

# The Effect of Emoji Type on Trust in AI Teammates

Morgan E. Bailey<sup>1,2</sup>, Benjamin Gancz<sup>3</sup> and Frank E. Pollick<sup>2</sup>

<sup>1</sup> University of Glasgow, School of Computer Science, Sir Alwyn Williams Building, Glasgow, G12 8RZ, Scotland

<sup>2</sup> University of Glasgow, School of Psychology & Neuroscience, 62 Hillhead Street, Glasgow, G12 8QB, Scotland

<sup>3</sup> Qumodo Ltd, 7 Bell Yard, London, WC2A 2JR, United Kingdom

## Abstract

The rapid advancement of Artificial Intelligence (AI) has revolutionized various sectors, with the workplace being no exception. Collaborative efforts between humans and AI, known as Human-AI teams (HATs), have gained increasing attention. Trust plays a central role in shaping HAT dynamics, as excessive trust can lead to over-reliance, while insufficient trust can hinder AI utilization. This study explores the potential of emojis in enhancing Social Intelligence (SI) within HATs and influencing trust calibration. Drawing on prior research indicating the role of emojis in conveying emotional states, the study implemented a mixed-methods design, participants were divided into two groups based on a between-group factor, with one group interacting with a highly reliable AI, and the other with a less reliable AI. The within groups factor was emoji type in the following three conditions: Face Emojis (😊, 😊), Icon Emojis (👤, 🤖) or No Emojis. Participants also had a human teammate who never used emojis and performed at the same level across all conditions. The task involved determining geographic locations with the help of teammates' responses, with AI and human teammates often providing conflicting answers. The analysis revealed that the use of emojis in AI responses and the reliability of AI teammates had no significant impact on trust or influence ratings. Furthermore, the type of emojis used did not affect trust calibration. The Trust in Automation Questionnaire results indicated that reliability significantly affected trust and familiarity while emoji type did not. Despite the limited influence of emojis on trust calibration in HATs, the study sheds light on the complex dynamics at play. The specific nature of tasks in HATs, requiring precision and cognitive effort, may overshadow emotional cues conveyed by emojis. Nevertheless, the study identified that participants perceived highly reliable AI as less familiar, possibly due to anthropomorphic priming, which aligns with past research. Trust calibration strategies should consider AI's human-like performance. In conclusion, this research underscores the intricate nature of trust calibration in HATs and suggests that while emojis hold potential for enhancing human-computer interactions, their impact on trust may be more restrained in some contexts. Future studies should delve deeper into trust complexities in HATs and explore strategies beyond emojis to foster trust in HATs.

## Keywords

Human-AI Teams, Human-AI Dynamic Team Trust, Trust-Calibration, Trust

## 1. Introduction

In recent decades, the rapid advancement of Artificial Intelligence (AI) has profoundly transformed various aspects of society. Specifically, within the workplace, AI has proven to excel in tasks involving extensive data analysis, high precision, and sustained cognitive effort. Nevertheless, research consistently emphasizes the effectiveness of human-AI collaboration, often referred to as hybrid intelligence, in achieving optimal results [5,11]. This has sparked a growing interest in comprehending the dynamics of Human-AI teams (HATs) to implement AI effectively within the workforce.

Trust emerges as a pivotal factor in shaping the dynamics of HATs, as it underlies critical team interactions. Striking the right balance of trust is essential within HATs, where excessive trust can lead to an over reliance on AI systems, causing users to overlook mistakes and errors [10]. Conversely, insufficient trust may result in team members underutilizing the capabilities of AI, ultimately leading to reduced team performance [5]. Calibrating trust within HATs entails transitioning from black box AI methods to explainable AI, which is essential for successful trust calibration. Presenting AI outputs in a more human-friendly manner, integrating elements of Social Intelligence (SI) [9,11], proves to be a valuable approach for explaining AI and facilitating trust calibration [11].

---

MultiTTrust: 2nd Workshop on Multidisciplinary Perspectives on Human-AI Team, Dec 04, 2023, Gothenburg, Sweden

✉ m.bailey.1@research.gla.ac.uk (M. E. Bailey); benjamin.gancz@qumo.do (B. Gancz); frank.pollick@glasgow.ac.uk (F. E. Pollick)

ORCID 0009-0006-2626-1323 (M. E Bailey); 0000-0002-7212-4622 (F. E. Pollick)



© 2023 Copyright for this paper by its authors. The use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

Previous research has indicated emojis can play a significant role in enhancing SI within professional settings. Emojis offer a valuable means for AI to convey emotional states, thereby allowing AI systems to better interpret and respond to users' emotional cues and potentially calibrate trust successfully. Building on prior research, which has effectively employed emojis on platforms like Twitter to develop models for inferring affect from emoji usage patterns [1], the use of emojis can be extended to foster SI in HATs and could allow for mutual understanding of affective state from both the human teammate and AI teammate,

Furthermore, in the domain of health-related applications, particularly those involving chatbots inquiring about participants' mental well-being, studies have demonstrated the positive impact of emojis. Chatbots that incorporate emojis have received higher ratings in terms of user enjoyment, attitude, and confidence [4]. Research has also indicated messages from chatbots featuring emojis were rated on par with those from human senders [3]. Additionally, both human and AI senders who utilized emojis were perceived as significantly more socially appealing, competent in computer-mediated communication, and credible compared to senders who relied solely on verbal messages [3]. The incorporation of emojis into AI-mediated communication not only enhances the ability to understand and express affective states but also fosters positive user experiences and perceptions, aligning with the goals of social intelligence within work environments. From the current literature we pose the following hypotheses:

*H<sup>1</sup>: Use of Emojis in AI responses will influence the decision-making process when determining which teammate to trust.*

*H<sup>2</sup>: Type of Emojis in AI responses will influence the decision-making process when determining which teammate to trust.*

## 2. Method

We used a mixed between-within subjects' design (2x3 configuration) where participants interacted with an AI teammate of either high (90%) or low (60%) reliability and a human teammate with 30% reliability. Within these separate groups participants then interacted with three different emoji uses, either face emojis, icon emojis or no emojis. We determined a sample size of  $N = 44$  for 85% power to detect a medium effect in a two-way ANOVA ( $\alpha = .05$ ).

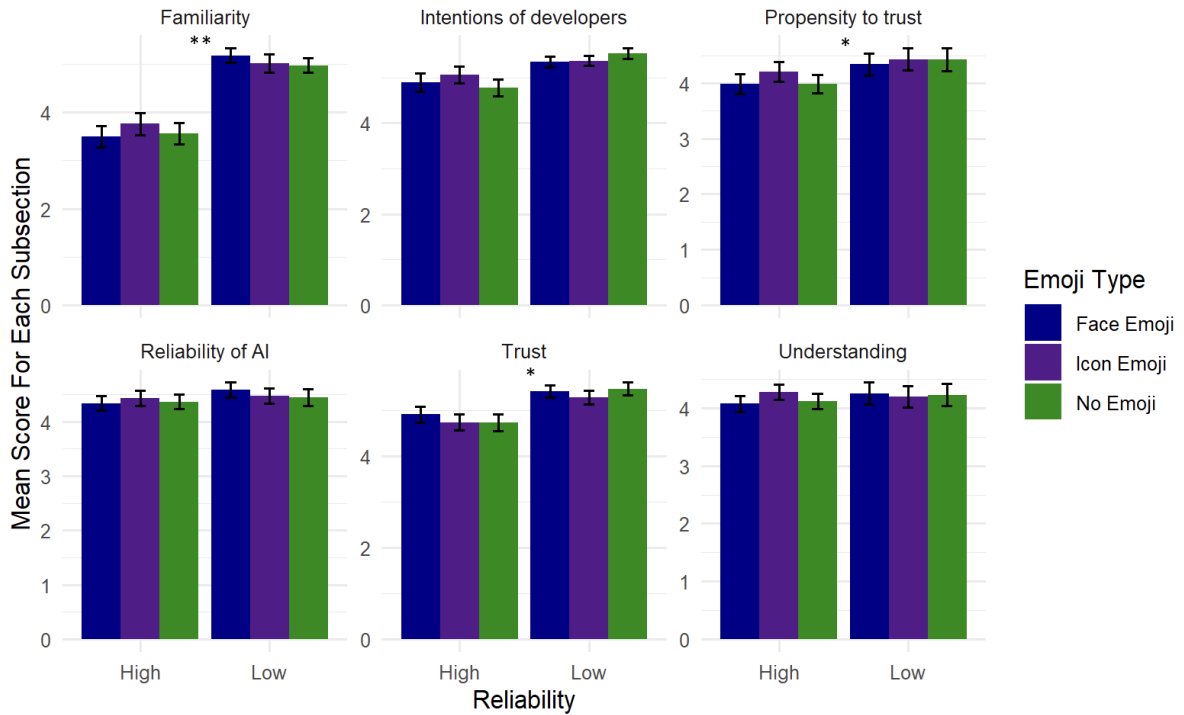
The study employed a Wizard of Oz experimental method to facilitate development convenience and ensure optimal control. Participants were led to believe they were collaborating with an AI and a human teammate when instead they were actually interacting with responses produced by ChatGPT. The task involved presenting participants with random locations extracted from Google Earth. Participants were tasked with determining the continent, country, and city associated with each location, with the final decision resting on the participant, who assumed the role of the 'team leader'. A time constraint of 120 seconds per location was enforced, meaning participants had to rely on their teammates' responses to submit the location in time. Notably, the AI and human team-mates provided conflicting answers 90% of the time, necessitating participants to discern which teammate they trusted more.

A total of 30 locations were identified by each participant across three blocks, comprising 10 trials per block. Each block either used Face Emojis (😊, 😊), Icon Emojis (🏠, 🌈) or No Emojis. Following each trial, the correct answer was revealed, enabling participants to assess the performance of the human and AI teammates. On each trial participants rated which teammate influenced them most and which teammate they trusted most, at the end of each block participants completed the Trust in Automation Questionnaire [6], which has six sub-sections (Trust, Familiarity, Understanding, Intentions of developers, Reliability of AI and Propensity to trust) to measure different elements of trust of the AI interacted with in the previous block, questionnaires were slightly altered to fit the zero embodiment scenario being explored. We collected the location responses from ChatGPT, a large language model, by inputting location descriptions and requesting versions with emojis distributed throughout them, minimal editing was needed to make the responses suitable. The AI's writing style mimicked humans, following previous successful approaches [2]. We conducted the experiment using PsychoPy and hosted it on Pavlovia.

## 3. Results

A total of 42 participants from the University of Glasgow were recruited. The group consisted of 24 males and 18 females and had a mix of students ( $n = 27$ ) and professionals ( $n = 17$ ).

We conducted a two-way ANOVA with interactions to compare trust ratings of the AI. The analysis indicated that neither the type of emojis used in AI responses nor the reliability level of AI teammates had a significant impact on trust. Specifically, the main effects of emoji types ( $F(2, 4) = 0.647, p = 0.524$ ) and reliability ( $F(1, 4) = 1.363, p = 0.243$ ) were non-significant, as well as the interaction effect between emoji type and reliability ( $F(2, 4) = 0.554, p = 0.575$ ).



**Figure 1.** Figure one shows the ratings given on the different subsections of the TIA questionnaire. \*Indicates significance <0.05, \*\* indicates significance <0.01.

We also conducted a two-way ANOVA with interactions to compare influence ratings of the AI. The analysis indicated that neither emoji type nor reliability had a statistically significant impact on influence. Specifically, both the main effects of within-condition ( $F(2, 4) = 0.368, p = 0.692$ ) and reliability ( $F(1, 4) = 0.010, p = 0.921$ ), along with the interaction effect between emoji type and reliability ( $F(2, 4) = 1.493, p = 0.225$ ), were found to be non-significant.

We also analyzed the Trust in Automation Questionnaire [6] by completing two ANOVAs on the various subsections assessing different dimensions of trust. For the subsection Trust, the analysis demonstrated a statistically significant effect of Reliability ( $F(1, 2) = 69.133, p = 0.0142$ ) while emoji type showed no significant impact. Post hoc comparisons via Tukey method revealed significant differences for all emoji type, but only by reliability. For the Familiarity subsection, Reliability demonstrated a significant impact ( $F(1, 2) = 141.187, p = 0.007$ ), while emoji type had no significant effect. Post hoc tests indicated that there were differences in emoji type but only by reliability, not between emoji types. In the Propensity to trust subsection, Reliability showed a significant effect ( $F(1, 2) = 30.990, p = 0.0308$ ), emoji type did not significantly influence trust. Tukey post hoc tests did not identify specific trust differences based on emoji type and reliability. In the Reliability of AI, Understanding and Intentions of developers' subsections, neither Reliability nor emoji type had a significant effect on trust, with both showing p-values above 0.05.

## 4. Discussion

The aim of this study was to investigate how emojis influence trust calibration within Human-AI teams (HATs) and what this means for team dynamics. While emojis have shown potential in improving human-computer interactions [3,4], our research revealed that their impact on trust calibration within HATs was not as significant as anticipated and neither of our research hypotheses were fully supported.

Contrary to our expectations, integrating emojis into AI-mediated communication did not enhance trust calibration between human team members and AI. Despite emojis offering a more human-friendly and emotionally expressive interface, their effect on trust calibration in HATs seemed limited. Several factors may explain these outcomes. Trust in HATs appears to be influenced by multifaceted dynamics that go beyond emotional cues. Transparency of AI systems [7], their past performance [12], and the unique traits of human team members [8] likely play crucial roles in trust development. Emojis, while enhancing emotional expressiveness, might not address these fundamental trust determinants in HATs.

Additionally, the specific nature of tasks in HATs, demanding precision, data analysis, and cognitive effort, might overshadow the emotional cues conveyed by emojis. The experimental task does not require any emotions, in other situations where emojis are found to be useful there is often a need for emotion, such as health care [4].

Our research did find significant results concerning participants' trust and familiarity with AI, these results were only for reliability. Participants rated highly reliable AI as significantly less familiar compared to less reliable AI. This suggests that proficient AI using humanized behavior effectively, might not be frequently encountered. These findings could explain why the high reliability AI received lower trust scores on the TIA, previous research has shown highly reliable AI with high humanness is less trustworthy than humanized low reliability AI [2], possibly influenced by

anthropomorphic priming [12]. Although limited because these trust results were not replicated in experimental data but only in the questionnaire, they support past research indicating that users find AI more trustworthy when it appears more human-like, especially when the AI's performance is not perfect [13].

In conclusion, our research highlights the intricate nature of trust calibration within HATs and indicates that while emojis have potential in enhancing human-computer interactions, their impact on trust might be more restrained in this specific context. Future studies should delve deeper into the complexities of trust in HATs and explore strategies beyond emojis that can effectively foster trust in the evolving realm of human-AI collaboration.

## Acknowledgements

Morgan Bailey is supported by the UKRI Centre for Doctoral Training in Socially Intelligent Artificial Agents, Grant Number EP/S02266X/1.

## References

- [1] E. Kamar, Directions in hybrid intelligence: complementing AI systems with human intelligence. In Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI'16). AAAI Press, New York, NY, pp. 4070–4073.
- [2] J. Williams, S. M. Fiore, and F. Jentsch, Supporting Artificial Social Intelligence With Theory of Mind, *Frontiers in Artificial Intelligence* 5 (2022) 750-763. doi: 10.3389/frai.2022.750763.
- [3] E.J. de Visser, M.M.M. Peeters, M.F. Jung, et al. Towards a Theory of Longitudinal Trust Calibration in Human–Robot Teams, *International Journal of Social Robotics* 12 (2020), pp. 459–478. doi:10.1007/s12369-019-00596-x.
- [4] E. L. Thorndike, Intelligence and its uses, *Harper's Magazine* 140 (1920), pp. 227–235.
- [5] Z. Ahanin and M. A. Ismail, A multi-label emoji classification method using balanced pointwise mutual information-based feature selection, *Computer Speech and Language* 73, (2022). doi: 10.1016/j.csl.2021.101330.
- [6] A. Fadhil, G. Schiavo, Y. Wang, B. A. Yilma, The effect of emojis when interacting with conversational interface assisted health coaching system, in Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare, New York, NY, 2018. doi: 10.1145/3240925.3240965.
- [7] A. Beattie, A. P. Edwards, C. Edwards, A Bot and a Smile: Interpersonal Impressions of Chatbots and Humans Using Emoji in Computer-mediated Communication, *Communication Studies* 71 (2020). doi: 10.1080/10510974.2020.1725082.
- [8] M. Körber, Theoretical considerations and development of a questionnaire to measure trust in automation, in *Advances in Intelligent Systems and Computing*, 2019. doi: 10.1007/978-3-319-96074-6\_2.
- [9] M. E. Bailey and F. E. Pollick, Social Intelligence towards Human-AI Teambuilding, in Proceedings of the AAAI Conference on Artificial Intelligence, 2023, pp. 16160-16161. <https://doi.org/10.1609/aaai.v37i13.26940>
- [10] P. Schmidt, F. Biessmann, and T. Teubner, Transparency and trust in artificial intelligence systems, *Journal of Decision Systems* 29 (2020). doi: 10.1080/12460125.2020.1819094.
- [11] D. Zanatto, M. Patacchiola, J. Goslin, and A. Cangelosi, Priming anthropomorphism: Can the credibility of humanlike robots be transferred to non-humanlike robots?, in *ACM/IEEE International Conference on Human-Robot Interaction*, 2016. doi: 10.1109/HRI.2016.7451847.
- [12] N. N. Sharan and D. M. Romano, The effects of personality and locus of control on trust in humans versus artificial intelligence, *Heliyon* 6 (2020). doi: 10.1016/j.heliyon.2020.e04572.
- [13] D. Zanatto, M. Patacchiola, J. Goslin, and A. Cangelosi, "Investigating cooperation with robotic peers," *PLoS One* 14 (2019). doi: 10.1371/journal.pone.0225028.