# A Comparative Evaluation of Formed Team Perception when in Human-Human and Human-Autonomous Teams

Chandni Murmu
cmurmu@clemson.edu
Clemson University
Clemson, South Carolina, USA

Gnaanavarun Parthiban
Clemson University
Clemson, USA

Konnor McDowell
Clemson University
Clemson, USA

Nathan McNeese
Clemson University
Clemson, USA

## ABSTRACT

In recent years, there has been a surge in studies on Human and Autonomous Agent (AA) teams (HAT) within Human-Computer Interaction (HCI). However, the current literature lacks unbiased evaluations of applied AA or AI teammates compared to human counterparts in HAT settings. Existing evaluations are often influenced by participants' prior experiences and expectations, rather than providing a current assessment. To address this gap, we conducted a single-blind preliminary study, assessing the perceptions of 10 participants in Human-Human and Human-AA teams using the Paladins Multiplayer Online Games (MPOG) platform.

## CCS CONCEPTS

• **Human-centered computing → Collaborative and social computing**.

## KEYWORDS

Human-AA team, Human-AI team, Human-computer interaction, Formed Perception, Autonomous Agent, AA Teammate

**ACM Reference Format:**
Chandni Murmu, Gnaanavarun Parthiban, Konnor McDowell, and Nathan McNeese. 2018. A Comparative Evaluation of Formed Team Perception when in Human-Human and Human-Autonomous Teams. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 3 pages. https://doi.org/XXXXXXX.XXXXXXX

## 1 INTRODUCTION

In the past few years, the HCI community has seen a massive increase in the development and implementation of AI or AA as teammates in HATs [11, 19, 27, 29, 30]. In these collaborative applications, the teams are often defined as consisting of at least one human cooperating with at least one AA [13]. The AA in such teams is a computer entity having a partial or high level of self-governance in terms of decision-making, adaptation, and communication [5, 15, 17]. It also performs shared tasks to achieve the same valued goals [3, 13, 22]. Existing perception studies, involving AI teammates, have focused on expectations [32], receptiveness and acceptance [1, 6, 24], expertise level [31], team cognition [14, 25], preference of learned or rule based agent [28], adaptive capabilities of AI affecting performance and cohesiveness [8], decision making strategy affecting team-efficacy [16], trust [2, 9], trust and ethics [26] and others.

In these studies, except [28], the participants are aware of their teammate being an AA [1, 14, 16, 25, 31] or surveys and interviews are conducted to collect thoughts on past and future aspects of an AA as a teammate [2, 6, 8, 9, 24, 26]. In both the scenarios, the studies and their outcomes are immensely influenced by the participants' perceptual history of what they thought AA teammates were and their future expectations on how they are supposed to be [6, 24, 32]. They contribute in paving the path for development and enhancement of future AA or AI teammates. However, the contemporary literature does not address the evaluation of the already applied AA or AI teammates when compared to a human teammate in the MPOG, or other areas. Further, any such comparisons made are biased by participants' past perception and future expectations of AA or AI teammates rather than the participants giving opinion of the current application. This exhibits a shortcoming in validation of the progress of application of AA or AI as a teammate in comparison to the human teammate.

## 2 PROPOSED STUDY DESIGN

We designed a within subject study using Paladins, a first person shooting (FPS) game and a semi-popular MPOG, to simulate two sessions of dyads of human-human (HH) teams and human-AA (HA) teams with participants being unaware of who or what their teammates were. Paladins' inbuilt AA was used for the experiment. Paladins is played through representation of players as avatar, which support in activating embodiment and presence psychology to facilitate trust [21] and also enables the AA teammate to be viewed as a team member and not as a tool [20] because it embodies a similar avatarized form as the human participants, exhibits interdependence [18] and can independently pursue courses of action through self governance [12]. Since, the participants are not aware or informed of who or what their teammates were, the opinions on their team, teammates and themselves would pertain purely on the game experience with their teammate.


*Conference acronym 'XX, June 03–05, 2018, Woodstock, NY*

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

## 2.1 Experiment Environment

We used Paladins for our experiment, creating four unique Steam accounts for gameplay. Each team had two selected hero avatars: Viktor and Ash. Custom game sessions were initiated using the "PLAY-CUSTOM" option with the "Onslaught" game type, chosen for its simplicity. Team size was set at "2 v 2" in the North American region, with "Open Draft" draft mode, no spectators, and disabled "Store" and "Loadouts" features. We selected the "Foreman's Rise" arena for its simplicity. A password was assigned to each session. To introduce the autonomous agent, researchers joined the game session briefly before exiting, enabling the level 10 autonomous agent to take over.

## 2.2 Experiment Design

Our within-subject experiment involved a 10-minute tutorial in Paladins for all participants, followed by two 10-minute gameplay sessions, or until one team collected all tickets to win. Participants were unaware of their teammate's identity in both sessions. To counterbalance within-subject effects, we alternated human and AA teammates. Each session consisted of two teams of two players. In the non-experimental side, both players were humans with Viktor and Ash avatars. In the experimental side, all participants used Viktor avatars, and their teammate, in both human and AA sessions, used the Ash avatar (Figure 1). The session with a human teammate served as our baseline, while the session with the AA teammate was our intervention.

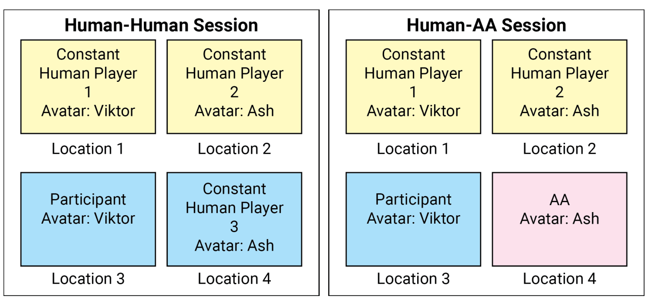


**Figure 1: A schematic diagram showing the details about the HH and HA sessions.**

## 2.3 Data Collection

Following the sessions, participants filled out a gameplay experience questionnaire and a team experience questionnaire [10]. Both included Likert scale and NASA Task Load Index (TLX) surveys [7]. Gameplay sessions were screen-recorded for analysis, capturing damage and shielding scores for participants and their teammates. A 10-15 minute interview gathered insights into participants' overall experiences, perceptions of teammates, game interface, controls, and team-related factors like communication, performance, shared mental models, and situation awareness. Two main metric types assessed the HH and HA teams: objective performance metrics (e.g., damage and shielding scores) and subjective or human-centered metrics (e.g., situation awareness, workload, interdependence). Objective metrics are task-dependent and primary for human-AA teams. Subjective metrics consider human-centered factors and provide quantitative insights, supporting a deeper understanding of perceptions.

## 2.4 Research Questions

Guided by our motivation and experiment design, our study addressed the following research questions; **RQ1**: What perceptions emerge from the game experience and performance in HH and HA teams in Paladins, without knowledge of their teammate's identity? **RQ1.a**: Specifically, how are cognitive load, performance, effort, frustration, interdependence, and situation awareness perceived? **RQ2**: How do these aspects contribute to the formed perceptions?

# 3 PILOT STUDY AND CONCLUSION

An initial data gathering involved 10 participants (1 female) [23]. The age range of these participants spanned from 18 to 32 years (M=24.6, SD=4.65).

Such a design was implemented to find out the perceptions formed in terms of cognitive load, performance, effort, frustration, interdependence and situation awareness, in the given condition. These aspects contributed to the perceptions the participants formed, during the gameplay experience with their assigned teammate. These perceptions were self-reported by the participants through the Likert scale and NASA TLX surveys. The results show that cognitive load had a negative impact on the experience, whereas low effort and frustration level, with a higher interdependence and situational awareness when with AA teammates showed positive experience. The key results in this study are as follows- (1) participants felt higher mental load and temporal demand when teamed up with AA teammates, (2) participants reported the same level of mental effort for both the sessions (human-human and human-AA), (3) participants expressed more frustration when teamed with a human, (4) participants felt that they performed better when teamed with an AA, (5) participants perceived that the AA teammate was more affected by them than the human teammate, and (6) participants reported higher perceived self-awareness and higher perceived teammate's awareness, when with AA teammates.

The current study stands out for its experiment design, wherein, the participants were blinded regarding the details of their teammate. This step up, unlike the previous works in this area, ensured unbiased perception of the participants whilst evaluating their teammate, irrespective of whether the teammate was an AA or a human. Based on this study, an interesting question can be raised about previous research works - for instance, would Demir et al. [4] obtain similar results (higher performance of the AA even with instabilities in coordination), if Demir et al. [4] were to conduct their same study with an exception of participants being unaware of their teammate, human or AA.

It would be worthwhile to re-conduct the study with a larger and more representative sample size and verify if the same outcome is obtained. In addition, the experiment design could be expanded to add another condition to have a 2 by 2 mixed-method study with one variable being characteristic of teammate (human and AA) and another variable could be teammate's visibility (blinded and aware), that is, HH and HA team and participant not aware of who/what their teammate is versus HH and HA team and participant aware of who/what their teammate is. This could enable validation of the outcome within the contextual area and could possibly provide significant results which could be very beneficial to the HCI community.

# REFERENCES


[1] Zahra Ashktorab, Q Vera Liao, Casey Dugan, James Johnson, Qian Pan, Wei Zhang, Sadhana Kumaravel, and Murray Campbell. 2020. Human-ai collaboration in a cooperative game setting: Measuring social perception and outcomes. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–20.

[2] Ying Bao, Xusen Cheng, Triparna De Vreede, and Gert-Jan De Vreede. 2021. Investigating the relationship between AI and trust in human-AI collaboration. (2021).

[3] Nancy J Cooke, Eduardo Salas, Janis A Cannon-Bowers, and Renee J Stout. 2000. Measuring team knowledge. *Human factors* 42, 1 (2000), 151–173.

[4] Mustafa Demir, Nancy J Cooke, and Polemnia G Amazeen. 2018. A conceptual model of team dynamical behaviors and performance in human-autonomy teaming. *Cognitive Systems Research* 52 (2018), 497–507.

[5] Mustafa Demir, Nathan J McNeese, and Nancy J Cooke. 2016. Team communication behaviors of the human-automation teaming. In *2016 IEEE international multi-disciplinary conference on cognitive methods in situation awareness and decision support (CogSIMA)*. IEEE, 28–34.

[6] Alexandra M Harris-Watson, Lindsay E Larson, Nina Lauharatanahirun, Leslie A DeChurch, and Noshir S Contractor. 2023. Social perception in Human-AI teams: Warmth and competence predict receptivity to AI teammates. *Computers in Human Behavior* (2023), 107765.

[7] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[8] Allyson I Hauptman, Beau G Schelble, Nathan J McNeese, and Kapil Chalil Madathil. 2023. Adapt and overcome: Perceptions of adaptive autonomous agents for human-AI teaming. *Computers in Human Behavior* 138 (2023), 107451.

[9] Keke Hou, Tingting Hou, and Lili Cai. 2023. Exploring Trust in Human–AI Collaboration in the Context of Multiplayer Online Games. *Systems* 11, 5 (2023), 217.

[10] Wijnand A IJsselsteijn, Yvonne AW De Kort, and Karolien Poels. 2013. The game experience questionnaire. (2013).

[11] Claire Liang, Julia Proft, Erik Andersen, and Ross A Knepper. 2019. Implicit communication of actionable information in human-ai teams. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.

[12] Joseph B Lyons, Sean Mahoney, Kevin T Wynne, and Mark A Roebke. 2018. Viewing machines as teammates: A qualitative study. In *2018 AAAI Spring Symposium Series*.

[13] Nathan J McNeese, Mustafa Demir, Nancy J Cooke, and Christopher Myers. 2018. Teaming with a synthetic teammate: Insights into human-autonomy teaming. *Human factors* 60, 2 (2018), 262–273.

[14] Nathan J McNeese, Beau G Schelble, Lorenzo Barberis Canonico, and Mustafa Demir. 2021. Who/what is my teammate? team composition considerations in human–ai teaming. *IEEE Transactions on Human-Machine Systems* 51, 4 (2021), 288–299.

[15] Joseph E Mercado, Michael A Rupp, Jessie YC Chen, Michael J Barnes, Daniel Barber, and Katelyn Procci. 2016. Intelligent agent transparency in human–agent teaming for Multi-UxV management. *Human factors* 58, 3 (2016), 401–415.

[16] Imani Munyaka, Zahra Ashktorab, Casey Dugan, J Johnson, and Qian Pan. 2023. Decision Making Strategies and Team Efficacy in Human-AI Teams. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–24.

[17] Christopher Myers, Jerry Ball, Nancy Cooke, Mary Freiman, Michelle Caisse, Stuart Rodgers, Mustafa Demir, and Nathan McNeese. 2018. Autonomous intelligent agents for team training. *IEEE Intelligent Systems* 34, 2 (2018), 3–14.

[18] Clifford Nass, Brian J Fogg, and Youngme Moon. 1996. Can computers be teammates? *International Journal of Human-Computer Studies* 45, 6 (1996), 669–678.

[19] Changhoon Oh, Jungwoo Song, Jinhan Choi, Seonghyeon Kim, Sungwoo Lee, and Bongwon Suh. 2018. I lead, you help but only with enough details: Understanding user experience of co-creation with artificial intelligence. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.

[20] Thomas O'Neill, Nathan McNeese, Amy Barron, and Beau Schelble. 2022. Human–autonomy teaming: A review and analysis of the empirical literature. *Human factors* 64, 5 (2022), 904–938.

[21] Samantha Reig, Jodi Forlizzi, and Aaron Steinfeld. 2019. Leveraging robot embodiment to facilitate trust and smoothness. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 742–744.

[22] Eduardo Salas, Terry L Dickinson, Sharolyn A Converse, and Scott I Tannenbaum. 1992. Toward an understanding of team performance and training. (1992).

[23] Peer M Sathikh, Tan Guan Yi, and Fang Zong Rui Dexter. 2021. Assessing indicators to testing human-AI interaction in au-tonomous vehicles: A pilot study with Omnipresent AI. (2021).

[24] Sarvesh Sawant, Rohit Mallick, Nathan McNeese, and Kapil Chalil Madathil. 2022. Mutually beneficial decision making in Human-AI teams: Understanding soldier's perception and expectations from AI teammates in human-AI teams. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 66. SAGE Publications Sage CA: Los Angeles, CA, 287–289.

[25] Beau G Schelble, Christopher Flathmann, Nathan J McNeese, Thomas O'Neill, Richard Pak, and Moses Namara. 2022. Investigating the Effects of Perceived Teammate Artificiality on Human Performance and Cognition. *International Journal of Human–Computer Interaction* (2022), 1–16.

[26] Beau G Schelble, Jeremy Lopez, Claire Textor, Rui Zhang, Nathan J McNeese, Richard Pak, and Guo Freeman. 2022. Towards Ethical AI: Empirically Investigating Dimensions of AI Ethics, Trust Repair, and Performance in Human-AI Teaming. *Human Factors* (2022), 00187208221116952.

[27] Sarah Sebo, Brett Stoll, Brian Scassellati, and Malte F Jung. 2020. Robots in groups and teams: a literature review. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–36.

[28] Ho Chit Siu, Jaime Peña, Edenna Chen, Yutai Zhou, Victor Lopez, Kyle Palko, Kimberlee Chang, and Ross Allen. 2021. Evaluation of human-AI teams for learned and rule-based agents in Hanabi. *Advances in Neural Information Processing Systems* 34 (2021), 16183–16195.

[29] Dakuo Wang, Elizabeth Churchill, Pattie Maes, Xiangmin Fan, Ben Shneiderman, Yuanchun Shi, and Qianying Wang. 2020. From human-human collaboration to Human-AI collaboration: Designing AI systems that can work together with people. In *Extended abstracts of the 2020 CHI conference on human factors in computing systems*. 1–6.

[30] Dakuo Wang, Justin D Weisz, Michael Muller, Parikshit Ram, Werner Geyer, Casey Dugan, Yla Tausczik, Horst Samulowitz, and Alexander Gray. 2019. Human-AI collaboration in data science: Exploring data scientists' perceptions of automated AI. *Proceedings of the ACM on human-computer interaction* 3, CSCW (2019), 1–24.

[31] Qiaoning Zhang, Matthew L Lee, and Scott Carter. 2022. You complete me: Human-ai teams and complementary expertise. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–28.

[32] Rui Zhang, Nathan J McNeese, Guo Freeman, and Geoff Musick. 2021. " An ideal human" expectations of AI teammates in human-AI teaming. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–25.