Title

Distributed statistical inference in social interaction networks

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Zubak, Yuliya

Dronavalli, Pranav

Chuang, Yun-Shiuan

et al.

Publication Date

2024

Peer reviewed

Distributed statistical inference in social interaction networks

Yuliya Zubak[1], **Pranav Dronavalli**[2], **Yun-Shiuan Chuang**[12], **Robert D. Hawkins**[1]
[1]Department of Psychology, University of Wisconsin-Madison
[2]Department of Computer Science, University of Wisconsin-Madison
{yzubak,dronavalli,chuang,rdhawkins}@wisc.edu

Abstract

Humans rely on our social networks to make more accurate inferences about the world. Yet it remains unclear how those inferences are shaped by the medium through which information is exchanged and beliefs are shared. In this paper, we report two experiments where participants ($N = 645$) were asked to make inferences about an unknown probability distribution based on limited private observations. They exchanged messages with neighbors in a small social network and were asked to update their beliefs over repeated rounds. We compared three conditions: a unidirectional message medium, a constrained slider medium, and an interactive chat. All groups were able to converge toward more accurate inferences, but their convergence rates varied across conditions in ways not well-captured by common models. We argue that computational models of collective behavior must move beyond the assumption of direct belief transmission to capture the complexities of sharing information through natural language.

Keywords: Social learning; statistical inference; collective behavior

Introduction

As individuals, humans must learn and make decisions based on relatively sparse and noisy observations. As social groups, however, we have access to a much larger pool of knowledge. We rely on communication to aggregate information across different agents and to collectively arrive at inferences, predictions, and decisions that integrate diverse experiences (Henrich, 2016; Gweon, 2021; Hawkins et al., 2023; Fränken, Valentin, Lucas, & Bramley, 2024). For instance, giving individual clinicians access to an information-sharing network significantly reduces medical errors (Centola et al., 2023). A longstanding question across the cognitive sciences concerns how exactly human groups manage to balance individual and social information (Galesic, Olsson, Dalege, van der Does, & Stein, 2021; Becker, Brackbill, & Centola, 2017). What cognitive mechanisms and social dynamics enable effective information sharing?

One fruitful approach to studying collective inference is to study agent-based models of opinion dynamics (Smaldino, 2023; Goldstone & Janssen, 2005; Golub & Jackson, 2010). These models simulate how networks of individual agents exchange and update beliefs through repeated social interactions. One classic example is a voter model (Axelrod, 1997), where agents copy the opinion of a randomly selected neighbor at each time step, leading to convergence or polarization across the network (Nowak, Szamrej, & Latané, 1990). Other models integrate more sophisticated cognitive factors like confirmation bias, where agents preferentially adopt opinions that match their existing views (Hegselmann & Krause, 2002) or social identity (Steiglechner, Smaldino, Moser, & Merico, 2023), where agents preferentially attend to members of their own group. These models may also be enriched by statistical inference mechanisms, as in Bayesian models of social learning (Bonawitz & Shafto, 2016; Tang & Chorus, 2019; Shafto, Goodman, & Frank, 2012)

While agent-based modeling provides a useful framework for deriving collective outcomes from assumptions about individual cognition, empirical measurements of collective behavior have lagged behind the proliferation of models (Moussaïd et al., 2017; Centola, 2010). Without quantitative data, it is unclear which models best capture the patterns observed in human groups and human social cognition. Most agent-based models assume direct transmission of scalar opinions, yet real-world interaction involves a *linguistic* bottleneck, requiring individuals to interpret verbal messages, extract relevant information, and map it back to their internal beliefs. That is, linguistic utterances have a complex, non-linear relationship with social cognition. Communicative acts are decoupled from internal belief updating (Moussaïd et al., 2018; Van Overwalle & Heylighen, 2006).

Controlled behavioral experiments are critical for revealing divergences from idealized theoretical dynamics. In this paper, we report results from a novel group inference task where participants exchanged messages in small social networks under varied conditions. By tracking how beliefs evolved over repeated interactions, we directly measured the pathways by which social information shapes collective knowledge. Our findings reveal that subtle factors in communication mechanics and network structure can accelerate or impair convergence to accurate beliefs. For example, introducing a constrained numerical format for messaging akin to classical models actually slowed or reversed the ability of the group to converge on the true value. These effects demonstrate the need to move beyond simplistic models of direct transmission and incorporate richer cognitive mechanisms of information exchange. More broadly, our work aims to provide a stronger empirical foundation for a new generation of models that capture the complexities of collective human inference.

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

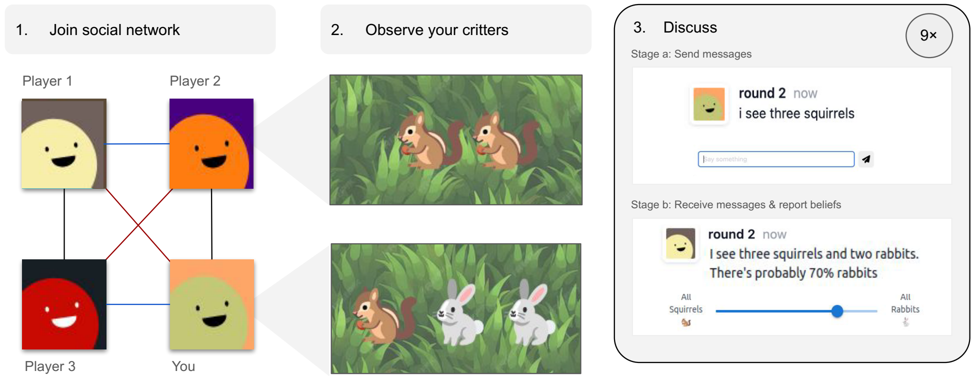

Figure 1: Participants were connected in groups of four and observed a private sample of wildlife (here, two squirrels, or one squirrel and two rabbits). On each trial of the experiment, they were asked to write a message to a neighbor and report their beliefs about the underlying distribution after receiving a message.

Experiment 1: Distributed statistical inference

Participants We recruited $N = 144$ participants from Prolific and assigned them to $N = 36$ groups of four. All participants were from the US, UK, or Canada, and were pre-screened as fluent English speakers. Active participants received a base pay of \$15 US per hour, and the experiment took approximately 15 minutes. We excluded six participants who failed to provide responses for 2 consecutive trials. We excluded six additional participants showing a deviation of more than 70% from their initial distribution on the first round, suggesting they misunderstood the direction of the scale on the slider.[1]

Stimuli Participants were asked to estimate the relative proportion of rabbits 🐰 vs. squirrels 🐿️ in the local wildlife population. We showed each participant a fixed set of critters out their virtual "window" as initial private information to be shared in subsequent rounds of social interaction. Private samples of critters for each participant in a given network were independently drawn from binomial distributions with the same fixed proportion p. In order to vary the total amount of observations across participants, we also placed a hyper-prior over the sample size N, giving each individual a unique sample size and yielding the following distribution:

$$P(\text{🐰}|p) = \mathbb{E}_N\left[f(\text{🐰}, N, p)\right] = \mathbb{E}_N\left[\binom{N}{\text{🐰}} p^{\text{🐰}}(1-p)^{\text{🐿️}}\right]$$

for $N \sim \text{Unif}\{0,9\}$. In other words, for each participant in a network, we first sampled an integer between 0 and 9 to be the total number of critters, and then we sampled the set of observations from a binomial distribution. This procedure allowed for variability both in the total amount of information and local proportions, while maintaining a constant underlying probability across the entire network. We set $p = 0.7$ for half of the networks and $p = 0.3$ for the other half.

[1] Data from the remaining participants in these networks were still included in analyses, although results were robust if we excluded all incomplete networks.

Procedure After reading the task instructions and passing a comprehension quiz, participants were directed to a reactive web application built with Empirica (Almaatouq et al., 2021). The task interface had two major components. On the left side of the screen, a group of critters (rabbits and squirrels) were presented out a 'window'. On the right side of the screen, a messaging interface was provided to communicate with other participants in the network and report beliefs about the true underlying distribution (see Figure 1).

On each round of the task, participants were paired with another participant in the group and given 30 seconds to send a message or messages about the underlying distribution. There was no limit to the number of messages they could send, but the communication modality was *unidirectional* so they did not receive any messages back during this period, reflecting the structure of information exchange typically used in agent-based models. Following the communication stage, they were shown the message(s) produced by their partner and asked to report their own updated beliefs about the underlying distribution using a slider. The slider ranged from no rabbits (100% 🐿️) to all rabbits (100% 🐰).

Participants proceeded to the next round after 30 seconds or as soon as all members of their network had submitted their responses. Dyadic pairings on each trial were rotated using a round-robin algorithm such that participants sent messages to, and received messages from, each partner exactly three times over the course of nine trials. We used this structure rather than a blocked design (i.e. repeated exchanges with a single partner before moving to the next partner) to accelerate the "mixing rate" of information in the network; a blocked design would be more similar to the interactive condition introduced in Experiment 2 below. Demographic information was collected in an exit survey following the final round.

Results

Communication improves inferential accuracy We begin by considering the extent to which groups are able to improve their collective accuracy through deliberation. We hypoth-

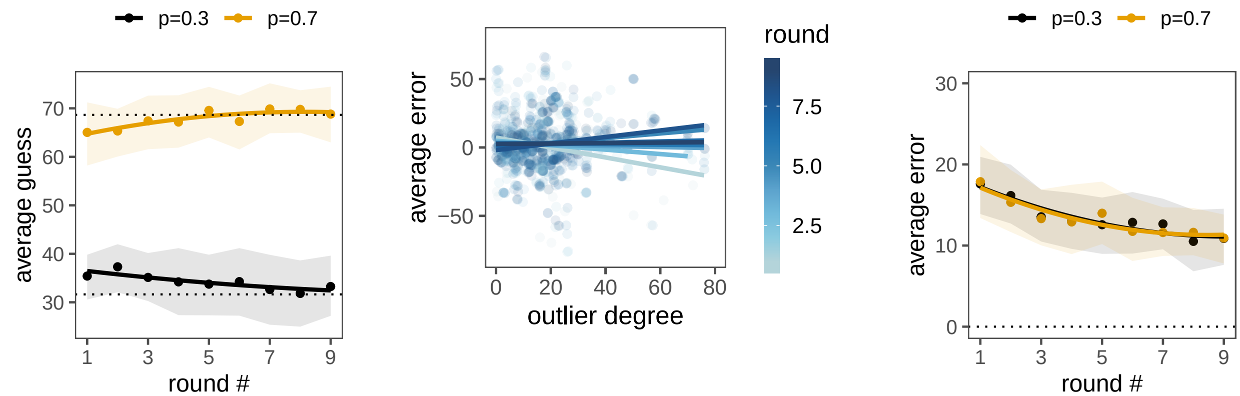

Figure 2: (Left) Participants' initial estimates tend toward the midpoint but as communication unfolds within the network, estimates approach the true latent probabilities p. Dotted lines represent empirical frequencies $\hat{p}$ in our sample. (Right) Average error decreases significantly for both conditions as a function of social information exchange. Participants with an outlier sample from the distribution (measured here as distance from game p) are more likely to significantly modify their responses, lowering error over time. Error ribbons are bootstrapped 95% CIs. (Middle) Participants whose private observations deviate more from the ground-truth proportions tend to adapt their guesses the most, eventually reaching an error rate closer to participants with more representative information.

esized that, although each participant received sparse local observations, the network as a whole would be able to aggregate their estimates over time to approach the true latent probability as seen in agent based models. We tested this hypothesis using a mixed-effects model predicting each slider rating (0 to 100) including fixed effects of round index (continuous; 1 to 9), latent probability condition (sum-coded; 0.3 vs. 0.7), and their interaction. Due to nested layers of clustered variation, we included intercepts and random effects of round index for each game as well as for each player within those games. We found a significant interaction, $b = -16.18$, $t(43.5) = -9.2$, $p < 0.001$, where participants in the $p = 0.7$ condition increased their guesses over time while participants in the $p = 0.3$ condition gradually decreased their guesses. In other words, participants tended to regularize their initial guesses closer to the midpoint of the scale before moving in toward the true rate as social information was acquired (see Figure 2, left).

Next, to more directly examine our hypothesis that a game's error would decrease over time as messages are exchanged, we constructed a second mixed-effects model instead predicting trial-level *error*: how far off each slider estimate was from the true probability. Because each game had a distinct distribution of critters that fluctuated around 0.7 or 0.3, we used the game-level empirical frequencies $\hat{p}_i$ = 🐰$_i$/(🐰$_i$ + 🐿$_i$) as our reference point, i.e. our dependent variable was $\varepsilon = |\hat{y}_i - \hat{p}_i|$. As predicted, we found a significant main effect of round index, $b = -63.38$, $t(71.6) = -4.86$, $p < 0.001$, suggesting that error decreases and estimates converge toward the empirical game-level frequency across both conditions (see Figure 2, right). We found no evidence that these slopes differed across conditions, $p = 0.62$. These findings were robust to additional exclusion criteria, such as removing players who happened to draw an empty sample (no critters at all), or removing games where more than one player was missing data due to inattention.

Outliers make bigger belief revisions Next, we turned to examine how individual belief updates vary as a function of private and social information. It is commonly observed in the collective behavior literature that agents make larger revisions to their estimates when they are more out of step with their neighbors (e.g. Becker et al., 2017). Here, we test the extent to which this effect replicates under linguistically-mediated communication. First, we calculated a measure of *outlier degree* for each participant, defined as the distance between the empirical proportion of rabbits in a participant's private sample $\hat{p}_{ij}$ and the empirical proportion in their game as a whole $\hat{p}_i$. We then constructed a linear mixed-effects model predicting participants' error magnitudes as a function of their outlier degree, round index, and the interaction between the two. All variables were z-scored, and we included maximal game-level random effects[2]

As predicted from prior work(Moussaïd, Kämmer, Analytis, & Neth, 2013), we found a significant interaction between outlier degree and round index on error, $b = 1.4, t(881) = 2.79, p = 0.0054$. Participants who happened to receive outlier samples (e.g. 2 rabbits and 1 squirrel in the condition where squirrels were more likely) initially had higher error in their estimates (light blue line) but the slope

[2] The estimated variance of the random interaction was near 0, so we removed it from the random effect structure.

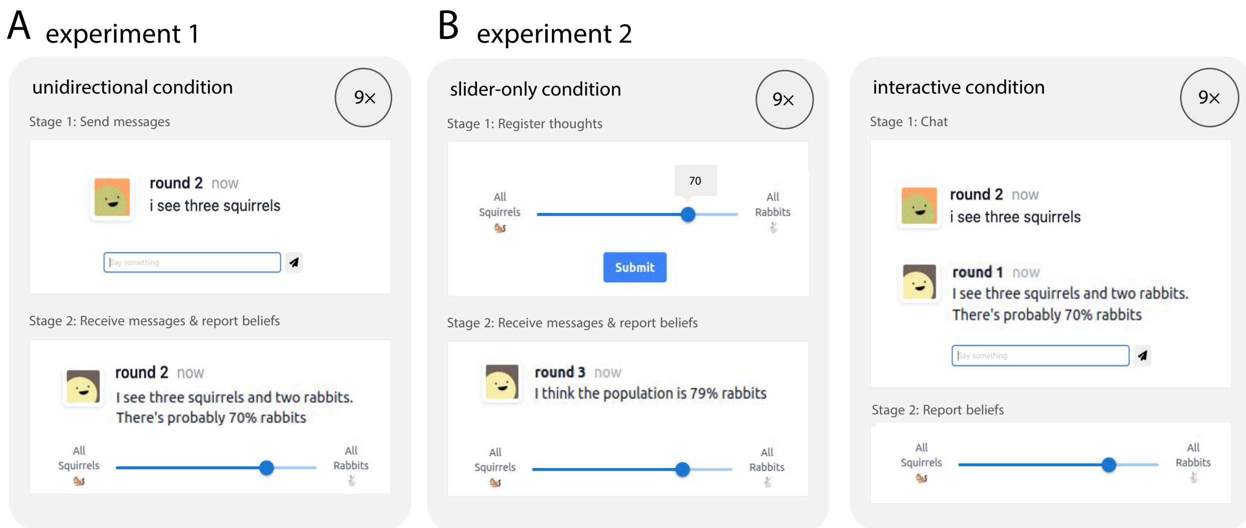

Figure 3: We compare three different communication modalities to examine their impact on group opinion dynamics. (A) In Experiment 1, we used a *unidirectional* interface where messages were asynchronously sent and received. (B) In Experiment 2, we added a 'slider-only' condition, which constrains input to a proportion slider, generating a pre-worded message, as well as an 'interactive' condition allowing unconstrained dyadic communication through the chat box.

of this relationship significantly decreased over time (darker blue lines), indicating that these participants were revising their guesses without other participants' estimates getting significantly worse (see Figure 2, middle). In other words, participants receiving social information that deviates strongly from their own observations tend to most significantly revise their beliefs.

Experiment 2: The effect of the communication medium

Our results so far suggest that groups effectively aggregate private knowledge into more accurate collective estimates by exchanging information. However, it remains unclear how the specific medium of communication impacts this process. Classic agent-based models (e.g. DeGroot, 1974) typically assume that agents have direct access to the underlying beliefs of neighboring agents. On one hand, we may expect the linguistic bottleneck that free-form messages must pass through may reduce the efficiency and accuracy of opinion transmission compared to directly sharing numerical representations, thus impairing convergence. On the other hand, the flexibility and richness of natural language may allow for new coordination strategies that improve on simplistic averaging models.

To test these accounts, we extended our paradigm from Experiment 1 with two additional conditions manipulating the communication medium: (1) a 'slider-only' condition resembling the classic assumption from agent-based models that others' estimates are directly accessible, and (2) an 'interactive' condition allowing free bidirectional communication. By comparing what makes groups more or less effective across these conditions, we aim to reveal how the *communication modality* of social networks shape their collective outcomes (Boyce, Hawkins, Goodman, & Frank, 2022).

Participants We recruited an additional $N = 420$ participants from Prolific and assigned them to $N = 105$ unique networks of size four (approximately $N = 25$ networks per condition). We used the same exclusion criteria as Experiment 1, filtering out 18 participants who initially used the slider inconsistently with the observations they were given.

Design The procedure was identical to Experiment 1, except for the communication interface. The *slider-only* condition constructed a close behavioural replica of classic agent-based models. Communication entirely took place through direct numerical data transmission of opinion reports. Rather than starting with a message-sending phase and then proceeding to a belief report phase, individuals were initially prompted to input their opinion using a slider. After all participants submitted their belief reports, they advanced directly to the next round, where they were shown the same slider interface but with an auto-generated message (e.g. "I think the population is 73% rabbits"). In other words, the two stages were merged, because providing a belief report is the same act as providing a message. Meanwhile, the *interactive* condition used the same two-stage control flow as Experiment 1, but instead of waiting until the report stage to receive messages, dyads were able to bidirectionally exchange messages in real time for the entire 30 second communication window, similar to a one-on-one text messaging conversation.

Results

Direct transmission of beliefs impairs group inference To evaluate the extent to which different communication modalities improve or impair the distributed inference abilities of participants in our task, we compared the average error

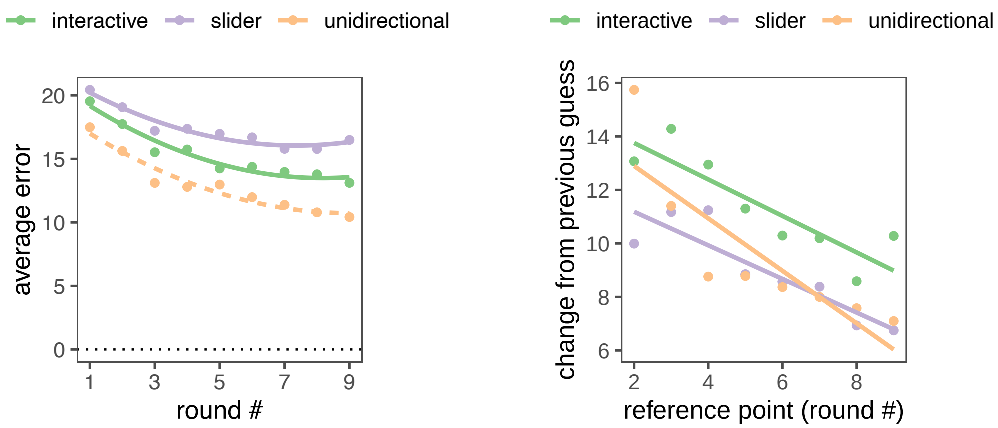

Figure 4: (left) Among our three conditions, groups using unidirectional messaging performed best (Experiment 1), groups using slider-only communication performed worst, and groups using interactive channels were in between (Experiment 2). (right) Participants in all three conditions change their reported beliefs less and less over the course of the study.

of each guess across conditions. As in Experiment 1, we calculated error as the distance between the slider value and the network's empirical frequency of rabbits $\hat{p}_i$. We constructed a mixed-effects model with fixed effects of condition (unidirectional vs. slider-only vs. interactive, sum-coded) and round index (1 through 9, centered), with nested random effects for each game and each player within that game.

First, we found a significant main effect of round index where error decreased over time in all conditions, $b = -6.9$, $t(134.3) = -7.11$, $p < 0.001$. More centrally for our hypothesis, however, we also found a main effect of condition, with the slider condition performing significantly more poorly at all time points than the average across conditions $b = 0.14$, $t(138.9) = 2.25$, $p = 0.028$, and the unidirectional condition performing significantly better than average, $b = -0.13$, $t(136.4) = -2.15$, $p = 0.034$. As seen in Figure 4, the interactive condition falls somewhere in between. Thus, we find some support for the hypothesis that, rather than serving as a bottleneck and impeding the direct transmission of belief states, language may in fact facilitate more accurate aggregation of information across networks. We consider potential reasons for this ordering in the Discussion below.

Direct transmission helps groups with more information Due to the fact that each participants' sample size N was randomized, some groups happened to receive to a larger amount of total information ('critters') than others. The increased number of total observations were distributed more or less unevenly across participants in the group. In this section we test the hypothesis that, even if groups in the *interactive* or *slider-only* conditions do not perform as well as the *unidirectional* groups overall, these alternative interfaces may simply be more sensitive to data sparsity and allow groups to perform quite well when sample sizes are large. To test whether total information impacted convergence differently across conditions, we compared a mixed-effects model including main effects of condition, sample size, and round index against a model also including the interaction between condition and sample size. As seen in Figure 5, the model with an interaction provided a significantly better fit in a nested model comparison, $\chi^2(4) = 11.48$, $p = 0.022$, suggesting that the relationship between information quantity and collective error depends on the communication medium. This effect suggests that interactive and slider-based modalities may be disproportionately affected by lower quality data, possibly due to the information format or working memory load. Follow-up analyses indicated that the effect of sample size on group error in Experiment 1 was significantly smaller than the effect for the interactive condition, $t(484) = 2.18$, $p = 0.03$, but we could not reject the null hypothesis for the slider condition, $t(485) = -0.688$, $p = 0.49$.

Participants' guesses stabilize over time A final reason for the observed difference in performance across conditions is the speed of convergence. Intuitively, information obtained later during learning should have less of an impact on beliefs than information obtained at the beginning; however, it was unclear whether this would differ as a function of the communicative channel. We measured the absolute difference between a player's guess on round k and their guess on $k+1$ and constructed a mixed-effects model with fixed effects of round index and condition and maximal random effect structure. First, we found a significant main effect of round index, $b = -110$, $t(168) = -5.93$, $p < 0.001$, indicating that participants were gradually stabilizing over time in all conditions; they were changing their responses less and less as the game went on. Additionally, we found a significant main effect between the slider condition and the interactive condition, $b = -2.3$, $t(160) = -2.14$, $p = 0.034$ where participants overall made much smaller changes in the slider-only condition.

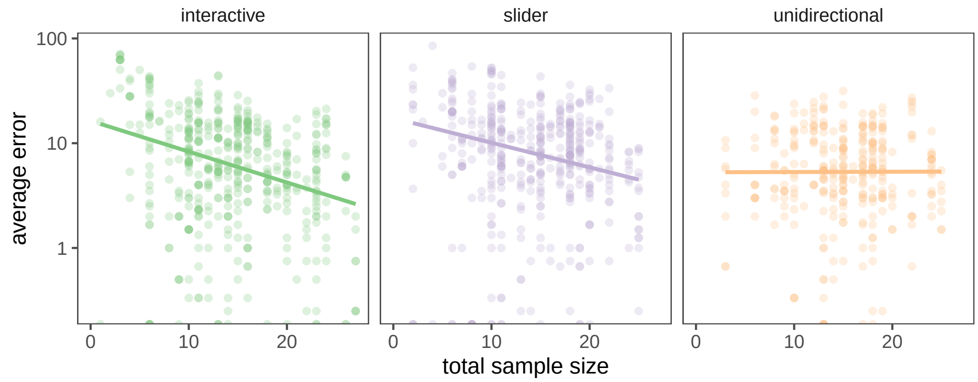

Figure 5: Unlike the unidirectional condition (Experiment 1), participants in the interactive condition and the slider-only condition (Experiment 2) performed relatively better when the group as a whole had access to a larger sample size of animals distributed across individuals in the group.

Discussion

Agent-based models of opinion dynamics are popular for a reason. They provide a simple interface for examining emergent behavior at the level of populations by grounding it in simple assumptions at the level of individuals. However, as these models grow in complexity, it becomes more critical to take a step away from idealized qualitative phenomena toward quantitative measurements of real human groups in controlled settings. In this paper, we considered a distributed inference task where a small social network collectively inferred a latent probability from limited individual observations. We found that, overall, groups were able to aggregate knowledge through communicating and successfully reduced the error of their estimates over repeated interactions.

Critically, we manipulated the *communication modality* used to exchange information, finding that a slider interface that enforces standard numerical assumptions of agent-based models was not as effective as natural-language interfaces. The improvement observed with text communication may arise from increased information density in communicative speech acts; in speech, the participants have the opportunity to communicate not only their own current beliefs, but also the beliefs of others, their confidence in these information sources, and the data itself, thus distinguishing their perception of the network from their beliefs about the latent value. It is possible that some of the disadvantage is attributable to presenting the slider-transmitted information as a percentage (Gigerenzer, Hertwig, Lindsey, & Hoffrage, 2000), but reported beliefs as percentages in all conditions, so this difficulty would likely affect all conditions equally.

One particularly surprising finding was that groups using a bi-directional messaging interface performed more poorly than groups using a uni-directional messaging interface despite allowing for higher-bandwidth deliberation. There are several potential explanations for this counterintuitive result. For one, the additional quantity of messages being exchanged may increase noise or working memory load, outweighing the benefits of bidirectional information flow. A related factor is the possibility of introducing interference between different conversations with different interlocutors, leading some information to be more socially salient or weighted more strongly than others; introducing stronger social cues and dissociable avatars might help to reduce interference. Finally, these unstructured dialogues may have introduced opportunities for disorganized, "stream-of-consciousness" arguments or mistrust about their interlocutor's reliability, undermining cooperative inference in the relatively short window of time allowed for discussion. Untangling these possibilities will require analysis of the language itself. But it is clear that simply allowing freer exchange does not automatically improve collective intelligence.

Another puzzling result is that, while error decreased in relative terms, participants also seemed to consistently over- or under-shoot the true latent value in absolute terms, resulting in unexplained residual error. It is possible that 9 rounds of interaction is not sufficient for convergence, or that participants simply have a positive response bias in their slider use. Another more intriguing explanation is that pieces of redundant social information were being double-counted. For example, participants could think they were being paired with another new partner on the 4th trial when in fact they were paired with their first partner again (Whalen, Griffiths, & Buchsbaum, 2018). These misunderstandings could create positive feedback loops that lead networks away from the ground truth.

Natural language facilitates a richer space of strategies than simply averaging opinions across neighbors. We argue that formal models going forward must move beyond direct transmission assumptions to capture this hidden layer of linguistic coordination and secondary information. By combining controlled behavioral experiments with computational modeling, we can reverse engineer the cognitive mechanisms that enable collectively intelligent behavior to emerge from local exchanges. More broadly, our work contributes to a cross-disciplinary understanding of how low-level linguistic processes scale up to shape the dynamics of collective reasoning and collective behavior.

References

Almaatouq, A., Becker, J., Houghton, J. P., Paton, N., Watts, D. J., & Whiting, M. E. (2021). Empirica: a virtual lab for high-throughput macro-level experiments. *Behavior Research Methods*, *53*(5), 2158–2171.

Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *Journal of conflict resolution*, *41*(2), 203–226.

Becker, J., Brackbill, D., & Centola, D. (2017). Network dynamics of social influence in the wisdom of crowds. *Proceedings of the national academy of sciences*, *114*(26), E5070–E5076.

Bonawitz, E., & Shafto, P. (2016). Computational models of development, social influences. *Current opinion in behavioral sciences*, *7*, 95–100.

Boyce, V., Hawkins, R., Goodman, N., & Frank, M. C. (2022). Two's company but six is a crowd: emergence of conventions in multiparty communication games. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 44).

Centola, D. (2010). The spread of behavior in an online social network experiment. *science*, *329*(5996), 1194–1197.

Centola, D., Becker, J., Zhang, J., Aysola, J., Guilbeault, D., & Khoong, E. (2023). Experimental evidence for structured information–sharing networks reducing medical errors. *Proceedings of the National Academy of Sciences*, *120*(31), e2108290120.

DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, *69*(345), 118–121.

Fränken, J.-P., Valentin, S., Lucas, C. G., & Bramley, N. R. (2024). Naïve information aggregation in human social learning. *Cognition*, *242*, 105633.

Galesic, M., Olsson, H., Dalege, J., van der Does, T., & Stein, D. L. (2021). Integrating social and cognitive aspects of belief dynamics: towards a unifying framework. *Journal of the Royal Society Interface*, *18*(176), 20200857.

Gigerenzer, G., Hertwig, R., Lindsey, S., & Hoffrage, U. (2000). Communicating statistical information. *Science*, *290*(5500), 2261-2262.

Goldstone, R. L., & Janssen, M. A. (2005). Computational models of collective behavior. *Trends in cognitive sciences*, *9*(9), 424–430.

Golub, B., & Jackson, M. O. (2010). Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, *2*(1), 112–149.

Gweon, H. (2021). Inferential social learning: Cognitive foundations of human social learning and teaching. *Trends in Cognitive Sciences*, *25*(10), 896–910.

Hawkins, R. D., Berdahl, A. M., Pentland, A. Tenenbaum, J. B., Goodman, N. D., & Krafft, P. (2023). Flexible social inference facilitates targeted social learning when rewards are not observable. *Nature Human Behaviour*, *7*(10), 1767–1776.

Hegselmann, R., & Krause, U. (2002). Opinion dynamics and bounded confidence models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, *5*(3), 1–2.

Henrich, J. (2016). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. princeton University press.

Moussaïd, M., Herzog, S. M., Kämmer, J. E., & Hertwig, R. (2017). Reach and speed of judgment propagation in the laboratory. *Proceedings of the National Academy of Sciences*, *114*(16), 4117–4122.

Moussaïd, M., Noriega Campero, A., & Almaatouq, A. (2018). Dynamical networks of influence in small group discussions. *PloS one*, *13*(1), e0190541.

Moussaïd, M., Kämmer, J., Analytis, P., & Neth, H. (2013). Social influence and the collective dynamics of opinion formation. *PLoS ONE*, *8*(11).

Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological review*, *97*(3), 362.

Shafto, P., Goodman, N. D., & Frank, M. C. (2012). Learning from others: The consequences of psychological reasoning for human learning. *Perspectives on Psychological Science*, *7*(4), 341–351.

Smaldino, P. (2023). *Modeling social behavior: Mathematical and agent-based models of social dynamics and cultural evolution*. Princeton University Press.

Steiglechner, P., Smaldino, P. E., Moser, D., & Merico, A. (2023). Social identity bias and communication network clustering interact to shape patterns of opinion dynamics. *Journal of the Royal Society Interface*, *20*(209), 20230372.

Tang, T., & Chorus, C. G. (2019). Learning opinions by observing actions: Simulation of opinion dynamics using an action-opinion inference model. *Journal of Artificial Societies and Social Simulation*, *22*(3).

Van Overwalle, F., & Heylighen, F. (2006). Talking nets: A multiagent connectionist approach to communication and trust between individuals. *Psychological Review*, *113*(3), 606.

Whalen, A., Griffiths, T. L., & Buchsbaum, D. (2018). Sensitivity to shared information in social learning. *Cognitive science*, *42*(1), 168–187.