

The Effect of Agent Reasoning Transparency on Automation Bias: An Analysis of Response Performance

Julia L. Wright¹(✉), Jessie Y.C. Chen¹,
Michael J. Barnes², and P.A. Hancock³

¹ US Army Research Laboratory, Orlando, FL, USA

{Julia.l.wright8.civ, Yun-sheng.c.chen.civ}@mail.mil

² US Army Research Laboratory, Ft. Huachuca, Sierra Vista, AZ, USA
michael.j.barnes.civ@mail.mil

³ University of Central Florida, Orlando, FL, USA
Peter.Hancock@ucf.edu

Abstract. We examined how the transparency of an agent's reasoning affected the human operator's complacent behavior in a military route selection task. Participants guided a three-vehicle convoy through a simulated environment in which they had a limited amount of information about their surroundings, all this while maintaining communication with command and monitoring their surroundings for threats. The intelligent route-planning agent, RoboLeader, assessed potential threats and offered changes to the planned route as necessary. RoboLeader reliability was 66 %, and the participant had to correctly reject RoboLeader's suggestion when incorrect. Access to RoboLeader's reasoning was varied across three conditions (no reasoning, reasoning present, and increased reasoning transparency), and each participant was assigned to one of the three conditions. Access to agent reasoning improved performance and decreased automation bias. However, when reasoning transparency increased, performance decreased while automation bias increased. Implications for presentation of reasoning information in operational settings are discussed in light of these findings.

Keywords: Human-agent teaming · Automation bias · Agent transparency · Complacency

1 Introduction

Soldiers interact with a wide variety of automated and robotic assets, both as part of their normal duties as well as to expand their scope of influence. Employing robotic assets to assist in their duties allows the Soldier to manage multiple tasks of increasing complexity. Such aids permit more effective multitasking and enhance performance on secondary tasks [1]. This is accomplished by reducing their workload and freeing cognitive resources [2]. These benefits are especially important when a Soldier is managing a team of robotic assets. Research has shown that a single operator managing multiple robotic assets can suffer performance decrements such as reduced situation

awareness (SA) and increased workload [3–5]. As the number of robots increase, so do these performance decrements [6]. To assist an operator managing a team of robots, an intelligent agent, RoboLeader (RL), has been developed [7]. RL acts as a mediator between the human operator and the team of subordinate robots. This gives the operator a single point of contact for such robotic assets. Several studies with RL have shown that using such an intelligent agent as a mediator for the robotic team improves both the operators' SA and task performance, while decreasing their perceived workload [8–10].

However, the use of an intelligent agent as a mediator between the operator and robotic team is not without problems. Supervisory control issues such as reduced SA and increased complacency [11] become apparent as the operator is further removed from the inner 'loop' of control [2, 12]. Previous RL studies have indicated that while the operator benefitted from reduced workload, their task performance or SA did not always improve concomitantly [8]. Indeed, increasing RL's level of assistance resulted in decreased performance for certain individuals [10]. While the addition of the intelligent agent can be a boon to an operator managing multiple tasks, it also creates the distance that makes effective supervision of the team more difficult. Often this "distance" results in the operator displaying automation bias in favor of agent recommendations. It remains unknown whether this bias is a result of the operator recognizing they do not have enough information to confidently override the agent suggestions when appropriate, or whether complacency is due to an operator's out-of-the-loop (OOTL) role. Increasing the transparency of the agent has been recommended as one way to reduce this distance, pulling the operator back into the inner loop of control [13]. One way to do this is to increase the operator's understanding of the agent's reasoning (i.e., why is the agent making this recommendation). As the users' understanding of the rationale behind a systems' behavior grows, the more accurate the users' calibration of their trust and reliance is expected to be [14–16].

1.1 Current Study

In order to distinguish between the latter two propositions, we engaged in the present study. Participants guided a convoy of robotic vehicles (an unmanned aerial vehicle [UAV], an unmanned ground vehicle [UGV], and a manned ground vehicle [MGV]) through a simulated environment with an intelligent agent (RoboLeader) assisting with the route planning task. As the convoy progressed, events would occur that would necessitate re-routing the convoy, and RL would recommend changes to the route as these events occurred. The participant had information from a variety of sources regarding the events, and had to either accept or reject RL's recommendation. In each scenario, RoboLeader suggested route changes 6 times, 4 times of which RL's suggestion was the correct choice. The participant was required to recognize and correctly reject an incorrect suggestions. In addition to the route selection task, participants maintained communications with command and monitored the area for threats. Operator complacency and trust across differing agent transparency conditions (explained below) were evaluated via performance and response time measures.

Agent transparency was manipulated by varying participants' access to RL's reasoning; participants were randomly assigned to one of three agent reasoning transparency

(ART) conditions. ART 1 was the Baseline—the agent notified the operator that a route revision was recommended, however, no reasoning for the suggestion was given (i.e., ‘Change to convoy path recommended’). ART 2 had the same notification process as in ART 1, but RL also explained its reason for the suggested route change (e.g., ‘Change to convoy path recommended. Activity in area: Dense Fog’). ART 3 had the same information as in ART 2, but RL also provided the time of report (TOR) (e.g., ‘Change to convoy path recommended. Activity in area: Dense Fog. TOR: 1 [h]’).

We hypothesized that access to agent reasoning would improve operator performance, reduce complacent behavior, and increase trust in the agent—but only to a degree, beyond which increased transparency of agent reasoning would negatively impact operator performance, increase complacent behavior, and reduce trust in the agent (i.e., $\text{ART 1} < \text{ART 2} > \text{ART 3}$). This hypothesis recapitulates an inverted [extended] U-shaped function often observed in operators in stressful conditions [22, 23]. Performance on the route planning task was evaluated using the number of correct acceptances and rejections of RL’s suggestions. Automation bias was evaluated by the number of incorrect acceptances. Distrust was evaluated objectively via incorrect rejections of RL’s suggestions, and subjectively via scores on a usability and trust survey. Decision time was expected to increase as access to agent reasoning increased: $\text{ART 1} < \text{ART 2} < \text{ART 3}$. Although RL’s messages were slightly longer in ARTs 2 and 3 than in ART 1, additional time was not expected to be required for reading the messages. Participants were expected to take longer to process the information and reach their decision, resulting in a longer decision time. Shorter response times may indicate less deliberation on the part of the operator before accepting or rejecting the agent recommendation. This could mean either positive automation bias or reduced task difficulty.

2 Method

2.1 Participants

Sixty participants (26 males, 33 females, 1 unreported, $\text{Min}_{age} = 18$ years, $\text{Max}_{age} = 32$ years, $M_{age} = 21.4$ years) from a large southern US university participated for either class credit or cash compensation.

2.2 Materials

Simulator. The simulator in this experiment was a modified version of the Mixed Initiative Experimental (MIX) Testbed [17], a distributed simulation environment developed for investigating how unmanned systems are used and how automation affects human operator performance. The RoboLeader algorithm was implemented on the MIX Testbed and had the capability of collecting information from subordinate robots, making tactical decisions, and then coordinating the activities of subordinate robots [7]. The Operator Control Unit (OCU) of the MIX Testbed (Fig. 1) was modeled after the Tactical Control Unit developed under the ARL Robotics Collaborative Technology Alliance. The simulation was delivered via a commercial desktop computer system, 22-inch monitor, standard keyboard, and three-button mouse.

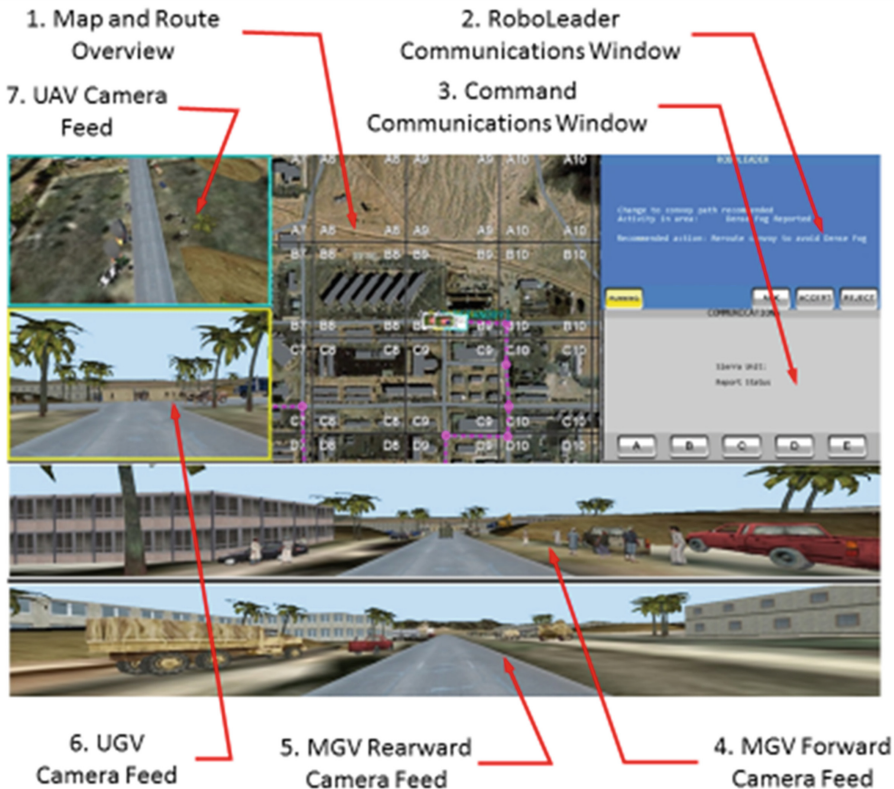


Fig. 1. Operator Control Unit (OCU), the user interface for convoy management. OCU windows are (clockwise from the upper center): 1. Map and Route overview, 2. RoboLeader communications window, 3. Command communications window, 4. MGV Forward 180° Camera Feed, 5. MGV Rearward 180° Camera Feed, 6. UGV Forward Camera Feed and 7. UAV Camera Feed.

Demographics. A demographics questionnaire was administered at the beginning of the training session and information on participant's age, gender, education level, computer familiarity, and gaming experience was collected.

2.3 Procedure

After signing the informed consent, participants completed a demographics questionnaire, a reading comprehension test, and a brief Ishihara Color Vision Test to screen for eligibility to participate in the experiment. Participants then received practice on their tasks. This training was self-paced and delivered by PowerPoint® slides. Participants were trained on the elements of the OCU, identifying map icons and their meanings, steps for completing various tasks, and completed several mini-exercises for practice. The training session lasted approximately 1.5 h. Participants were assessed on their

proficiency on the required tasks before proceeding to the experimental session, and those that did not achieve at least 90 % on the assessments were dismissed.

The experimental session lasted about 2 h and began immediately after the training session. Participants were randomly assigned to an Agent Reasoning Transparency (ART) condition (ART 1, ART 2, or ART 3). Each experimental session had three scenarios, each lasting approximately 30 min. The scenario order and ART were counterbalanced across participants.

Participants guided a convoy of three vehicles (their MGCV, a UAV, and a UGV) along a predetermined route through a simulated urban environment. As the convoy proceeded, events occurred which may have necessitated altering the route. Events and their associated area of influence were displayed on the map. The operator either accepted the suggestion or rejected it and kept the convoy on its original path. RoboLeader (RL) suggested a potential route revision six times per session. Two of these suggestions were incorrect, requiring the participant to correctly reject the suggestion. Once RL suggested a route revision, participants had 15 s to acknowledge the suggested change before RL automatically continued along the original route. Once acknowledged, the vehicles paused until the participant either agreed with or rejected RL's suggestion.

Participants maintained communication with 'command' via a text feed directly below RL's communication window. Incoming messages (from either source) appeared approximately every 30 s. Communications from command included messages directed at other units, which the participant should disregard, and requests for information, which required a response. Each mission contained 12 information updates, two of which would result in the need to override RL's route recommendation.

While moving through the environment participants' maintained local security surrounding their MGCV by monitoring the MGCV and UGV indirect-vision displays and detect threats (armed civilians) in the immediate environment. Participants identified threats by clicking on the threat in the window using the mouse, and received no feedback on this task. There were unarmed civilians and friendly dismounted soldiers in the simulated environment to increase the visual noise present in the threat detection tasks. Following completion of all three scenarios participants were debriefed and any questions they had were answered by the experimenter.

2.4 Experiment Design and Performance Measures

The study was a between-subject experiment with Agent Reasoning Transparency (ART) as the independent variable. Dependent measures were route selection performance score, automation bias score, distrust score, and decision time.

Data was analyzed using planned comparisons ($\alpha = .05$). Specifically, ART 1 was compared to ART 2, ART 2 to ART 3, and ART 1 to ART 2 + 3, unless otherwise noted. Omnibus ANOVAs ($\alpha = .05$) are also reported.

Performance Score. Total correct acceptances and rejections were summed across all missions. The range for this score is 0 (no correct rejects or accepts) to 18 (all suggestions correctly accepted or rejected).

Automation Bias. Twice each mission RoboLeader made a suggestion that should be rejected. Participants scored 1 point for each incorrect acceptance and these were summed across all missions. Higher scores indicate higher automation bias. The score range for this measure is 0 – 6.

Distrust. Four times each mission RoboLeader made a suggestion that should have been correctly accepted. Participants scored 1 point for each incorrect rejection and these were summed across all missions. Higher scores indicate greater distrust. The score range for this measure is 0 – 12.

Decision Time. Decision time was measured as time between alert acknowledgment and route selection, and averaged across missions.

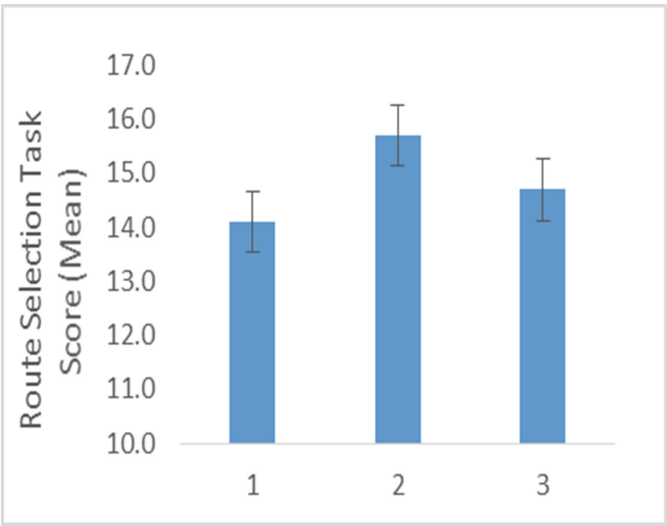


Fig. 2. Average Route selection task score by agent reasoning transparency level. Bars denote SE.

3 Results

3.1 Route Selection Task Performance

There was no significant effect of ART on the route selection task scores, $F(2,57) = 2.00$, $p = .145$, $\omega^2 = .03$ (Fig. 2). Planned comparisons revealed mean performance scores were slightly higher in ART 2 ($M = 15.70$, $SD = 2.23$) than in ART 1 ($M = 14.10$, $SD = 2.59$), $t(57) = 1.98$, $p = .053$, $r_c = 0.25$. There was no significant difference in performance between ART 2 and ART 3 ($M = 14.70$, $SD = 2.81$), $t(57) = -1.24$, $p = .221$, $r_c = 0.16$. The hypothesis was partially supported, as the medium-large effect size between ARTs 1 and 2 indicates that the addition of agent reasoning did improve route selection. Scores in ART 3 were lower than those in ART 2, however this difference was not significant, indicating that performance in these two conditions was essentially the same.

3.2 Automation Bias

Evaluating Automation Bias scores between ART conditions, there was a violation of the homogeneity of variance assumption. As such, Welch's correction has been reported, and contrast tests did not assume equal variance between conditions. There was a significant effect of ART on automation bias, $F(2,34.8) = 7.96$, $p = .001$, $\omega^2 = .14$ (Fig. 3). Mean automation bias scores were lower in ART 2 ($M = 1.14$, $SD = 1.28$) than in ART 1 ($M = 3.25$, $SD = 2.27$), $t(57) = -3.63$, $p = .001$, $r_c = 0.55$, and ART 3 ($M = 2.65$, $SD = 2.32$), $t(57) = 2.55$, $p = .016$, $r_c = 0.43$. Overall, automation bias scores were significantly lower when agent reasoning was provided, $t(57) = -2.31$, $p = .028$, $r_c = 0.38$. The hypothesis was supported, since access to agent reasoning did reduce automation bias in a low information environment, and increased transparency of agent reasoning began to overwhelm participants, resulting in increased automation bias.

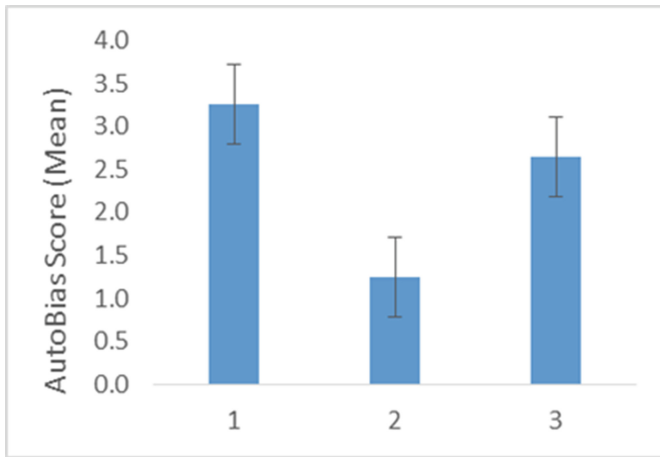


Fig. 3. Average Automation bias scores by agent reasoning transparency level. Bars denote SE.

Automation bias could also be indicated by reduced decision time on the route selection task. We hypothesized that decision time would increase as agent reasoning transparency increased, as participants should require additional time to process the extra information. Thus, reduced time could indicate less time spent in deliberation, which could be an indication of automation bias.

There was not a significant effect of ART on elapsed decision time, $F(2,57) = 1.51$, $p = .230$, $\omega^2 = .02$ (Fig. 4). Mean automation bias scores were lower in ART 2 ($M = 2787.58$, $SD = 1055.09$) than in ART 1 ($M = 3530.32$, $SD = 1567.98$), $t(57) = -1.74$, $p = .088$, $r_c = 0.22$, and ART 3 ($M = 3176.34$, $SD = 1383.57$), $t(57) = 0.91$, $p = .367$, $r_c = 0.12$. Overall, decision times were lower when agent reasoning was provided, but not significantly so, $t(57) = -1.49$, $p = .144$, $r_c = 0.19$.

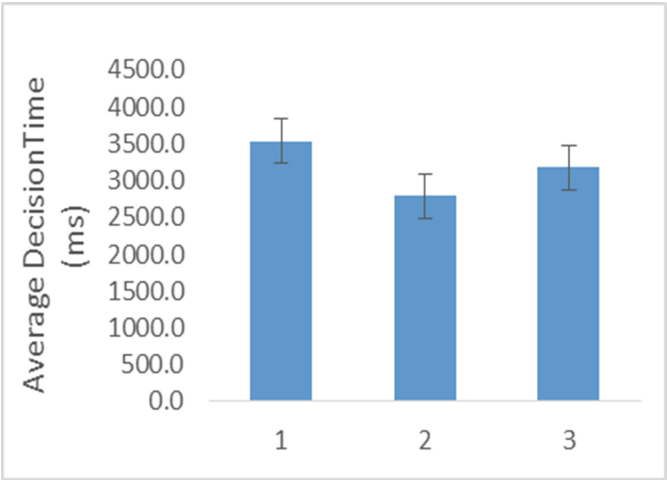


Fig. 4. Average decision time by agent reasoning transparency level. Bars denote SE.

3.3 Distrust Score

Evaluating Distrust (incorrect rejection of the agent suggestion) there was no significant effect of ART on distrust scores, $F(2,57) = 0.28$, $p = .756$, $\omega^2 = .02$, (Fig. 5). Planned comparisons revealed distrust scores were slightly higher in ART 2 ($M = 1.05$, $SD = 1.15$) than in ART 1 ($M = 0.80$, $SD = 1.36$), $t(57) = 0.52$, $p = .606$, $r_c = 0.07$, and ART 3 ($M = 0.80$, $SD = 1.36$), $t(57) = -0.73$, $p = .470$, $r_c = 0.10$, however these differences were not significant.

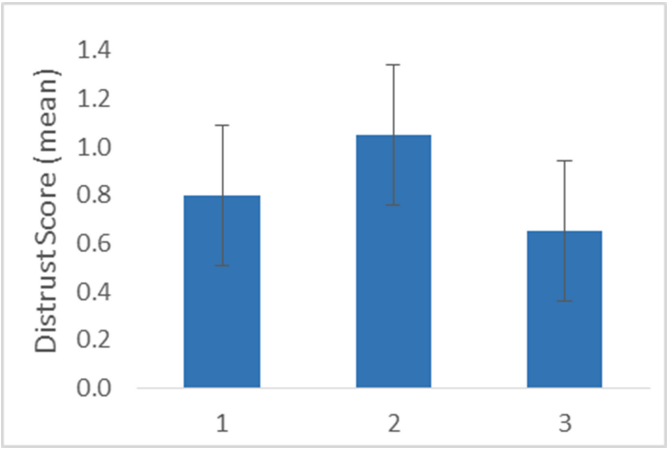


Fig. 5. Average Distrust scores by agent reasoning transparency level. Bars denote SE.

3.4 Usability and Trust Survey

There was not a significant effect of ART on Trust score, $F(2,57) = 2.52$, $p = .089$, $\omega^2 = .05$, (Fig. 6). There was also a significant curvilinear trend to the data, $F(1,57) = 4.15$, $p = .046$, $\omega^2 = .05$. Planned comparisons revealed that trust scores in ART 2 ($M = 54.40$, $SD = 10.23$) were slightly lower than in ART 1 ($M = 58.55$, $SD = 8.28$), $t(57) = -1.29$, $p = .202$, $r_c = 0.17$, and significantly lower than ART 3 scores ($M = 61.60$, $SD = 11.72$), $t(57) = 2.24$, $p = .029$, $r_c = 0.28$. These findings did not support the hypothesis, as ART 2 had the lowest Trust scores while ART 3 had the highest.

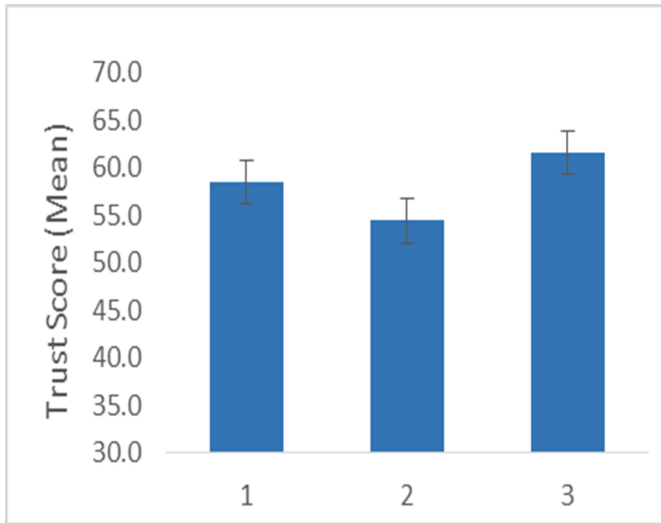


Fig. 6. Average Trust scores by agent reasoning transparency level. Bars denote SE.

There was a significant effect of ART on Usability scores, $F(2,57) = 5.11$, $p = .009$, $\omega^2 = .12$, (Fig. 7). There was also a significant curvilinear trend to the data, $F(1,57) = 9.96$, $p = .003$, $\omega^2 = .13$. Pairwise comparisons showed that Usability scores in ART 2 ($M = 40.75$, $SD = 6.60$) were significantly lower than those in either ART 1 ($M = 46.75$, $SD = 5.33$), $t(57) = -2.98$, $p = .004$, $r_c = 0.37$, or ART 3 ($M = 45.75$, $SD = 7.03$), $t(57) = 2.49$, $p = .049$, $r_c = 0.31$. Overall, Usability scores were significantly lower when agent reasoning was present than when it was not, $t(57) = -2.01$, $p = .049$, $r_c = 0.26$.

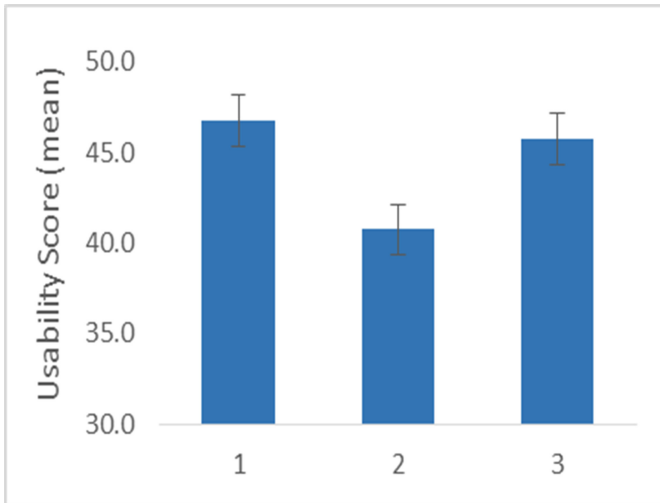


Fig. 7. Average Usability scores by agent reasoning transparency level. Bars denote SE.

4 Discussion

The goal of this study was to examine how the level of transparency of an intelligent agent’s reasoning process affected response complacency. Participants supervised a three-vehicle convoy as it traversed a simulated environment and re-routed the convoy when needed with the assistance of an intelligent agent, RoboLeader (RL). When the convoy approached a potentially unsafe area, RL would recommend re-routing the convoy. Each participant was assigned to a specific level of agent reasoning transparency (ART). The reasoning provided as to why RL was making the recommendation differed among these levels. ART 1 provided no reasoning information—RL notified that a change was recommended without explanation. The type of information the agent supplied varied slightly between ARTs 2 and 3. This additional information did not convey any confidence level or uncertainty but was designed to encourage the operator to actively evaluate the quality of the information rather than simply respond. Therefore, not only was access to agent reasoning available, but the impact of the type of information the agent supplied could be examined also.

Performance on the route selection task was evaluated via correct rejections and acceptances of the agent suggestion. An increased number of correct acceptances and rejections, and reduced response times were all indicative of improved performance. Route selection performance was hypothesized to improve with access to agent reasoning and then decline as agent reasoning transparency increased, and this hypothesis was partially supported. Performance did improve when access to agent reasoning was provided. However, increased transparency of agent reasoning did not result in a performance decrement.

Complacent behavior was examined via primary (route selection) task response, in the form of automation bias (i.e., incorrect acceptances of RL suggestions). As predicted,

access to agent reasoning reduced incorrect acceptances, and increased reasoning transparency increased incorrect acceptances. Complacent behavior was highest when no agent reasoning was available. When the transparency of agent reasoning was increased to its highest level, complacent behavior increased to nearly the same level as in the no-reasoning condition. This pattern of results indicated that while access to agent reasoning in a decision-supporting agent can counter automation bias, too much information results in an out-of-the-loop (OOTL) situation and increased complacent behavior.

Similar to previous findings [16], access to agent reasoning did not increase response time. In fact, decision times were reduced in the agent reasoning conditions, even though the agent messages in the reasoning conditions were slightly longer than in the no reasoning condition and required slightly more time to process. Similar studies have suggested that a reduction in accuracy with consistent response times could be attributed to a speed-accuracy trade-off [18]. However, the present findings indicated that may not be the case. We saw an increase in accuracy with no accompanying increase in response time (hence no trade-off). What appears to be more likely is that not only does the access to agent reasoning assist the operator in determining the correct course of action, but the type of information the operator receives also influences their behavior.

The objective measure of operator trust indicated no difference in trust due to agent reasoning transparency. Subjective measures indicated agent reasoning had no effect on operator trust and reduced the usability evaluation. Increased transparency of agent reasoning resulted in increased trust and usability ratings, however there was no associated improvement in performance. Interestingly, operators reported highest trust and usability in the conditions that also had the highest complacency, and lowest in the condition that had the highest performance.

In all conditions, the operator received all information needed to correctly route the convoy without the agent's suggestion. In the Baseline condition, the operators demonstrated a clear bias for the agent suggestion. With a moderate amount of information regarding the agent reasoning, the operators were more confident in overriding erroneous suggestions. In the highest reasoning transparency condition, operators were also given information regarding when the agent had received the information (i.e. its recency), and while this information did not imply any confidence or uncertainty on the part of the agent, such additional information appeared to create ambiguity for the operator. This encouraged them to defer to the agent's suggestion.

5 Conclusion

The findings of the present study are important for the design of intelligent recommender and decision-aid systems. Keeping the operator engaged and in-the-loop is important for reducing complacency, which could allow lapses in system reliability to go unnoticed. To that end, we examined how agent transparency affected operator complacent behavior, as well as task performance, and trust. Access to agent reasoning was found to be an effective deterrent to complacent behavior when the operator has limited information about their task environment. Contrary to the position adopted by Paradis et al. [19], operators do accept agent recommendations even when they do not

know the rationale behind the suggestions. In fact, the absence of agent reasoning appears to encourage automation bias. Access to the agent's reasoning appears to allow the operator to calibrate effectively their trust in the system, reducing automation bias and improving performance. This outcome is similar to findings previously reported by Helldin et al. [20] and Mercado et al. [16]. However, the addition of information that created ambiguity for the operator again encouraged complacency, resulting in reduced performance and poorer trust calibration. Prior work has shown that irrelevant or ambiguous information can increase workload and encourage complacent behavior [11, 21], and these findings align with those. As such, caution should be exercised when considering how transparent to make agent reasoning and what information should be included.

References

1. Parasuraman, R., Molloy, R., Singh, I.L.: Performance consequences of automation-induced complacency. *Int. J. Aviat. Psychol.* **3**(1), 1–23 (1993)
2. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: A model for types and levels of human interaction with automation. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **30**(3), 286–297 (2000)
3. Chen, J.Y.C., Durlach, P.J., Sloan, J.A., Bowens, L.D.: Human robot interaction in the context of simulated route reconnaissance missions. *Mil. Psychol.* **20**(3), 135–149 (2008)
4. Wang H, Lewis M, Velagapudi P, Scerri P, Sycara K.: How search and its subtasks scale in N robots. In: *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction*, pp. 141–148. ACM Press, 9–13 March 2009, La Jolla, CA, New York (NY) (2009)
5. Wang J, Wang, H, Lewis M.: Assessing cooperation in human control of heterogeneous robots. In: *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*, pp. 9–15. ACM Press, 12–15 March 2008, Amsterdam (The Netherlands). New York (NY) (2008)
6. Lewis, M.: Human interaction with multiple remote robots. *Rev. Hum. Factors Ergon.* **9**(1), 131–174 (2013)
7. Chen, J.Y.C., Barnes, M.J., Qu, Z.: RoboLeader: A surrogate for enhancing the human control of a team of robots. Aberdeen Proving Ground (MD): Army Research Laboratory (US), Report No.: ARL–MR–0735, January 2010. <http://www.dtic.mil/get-tr-doc/pdf?AD=ADA514855>
8. Chen, J.Y.C., Barnes, M.J.: Supervisory control of multiple robots: effects of imperfect automation and individual differences. *Hum. Factors* **54**(2), 157–174 (2012)
9. Chen, J.Y.C., Barnes, M.J.: Supervisory control of multiple robots in dynamic tasking environments. *Ergonomics* **55**, 1043–1058 (2012)
10. Wright, J.L., Chen, J.Y.C., Quinn, S.A., Barnes, M.J.: The effects of level of autonomy on human-agent teaming for multi-robot control and local security maintenance. Aberdeen Proving Grounds (MD): Army Research Laboratory (US), Report No.: ARL–TR–6724, November 2013. www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA595105
11. Chen, J.Y.C., Barnes, M.J.: Human-agent teaming for multirobot control: A review of human factors issues. *IEEE Trans. Hum. Mach. Syst.* **44**(1), 13–29 (2014)

12. Wickens, C.D.: Designing for situation awareness and trust in automation. In: Proceedings of the International Federation of Automatic Control (IFAC) Conference, Baden–Baden, Germany. Amsterdam (The Netherlands), Elsevier Ltd., pp. 174–179, 27–29 September 1994
13. Chen, J.Y.C., Procci, K., Boyce, M., Wright, J., Garcia, A., Barnes, M.J.: Situation awareness–based agent transparency. Aberdeen Proving Ground (MD): Army Research Laboratory (US), Report No.: ARL–TR–6905, April 2014. http://www.arl.army.mil/www/default.cfm?technical_report=7066
14. Lee, J.D., See, K.A.: Trust in automation: designing for appropriate reliance. *Hum. Factors* **46**(1), 50–80 (2004)
15. Lyons, J.B.: Being transparent about transparency: a model for human–robot interaction. Papers from the 2013 AAAI Spring Symposium Series, 25–27 March 2013, Stanford, CA. Palo Alto (CA). The AAAI Press, pp. 48–53 (2013)
16. Mercado, J.E., Rupp, M., Chen, J.Y.C., Barber, D., Procci, K., Barnes, M.J.: Intelligent agent transparency in human-agent teaming for multi-UxV management. *Hum. Factors* **58**, 401–415 (2016). (in press)
17. Barber, D., Davis, L., Nicholson, D., Finkelstein, N., Chen, J.Y.C.: The Mixed Initiative Experimental (MIX) testbed for human robot interactions with varied levels of automation. In: Proceedings of the 26th Army Science Conference, ASC 2008, Orlando, FL, 1–4 December 2008, U.S. Dept. of Army, Washington, DC (2008)
18. Wickens, C.D., Clegg, B.A., Vieane, A.Z., Sebok, A.L.: Complacency and automation bias in the use of imperfect automation. *Hum. Factors* **57**(5), 728–739 (2015)
19. Paradis, S., Benaskeur, A., Oxenham, M., Cutler, P.: Threat evaluation and weapons allocation in network-centric warfare. In: 8th International Conference on Information Fusion, vol. 2, pp. 1078–1085. IEEE, 25 July 2005
20. Helldin, T., Ohlander, U., Falkman, G., Riveiro, M.: Transparency of automated combat classification. In: Harris, D. (ed.) EPCE 2014. LNCS, vol. 8532, pp. 22–33. Springer, Heidelberg (2014)
21. Westerbeek, H., Maes, A.: Route-external and route-internal landmarks in route descriptions: Effects of route length and map design. *Appl. Cogn. Psychol.* **27**(3), 297–305 (2013)
22. Hancock, P.A., Warm, J.S.: A dynamic model of stress and sustained attention. *Hum. Factors: J. Hum. Factors Ergon. Soc.* **31**(5), 519–537 (1989)
23. Yerkes, R.M., Dodson, J.D.: The relation of strength of stimulus to rapidity of habit-formation. *J. Comp. Neurol. Psychol.* **18**, 459–482 (1908)