2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)
Banff Center, Banff, Canada, October 5-8, 2017

Effects of Agent Transparency and Communication Framing on Human-Agent Teaming

Ryan W. Wohleber
Institute for Simulation & Training
University of Central Florida
Orlando, FL, USA
rwohlebe@ist.ucf.edu

Kimberly Stowers
Culverhouse College of Commerce
University of Alabama
Tuscaloosa, AL, USA
kstowers@cba.ua.edu

Jessie Y.C. Chen &
Michael Barnes
Human Research & Engineering Directorate
U.S. Army Research Laboratory
Aberdeen Proving Ground, MD, USA
{yun-sheng.c.chen.civ; michael.j.barnes.civ}@mail.mil

***Abstract*—Communication of the inner workings of an intelligent agent, or agent transparency, is important for human operators' calibration of their trust in the agent. Recent work has focused on methods of communicating this information in the context of agent-assisted mission planning. This investigation looked at the framing of agent communication and how it might interact with level of agent transparency. Results showed that critical framing improved operator's perception and trust of the agent regardless of the agent's transparency level. Further, findings suggest that the effects of framing are especially pronounced in low transparency applications.**

Keywords—transparency; trust; framing; human-robot teaming

I. Introduction

As advances in automation begin to make single operator control of multiple unmanned assets viable, work is underway to address the human performance aspect [1]. One popular approach to helping the human operator manage multiple assets employs an intelligent agent (IA)-based teammate to assist with higher level tasking, such as mission planning [1]. Recent work has demonstrated the importance of IA transparency for calibration of operator trust in the agent [2]. The current effort looks at another important factor in transparency communication: communication style and framing of information. There is now evidence that increasing agent transparency helps operators to calibrate trust and improves operator performance in simulated mission planning [2]. What is less clear is how agents' approach to communicating this information might affect these outcomes. The current effort explores the effects of agent transparency and communication framing on operators' perceptions of the agent and decision making with agent assistance.

A. Agent Transparency

The communication of information about an automated system's inner workings—including what it is doing and why—has generally been described as agent transparency [3]. To facilitate this communication in a manner consistent with models of human situation awareness, Chen and colleagues [3] developed a model of Situation awareness-based Agent Transparency (SAT). This model holds that transparency can be operationalized by mirroring human situation awareness—as perception, comprehension, and projection [4]. At level one of SAT (L1), the agent communicates its status, goals, and understanding of current environmental conditions (*perception*). At level two (L2), the agent communicates its reasoning for executing or recommending specific actions (*comprehension*). At level three (L3), the agent communicates its expectations of future states (e.g., likelihood of mission success; *projection*). Projection uncertainty (U) may also be communicated. Research using the SAT model has indicated that increasing agent transparency can improve performance in human-agent teams across several contexts [3]. What remains unclear is how an agent's attitude or style of communication may alter the effect of transparency.

B. Agent Communication Framing

Even when not perceptible, affective components of human to human communication can influence perception of communicated information [5]. This is also true of machine to human communication as humans respond to machines in a manner similar to that which they respond to other humans [6]. For example, Parasuraman and Miller found that machine etiquette had important consequences for human perception and performance; automated systems that were non-interruptive and patient elicited higher trust and greater likelihood of successful fault diagnosis compared to systems that were interruptive and impatient [7].

Relatedly, agent transparency communication that draws attention to certain types of information could be thought of as a type of framing, or structuring of information such that it leads to some type of bias. An example of this is attribute framing, in which an attribute of an object is described in either positive or negative proportions (e.g. the glass is half empty or half full; [8]). Historically, it is known that framing affects evaluations of objects (c.f., [9]), but how does the framing of something abstract, such as choice parameters, affect decision making? In the context of our study, where a human operator is making a decision (i.e. "play-calling" in the context of multi-robot management; see Method) based on a set of parameters presented by an IA, positive or negative framing of the operator's decision by the agent may affect the human operator's trust and perception of the agent.

C. Application

While recent work indicates that high transparency communication results in more effective operator trust calibration [2], increased transparency may also cause some problems. For example, when automation is reliable but makes errors in judgment uncharacteristic of that which a human teammate might make, human operators may disuse the automation, even if they know the automation's performance is more reliable than their own [10]. This has been attributed to the fact that information inconsistent with expectation is likely to be well remembered; thus automation mistakes may have undue weight relative to correct automation decisions [10].

Higher automation transparency risks revealing anomalous judgement inconsistencies. However, enough transparency can also help the operator to better understand why the misjudgment occurred. Dzindolet et al.'s third study revealed that information which served to explain automation error was instrumental in trust recovery following automation error and increased trust and reliance in the automation overall [10]. Unfortunately, the reliance regained by explaining error was not always discerning; participants tended to trust reliable and unreliable aids indiscriminately and regardless of the nature of errors the automation tended to commit. Further, time may not always be available for operators to review increased transparency information. Under these circumstances, availability of large amounts of transparency information may contribute to feelings of overload. Work has suggested that people tend to hold onto control when they become overloaded, even when offloading to automation would improve performance (e.g., [11]). Thus, it is important to understand how best to communicate attenuated levels of transparency information (e.g., SAT L1+2, but not L3 or uncertainty) [3].

D. Goal

This current effort is part of the IMPACT program [12], supported by U.S. Department of Defense's Autonomy Research Pilot Initiative. The objective of IMPACT is to investigate issues associated with flexible play-calling, global cooperative control and local adaptive/reactive capability, and bidirectional human-autonomy interaction in military mission contexts [12]. The current study focuses on the human factors aspects of the agent's transparency and communication framing in the context of human-autonomy teaming via IMPACT technologies and capabilities.

The overall goal of the present study was to understand the interaction between level of agent transparency communication, according to the SAT model [3], and the agent's framing of communication. In line with Dzindolet et al.'s findings [10], we expect trust and evaluation of the agent to be higher with a high transparency interface than with a low transparency interface. When the agent is more transparent, and critical of the participant's plan decisions (critical framing), it will be perceived better and trusted more than a complimentary agent as it highlights reasons for error. On the other hand, we expect that a more opaque yet complimentary agent will increase trust in the agent by helping to counteract the overweighing of faults.

II. METHOD

A. Participants

Twenty-nine students from an American university were recruited for cash payment. Data were analyzed for 26 (17 men, 12 women, M_{age} = 20.03, SD_{age} = 2.09). Three were omitted from analysis due to technical issues.

B. Design and Apparatus

This experiment involved a 2x2 mixed design with agent transparency as the within-subjects independent variable and communication framing as the between-subjects independent variable. Agent transparency was tested at two levels: (a) L1+2: containing reasoning information, and (b) L1+2+3+U containing reasoning and projection *with* projection uncertainty information. Communication framing was tested as two contrasting attitudes from the agent: (a) Critical: highlighting a parameter of the chosen plan that is not satisfied, and (b) Complimentary: highlighting a parameter of the chosen plan that is optimal. The user interface varied per condition by showing corresponding pieces of SAT-level information on a map display, in text, and on a sliding bar scale (Fig. 1). Prior to the experimental trials, participants received about 1 hour of training. The experiment was divided into 2 blocks of 8 missions. Transparency order and communication framing were counterbalanced within sets of four participants, within which the scenarios where the agent's recommendations were correct and incorrect were held constant. The choice of correct and incorrect scenarios was randomized for each set but kept the 5 correct and 3 incorrect ratio described below.

Using the interface presented in Fig. 1, participants were presented with two plan options to complete each mission. Both plans were viable and were balanced in their fulfillment of the mission needs, but optimal plans prioritized the commander's intent and addressed more mission goals and needs. The agent always recommended one plan over the other, and the participant's task was to select the best plan based on current mission requirements. After the participant made the initial selection, an update informed the participant and agent of an aspect of the scenario which had evolved after the initial plans were generated (Fig. 1; faded in image at the bottom). Once the update was acknowledged, the agent re-assessed the plans and informed the participant of a parameter that had either *remained or become suboptimal* (criticizing condition; e.g., "Based on this update, the time (speed) aspect of the current plan selection (A) remains suboptimal. Please consider this when reevaluating plan selection") or had *remained or become optimal* (complimentary condition). The participant confirmed or changed the plan based on the new information. While the transparency manipulation applied throughout the entire interaction, the communication framing manipulation only applied post-update. Reliability of the agent

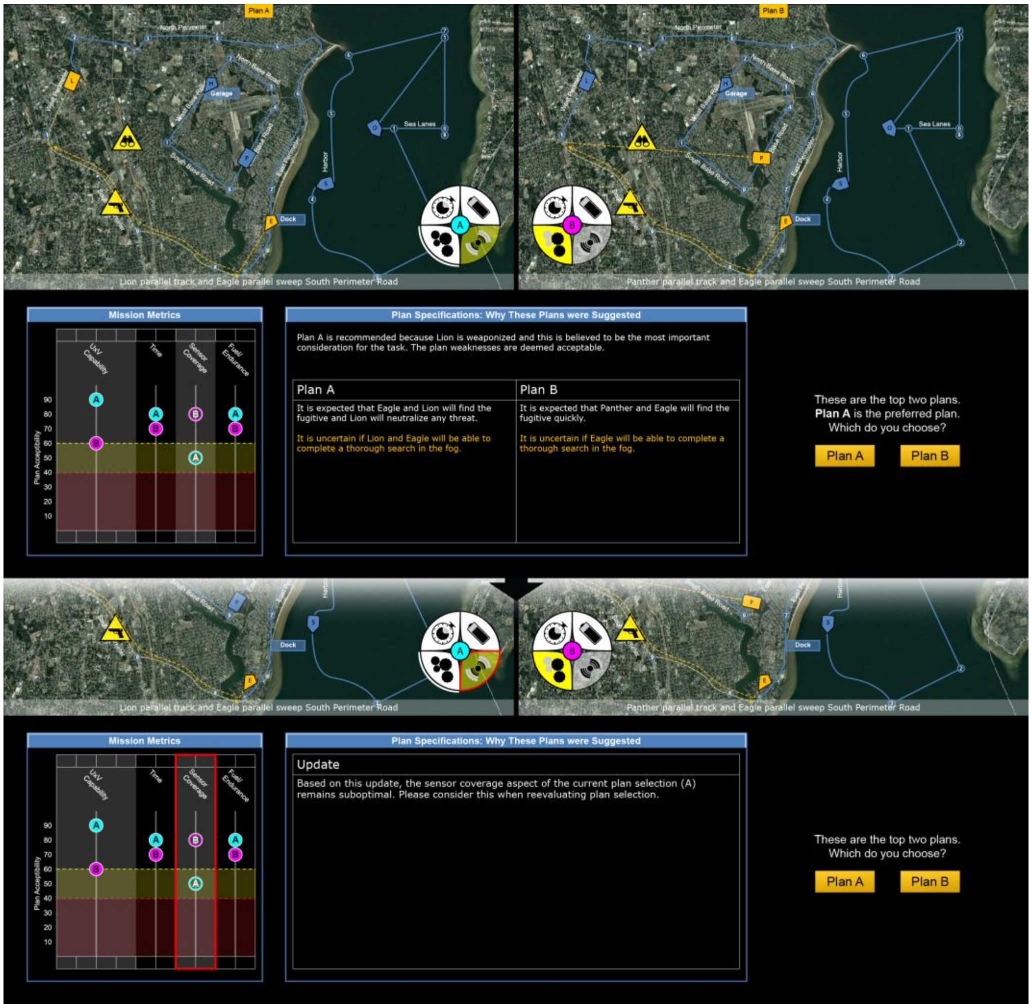

Fig. 1. Multiple-heterogeneous UxV control mission plan decision (top) and mission update decision (bottom; faded in) screenshots.

was held constant such that its initial recommendation was correct for 5 of every 8 scenarios. The agent's post-update assessment was correct in all scenarios (though no actual recommendation was made for the post-update plan decision). Participants assessed their trust and perception of the agent after each block of scenarios. Within each block, participants also filled out items for usefulness and reliance (described in Sect. II-D) following each of the eight initial mission plan decisions.

C. Performance Measures

1) Operator Task Performance: Task performance was operationalized as the number of scenarios within each block that the participant chose the best option (as defined above).

2) Agreement: Agreement was the number of scenarios within each block that the participant agreed with the agent's recommended mission plan selection (always Plan A).

D. Subjective Measures

1) Trust Survey 1: This recently developed trust measure *[13]* measures human-machine trust in a way consistent with human-human trust. This trust survey was developed based on the dimensions of trust defined by Mayer and colleagues *[14]*. Participants respond to statements about their willingness to trust the agent using a 7-point Likert scale anchored by "strongly disagree" (1) to "strongly agree" (7).

2) Trust Survey 2: This 12-item measure is a modified version of Jian, Bisantz, and Drury's *[15]* Checklist for Trust between People and Automation. Participants responded to items regarding two aspects of the agent: 1) integrating and displaying analyzed information, and 2) suggesting and making decisions *[16]*. Responses made on a 7-point Likert scale anchored by "not at all" (1) and "extremely" (7).

3) Perceived Agent Aptitude: This measure of operator trust consisted of four trust-related items of the Technology Acceptance Model *[17]*. Participants responsed to these items with regard to the two aspects of the agent describe above for Trust Survey 2 using the same 7-point scale.

4) Perceived Agent Reliability: Participants rated their perception of the agent's reliability, using a slider scaling from 0% to 100%.

5) Scenario Specific Perceptions: Two measures were administered after each initial mission plan decision. The first asked participants to rate the usefulness of each display that communicated agent transparency information (map display, text display, and sliding bar scale display). The second asked participants to rate their reliance on each display. They responded using a 7-point Likert-type scale anchored by “Not at All” (1) and “Completely” (7) which overlayed each of the three displays. All response to all components were averaged for comparisons between blocks of scenarios.

III. Results

To investigate the influence of both transparency and framing, as well as to account for order effects, a series of 2 (low & high transparency) x 2 (complimentary & critical framing) x 2 (transparency block order) mixed model ANOVAs were run with transparency as the within-subjects variable. We reported here only main effects and interactions for transparency and framing; including order in the analysis is important to account for order effects, but is not of specific interest here. Correlations were then run to help with interpretation. In light of relatively low statistical power, we reported trends with a p-value greater than .05 but less than .10.

A. Performance

There were no significant task performances differences. However, for agreement with the agent, there was a main effect for transparency, $F(1,26) = 4.28$, $p = .049$, $\eta^2_p = .141$, as well as an interaction between transparency and framing, $F(1,26) = 5.35$, $p = .029$, $\eta^2_p = .171$. Agreement with the complimentary agent was consistent between transparency conditions. In contrast, agreement with the critical agent was higher in the low transparency automation condition than in the high transparency condition (see Fig. 2).

B. Trust

Trust results were mixed. Trust Survey 1 revealed no significant findings, but a main effect trend for framing $F(1,26) = 3.97$, $p = .057$, $\eta^2_p = .132$. Participants were more trusting when the agent was critical (Fig. 3). With regard to trust in the agent’s ability to integrate and display analyzed information, Trust Survey 2 revealed a significant main effect for transparency, $F(1,26) = 6.49$, $p = .017$, $\eta^2_p = .200$. There were no other significant findings, but a trend suggests the possibility of an interaction between transparency and framing were there more statistical power, $F(1,26) = 3.40$, $p = .077$, $\eta^2_p = .116$. Participants were relatively distrustful of the low transparency complimentary agent (Fig. 4). There were no significant findings from Trust Survey 2 with regard to the agent’s ability to suggesting and making decisions.

C. Perceptions of Agent

There were main effects of both transparency, $F(1,26) = 4.89$, $p = .036$, $\eta^2_p = .158$, and a non-significant interaction trend between the variables, $F(1,26) = 3.93$, $p = .058$, $\eta^2_p = .131$, for perceptions of agent aptitude with regard to integrating and displaying information. Regarding perceptions of agent aptitude in suggesting plan decisions, there were also main effects for transparency, $F(1,26) = 5.43$, $p = .028$, $\eta^2_p = .173$, and framing, $F(1,26) = 4.78$, $p = .038$, $\eta^2_p = .155$, but an interaction was not significant $p > .10$. In both cases, participants perceived the agent to be more apt when transparency was high. Perceptions of agent aptitude when integrating and displaying information were higher when the agent framed the update plan critically (Fig. 5). For suggesting decisions, the interaction was driven by the difference between the low transparency complimentary condition and the other three conditions. There were no significant (or trends of) differences between condition in perceived automation reliability ($p > .10$).

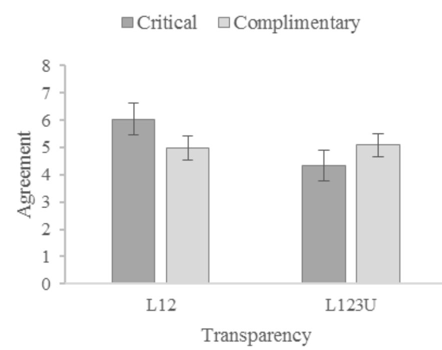

Fig. 2. Mean number of scenarios per block participants agreed with automation. Bars are standard error.

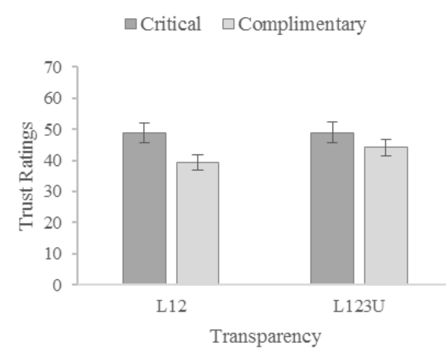

Fig. 3. Trust Survey 1 ratings (out of 70). Bars are standard error.

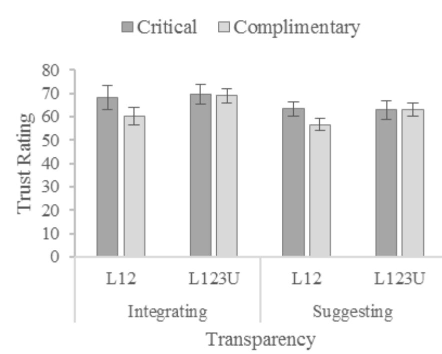

Fig. 4. Trust Survey 2 ratings (out of 84) for trust in the automation's ability to integrate and display analyzed information, and to suggest and make decisions. Bars are standard error.

D. Post Scenario Ratings

There was a strong main effect of transparency, $F(1,26) = 49.59$, $p < .001$, $\eta^2_p = .656$, and of framing $F(1,26) = 11.87$, $p = .002$, $\eta^2_p = .314$, for ratings of automation interface usefulness; participants found the high transparency and critical agent configurations to be more useful, respectively (see Fig. 6). For reliance, there were also main effects for transparency, $F(1,26) = 46.82$, $p < .001$, $\eta^2_p = .643$, and framing, $F(1,26) = 12.56$, $p = .002$, $\eta^2_p = .326$. Participants reported more reliance when transparency was high and the update assessment was critical in nature (Fig. 7).

E. Correlational Analyses

To help put the above findings into context, a series of correlations were run for each condition between agreement with the agent and the trust and agent perception measures (see Table I). Except for agent reliability, trust and perception of the agent were generally more highly correlated with agreement in the low transparency condition.

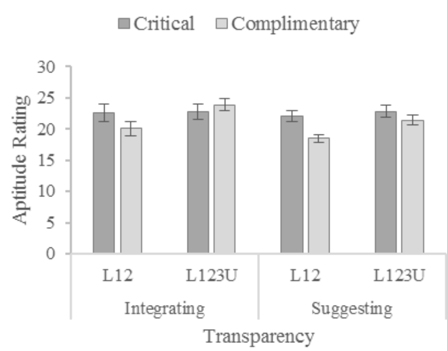

Fig. 5. Automation aptitude ratings (out of 28) with regard to integrating and displaying analyzed information, and suggesting and making decisions. Bars are standard error.

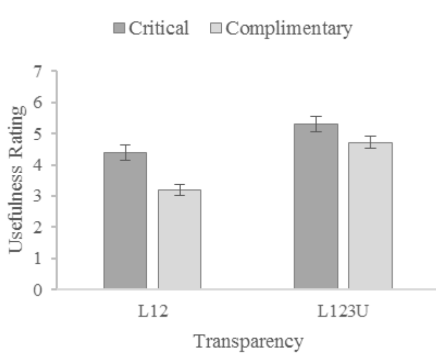

Fig. 6. Ratings of usefulness of the automation's transparency information interface. Bars are standard error.

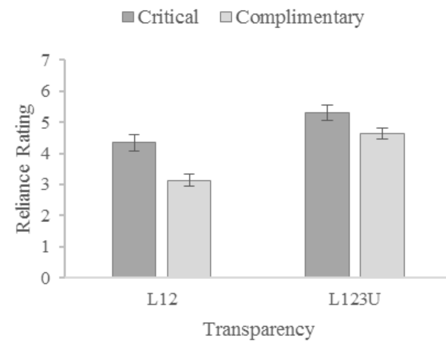

Fig. 7. Reliance on the automation's transparency information interface.

TABLE I.

CORRELATION OF AGREEMENT WITH TRUST AND AUTOMATION PERCEPTION

	L12	U
Trust Survey 1	.41*	.31
Trust Survey 2 (Integrating)	.12	-.11
Trust Survey 2 (Suggesting)	.43*	.20
Agent Aptitude (Integrating)	-.07	-.01
Agent Aptitude (Suggesting)	.21	.03
Agent Reliability	.47**	.45*
Reliance	.39*	-.06
Usefulness	.38*	-.10

* $p < .05$, ** $p < .01$

IV. DISCUSSION

Our predictions centered around two assumptions about the nature of our tasking paradigm. First, the low transparency condition was thought to provide enough information (decision, and clues about reasoning) to allow participants to identify flaws in the agent reasoning, but not enough to draw strong conclusions about why those flaws existed. Accordingly, outcome predictions were based on the results of Dzindolet and colleague's [10] study 2, which had similar features (decision knowledge, but no accuracy rate feedback). The second assumption drew parallels between our high transparency condition and some of the features of Dzindolet et al.'s study 3. Specifically, the high transparency agent communicated uncertainty about its projections, which highlighted the weaknesses of the recommendations it generated just as the participants in Dzindolet et al.'s study 3 received information about the weaknesses of the automation they used. Using these parallels as a basis, we believed that agreement, trust, and perceptions of the agent would be higher in the high transparency condition in agreement with Dzindolet et al.'s outcomes (study 2 vs study 3).

The results told a different story. With the notable exception of agent reliability perceptions, perceptions of the agent increased with transparency as predicted. Trust Survey 2 also yielded an increase with transparency regarding integration and displaying of information. This increase only emerged for the complimentary agent, which was trusted less in the low transparency condition. Agreement outcomes were the reverse of our expectations: participants agreed more with the low transparency agent. As with trust this outcome affected only one communication framing style: critical. This agreement outcome appears to resemble the positivity bias toward automation in Dzindolet and colleague's study 1 more than the high self-reliance outcomes found in study 2. Trust and agreement outcomes taken together suggest that trust and agreement in a low transparency agent might be more vulnerable to the influence of affective aspects of communication. This is important as the correlational analyses showed that agreement was related to participants' trust and perception of the agent to a much greater extent in the low transparency condition than in the high transparency condition.

Regarding communication framing, we expected that when transparency was high, the critical agent would be trusted more, perceived more positively, and agreed with more frequently. This hypothesis was based on the idea that criticism of plan options, which were often selected by the participant in agreement with the agent's recommendation, would highlight the weaknesses in agent's abilities, conveying a sense of self-awareness of the imperfections in the plans it recommends. Indeed, the critical agent was perceived as having more aptitude, reliability, and usefulness than the complimentary agent when transparency was high. In the low transparency condition, we expected the complimentary agent to receive higher trust, perception ratings, and agreement, as the positive affect (and even camaraderie) induced by the complimentary agent was expected to help lessen the tendency to overweigh anomalous agent faults. Contrary to this expectation, it was the critical agent that was more trusted, perceived more positively, and agreed with more frequently. It is possible that overemphasis on errors was not at issue because the level of transparency did not afford high enough rates of error detection in agent logic - the agent may even have benefitted from a positivity bias). Also, participants may have been risk averse and more comfortable with an agent that called out problems.

This study provided a preliminary look at how agent framing regarding the fitness of past decisions to evolving conditions in multi-UAV control affects operators' judgements of the agent. However, our interpretation of the agreement findings should be considered with caution. Although trust and perceptions of the agent were more highly related to agreement in the low transparency condition than in the high transparency condition, it is unclear how much agreement in each case was misuse. A future effort must look at the different components of agreement (e.g., [2]) to determine whether it is beneficial that the role of trust and perceptions of the agent in compliance behavior decreases as transparency increases. An additional limitation is the manipulation of framing. Our operationalization of communication framing is more complex than the "Glass half empty / half full" scenario traditionally presented in framing research (see [8]). By changing the way communication was framed, we changed the information itself (by focusing on different parameters as positive or negative). Future research might explore results of framing that uses consistent information.

ACKNOWLEDGEMENT

This research is supported by the U.S. Department of Defense's Autonomy Research Pilot Initiative. Our views and conclusions do not represent the official policies or position, either expressed or implied, of the U.S. Army Research Laboratory, or U.S. Government. We thank contributors Olivia Newton, Daniel Barber, Jonathon Harris, Jack Hart, Alexis San Javier, Harrison Spellman, Gloria Calhoun, and Mark Draper.

REFERENCES

[1] J. Y. C. Chen and M. J. Barnes, "Human-agent teaming for multirobot control: A review of human factors issues," *IEEE Trans. Hum.-Mach. Syst.*, vol. 44, no. 1, pp. 13–29, Feb. 2014.

[2] J. E. Mercado, M. A. Rupp, J. Y. Chen, M. J. Barnes, D. Barber, and K. Procci, "Intelligent agent transparency in human–agent teaming for Multi-UxV management," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 58, no. 3, pp. 401–415, 2016.

[3] J. Y. C. Chen, S. G. Lakhmani, K. Stowers, A. R. Selkowitz, J. Wright, and M. J. Barnes, "Situation awareness-based agent transparency and human-autonomy teaming effectiveness," *Theor. Issues Ergon. Sci.*, in press.

[4] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 37, no. 1, pp. 32–64, Mar. 1995.

[5] L. Donohew, H. E. Sypher, and E. T. Higgins, Eds., *Communication, social cognition, and affect*. Hillsdale, N.J: L. Erlbaum Associates, 1988.

[6] B. Reeves and C. Nass, The media equation: how people treat computers, television, and new media like real people and places, 1. paperback ed., [Reprint.]. Stanford, Calif: CSLI Publ, 1996.

[7] R. Parasuraman and C. A. Miller, "Trust and etiquette in high-criticality automated systems," *Commun. ACM*, vol. 47, no. 4, pp. 51–55, 2004.

[8] I. P. Levin, S. L. Schneider, and G. J. Gaeth, "All frames are not created equal: A typology and critical analysis of framing effects," *Organ. Behav. Hum. Decis. Process.*, vol. 76, no. 2, pp. 149–188, Nov. 1998.

[9] I. P. Levin and G. J. Gaeth, "How consumers are affected by the framing of attribute information before and after consuming the product," *J. Consum. Res.*, vol. 15, no. 3, p. 374, Dec. 1988.

[10] M. T. Dzindolet, S. A. Peterson, R. A. Pomranky, L. G. Pierce, and H. P. Beck, "The role of trust in automation reliance," *Int. J. Hum.-Comput. Stud.*, vol. 58, no. 6, pp. 697–718, Jun. 2003.

[11] J. Lin *et al.*, "Automation reliability and other contextual factors in multi-UAV operator selection," *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 60, 2016.

[12] M. Draper et al. "Intelligent Multi-unmanned vehicle Planner with Adaptive Collaborative control Technologies (IMPACT),". *Proc. Int. Symp. Aviation Psychology*, 2017.

[13] J. B. Lyons, K. S. Koltai, N. T. Ho, W. B. Johnson, D. E. Smith, and R. J. Shively, "Engineering trust in complex automated systems," *Ergon. Des. Q. Hum. Factors Appl.*, vol. 24, no. 1, pp. 13–17, Jan. 2016.

[14] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An integrative model of organizational trust," *Acad. Manage. Rev.*, vol. 20, no. 3, pp. 709–734, 1995.

[15] J.-Y. Jian, A. M. Bisantz, and C. G. Drury, "Foundations for an empirically determined scale of trust in automated systems," *Int. J. Cogn. Ergon.*, vol. 4, no. 1, pp. 53–71, Mar. 2000.

[16] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation," *Syst. Man Cybern. Part Syst. Hum. IEEE Trans. On*, vol. 30, no. 3, pp. 286–297, 2000.

[17] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Q.*, vol. 13, no. 3, p. 319, Sep. 1989.