

Examining the Role of AI Information-Sharing on Trust in Human-AI Teams

Proceedings of the Human Factors and Ergonomics Society Annual Meeting
1–6

Copyright © 2025 Human Factors and Ergonomics Society
DOI: 10.1177/10711813251367726
journals.sagepub.com/home/pro



Beau G. Schelble¹, Gracy Filer¹, Christopher Flathmann²,
and Kapil Chalil Madathil²

Abstract

Creating reliable human-AI teams with strong team cognition requires a deeper understanding of how humans perceive and react to various information-sharing tactics employed by an AI teammate. The current study details a factorial survey investigating how participants' trust in their AI and human teammates changes across various forms of AI information sharing. These information-sharing attributes included backup behavior, explainability, team member statuses, intra- and extra-team changes, augmenting team memory, and control. The study found that participants' trust in their AI teammate benefited most from information about changes inside and outside the team, but all conditions outperformed the control. Further, the back-up behavior information negatively affected the participants' trust in their human teammate as the AI corrected the human teammates' actions. These findings demonstrate the utility of a specific form of situation awareness information, while highlighting that corrective back-up behavior information from an AI may harm perceptions between human teammates.

Keywords

trust, human-AI teaming, situation awareness, information sharing, team cognition

Introduction

The current state of artificial intelligence (AI) technology has led researchers to explore the potential for AI to operate alongside humans as teammates. As a teammate, one of the core affordances provided by AI, as opposed to human teammates, is its ability to constantly sense and interpret copious amounts of information surrounding the team (Fui-Hoon Nah et al., 2023). Indeed, AI teammates can process significantly more data than their human counterparts, enabling them to assess a team's environment and make critical decisions rapidly (Westby & Riedl, 2023). However, beyond their decision-making, AI teammates are also in a position to actively use this information to support their human teammates' team cognition through contributions to situational awareness and decision-making. Indeed, whether it is promoting situational awareness, performing backup behaviors, or explaining a team's current environment, AI teammates present numerous ways to utilize their information-gathering and sharing capabilities to benefit teamwork. As human-AI teams become more prevalent, examining the trust between the two partners is crucial to ensure that the teamwork is beneficial.

One construct within the broader context of emergent states in teaming that is crucial for enabling AI teammates to contribute to teaming is trust, defined as one party's

willingness to expose themselves to the actions of another (Mayer et al., 1995). If AI teammates are ever going to become effective contributors to team cognition constructs, such as team or shared situation awareness, shared mental models, and transactive memory systems, then humans must trust the contributions they are actively providing. This need stems from trust influencing teammates' perception of one another and impacting how humans act upon AI insights (Choung et al., 2023). If humans do not trust AI, they may dismiss valuable suggestions, which could be detrimental to team performance (Parasuraman & Riley, 1997). On the other hand, over-trust in AI can lead to overreliance, reduced performance, and a lack of accountability, among other detriments. As such, it is essential to understand the trust dynamics among human-AI teams to ensure optimal team performance. Ensuring optimal performance involves properly calibrating trust in the information-sharing AI teammates engage in to contribute to team cognition.

¹The University of Tennessee, Knoxville, USA

²Clemson University, SC, USA

Corresponding Author:

Beau G. Schelble, Industrial & Systems Engineering, The University of Tennessee, Knoxville, 515 John Tickle Building, 851 Neyland Dr, Knoxville, TN 37996, USA.

Email: bschelbl@utk.edu

Information-sharing is the cornerstone of team cognition development (Cannon-Bowers et al., 1993; Cooke et al., 2013), and AI teammates possess technical advantages that can support and advance team cognition in novel ways. AI teammates also generate numerous additional team cognition and informational needs for effective operation and coordination, introducing concepts such as AI transparency and explainability (Endsley, 2023; Mercado et al., 2016). This need is especially critical when examining AI contributions to team cognition constructs. However, very little existing literature explicitly examines how different forms of information-sharing by an AI teammate affect trust within the human-AI team. Given the known differences between human-AI and human-only teams (Schelble et al., 2022), it is necessary to investigate how these differences manifest to develop more robust and effective collaborative AI systems. As such, the current study is motivated to address the following research questions (RQs):

RQ1: *Do specific types of information that an AI teammate shares affect humans' trust in it?*

RQ2: *Can specific forms of information-sharing made by an AI teammate affect trust between human teammates?*

RQ3: *Does an AI teammate's interpretation of the information it shares affect trust within a human-AI team?*

Examining the above RQs will contribute to the field of human factors and human-AI teaming in three ways. First, the study provides the first evaluation of how different information-sharing strategies influence humans' dispositional trust in their AI teammate. This emphasis thus enables a deeper understanding of the pros and cons of developing AI to contribute to specific concepts within team cognition. Second, studying the effect of AI statements on humans' dispositional trust in their fellow human teammates extends the field's understanding of how AI teammates can influence perceptions between human teammates. Third, the study advances the concept of AI teammates making active contributions to teamwork through interpretation and direction, positioning AI to contribute to leadership, which is a budding area of human-AI teaming research.

Background

Ensuring the effectiveness of information-sharing within human-AI teams for team cognition requires the intelligent design of AI teammates. Team cognition is an umbrella term for an emergent state describing the collective cognition shared among all team members (Cooke et al., 2013; Mohammed et al., 2017). Team cognition encompasses well-known constructs, such as shared/team situation awareness, shared mental models, and transactive memory systems, all constructs that human-AI teams struggle to develop (Endsley, 2023). Developing AI that can leverage its unique advantages to contribute toward, advance, and support team

cognition is likely necessary to address the shortcomings in these constructs that human-AI teams currently face. As such, developing AI to engage in information-sharing that makes these contributions is a potential avenue, as demonstrated in prior AI explainability and transparency studies (Mercado et al., 2016). While numerous forms of information exist for AI teammates to share within their team, such as information about the environment, the team, and progress toward goals, their human teammates may not be fully receptive to certain forms of information-sharing from AI teammates (Choung et al., 2023; Flathmann et al., 2024). Indeed, the massive amount of information obtained by AI teammates could quickly overwhelm human teammates, even if it is shared to support the completion of their goals (Cuevas et al., 2007). Furthermore, humans may have varying levels of trust in certain types of information-sharing from AI teammates, as they may perceive AI as less capable of sensing and sharing specific kinds of information. In turn, it may inevitably be a detriment if humans are not properly calibrated to this AI information-sharing (Parasuraman & Riley, 1997). While trust is a highly studied topic in human factors and human-AI teaming specifically, it is necessary to consider when designing AI systems to function as teammates actively contributing to a complex teaming construct, such as team cognition. Thus, empirical research is needed to understand what information-sharing from AI teammates is trusted by humans and the impact that information-sharing may have on trust.

Based on this gap, this paper reports on an empirical study that quantifies the impact of intentional AI information sharing on perceived trust in AI and human teammates. In particular, this study examines the effects of six information-sharing methods commonly employed in human teams when presented by an AI teammate on trust. Further, this study explores the benefits of AI in providing interpretation and information on both trust outcomes. By examining the trust in AI and human teammates, this study can identify which information-sharing methods humans are most receptive to from AI teammates and their effect on human-human trust.

Methods

The current study employed an online factorial survey, a methodology that involves experimental manipulations by modifying text or video-based scenario descriptions (Jasso, 2006; Li et al., 2023). The factorial survey method has seen significant success over the last decade, as it enables the collection of perceptions of human-AI interaction in response to complex or challenging situations (Li et al., 2023). It has also achieved explicit success in the field of human-AI teaming across various application spaces (Flathmann et al., 2024; Schelble et al., 2025). Utilizing this approach. The current study was designed to manipulate the information shared by the AI teammate and determine whether the AI teammate

provided an interpretation of the information. This resulted in a mixed factorial 6 (AI Information-Sharing: situation awareness of Team Member Information, situation awareness of Intra/Extra Team Information Changes, Back-Up Behavior Information, Augmenting Team Memory Information, Explainability Information, Control) \times 2 (AI Interpretation: Yes, No) with AI interpretation conducted between-subjects and AI information-sharing type conducted within-subjects.

Participants

An a priori power analysis was conducted ($\beta = .85$, $\eta_p^2 = .10$) using the study's mixed 6×2 design in GPower 3.1 (Faul et al., 2009) to ensure adequate power was achieved. This power analysis indicated that at least 139 participants were necessary to run the analysis effectively. As such, the study recruited a total of 173 participants; however, 22 participants returned an incomplete survey, and one participant failed at least two out of the four attention checks. These participants were excluded from the data set, resulting in unequal cell sizes. To address this, participants were recruited randomly until the cell sizes were equal. This resulted in a final data set comprising 156 participants, with an average age of 32.28 ($SD = 9.06$). Of the 156 participants, 121 self-reported as male, 29 as female, five as non-binary, and one chose not to disclose. All participants were recruited from the Prolific online research recruitment platform, which offers monetary incentives to individuals in return for completing research studies (Palan & Schitter, 2018). Participants were unable to participate unless they self-reported at least some weekly experience with video games to ensure familiarity with team-based settings alongside AI. The survey took an average of 20 min, and participants were compensated \$8.00 per hour for their time.

Human-AI Teaming Scenario Vignettes and Manipulations

Utilizing the factorial survey approach, the current study developed six text-based scenarios of a human-AI team consisting of the participant (Teammate A), one additional human teammate (Teammate B), and one AI teammate (Teammate C) playing a paintball eSports game against an opposing team. Specifically, the scenario was a game of capture the flag between two teams of three individuals that read as follows: *“For the rest of this survey, you will be shown multiple scenarios and asked questions about each scenario. In the following scenarios, you will be a member of a human-AI team playing an online paintball capture the flag video game, and you will be asked about how six different AI teammates and their information-sharing affect your perceptions of your team and the situation described. Specifically, capture the flag is where two teams each have a flag located in their home base, and the objective is to*

steal the other team's flag and bring it safely back to your base. Players can be knocked out of the game if they are tagged with a paintball fired by the opposing team. You and your two teammates must go up against three other players, successfully get past their defenses, steal the opposing team's flag, and then return it to your team's base without being eliminated by enemy paintballs.” This scenario has also been used in prior related human-AI teaming studies (Schelble, 2023; Schelble et al., 2025).

Participants would read six different vignettes set within this eSports paintball capture the flag scenario, as each vignette applied to a different AI information-sharing condition. The AI information-sharing manipulation consisted of information relevant to the eSports paintball capture the flag scenario and was specific to the six levels of the manipulation. The six manipulations covered the following AI information-sharing strategies: (1) the situation awareness of team members' information-sharing condition communicated the location of each teammate and the amount of resources they possessed, (2) situation awareness of intra/extra team information communicated what happened to teammates and what happened to the opposing team, (3) back-up behavior conveyed how Teammate B might improve their performance within the team based on past actions, (4) augmenting team memory communicated each teammate's specific skills, provided a map, and reminded the team of time remaining, (5) explainability communicated the AI teammates' reasoning for suggesting a particular course of action, (6) the control level provided no extra information. The order of the six within-subjects levels was presented randomly, and the names used for the AI teammate and Teammate B were changed across all six vignettes. The names for the AI were Greek alphabet letters (i.e., Sigma, Iota), and the human teammate names were unisex (i.e., Chandler, Logan) to help control for any potential confound.

The between-subjects manipulation of AI interpretation was applied across all six vignettes, and the AI made an interpretation and provided direction to the team. The interpretation given by the AI teammate was as follows: *“You and I will provide cover and support to Chandler (Teammate B) while they move forward for the flag.”* In this context, the interpretation manipulation involved the AI teammate suggesting the best way to coordinate based on the information it shared with the team (the AI's information-sharing attribute).

Procedure

Participants who chose to participate in the study were first presented with an informed consent document that they had to read and agree to before proceeding with the survey. If they decided to participate in the study after reading the informed consent document, they answered a series of demographic questions and were randomly assigned to one of the

between-subjects conditions. After completing the initial series of questionnaires, the eSports paintball capture the flag scenario was described to them before viewing the first text-based vignette. Furthermore, before each text-based vignette, participants were provided with a description of the information-sharing attribute of the AI they were about to work with, ensuring they all had a consistent understanding of the AI's information-sharing attribute before reading the associated vignette. After reading each vignette, participants rated their perceived trust in the human and the AI teammate. Participants completed this process for six vignettes, and once all six were completed, the study was finished. The study took participants about 20 min.

Measures

Trust in the AI Teammate. Dispositional trust in the AI teammate was elicited using a single-item measure evaluated on a seven-point Likert scale ranging from “Strongly Disagree” to “Strongly Agree.” The item read “*I trust Alpha.*” The name of the AI teammate (shown in bold) was changed to match the name of the AI teammate in the relevant vignette.

Trust in the Human Teammate. Dispositional trust in the human teammate was also elicited using a single-item measure evaluated on the same seven-point Likert scale ranging from “Strongly Disagree” to “Strongly Agree.” The item read “*I trust Alex.*” The name of the human teammate (Teammate B, shown bolded) was changed to match the teammate's name in the relevant vignette.

Results

Trust in the Human Teammate

A 6×2 mixed ANOVA was conducted to assess the effect of AI interpretation (between-subjects) and information-sharing type (within-subjects) on participants' trust in their human teammate. The main effect of information-sharing type on participants' trust in their human teammate was found to be significant ($F(5, 924) = 34.99, p < .001, \eta_p^2 = .16$; Figure 1). All assumptions for ANOVA tests were checked and met. Post-hoc tests using Tukey's HSD correction revealed that the participants' trust in their human teammate with the situation awareness of team members' AI ($M = 4.98, SE = 0.10$) was significantly higher than their trust in the back-up behavior AI ($M = 3.94, SE = 0.10$).

The situation awareness of intra/extra team information AI condition ($M = 5.28, SE = 0.10$) also had significantly higher trust in their human teammates than the back-up behavior AI condition. Furthermore, the back-up behavior AI condition had significantly lower trust in the human teammate than the augmenting team memory AI ($M = 5.35, SE = 0.10$), the explainability AI ($M = 5.18, SE = 0.10$), and the control AI ($M = 4.94, SE = 0.10$). Lastly, the augmenting team memory

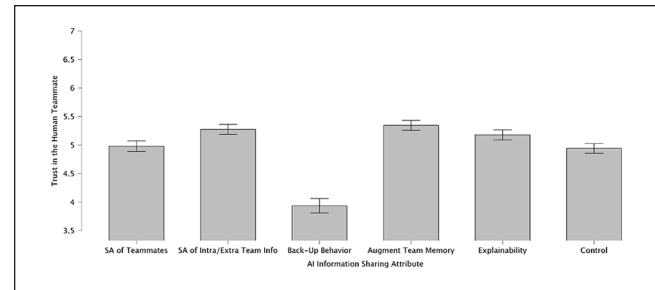


Figure 1. Trust in the human teammate. Error bars represent standard error.

AI saw participants trust their human teammate significantly more than the control AI.

Trust in the AI Teammate

A 6×2 mixed ANOVA was conducted to assess the effect of AI interpretation (between-subjects) and information-sharing type (within-subjects) on participants' trust in their AI teammate. The main effect of information-sharing type on participants' trust in the AI teammate was significant ($F(5, 924) = 30.68, p < .001, \eta_p^2 = .14$; Figure 2). All assumptions for ANOVA tests were checked and met. Tukey's HSD post-hoc tests found that the situation awareness of team members AI ($M = 4.97, SE = 0.10$) engendered significantly less trust from participants than the situation awareness of intra/extra team information AI ($M = 5.57, SE = 0.10$) and the explainability AI ($M = 5.63, SE = 0.10$). However, participants still trusted it more than the control AI ($M = 4.14, SE = 0.10$).

The situation awareness of intra/extra team information AI had significantly higher levels of participant trust than the back-up behavior AI ($M = 5.09, SE = 0.10$) and the control AI. The back-up behavior AI was trusted significantly less by participants than the explainability AI, but more than the control AI. Lastly, the augmenting team memory AI ($M = 5.36, SE = 0.10$) and the explainability AI were both trusted significantly more than the control AI.

Discussion

These analyses demonstrate that some information-sharing from an AI teammate was more effective than none in terms of trust in the AI teammate, as indicated by the control AI condition. Additionally, the effect size on trust in the AI teammate was large, suggesting that when implementing AI with information-sharing features that *effectively* contribute toward situation awareness, potential developers will likely enhance trust in the AI teammate. Alternatively, the measure of how participants felt toward their human teammate was also significantly influenced by the information-sharing attributes of the AI, with trust toward the human teammate dropping significantly when the AI teammate provided

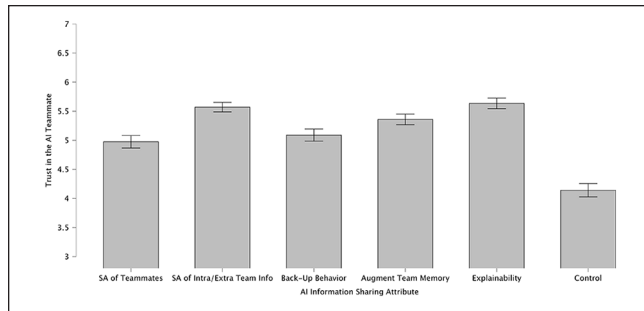


Figure 2. Trust in the AI teammate. Error bars represent standard error.

corrective backup behavior information directed toward the other human teammate (Teammate B). Additionally, it was found that the augmenting team memory AI also increased participants' trust in the other human teammate, possibly due to the participants seeing increased utility from their human teammate with the extra information provided by the AI teammate.

The first major takeaway from the current study is that certain information shared by AI teammates can alter human teammates' perceptions of their fellow human teammates. As seen in the analysis of participants' trust in the human teammate, the information shared by the back-up behavior AI had a surprisingly strong adverse effect on participants' perceived attitude toward their fellow human teammate. The AI teammate provided corrective back-up behavior information in that specific vignette, advising the teammate to wait for covering fire before moving forward with the team next time (stating they had not in the past). As such, AI designers should be cautious about what their AI teammates say to human teammates in an open setting (possibly using private direct chats with that specific teammate instead), and future research is necessary to determine the extent of this effect. The second critical finding was the importance of explainability and situational awareness of intra- and extra-team information changes to participants' perceptions of trust in the AI teammate. These findings align with previous human-AI interaction literature (Shin, 2021) and extend it to human-AI teaming. Information on AI explainability and environmental updates appears to address a unique need within human-AI teams, as situational awareness information may help mitigate uncertainties associated with AI collaboration (Endsley, 2023).

Limitations and Future Work

Some limitations of the study should be considered when interpreting the results of the current study. Specifically, the scenario was conducted using a low-fidelity text-based vignette, which means the results may not readily generalize to physical scenarios. As such, future work should aim to conduct similar studies examining AI information-sharing

attributes in both real-world and simulated laboratory settings. Furthermore, the study did not find any effects of AI interpretation, which could be due to the interpretation provided by the AI not being impactful enough. Any future work studying AI support for decision-making and leadership should continue to explore how information interpretation by AI in support of decision-making influences trust, especially between human teammates.

Conclusion

The current study details how specific AI information-sharing attributes can significantly affect humans' trust in them and their fellow human teammates. AI information-sharing was also a significant improvement in all cases for trust in the AI, demonstrating the importance of developing future collaborative AI systems that can contribute meaningfully to team cognition. Given the current struggles in supporting team cognition and complex coordination in these human-AI teams, it is more necessary than ever to leverage the inherent strengths and technical advantages that AI systems bring to the table. As such, the current study demonstrates that AI teammates can make these contributions, and human teammates do trust them. However, there are caveats as information related to explainability and situation awareness outperforms other forms of information. Further, backup behavior information about the human teammate's mistakes was detrimental to the trust between human teammates. This information should be considered when developing collaborative AI systems to enhance the team cognition of human-AI teams.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Beau G. Schelble  <https://orcid.org/0000-0003-3704-697X>

Christopher Flathmann  <https://orcid.org/0000-0002-5448-2610>

References

- Cannon-Bowers, J., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N. J. Castellan (Ed.), *Individual and group decision making: Current issues* (pp. 221–246). Lawrence Erlbaum Associates, Inc.
- Choung, H., David, P., & Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human-Computer Interaction*, 39(9), 1727–1739. <https://doi.org/10.1080/10447318.2022.2050543>
- Cooke, N. J., Gorman, J. C., Myers, C. W., & Duran, J. L. (2013). Interactive team cognition. *Cognitive Science*, 37(2), 255–285.

- Cuevas, H. M., Fiore, S. M., Caldwell, B. S., & Strater, L. (2007). Augmenting team cognition in human-automation teams performing in complex operational environments. *Aviation, Space, and Environmental Medicine*, *78*(5 Suppl), B63–B70.
- Endsley, M. R. (2023). Supporting human-AI teams: Transparency, explainability, and situation awareness. *Computers in Human Behavior*, *140*, 107574.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Flathmann, C., Schelble, B. G., McNeese, N. J., Knijnenburg, B., Gramopadhye, A. K., & Chalil Madathil, K. (2024). The purposeful presentation of AI teammates: Impacts on human acceptance and perception. *International Journal of Human-Computer Interaction*, *40*(20), 6510–6527. <https://doi.org/10.1080/10447318.2023.2254984>
- Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, *25*(3), 277–304. <https://doi.org/10.1080/15228053.2023.2233814>
- Jasso, G. (2006). Factorial survey methods for studying beliefs and judgments. *Sociological Methods & Research*, *34*(3), 334–423. <https://doi.org/10.1177/0049124105283121>
- Li, T., Vorvoreanu, M., Debellis, D., & Amershi, S. (2023). Assessing human-AI interaction early through factorial surveys: A study on the guidelines for human-AI interaction. *ACM Transactions on Computer Human Interaction*, *30*(5), 1–69:45. <https://doi.org/10.1145/3511605>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709–734. <https://doi.org/10.2307/258792>
- Mercado, J. E., Rupp, M. A., Chen, J. Y. C., Barnes, M. J., Barber, D., & Procci, K. (2016). Intelligent agent transparency in human-agent teaming for multi-UxV management. *Human Factors*, *58*(3), 401–415. <https://doi.org/10.1177/0018720815621206>
- Mohammed, S., Hamilton, K., Sánchez-Manzanares, M., & Rico, R. (2017). Team cognition: Team mental models and situation awareness. In E. Salas, R. Rico, & J. Passmore (Eds.), *The Wiley Blackwell handbook of the psychology of team working and collaborative processes* (pp. 369–392). Wiley.
- Palan, S., & Schitter, C. (2018). Prolific. Ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27.
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*(2), 230–253.
- Schelble, B. G. (2023). *Leveraging artificial intelligence for team cognition in human-AI teams* [PhD Thesis, Clemson University]. <https://search.proquest.com/openview/ea8359f17363790a96591fd4693d7c5d/1?pq-origsite=gscholar&cbl=18750&diss=y>
- Schelble, B. G., Flathmann, C., Macdonald, J. P., Knijnenburg, B., Brady, C., & McNeese, N. J. (2025). Modeling perceived information needs in human-AI teams: Improving AI teammate utility and driving team cognition. *Behaviour & Information Technology*, *44*, 2069–2092. <https://doi.org/10.1080/0144929x.2024.2396476>
- Schelble, B. G., Flathmann, C., Musick, G., McNeese, N. J., & Freeman, G. (2022). I see you: Examining the role of spatial information in human-agent teams. *Proceedings of the ACM on Human-Computer Interaction*, *6*(CSCW2), 1–27.
- Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, *146*, 102551.
- Westby, S., & Riedl, C. (2023). Collective intelligence in human-AI teams: A Bayesian theory of mind approach. *Proceedings of the AAAI Conference on Artificial Intelligence*, *37*(5), 6119–6127. <https://ojs.aaai.org/index.php/AAAI/article/view/25755>