Human Factors and Simulation, Vol. 30, 2022, 44–51
https://doi.org/10.54941/ahfe1001490

Impact of Camera Perspective and Image Throughput on Human Trust of a Quadrupedal Robot Scout

Ralph Brewer[1], Zachary Guyton[2], Tyler Long[2], Angela Vantreese[2], Mason Russell[1], Chad Kessens[1], and and Ericka Rovira[2]

[1]US DEVCOM Army Research Laboratory Aberdeen Proving Ground, MD 21005, USA
[2]United States Military Academy West Point, NY 10996, USA

ABSTRACT

The objective of this study is to understand user perceptions of robot behaviors. Specifically, we are interested in the possible effects of providing the user with different camera perspectives and with regular snapshots versus a continuous camera feed in the context of a small-unit military operation. The study will employ a mixed 2 (camera perspective: 1st person vs over the shoulder 3rd person) x 2 (camera feed: snapshots vs continuous) factorial design, with participants viewing a robot performing military tasks in both rural and urban operational settings. After viewing the robot's performance, participants will answer performance questions based on the context of the military mission, as well as questionnaires that measure trust in the autonomous system. Dependent variables include performance outcomes from tactical performance questions and subjective results of the trust questionnaires. Data from participants will be analyzed with a 2x2 between subjects ANOVA. We anticipate that the findings will suggest that a third person perspective and continuous camera feed will result in the highest trust and best performance outcomes.

Keywords: Spot robot, Trust, Cohesion, Perception, Manned-Unmanned teaming, Human-Robot interaction

INTRODUCTION

As the Army continues to modernize, future combat operations will rely on robot technology to enhance troop performance more than ever. Although the military has recently prioritized the use of robotics and autonomous systems, the practice is nothing new. The U.S. military has relied on autonomous systems for reconnaissance since the 1990s. For example, unmanned aerial vehicles such as the MQ-1B Predator and the MQ-9 Reaper are remotely piloted, reducing risk of harm to Airmen while allowing them to conduct missions worldwide. However, the U.S. is not alone in these pursuits, and our adversaries are also relying more heavily on robotics and autonomous systems (US Army Training and Doctrine Command, 2017). The US Army has a need to integrate Robotic and Autonomous System (RAS) technology into its ground maneuver elements to meet the future demands of warfare. According to the 2017 US Army RAS Strategy, robots are expected to contribute heavily towards the Department of Defense in the future by reducing the number of US military casualties, increasing decision speed, and performing missions that human Soldiers cannot.

Although robot technology has the potential to facilitate decision-making while also mitigating risk on the battlefield, Soldiers may have difficulties using it. This is especially the case with Soldiers who have limited experience with autonomous systems. Some of the biggest concerns lie in whether Soldiers can understand robotic behavior well enough to recognize when the robot is performing as it should or if it is malfunctioning. Being able to assess a robot's performance will impact how well a squad will collaborate with an autonomous teammate in the field. If Soldiers incorrectly perceive a robot to be misbehaving, they may waste time trying to fix a non-existent problem. If the robot can pick itself up after falling, then the Soldiers can focus on their mission. However, if the Soldiers cannot tell whether the robot is able to successfully pick itself up, they may try to help the robot anyway. Not only would this take up the Soldiers' time, energy, and attention in the moment, but further it could undermine the Soldier's trust in the technology, diminishing the propensity for future robot use. More consequential, Soldiers who cannot accurately understand robot behavior could inadvertently expose themselves to enemy Soldiers, radars, or sensory systems causing injury or even death.

The reach of autonomous robotic systems goes well beyond the military. Understanding and optimizing autonomous robotic systems will have profound impacts on other domains such as emergency services, transportation, healthcare, manufacturing, and even construction. Research must continue to explore autonomous robotic systems to maximize our ability to improve performance, safety, and satisfaction of these technologies. This paper looks to discover to what degree, do camera perspective and type of camera-feed influence trust and human perceptions of autonomous robots.

BACKGROUND: FACTORS THAT IMPACT PERFORMANCE AND TRUST IN HRI

Human-robot team performance extends beyond technical capability and autonomy to include human perception of the team dynamic. Specifically trust, situation awareness, and communication have been critical research areas that can impact team operations.

As robots become more and more intelligent they will accomplish more without direct interaction with the user. Understanding what the robot is doing and trying to accomplish with allow proper mental models to be formed by the human teammates (Hancock et al, 2011). The difficulty here is how to measure trust and understand the impact it has on performance (for a full review see Schaefer et al., 2021a; 2021b). However, a rule by thumb to follow is that a robot that has transparent behaviors and decision-making allows the user to see its form and function and gain a better understanding and thereby develop trust (Chen et al., 2014).

In the context of military operations, Soldiers in a tactical operations center, or even on the ground, need to monitor the performance of robots remotely to maintain team and task situation awareness. This often involves observing the robot through a camera that displays video feed or pictures to the user; however, multiple robots providing continuous camera feeds will strain communications bandwidth, drain battery life, and overload human cognition. Having the camera take a snapshot of the field and send it within a certain time interval could reduce these effects but might also affect the user's ability to understand the scene and/or the robot's activity. Previous research by Wildemuth et al. (2003) measured participants' visual comprehension when viewing video footage that only shows every nth frame. The study indicates that showing 1 out of every 64 snapshot frames was enough for participants to gain an understanding of the video. However, it is unclear how this relates to HRI. For instance, would continuous feed imbue trust at the expense of data overload for the human operator and power depletion of the robot?

Additionally, the camera point of view may also impact the user's understanding of the robot's performance in turn, influencing individual's trust in the robot's competency. Pazuchanics (2006) conducted a study to determine whether a first-person perspective or third-person perspective would improve the performance of users when navigating an uninhabited ground vehicle. The results suggest that the users performed some aspects of navigation better when viewing the ground vehicle through the third person perspective. Another study conducted by Kallinen et al. (2007) found that using a first-person point of view was more immersive for participants playing a video game, however performance data was not evaluated. While both of these studies were completed over a decade, understanding how these methods could communicate advances in technology within an HRI environment impact robot perceptions and trust.

RESEARCH QUESTIONS AND HYPOTHESES

Previous research investigating human-robot interaction has primarily focused on distinct factors that contribute to positive and negative human-robot performance and trust. To date, little research has focused on human users viewing robots from different camera perspectives or viewing robots with continuous camera feed or snapshots with regards to trust.

We seek a firmer understanding of how, and to what degree, do camera perspective and type of camera-feed influence trust and human perceptions of autonomous robots. To answer this, we will show each participant two video feeds from Spot, a quadruped robotic platform, as it completes military tasks in a wooded and an urban environment. Both videos will be either from a first-person or third-person video perspective. One video will be continuous, while the other will display snapshots changing every two seconds.

We have three hypotheses for this experiment. First, the third person point of view will produce the best performance, highest trust, highest perceived competence and safety. Previous literature suggests that humans display higher performance when viewing from a third person perspective (Pazuchanics, 2006). Second, the continuous camera-feed will produce the best performance, highest trust, highest perceived competence and safety. While snapshots can portray how the robot behaves, the continuous feed will

Figure 1: Spot Dynamic Robot System (Boston Dynamics Support Center, 2021), left image is CAD drawing of Spot, right image is Spot shown in wooded environment.

allow human users to observe the robot's behavior more accurately. Third, the decrements to performance, trust, and perceived confidence seen in the two-second snapshot condition will not be profound enough to eliminate this condition from certain use-cases. For example, in low bandwidth or multi-robot environments, the two-second snapshot mode will still be good enough to support human operators and may be the only feasible operating mode available.

EXPERIMENTAL DESIGN

For the experimental design, a mixed 2 (viewing perspective: 1st person vs 3rd person) x 2 (video feed: continuous vs snapshot) factorial design was chosen. The between-subjects variable is the camera perspective, and the within subjects variable is the camera feed. The two independent variables are viewing perspective and video feed. Dependent variables include performance, trust, and measures of human perception of the experimental robots.

Spot (Figure 1) is a dog-like robotic platform which computes live action mechanical and visual feedback into ambulatory quadrupedal motion. Its length is 110 cm, height of 61 cm, width of 50 cm, and weight of 32.7 kg with the battery. Its maximum speed is 1.6 meters/second and can maintain a 90-minute active run time with the capability of 3-hour standby duration (Boston Dynamics Support Center, 2021). For this experiment, we mounted a GoPro camera to Spot's rail system at a nominal 77.8 cm height off the ground and at a nominal 10 degree downward pitch angle 70 cm behind the nominal first person camera which was mounted to the underside 44.7 cm from the ground and 7.3 cm from the front edge.

METHODS

Our participants will include approximately 40 undergraduate students aged 18-22 who have completed basic military training and are currently Cadets at the United States Military Academy. They will use a Windows 10 Pro 64-bit Dell Latitude 5490 laptop, to perform the experiment. All materials

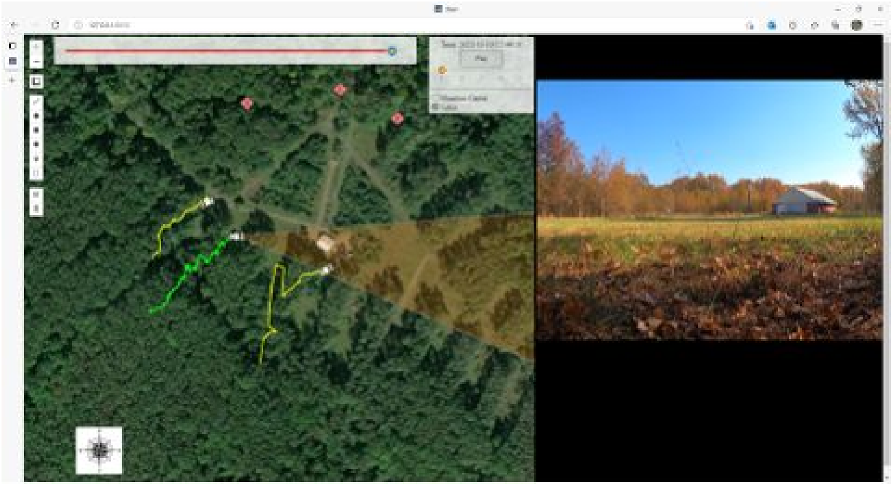

Figure 2: Graphical User Interface.

will be run from the desktop to ensure a uniform user experience and limit the possibility of inadvertent video lag that may otherwise occur through streaming.

The participants will first view seven training videos designed to level the understanding of Spot's capabilities for all who partake in the experiment. The training includes clips of Spot walking upstairs, walking over uneven terrain, and crouching down for cover from a point of view of an observer of Spot. The independent variable we are particularly focused on is the camera feed type, which will include the live-feed video from Spot (30 frames/sec, 1920×1080 resolution) or the snapshot design. The snapshot design is a simple down-sampling of the continuous feed, with 1 still frame of the video from Spot's current location updated every 2 seconds. The 2 second frame rate is used consistently throughout the entire experiment for the snapshot condition. This frame rate was selected based off a pilot study conducted with 9 individuals who rated the three different frame rates based off "easy to view" and "understandability" on a 7-point Likert scale. The results of this pilot study showed that the group preferred the 2 seconds frame rate over the 1 second and 3 second frame rates respectively in both categories.

A graphical user interface (Figure 2) was designed to highlight the movement of Spot on a 2D map on the left side, while allowing the user to watch the streaming video on the right. Spot is shown on the map with a green trail, and the beige triangle highlights the field of view of the camera. Two other notional robot teammates are shown with yellow trails. During the mission, friendly and/or enemy icons are displayed on the map, providing information the user will need to answer questions.

Two military scenarios will be used in this experiment. Participants play the role of military commander, observing operations from a tactical operations center. The first scenario will be crossing an unsecured area in the woods while staying hidden in the brush. In this context, there will be known enemies in the area of operations across the wood line in an identifiable building.

Spot and team will covertly cross the tall grass and wooded area while slowly moving in the direction of the building. The second scenario is in an urban environment. Spot will cross a rocky patch of open ground and enter a building with two non-combatants inside. Both videos were filmed at the Robotics Research Collaboration Campus, an Army Research Laboratory testing facility in Middle River, Maryland, and are always shown in the same order.

MEASURES & PROCEDURES

Our measurements will derive from the answers to the subjective questionnaires at the conclusion of each video. These questionnaires will test multiple factors in different domains of human-automation interaction. Participants will arrive for this in-person study and first provide informed consent.

Participants will then complete the Negative Attitude toward Robot Scale (NARS;Nomura et al,. 2008) with the goal being to reveal any underlying bias of robots and robotic behaviors prior to the experiment. Using 14 questions it asks the users about attitudes and preferences with robots to gauge any prior negative outlooks about robots.

The Visuospatial paper folding task entails managing visual and spatial information in working memory (Castro-Alonzo, 2019). This is given prior to viewing the experimental videos with the goal to establish each participant's natural visuospatial abilities. The instructions lay out what the user must remember in folding a paper properly when given a sequence of dots on it. After instructions, the participant will be given 10 multiple choice questions where they must choose the correct pattern of folding the paper based on the dots shown (Linn, 1985; Kyllonen, 1984).

The trust perception questionnaire is given prior to the start of the study and then again after each experimental video to gauge a participant's change in the trust they have of the robotic platform. This scale involves 14 items selected to measure trust, graded on a scale of 0–100% (Schaefer, 2016).

Also following each experimental video there are performance-based questions and the Human Computer Trust Scale (HCTS). Performance questions will assess the participants' understanding of the situation as it relates to a military mission. We can assess situation awareness and comprehension ability from each of the independent variables. The questions were constructed and approved by the research team and used in a small pilot study to confirm its validity. The questionnaire has 25 questions is divided into 5 factors: perceived reliability, perceived technical competence, perceived understandability, faith, and personal attachment for a subjective measurement of "cognition based" and "affective-based" trust (Madsen and Gregor, 2000).

The experiment will conclude with two post-experimental questionnaires, the Godspeed and the Psychological Closeness Godspeed provides feedback on the subjective views that the participant has on the characteristics and performance of Spot on a scale between two antonyms. The five categories of Godspeed are anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety (Carpinella, 2017). The Psychological Closeness

Scale (Salem et al., 2015). aims to measure the degree of closeness participants feel toward the SPOT robot (e.g., commonality with the robot?).

CONCLUSION

Data collection for this experiment will occur between January 2022 – March 2022. Results will be forthcoming and published prior to the AHFE 13th annual meeting in July 2022. Collection of these data will verify our hypotheses that the third person point of view and continuous camera feed will produce the best performance, highest trust, highest perceived competence and safety. Understanding the best camera views and feeds from the robot will continue to foster trust between Soldiers and their autonomous teammates.

REFERENCES

Boston Dynamics Support Center. (2021). Bostondynamics.Com. Retrieved December 6, 2021, from https://support.bostondynamics.com/s/article/Robot-specifications

Carpinella, C. M., Wyman, A. B., Perez, M. A., & Stroessner, S. J. (2017). The Robotic Social Attributes Scale (RoSAS): Development and validation. *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction.*

Chen, J. Y., Procci, K., Boyce, M., Wright, J., Garcia, A., & Barnes, M. (2014). *Situation awareness-based agent transparency*. (No. ARL-TR-6905).

Gregor, M., Madsen, S. (2000). Measuring Human-Computer Trust. *Measuring Human-Computer Trust.*

Hancock, P. A., Billings, D. R., & Schaefer, K. E. (2011). Can you trust your robot?. *Ergonomics in Design*, *19*(3), 24–29.

Kalline, K., Salimen, M., Ravaja, N., Kedzior, R., Saaksjarvi, M. (2007). *Presence and emotion in computer game players during 1st person vs. 3rd person playing view: evidence from self-report, eye-tracking, and facial muscle activity data.* Psu.Edu; Center for Knowledge and Innovation Research, Helsinki School of Economics and Swedish School of Economics, and Business Administration.

Kyllonen, P. C., Lohman, D. F., & Snow, R. E. (1984). Effects of aptitudes, strategy training, and task facets on spatial task performance. *Journal of Educational Psychology*, *76*(1), 130.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child development*, 1479–1498.

Nomura, T., Kanda, T., Suzuki, T., and Kato, K. (2008, April). Prediction of Human Behavior in Human-Robot Interaction Using Psychological Scales for Anxiety and Negative Attitudes Toward Robots. *IEEE Transactions on Robotics, 24*(2), 442–451.

Pazuchanics, S. (2006) The Effects of Camera Perspective and Field of View on Performance in Teleoperated Navigation. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting.*

Salem, M., Lakatos, MG., Amirabdollahian, F., Dautenhahn, K. (2015). "Would You Trust a (Faulty) Robot? Effects of Error, Task Type and Personality on Human-Robot Cooperation and Trust," 2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2015, pp. 1-8.

Schaefer, K. E., Perelman, B. S., Gremillion, G. M., Marathe, A. R., & Metcalfe, J. S. (2021). A roadmap for developing team trust metrics for human-autonomy teams. In *Trust in Human-Robot Interaction* (pp. 261-300). Academic Press.

Schaefer, K. E., Perelman, B., Rexwinkle, J., Canady, J., Neubauer, C., Waytowich, N., ... & Marathe, A. (2021). Human-Autonomy Teaming for the Tactical Edge: The Importance of Humans in Artificial Intelligence Research and Development. In *Systems Engineering and Artificial Intelligence* (pp. 115–148).

Schaefer, K. (2016). Measuring trust in human robot interactions: Development of the "trust perception scale-HRI. *Robust Intelligence and Trust in Autonomous Systems*. 191–218.

U.S. Army Training and Doctrine Command. (2017). *2017 U.S. Army Robotics and Autonomous Systems Strategy*. https://www.tradoc.army.mil/wp content/uploads/2020/10/RAS_ Strategy.pdf

Wildemuth, B. (2004). The Effects of Domain Knowledge on Search Tactic Formulation. *JASIST, 55,* 246–258.