Inferring the internal structure of social collectives

Isaac Davis (isaac.davis@yale.edu), Yarrow Dunham (yarrow.dunham@yale.edu),
Julian Jara-Ettinger (julian.jara-ettinger@yale.edu)
Department of Psychology, Yale University. 2 Hillhouse Avenue, New Haven, CT 06520 USA

Abstract

We investigate how humans infer the rich internal structure of social collectives from patterns of interactions between agents. We propose a computational model of this process which integrates a domain-general statistical learning mechanism with, domain-specific knowledge about social contexts (i.e.: "intuitive sociologies"). We test our model in two experiments where participants observe a sequence of animated interactions between agents, and then assign the agents to groups according to their role or type within the social collective. Crucially, the two experiments depict different types of social interactions which reflect different types of underlying social structures. The patterns of correspondence between model predictions and human data support our account, and demonstrate the importance of both general statistical learning and specific social knowledge when reasoning about social collectives.
Keywords: Social Inference; Intergroup Cognition; Computational Modeling

Introduction

Imagine you are starting a new job in a large building with several floors of offices. At your first day of work, you notice some regularities about the people around you. For example, you notice some people wearing formal suits and working in large offices, some wearing company uniforms and working in smaller cubicles, and others wearing more diverse, casual clothing and regularly moving between offices. From these observations, you begin to build initial representations of your new social environment centered around the notion of a group. You might conclude, for example, that the office building houses several different companies, each with its own group of employees.

Our tendency to categorize people into groups is a fundamental aspect of human social cognition. A wealth of research has shown that this tendency begins to develop in early infancy and plays a crucial role in social cognition throughout childhood (Jin & Baillargeon, 2017); it occurs spontaneously, even when groups are only indicated by arbitrary markers or labels (Dunham, 2018); and it strongly influences our expectations about the characteristics and behaviors of those around us (Kawakami et al., 2021).

Despite their importance, simple group representations are often insufficient for navigating social environments. While it certainly can be important to categorize people according to which company they work for, it is also important to recognize the rich *internal* structure of each group. For example, certain people may have more authority to tell others what to do (e.g.: managers), and others may have specialized expertise (e.g.: IT support). Even if a manager and IT specialist belong to the same group (i.e. work at the same company), it may be inappropriate to interact with the manager and IT support in the same way. Understanding this rich internal structure of groups is critical for embedding ourselves into all kinds of social environments, from large scale institutions like governments and corporations to more transient organizations like sports teams or classrooms.

In this paper, we develop a computational account of the human capacity to infer the rich internal structure of social collectives from patterns of interactions between agents. We approach the problem as a form of latent structure inference: given some observed data (in this case, interactions between agents), what is the underlying social structure that best explains the data? Importantly, these latent structures can involve many different structural forms and relations, such as hierarchies, cliques, task groups, expertise groups, or combinations of these types. In the office setting, for example, we might group agents according to their relative positions in the company hierarchy, and the agents over which they have authority. With these groupings come certain expectations about how the agents will interact with each other (e.g.: that lower-level employees are more likely to fulfill orders from higher-level managers). We might also notice that certain employees regularly engage in friendly chit-chat or invite each other out for drinks, which suggests a different kind of grouping into social cliques, and carries its own set of expectations about agent behavior.

We propose that humans achieve this flexible inference by leveraging two different mechanisms. The first is a domain-general statistical inference mechanism for extracting latent structures from observable data patterns (Wood et al., 2012; Mansinghka et al., 2012). Previous work has used similar structure-learning algorithms to explain human judgments in a variety of non-social categorization tasks (Austerweil & Griffiths, 2008; Griffiths & Ghahramani, 2005), and has more recently been applied to social domains as well (Gershman et al., 2017; Gershman & Cikara, 2020). However, while much of this work focuses on grouping objects or agents according to perceived similarity, we focus on grouping agents based on how they interact with each other, which is often a stronger signal of latent internal structure than perceived similarity (e.g.: it would likely be inappropriate to treat a CEO

and a summer intern identically, even if they share very similar taste in music and movies). To this end, the second component of our model is a set of domain-specific "intuitive sociologies" (Mahalingam, 2007; Shutts & Kalish, 2021) that capture our commonsense expectations about human social behavior within different social contexts.

Now consider the problem of inferring the organizational hierarchy and social cliques which reflect the operation of two different intuitive sociologies. An observer fluent in them can not only make inferences from known structure (e.g. who will give orders to whom; who will invite whom for drinks) but can also use observed interactions (e.g. who actually gives orders to whom; who actually invites whom to drinks) to infer the most plausible structures. Our goal is to model this process of structure learning in both human subjects and a computational model.

We evaluate our model with a novel experimental paradigm in which subjects watch animated videos of agents interacting with each other, then group the agents according to their role or type within the depicted collective. Importantly, we design stimuli to depict two different kinds of social interactions, each of which reflects a different kind of underlying social structure. This allows us to investigate the role and importance of the "intuitive sociology" component of our model, and how human inferences change depending on the type of interactions depicted.

Computational Framework

Our framework seeks to explain how humans infer the latent structures underlying social collectives from patterns of interactions between agents. At a high level, we model the problem as a process of inferring a latent structure S which best explains a set of observed interactions D. In our model, D consists of a set of pair-wise interactions $d = (i, j, a)$, each specifying the agent i who initiated the interaction, the recipient j of the interaction, and the recipient's answer a to agent i.

While our framework is general enough to be applied to a variety of social domains, as alluded to above we focus on two domains that we test in our experiment: company hierarchies and friendship cliques. In these contexts, an interaction represents a work demand or a social invitation (for hierarchies and friendships, respectively), and the response consists of a binary "yes/no," indicating whether j decides to fulfill i's demand or accept i's invitation (depending on the interaction type).

Given the observed data, our model infers the social structure according to Bayes' rule: $p(S|D) \propto p(D|S)p(S)$. The first term on the right-hand side $P(D|S)$ is the likelihood of the data given the true structure S, which is induced by the observer's intuitive sociology, and the second term $P(S)$ is a prior distribution over social structures. We explain the derivations of these terms in greater detail below.

Formalizing intuitive sociology ($P(D|S)$)

We represent an intuitive sociology as a set of social types (i.e.: groups or roles that people can occupy) and a set of norms or expectations about how people interact with one another within and between types. Formally, we express these expectations as probabilistic functions encoding the likelihood of certain kinds of interactions occurring between agents of particular types. This probabilistic function specifies three key terms:

1. The probability $P_{init}(i|S)$ that agent i initiates an interaction, given the social structure (and agent i's role in that structure). This probability is proportional to the number of agents subordinate to agent i for work-demands, and proportional to the number of agents in the same friendship clique for social invitations (plus a noise term, ensuring a non-zero probability that any agent might initiate an interaction).

2. The probability $P_{rec}(j|S,i)$ that agent j is the recipient of an interaction, given the initiator. This is determined by two parameters, $\beta_{low} < \beta_{high}$. For work-demands, $P_{rec}(j|S,i) \propto \beta_{high}$ if j is subordinate to i and β_{low} otherwise. For social invitations, $P_{rec}(j|S,i) \propto \beta_{high}$ if j is in the same friendship clique as i and β_{low} otherwise. We place a uniform prior over $\beta_{high} \in (0,1)$ and a uniform prior over $\beta_{low} \in (0, \beta_{high}/2)$ and integrate these parameters out of our final computations.

3. The probability $P_{ans}(r|S, init = i, rec = j)$ of agent j's response r to agent i. Similar to above, we define a high and low parameter $\gamma_{low} < \gamma_{high}$, and set $P_{ans}(yes|S, init = i, rec = j) \propto \gamma_{high}$ or $P(yes|S, init = i, rec = j) \propto \gamma_{low}$ depending on whether i has authority over j (for work-demands) or i and j are in the same clique (for social invitations). We integrate these parameters out of the model in the same fashion as β_{high} and β_{low}.

Given these three terms, the likelihood term for each sociology is equal to

$$P(D|S) = \prod_{d=(i,j,a)\in D} P_{init}(i|S)P_{rec}(j|S,i)P_{ans}(a|S,i,j) \quad (1)$$

where the terms on the right-hand-side are computed as described above for each sociology.

Prior over social structures ($P(S)$)

The second component of our model is a prior distribution $P(S)$ over social structures. We define a social structure as a 2-tuple $S = \{C, E\}$ where $C : A \rightarrow T$ is a mapping of agents $a \in A$ onto social types $t \in T$, and E is a set of typed edges between clusters indicating inter-cluster relations (e.g.: E may be an authority relation indicating which groups of agents have authority over which other groups).

An important aspect of our framework is that it should be able to learn novel social structures that it has never encountered before. To achieve the requisite flexibility, we use an

infinite mixture model (Rasmussen et al., 1999) to define a prior distribution over structures of arbitrary complexity. To this end, we use a Chinese Restaurant Process (CRP) prior, which defines a probability distribution over cluster assignments $P(C)$ with an unbounded number of possible clusters, allowing the model to infer the appropriate number of clusters directly from the data. We then define a prior distribution $P(E|C)$ over edges given clusters by sampling a biased coin-flip for each cluster pair, to determine whether an edge exists between those clusters (the edge bias parameter is then integrated out with a Beta$(2,1)$ prior). This yields a prior distribution $P(S) = P(C)P(E|C)$ over social structures with an unbounded number of potential clusters.

Inference

The main application of our model is to infer a social structure $S = \{C, E\}$ from data D depicting a sequence of interactions among a fixed set of agents. At a computational level, this requires computing the posterior distribution $P(S|D)$ according to Bayes' rule as described above. In practice, this distribution is intractable to compute exactly, as it ranges over an infinite set of potential structures. We therefore approximate this distribution using a Metropolis-Hastings Markov Chain Monte Carlo (MCMC) algorithm with Gaussian proposal distributions, taking 30,000 samples and retaining every 6th sample, for a final approximation consisting of 5,000 samples. From these samples, we compute the average adjacency matrix $\bar{M} = \{m_{i,j}\}$, where $m_{i,j}$ indicates the proportion of samples in which agents i and j occupy the same cluster.

Experiments

We conducted two studies, depicting two different types of social interaction between agents: the "friendship" trials (Study 1a, trials F1-F5) depicted agents inviting other agents to socialize after work, while the "authority" trials (Study 1b, trials A1-A5) depicted agents giving work-related commands to other agents.

Participants

For each study, we recruited 40 adult participants with US-based IP addresses via Amazon Mechanical Turk. 2 participants were excluded from Study 1a after failing one or more comprehension checks, leaving N=38 participants (mean age=38.9, SD=11.4); 6 participants were excluded from Study 1b, leaving N=34 (mean age=38.5, SD=11.4).

Stimuli

Each study comprised 5 trials. In each trial, participants were shown an animated video depicting a sequence of 7 interactions between 5 agents who work in the same office. Each interaction contained three parts: first, the initiator of the interaction is shown moving from their starting position to the intended recipient. Then, a speech bubble appears next to the initiator depicting one of two symbols, corresponding to the two interaction types (a martini glass with a question mark for social invitations, or an envelope in a mailbox with an exclamation mark for work-related orders). Finally, the recipient of the interaction responded with a speech bubble depicting a "thumbs-up" and green check-mark (indicating "yes") or a "thumbs-down" and red X (indicating "no"). See figure 1 for examples of animated interactions.

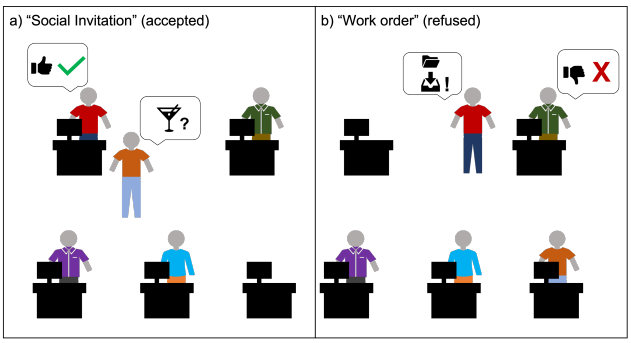
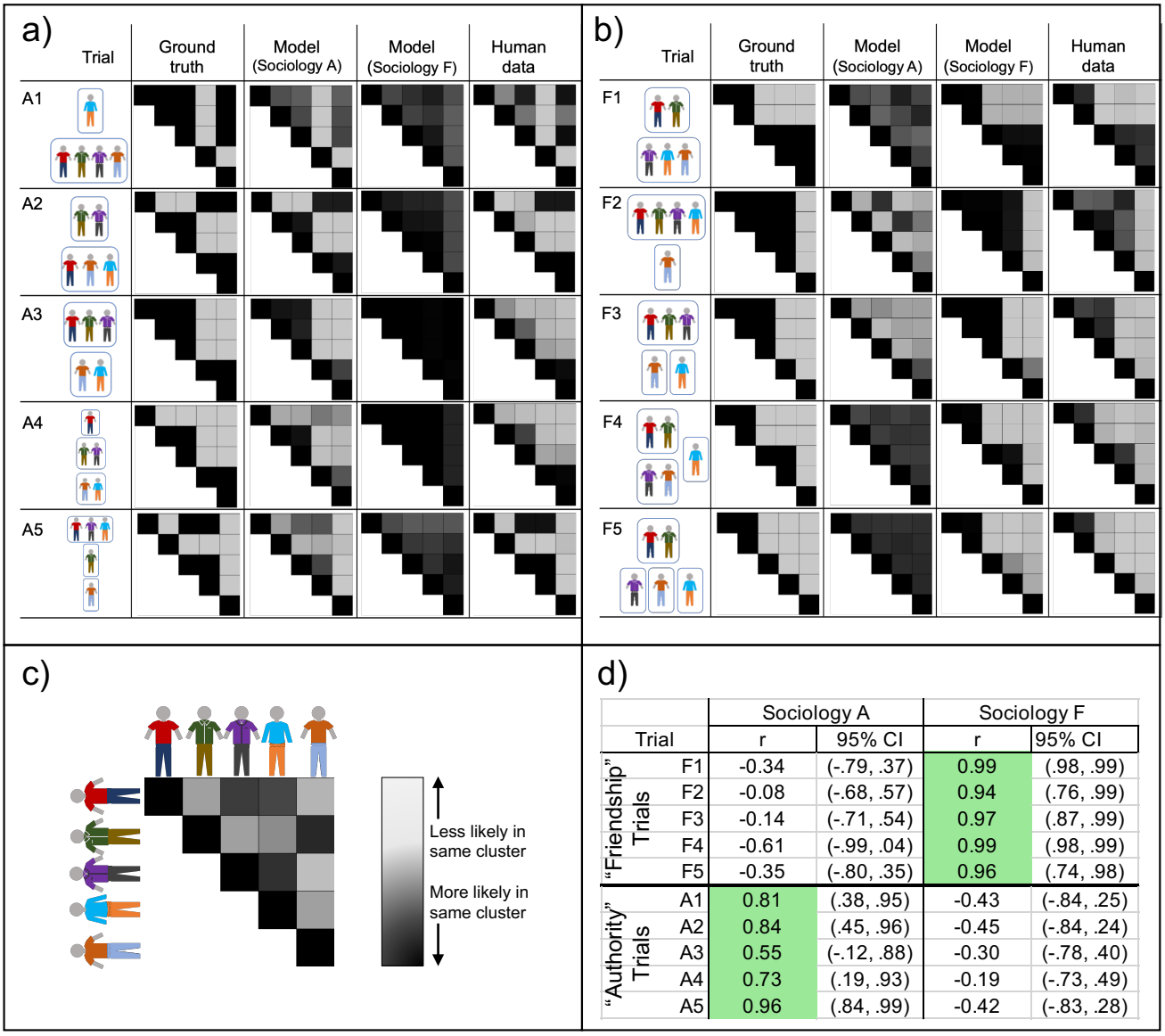


Figure 1: Example of interactions as presented in the stimuli. Panel a) depicts Orange giving an invitation to Red, who then accepts the invitation. Panel b) depicts Red giving a work-related order to Green, who then refuses to do the order.

To generate the animated videos, we first chose, for each trial, a "ground-truth" social structure of the appropriate type, each containing between 2 and 4 distinct clusters of agents. For each structure, we then sampled a sequence of 7 interactions between agents using the associated generative model. Because of the relatively small number of interactions depicted, a particularly noisy data set may lead to uninformative inferences (for both the model and participants). To select for informative data sets, we re-sampled each sequence until one was generated that met the following criteria: 1) it must include at least one "rejected" interaction, 2) it cannot depict the same interaction (between the same two agents) more than twice, and 3) when applied to the sequence, the inference model must infer a clustering that is at most one reassignment away from ground truth (i.e. at most one agent assigned to an incorrect cluster). After sampling an acceptable interaction sequence for each trial, we converted the sequence into a short animation (approximately 30 seconds each) using an application coded in Processing.

Procedure

Participants in each study were first shown a series of instructions explaining how to interpret the animations and interactions. In Study 1a, participants were told that people in the office will sometimes invite each other out after work, and that people may accept or reject these invitations. In Study 1b, participants were told that there were several different roles people could occupy in the company, including upper managers, middle managers, and low-level employees. Participants were further told that managers were more likely to give orders and less likely to agree to orders, while lower-level employees were more likely to receive orders and more



		Sociology A		Sociology F	
	Trial	r	95% CI	r	95% CI
"Friendship" Trials	F1	-0.34	(-.79, .37)	0.99	(.98, .99)
	F2	-0.08	(-.68, .57)	0.94	(.76, .99)
	F3	-0.14	(-.71, .54)	0.97	(.87, .99)
	F4	-0.61	(-.99, .04)	0.99	(.98, .99)
	F5	-0.35	(-.80, .35)	0.96	(.74, .98)
"Authority" Trials	A1	0.81	(.38, .95)	-0.43	(-.84, .25)
	A2	0.84	(.45, .96)	-0.45	(-.84, .24)
	A3	0.55	(-.12, .88)	-0.30	(-.78, .40)
	A4	0.73	(.19, .93)	-0.19	(-.73, .49)
	A5	0.96	(.84, .99)	-0.42	(-.83, .28)

Figure 2: Summary of study results. Panels a) and b) show a comparison of participant adjacency matrices against inference model and the ground truth. Rows correspond to trials. Leftmost column shows the ground-truth structure used to generate stimulus data for each trial. Second column shows binary adjacency matrices corresponding to the ground-truth structure. Third and fourth columns show average adjacency matrices inferred by the computational model using the "Authority Sociology"(A) and "Friendship Sociology" (F) modules, respectively. Fifth column shows average adjacency matrix induced by participant responses. Panel c) shows a legend for interpreting adjacency matrices: each cell depicts the probability that the two corresponding agents share a cluster. Darker colors indicate higher probabilities. Panel d) shows Pearson correlations and 95% confidence intervals between participant responses and model predictions generated with with sociology module A (first two columns) and F (Second two columns). Positive model correlations are highlighted in green

likely to agree to fulfill the orders they receive. After the initial instructions, participants were given two chances to pass a 4-question comprehension check ensuring that they were able to correctly interpret the stimuli. Participants who failed at least one question on both attempts were excluded from the study.

Participants who passed the comprehension check were then shown all five trials for the study in a random order. In each trial of each study, participants first watched the 30 second animation, and were then asked which agents were friends with each other (for Study 1a) or which agents had the same role at the company (for Study 1b). Participants responded by dragging and dropping 5 icons, corresponding to the 5 agents, into at most 5 different groups on the computer screen.

Results

As pre-registered[1], we first computed, for each participant in each trial, the 5×5 binary adjacency matrix induced by the participant's grouping. We then averaged these matrices across participants, yielding, for each trial, a 5×5 real-valued matrix whose i, jth entry equals the fraction of participants who placed agents i and j in the same cluster for that trial. To compare against model predictions, we applied the inference model to the same data set of interactions used for each trial two times, once with each sociology module. For each application, we generated 5,000 samples from the posterior distribution $P(S|D)$ using a Metropolis-Hastings algorithm, and computed the average adjacency matrix induced by these samples.

To compare model predictions against human data, we isolated the 10 independent, non-trivial parameters of the adjacency matrices[2] and computed, for each trial, the Pearson correlation between the average matrix induced by the human data and the average matrix outputted by the model. Correlations for all 10 trials across both studies are shown in Figure 2d.

As an additional means of visualizing the data, we generated four color-coded adjacency matrices for each trial, corresponding to the ground truth structure, the average structure inferred using the Authority sociology and the Friendship sociology, and the average structure induced by human data (Figures 2a-2c). This qualitative visualization is an important supplement to the quantitative measures of fit, as small qualitative changes to inferred structure (e.g.: moving a single agent from one cluster to another) can result substantial drops in the correlation between adjacency matrices.

Our results support both a) that participants inferred the underlying social structures in a manner that tracks with model predictions and b) that participants are sensitive to the type of interaction being depicted in a way that reflects the two "intuitive sociologies" coded in our model. For the Friendship trials (F1-F5), participant judgments were strongly and significantly correlated with model predictions generated using the Friendship sociology, and negatively correlated or uncorrelated with predictions generated using the Authority sociology. Results from the Authority trials were more varied in degree of fit but showed the same overall pattern: participant judgments were strongly correlated with the Authority predictions in all but one trial, and were negatively correlated with the Friendship predictions in all five trials.

The wider range in degrees of fit for the Authority trials is expected to some degree, as "authority" is, in several ways, an inherently more complex social dynamic than friendship. First, friendship is generally a symmetrical relation, while authority is generally anti-symmetrical: if A is friends with B, it usually follows that B is friends with A, but if A has authority over B, it usually follows that B does *not* have authority over A. Similarly, it is often easier to generalize friendship across agents than authority: for example, if A is friends with B and C, then it is more likely that B and C are friends as well. However, if A has authority over both B and C, it may also be the case that B has authority over C, or C has authority over B, or neither has authority over the other, none of which are directly supported by A's authority over B and C. Thus, given the relative sparsity and noisiness of the stimuli, it is unsurprising that participant responses were overall more consistent in the Friendship trials than the Authority trials. That said, the qualitative comparison in figures 2a & 2b suggests that much of the discrepancy between participant responses and model predictions (in the Authority trials) is degree of confidence: that is, for most of the authority trials, average participant responses yielded qualitatively similar adjacency matrices to the corresponding model inference, but differed in the degree of certainty in each pair-wise adjacency.

Discussion

The tendency to assign people to social categories and groups is a fundamental aspect of human social cognition, and exerts a major influence on our behavior towards and expectations of those around us (Dunham, 2018; Jin & Baillargeon, 2017; Kawakami et al., 2021). Much existing research on this subject focuses on simple group representations (e.g.: ingroup versus outgroup), effectively treating each group as an unstructured container of individuals. However, there are many contexts in which these simple group representations are insufficient for navigating social environments. When a child joins a team sport during recess, for example, it is often not sufficient to simply recognize which children are on which team: the child must also recognize the different roles or positions that exist within the team, and how these roles or positions influence the behavior and expectations of the children occupying them. Understanding this rich *internal structure* is often critically important for embedding ourselves into social

[1] osf.io/d25pz

[2] Since an adjacency matrix is symmetric about its diagonal, we use only the top half of each matrix for our correlation computations. Furthermore, the diagonal will always consist of all 1's, since each agent is always in the same cluster as itself. This leaves the 10 independent, non-trivial parameters above the diagonal of each matrix

collectives.

In this project, we proposed and tested a computational account of the human capacity to infer this internal social structure. Our account leverages two core mechanisms: a domain-general mechanism for extracting latent structures from observable data (Wood et al., 2012; Mansinghka et al., 2012), and a set of domain-specific expectations about the behavior of agents in different social contexts, i.e.: an intuitive sociology (Mahalingam, 2007; Shutts & Kalish, 2021). Our experiments provide converging support for our computational account. Across both experiments, the average latent structures inferred by participants tracked with model predictions in terms of the number of distinct clusters, as well as the assignment of agents to clusters. This supports our hypothesis that humans leverage something like a statistical latent-structure learning mechanism for reasoning about the internal structure of social collectives. Furthermore, by comparing participant responses against model predictions generated using *both* of our sociology modules, we demonstrated that participants draw on different sets of expectations depending on the type of interactions being depicted, and that these expectations are similar to the two sociologies coded in our model. In particular, participants generally expected social invitations to reflect a symmetrical relationship between agents (i.e.: if A invites B, then A is friends with B and B is therefore also friends with A), but expected work-related commands to reflect an asymmetrical relationship (i.e.: if A gives an order to B, then A has authority over B but B does not have authority over A). Thus, our model demonstrates how the abstract social dynamics encoded by an intuitive sociology can be concretely implemented in a group of agents, and how we can leverage our expectations about these dynamics to infer the underlying social structure.

While our initial results are promising, they also contain important nuances that necessitate further investigation. In general, participant responses showed much greater variability, and larger discrepancies with model predictions, in the Authority trials than in the Friendship trials. To some degree this increased variability was expected, as authority is an anti-symmetric relation and in several ways more complex than friendship. Another possible source of these discrepancies lie in how participants and the model each interpret conflicting signals in the data. For example, if A gives an order to B, that signals that A has authority over B, but if B refuses to fulfill A's order, that signals that A lacks authority over B. Thus, an interaction in which B refuses an order from A has two potential and mutually exclusive explanations: either A lacked the authority to order B in the first place, or B refused an order that they should, in fact, have fulfilled. In our model, the probability of each interpretation is controlled by a separate parameter, and while these parameters are independent from each other, they are both drawn from the same prior distribution. It is therefore possible that human participants consistently interpreted one signal more strongly than the other. For example, if participants believe that refusing an order from someone with authority over you is much less likely than giving an order to someone over whom you have no authority, the average responses would be systematically biased in a way that reflects this asymmetry. Thus, further investigation is required to determine whether the discrepancies in the Authority trials are due to a similar asymmetry in human expectations about authority.

A further limitation of the current work is that the experimental stimuli depict only one type of interaction at a time, but in real-world social environments, we often observe multiple kinds of interactions, reflecting multiple types of underlying relations and structures. For example, one may have several bosses at work but also be friends with one of their bosses outside of work. This may lead to more nuanced patterns of interactions: one may interact differently with a boss who is also a personal friend than with a boss who is not a personal friend, and may also interact differently with a friend who is a boss than a friend who has the same role at the company. Our model can be extended to do inference over data depicting multiple types of interaction, and to simultaneously infer multiple underlying social structures reflected in the different interactions (e.g.: simultaneously inferring who is friends with each other *and* who has authority over each other).

Finally, there are many ways in which people can signal their affiliation with each other without *directly* interacting with each other. One fairly well-studied example of this are social choices, which are choices that directly affect someone other than just the chooser. For example, if there is a limited supply of resources (e.g.: snacks in the break room), then my choice to take some of those resources directly affects the options available to anyone else with access to the same resources. Previous work has leveraged situations like these to show that people make strong inferences about affiliations between individuals (e.g.: how much A cares about B) based on the social choices of one individual (e.g.: which snacks A takes for themself, and which ones they leave for B) (Davis et al., 2021; Jern et al., 2017; Van Doesum et al., 2013). This suggests a potentially fruitful extension to our model, by enabling it to infer affiliations between agents based on both direct interaction *and* indirect actions like social choices, and extrapolate the broader social structure from both sources of information.

To conclude, we proposed a computational account of the human ability to infer the latent structure underlying human social collectives from observed patterns of interactions between agents within the collective. Our account leverages both domain-general statistical learning mechanisms and domain-specific expectations about social dynamics. We demonstrated that people infer these latent social structures in a fashion that reflects both a general statistical inference process and context-specific expectations about social behavior. This constitutes an important step towards a unified computational account of inter- and intra-group cognition.

References

Austerweil, J. L., & Griffiths, T. L. (2008). Analyzing human feature learning as nonparametric bayesian inference. In *Nips* (pp. 97–104).

Davis, I., Carlson, R. W., Dunham, Y., & Jara-Ettinger, J. (2021). Reasoning about social preferences with uncertain beliefs.

Dunham, Y. (2018). Mere membership. *Trends in Cognitive Sciences*, *22*(9), 780–793.

Gershman, S. J., & Cikara, M. (2020). Social-structure learning. *Current Directions in Psychological Science*, *29*(5), 460–466.

Gershman, S. J., Pouncy, H. T., & Gweon, H. (2017). Learning the structure of social influence. *Cognitive Science*, *41*, 545–575.

Griffiths, T. L., & Ghahramani, Z. (2005). Infinite latent feature models and the indian buffet process. In *Nips* (Vol. 18, pp. 475–482).

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

Jin, K.-s., & Baillargeon, R. (2017). Infants possess an abstract expectation of ingroup support. *Proceedings of the National Academy of Sciences*, *114*(31), 8199–8204.

Kawakami, K., Hugenberg, K., & Dunham, Y. (2021). Perceiving others as group members: Basic principles of social categorization processes.

Mahalingam, R. (2007). Essentialism, power, and the representation of social categories: A folk sociology perspective. *Human Development*, *50*(6), 300–319.

Mansinghka, V., Kemp, C., Griffiths, T., & Tenenbaum, J. (2012). Structured priors for structure learning. *arXiv preprint arXiv:1206.6852*.

Rasmussen, C. E., et al. (1999). The infinite gaussian mixture model. In *Nips* (Vol. 12, pp. 554–560).

Shutts, K., & Kalish, C. W. (2021). Intuitive sociology. In *Advances in child development and behavior* (Vol. 61, pp. 335–374). Elsevier.

Van Doesum, N. J., Van Lange, D. A., & Van Lange, P. A. (2013). Social mindfulness: skill and will to navigate the social world. *Journal of Personality and Social Psychology*, *105*(1), 86.

Wood, F., Griffiths, T., & Ghahramani, Z. (2012). A non-parametric bayesian method for inferring hidden causes. *arXiv preprint arXiv:1206.6865*.