, and materials are available on OSF.

Participants 50 U.S. participants (mean age = 36 years, range = 20-77 years; 36 female, 14 male) were recruited from Prolific. One participant was excluded from the analysis due to a self-reported attention score of less-than 8 out of 10 and two outliers in their responses (defined as 2.5 standard deviations away from the mean in any of the 30 trials) as preregistered. This resulted in a final sample of 49 participants.

Stimuli The stimuli consisted of 30 static images of an illustration of a meeting room with a desk, a set of chairs around the table, an entrance, and two agents, a yellow one and a purple one (see Fig. 2 for examples). Yellow was always seated in one of the chairs. Purple appeared seated in one of the chairs, along with a trajectory that indicated how they reached that seat from the entrance. Each image depicted a scenario where Purple entered a meeting room and decided where to sit while Yellow was already seated.

To create a rich space of trials, we used a combinatorial design. We started by selecting three initial regions where Yellow could be sitting: (1) near the entrance, (2) far from the entrance, and (3) in the middle of the room. We then selected five different seating choices for Purple to sit in: (1) closest to the entrance, (2) closest to Yellow, (3) farthest from Yellow, and (4-5) two possible intermediate distances from Yellow. For each of the three initial regions where Yellow could be seated, we selected two possible seats in this region (e.g., one along the vertical row and one along the horizontal row). This results in a total of 30 (3x5x2) possible theoretical configurations.

Additionally, the stimuli were designed to ensure that there are clusters of trials (two sets with 4 trials, and two with 6 trials) that controlled for the distance between the two agents so that we could better evaluate alternative heuristics.

Procedure Participants were first familiarized with the seating scenario setup through a cover story. Participants were told that they would see events where a protagonist, Purple, arrived in a meeting room. Another agent, Yellow, had already arrived and seated. Participants were then told that Purple's seating choice would affect the probability that Yellow would initiate a conversation before the meeting started. Thus, Purple would choose a seat based both on how far they had to walk and on how they felt toward Yellow. After reading the cover story, participants were asked six simple comprehension check questions to ensure they understood the logic of the task (the cover story and 30 stimuli trials are available on the OSF page). Participants had to correctly answer each comprehension question, and pass a reCAPTCHA test to proceed to the test trials. Participants who failed one of the comprehension checks were asked to review the cover story and were given unlimited attempts to answer the comprehension check question. They could only proceed to the next question if they correctly answered the previous one, or they could choose to exit this study.

In the test phase, participants were presented with all 30 trials in a randomized order. In each trial, participants viewed a static image of the meeting room with Yellow's seat and Purple's choice. Participants were asked to answer "How does Purple feel about Yellow?" using a slider, with one end representing *Strongly Dislikes* (coded as -7) and the other end representing *Strongly Likes* (coded as 7).

At the end of the 30 trials, participants were asked to explain what strategy they used in the task, and they were asked "How intuitive do you find the following statement": (1) "The farther away you sit from someone, the less likely they are to talk to you." (2) "Imagine you are in the same meeting room setting as shown in the study. If you are speaking with someone, the farther away they are sitting, the harder it is to maintain a conversation." Participants rated the strength of their intuitions using a Likert scale from 1 to 7.

Results We begin by comparing our results to our main models, demonstrating how they successfully capture participant judgments; we then evaluate participant judgments against the ablated model to test the necessity of key assumptions, and finally conclude by analyzing whether a simpler distance heuristic can explain the observed nuanced patterns.

Comparison to main models

Participant judgments were z-scored within participants and averaged. We first began by comparing participant responses with the three variations of our core models (each model prediction is also z-scored).

As shown in Figure 3, our three core models highly correlated with average participant judgments: combined model ($r = 0.86$; 95% CI: 0.75-0.93), probability model ($r = 0.90$; 95% CI: 0.82-0.95), and cost model ($r = 0.93$; 95% CI: 0.89-0.97). As the 95% bootstrapped confidence interval does not cross zero, we interpret this as evidence that our model meaningfully captures participant responses, thereby validating our first prediction.

Figure 2 highlights key example trials illustrating participant intuitions and how our models capture these graded inferences. In Fig. 2a, Purple takes the seat closest to the entrance, which is also near Yellow. Both participants and our models show uncertainty about whether this reflects Purple's positive feelings toward Yellow or merely convenience, suggesting ambiguous judgments. A similar case in Fig. 2b leans more toward a disliking inference but still shows uncertainty. In contrast, Fig. 2c and 2d depict scenarios where Purple crosses the entire room to sit. Here, both participants and our models strongly infer that Purple likes Yellow (Fig. 2c) or dislikes Yellow (Fig. 2d), with lower uncertainty.

Comparison to ablated model

Having established that our models accurately captured participant judgments, we next evaluated whether these judgments could be explained by simpler, alternative accounts. Specifically, we tested whether the assumptions that (1)*interaction difficulty* and (2)*interaction probability* depend on the distance between the two agents were necessary for our models to explain participant judgments. The ablated model, which removed both assumptions, showed a correlation of $r = 0.07$ (95% CI: -0.27-0.45) and performed reliably worse than our three main models (Figure 2), suggesting that the explanatory power of our main models did depend on assumptions about how seating affects the probability and cost of social interactions occurring.

Comparison to distance heuristics Having established that our main models capture participant judgments – and that they critically rely on an expectation of social interactions – we next explored whether people might instead be using a simple heuristic to make their judgments. Specifically, as research on social perception has argued that physical proximity is a low-level cue to social relations, we considered whether participants might be specifically relying on distance between the two agents alone.

Figure 3e show the fit between the distance heuristic and participant judgments. Overall, we found a high correlation between the two ($r = 0.91$; 95% CI: 0.84-0.96). Moreover, this correlation is not significantly different from the correlation we found between participants and our main models ($\Delta_{DistanceModel-CostModel} = -0.001$(95% CI: -0.23-0.22), $\Delta_{DistanceModel-ProbablityModel} = 0.001$(95% CI: -0.26-0.24), $\Delta_{DistanceModel-CombinedModel} = -0.001$(95% CI: -0.29-0.27)).

At first sight, this suggests that people might indeed be using a simple distance heuristic, because this strategy has equally good explanatory power, but provides a simpler explanation and is computationally cheaper.

However, this analysis offers only a coarse evaluation of overall model fit. While our main accounts agree that distance should matter in people's inferences, because it affects social interactions, they also posit a more fine-grained structure: that people are further sensitive not only to the final distance between agents, but also to the trajectory and effort involved in reaching that seat, relative to the entrance.

Because we expected this overall high numerical correlation, our stimuli was designed to ensure we had clusters of tri-

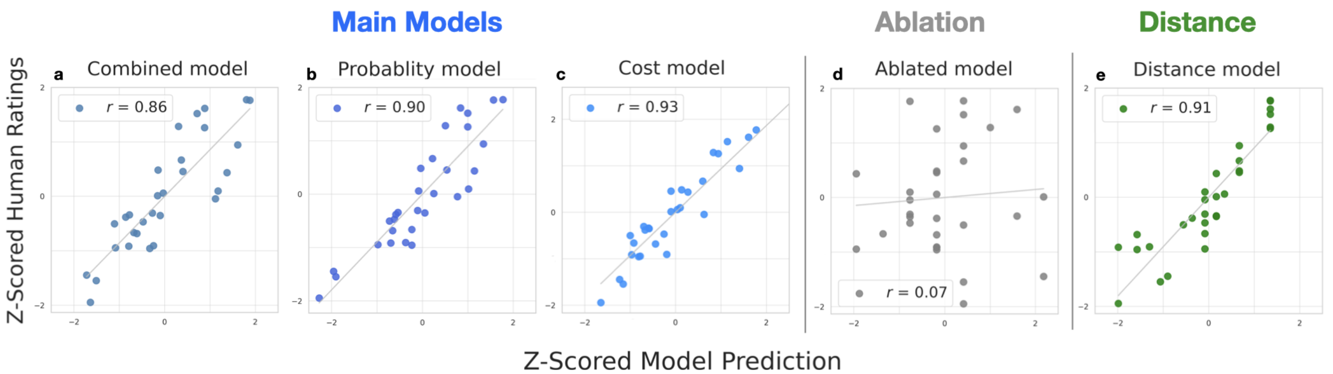

Figure 3: Comparing each model predictions and mean human judgments (N=49) across all 30 trials. Pearson correlation coefficient r: a) Combined model $r = 0.86$, b) Probability model $r = 0.90$, c) Cost model $r = 0.93$, d) Ablated model $r = 0.07$, and e) Distance heuristics $r = 0.91$.

als where the distance was matched, but the effort that Purple exerts is different. This can be visualized in Figure 3e as the four columns of data, approximately on values -0.07, 0.15, 0.61, 1.23. In each of these cases, the column of data represents a set of trials where the distance heuristics makes the same prediction because the distance is matched, but humans make different judgments in each trial (therefore showing a vertical pattern).

Figure 4 shows one of the four trial sets designed with this property. Here, the distance heuristic makes the same prediction across conditions because the distance is matched (visualized as the green horizontal line). However, participants and our main models make different judgments in each trial, adjusting their inferences based on the effort that Purple exerted. To analyze this formally, we calculated the correlation between participants and our model, for each of these four clusters. If the distance heuristic were correct, then any variation from the flat line should be participant noise. In contrast, if our main models are correct, they should predict the variability across these trials that the distance heuristic cannot.

As predicted, all four clusters showed positive correlations with our main models. Across the 12 correlations (4 clusters against 3 models), values ranged from 0.77 to 0.99, with an average correlation of 0.90. This suggests that while the distance heuristic captures the coarse patterns of inferences, it fails to have the fine-grained sensitivity that people show to the specific path and effort Purple takes to a seat (independent of the distance), which our main models successfully capture.

Distinguishing between our main models

Our results so far have established that the core models best explain human judgments, both in terms of overall correlation and in capturing fine-grained variation in human intuitions. We next sought to distinguish among these three models. To do so, we began by comparing pairwise correlation differences and found that the models were overall statistically indistinguishable: $\Delta_{CombinedModel-ProbablityModel} = -0$ (95% CI: -0.09-0.09), $\Delta_{CombinedModel-CostModel} = -0.001$ (95% CI: -0.15-0.16), $\Delta_{ProbablityModel-CostModel} = 0$ (95% CI: -0.18-0.18). This suggests that, at the group level, participant intuitions about seating choices can be equally well explained by either the intuition that distance affects the *probability* of a conversation happening or the *difficulty* of having one. Moreover, combining both sets of intuitions does not improve model fit.

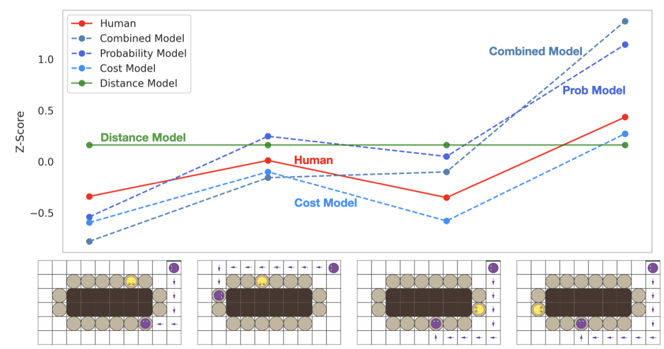

Figure 4: One of the four cluster trials where distance heuristics makes the same prediction but human participants and three main models make graded judgments in each scenario and highly align with each other.

To explore this further, we analyzed subject-level responses to determine whether one model fits most participants better, or whether different participants are best fit by different models. Overall, 73.5% of participants were best fit by the Cost model (n=36/49), followed by the Probability model (24.5% of participants; n=12/49), and the Combined model was the best fit for only one participant (2%) (Fig. 5).

Interestingly, however, participants' explicit reports of the two key intuitions – conversation difficulty (cost) and interaction probability (probability) – were both high and comparable. On average, participants rated the cost assumption

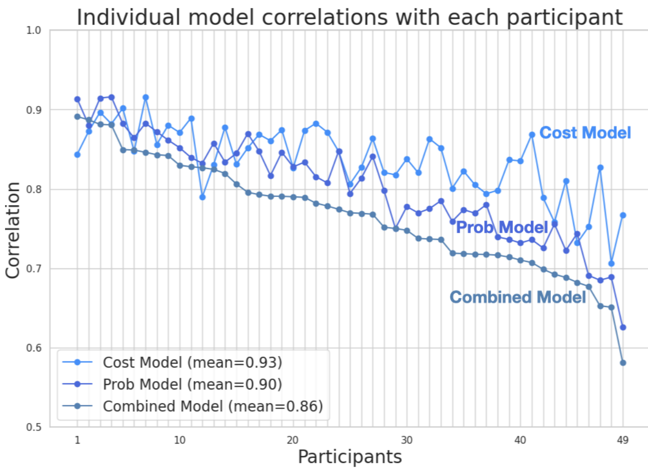

Figure 5: Subject-level model correlations. The x-axis shows participants; the y-axis shows correlations with main models.

as intuitive at 6.47 out of 7, and the probability assumption at 6.16 out of 7 on an intuition scale. However, since these ratings were provided after having completed the task, it is therefore possible that participants found both statements intuitive once they were explicitly asked about them, but during the task (before we had asked them explicitly), they might have been relying more heavily on the cost assumption.

Discussion

In this paper, we studied how people make graded inferences about how much one person like another based on minimal social cues, such as seating choices. To test this, we designed a set of closely related computational models, each implementing different assumptions about the underlying processes that drive these inferences. At the population level, our main models successfully explained participant judgments. While a simpler distance heuristic also accounted for some patterns, sensitivity analyses revealed that this heuristic failed to predict detailed patterns in participants' intuitions, which the main models captured more effectively (Fig. 2-4).

Our main models explored two possible causes for why people might choose to sit close to one another. One assumption is that sitting closer makes it more inviting for the other person to initiate a conversation. The other assumption is that sitting closer makes it easier to talk to each other. While these two possibilities could not be distinguished at the population level, subject-level analyses revealed that most participants aligned most closely with the Cost model, which emphasizes the ease or difficulty of conversation.

Our findings reveal that people can go beyond qualitative judgments about ”liking” or ”disliking” to make graded inferences about the degree of social attitudes. This is consistent with other research showing that humans make fine-grained inferences about others' mental states (Jern, Lucas, & Kemp, 2017; Baker et al., 2017).

These findings contrast with recent proposals suggesting that inferences about social affiliations might be implemented directly from a bottom-up process in the visual system (Malik & Isik, 2023; McMahon & Isik, 2023), and hence not driven by a top-down process. While these accounts posit that social relationships are perceived automatically, our results indicate that graded patterns of inferences likely require top-down cognitive processes. However, we do not see these views as mutually exclusive but rather complementary. Here, we offer a few possibilities for how this might be the case.

One possibility is that visual system / perception provides a rapid, initial estimate of the social relation, which is then further refined by higher-order processes, such as Theory of Mind. For instance, people might quickly sense whether there is a social relation and, when they detect one, refine the inferences using a generative model. This would allow us to avoid false positives. For example, seeing two people facing each other in close proximity might typically suggest a social relation – but not if they have accidentally collided as they walk around the corner. If these inferences could not be further refined and disambiguated by Theory of Mind, we would constantly make incorrect inferences in these edge cases. Moreover, computational models of this capacity have found that perception produces categorical judgments, but not the fine-grained, graded inferences observed in our study. Thus, it is plausible that perception serves as a form of ”amortized inference”, providing a quick initial categorical guess about social affiliation that is then refined by generative models to account for more complex scenarios.

Our study has a few limitations. First, our study assumed that Purple's seating decision was based solely on their intrinsic social attitude toward Yellow. In real-world contexts, however, people's decisions are also shaped by social norms and concerns about how one's actions might be perceived by others. For instance, if Yellow is sitting right next to the entrance, Purple might be less willing to go sit on the other end because this would make it obvious that they are actively avoiding Yellow. Conversely, even if Purple really likes Yellow, they might refrain from going all the way cross the room and sit directly next to them to avoid making their intentions overly obvious. We are currently exploring these dynamics in ongoing work, aiming to extend our framework and provide further insight into how people incorporate their social goals with impression management.

Another limitation of our work is that we focused on unidirectional inferences (e.g., Purple's social attitude toward Yellow), but social interactions are inherently bidirectional. We are currently exploring how individuals infer mutual attitudes and interpret dyadic interactions from such minimal cues.

Altogether, our work demonstrates that people can make graded judgments about social relationships from very minimal events, such as a single seating choice. This work highlights the richness and flexibility of human social reasoning, showing how we form quick yet sophisticated impressions of social relations and affiliations in every social interactions.

Acknowledgments

We thank Ilker Yildirim for a helpful discussion and members of the Computational Social Cognition Lab for helpful feedback. This work was supported by NSF award BCS-2045778.

References

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*(4), 0064.

Flavell, J. C., Over, H., Vestner, T., Cook, R., & Tipper, S. P. (2022). Rapid detection of social interactions is the result of domain general attentional processes. *Plos one*, *17*(1), e0258832.

Isik, L., Mynick, A., Pantazis, D., & Kanwisher, N. (2020). The speed of human social interaction perception. *NeuroImage*, *215*, 116844.

Jern, A., Derrow-Pinion, A., & Piergiovanni, A. (2021). A computational framework for understanding the roles of simplicity and rational support in people's behavior explanations. *Cognition*, *210*, 104606.

Jern, A., Lucas, C. G., & Kemp, C. (2017). People learn other people's preferences through inverse decision-making. *Cognition*, *168*, 46–64.

Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial intelligence*, *101*(1-2), 99–134.

Kudrnova, V., Spelke, E. S., & Thomas, A. J. (2024). Infants infer social relationships between individuals who engage in imitative social interactions. *Open Mind*, *8*, 202–216.

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

McMahon, E., & Isik, L. (2023). Seeing social interactions. *Trends in Cognitive Sciences*.

Papeo, L., & Abassi, E. (2019). Seeing social events: The visual specialization for dyadic human–human interactions. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(7), 877.

Papeo, L., Stein, T., & Soto-Faraco, S. (2017). The two-body inversion effect. *Psychological science*, *28*(3), 369–379.

Pitcher, D., & Ungerleider, L. G. (2021). Evidence for a third visual pathway specialized for social perception. *Trends in cognitive sciences*, *25*(2), 100–110.

Thiele, M., Hepach, R., Michel, C., & Haun, D. (2021). Infants' preference for social interactions increases from 7 to 13 months of age. *Child Development*, *92*(6), 2577–2594.

Thomas, A. J., Woo, B., Nettle, D., Spelke, E., & Saxe, R. (2022). Early concepts of intimacy: Young humans use saliva sharing to infer close relationships. *Science*, *375*(6578), 311–315.