



AFRL-RH-WP-TR-2020-0090

HUMAN MACHINE TEAMING (HMT): TRUST CUES IN COMMUNICATION AND BIAS TOWARDS ROBOTIC PARTNERS

April Rose Panganiban, Ph.D.

Michael D. Long
AFRL/711 HPW/ RHWS

Gerald Matthews, Ph.D.

University of Central Florida
IST, SIMITAR Laboratory

September 2020

Final Report

DISTRIBUTION A. Approved for public release; distribution unlimited.

**AIR FORCE RESEARCH LABORATORY
711TH HUMAN PERFORMANCE WING
AIRMAN SYSTEMS DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE OH 45433
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE**

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 88th Air Base Wing Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (<http://www.dtic.mil>).

AFRL-RH-WP-TR-2020-0090 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//signed//

APRIL ROSE PANGANIBAN, Ph.D.
Work Unit Manager
Collaborative Interfaces and Teaming Branch
Warfighter Interactions and Readiness Division
Airman Systems Directorate
711th Human Performance Wing
Air Force Research Laboratory

//signed//

TIMOTHY S. WEBB, Ph.D., DR-IV
Chief, Collaborative Interfaces and
Teaming Branch
Warfighter Interactions and Readiness Division
Airman Systems Directorate
711th Human Performance Wing
Air Force Research Laboratory

//signed//

LOUISE A. CARTER, Ph.D., DR-IV
Chief, Warfighter Interactions and Readiness Division
Airman Systems Directorate
711th Human Performance Wing
Air Force Research Laboratory

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

| | | | | | |
|--|--------------------|--------------------------------|-----------------------------------|---|--|
| 1. REPORT DATE (DD-MM-YYYY) 09/10/2020 | | 2. REPORT TYPE Final | | 3. DATES COVERED (From - To) 10/2014-10/2020 | |
| 4. TITLE AND SUBTITLE Human Machine Teaming (HMT): Trust Cues in Communication and Bias Towards Robotic Partners | | | | 5a. CONTRACT NUMBER In-House | |
| | | | | 5b. GRANT NUMBER 16RHCOR367 | |
| | | | | 5c. PROGRAM ELEMENT NUMBER | |
| 6. AUTHOR(S) *April Rose Panganiban, PhD., ** Gerald Matthews, Ph.D. *Michael D. Long | | | | 5d. PROJECT NUMBER | |
| | | | | 5e. TASK NUMBER | |
| | | | | 5f. WORK UNIT NUMBER HOPZ | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) *711 HPW/RHWS 2210 Eighth Street Wright-Patterson AFB, OH 45433 | | | | 8. PERFORMING ORGANIZATION REPORT NUMBER | |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Materiel Command Air Force Research Laboratory 711th Human Performance Wing Airman Systems Directorate Warfighter Interactions and Readiness Division Collaborative Interfaces and Teaming Branch Wright-Patterson AFB, OH 45433 | | | | 10. SPONSOR/MONITOR'S ACRONYM(S) | |
| | | | | 11. PONSOR/MONITOR'S REPORT NUMBER(S) AFRL-RH-WP-TR-2020-0090 | |
| 12. DISTRIBUTION/AVAILABILITY STATEMENT Distribution A. Approved for public release distribution unlimited | | | | | |
| 13. SUPPLEMENTARY NOTES 88ABW-2020-3084; Cleared 7 Oct 2020 | | | | | |
| 14. ABSTRACT This technical report outlines the aims and results of three areas of research related to Trust and Human-Machine Teaming (HMT): teammate-focused agent communications (Study 1), visual socio-emotional cues (Study 2), and human biases in response to machine information processing capabilities (Studies 3 and 4). Study 1 explored the manipulation of social factors from a robotic partner in a study context of a, minimalistic, F-16 mission study. Study 2 followed a similar vein of manipulating socio-emotional cues from a computer agent in an Intelligence, Surveillance and Reconnaissance (ISR) task. The final two studies observed biases towards machine partner processing. The summarized work addresses two key results related to designing HMT interactions: existing biases towards the role of the machine teammate and the ability to perceive teammate actions from a machine partner. How machines react to human partners conveys teaming information. In contrast, what the machine is called upon to analyze may stir negative biases towards robotic partners which can harm the HMT dynamic. | | | | | |
| 15. SUBJECT TERMS Trust, communication, human-machine teaming, mental models, individual differences | | | | | |
| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | April Rose Panganiban |
| U | U | U | SAR | 30 | 19b. TELEPHONE NUMBER (Include area code) |

TABLE OF CONTENTS

| | |
|--|----|
| List of Figures | ii |
| List of Tables | ii |
| 1.0 SUMMARY | 1 |
| 2.0 INTRODUCTION | 3 |
| 2.1 Trust and Transparency in HRT | 4 |
| 2.2 Promoting Trust through Affect Management | 6 |
| 2.3 F-16 Study: Teaming with Autonomous Collaborative Technology (TACT)..... | 8 |
| 2.4 ISR Study: Human-Autonomy Teaming (HAT)..... | 9 |
| 2.5 ROTA Study..... | 9 |
| 3.0 METHODS | 10 |
| 3.1 F-16 Study: TACT | 10 |
| 3.2 ISR Study: HAT | 10 |
| 3.2.1 Difficulty Parameters | 11 |
| 3.2.2 Avatar Expression..... | 11 |
| 3.3 Robot Threat Assessment (RoTA) Scale Studies..... | 13 |
| 3.3.2 Simulation Study..... | 14 |
| 3.3.2.1 Simulation Environment: | 14 |
| 4.0 RESULTS AND DISCUSSION | 16 |
| 4.1 F-16 Study | 16 |
| 4.2 ISR Study. | 17 |
| 4.3 ROTA Studies | 19 |
| 4.3.1 Mental Models Study..... | 19 |
| 5.0 CONCLUSION..... | 20 |
| 6.0 REFERENCES | 21 |
| 7.0 LIST OF ABBREVIATIONS, ACRONYMS AND SYMBOLS | 25 |

LIST OF FIGURES

| | | |
|-----------|---|----|
| Figure 1. | Leite et al. (2013) Chart of Empathetic Robotic Partner’s Expressions. | 8 |
| Figure 2. | Example of a Trial (Psychological Analysis)..... | 15 |

LIST OF TABLES

| | | |
|----------|--|----|
| Table 1. | Adaptation of Lyons’ (2013) Conceptualization of HRI Transparency Separated by Signaling Type..... | 6 |
| Table 2. | Potential Dimensions of Individual Differences in Mental Models for Interacting with Advance Robots..... | 6 |
| Table 3. | Leite et al. (2013) Interpretation of Robotic Partner’s Expectation Compared with Operator Behavior Driving Partner’s Displayed Expressions and Communication during a Chess Match | 8 |
| Table 4. | All Images are Examples, Based on FaceGen Demo Version. | 12 |
| Table 5. | Summary of Support for Hypotheses | 17 |

1.0 SUMMARY

This technical report outlines the aims and results of three areas of research related to Trust and Human-Machine Teaming (HMT): teammate-focused agent communications (Study 1), visual socio-emotional cues (Study 2), and human biases in response to machine information processing capabilities (Studies 3 and 4). Study 1 explored the manipulation of social *factors* from a robotic partner in a study context of a minimalistic, F-16 mission study. This study applied cues of benevolence and partnering in an Air Force context to positively impact subjective ratings of cognitive/emotional factors such as workload, teaming and trust. Such research is dominant in the civilian sector of robotics and computer aiding agents but *minimally* applied to the military domain of autonomy in mission contexts. The findings support the benefit of designing an autonomous wingman to display benevolent intent in a task environment, seen as a reduction of reported subjective workload and increased feelings of team cohesion, specifically in team *collaboration*.

The second study followed a similar vein of manipulating socio-emotional cues from a computer agent in an Intelligence, Surveillance and Reconnaissance (ISR) task. In this study, cues were provided via facial expressions which were noticed by individuals with higher self-reported attentional control and reported trust in machine partners. Additionally, individuals perceived higher team collaboration when the machine partner was affectively responsive, seen as emoted concern when Unmanned Aerial Vehicles (UAV) could sustain damage.

In contrast to the manipulation of machine partner output, the final two studies examined biases towards machine partner processing. A survey performed in Amazon's Mechanical Turk, consisting of various human-machine threat detection scenarios in which a robotic partner processes data and advises on the threat, revealed that individuals differentiate between information processing which results in psychological as opposed to purely physical evaluations.

This work provides insight into the difficult issue of transitioning machine partners from tools to teammates. Improvements in technology allow a machine partner to carryout teamwork behaviors of coordination and task reallocation (O'Neill, McNeese, Barron & Schelbe, 2020) however this raises its autonomy level and potentially increases workload, lengthens decision-making time, reduces situation awareness and encumbers informed trust (Endsley, 2015). Additionally, whereas the operator's trust assessment was in information, her or his teammates, and the technology, now the machine system is the source of information, the processing unit for this information and also a supportive teammate (Endsley, 2015). In the current set of basic experimental work, the effect of displaying social, teammate behavior did positively raise perceptions of team behavior, but with distinct differences in trust. Taken together, these findings inform HMT researchers that design of communication style can elicit teammate responses from a human operator toward its machine partner. More specifically, communicating awareness of the impact of one's performance on the team is beneficial to perceived workload and is attainably demonstrated in the F-16 study. Attention towards social cues are perceivable, however, any benefit for state factors and team attitudes may depend on biases towards robotic partners, in general, the type of analysis the partner is performing. And so, these studies also

point towards the future of HMT assessment where individual biases towards robots should be carefully considered when interpreting the outcomes of any HMT interaction. Negative attitudes towards technology may prevent the perception of teammate qualities in a robotic partner and may affect trust in that partner. HMT trust measures may need to be constructed to differentiate the social as opposed to performance aspects of the robotic partner, which may map onto differences between team and task-related attitudes toward a teammate.

2.0 INTRODUCTION

There is an extensive literature on effective human-human teaming (for review see Cannon-Bowers & Bowers, 2011; Marks, Mathieu & Zaccaro, 2001) which may inform research on Human-Robot Teaming (HRT). A key factor is the extent to which team members share similar cognitions (Salas, Cooke & Rosen, 2008). Such cognitions include not only knowledge of the task but also social cognitions (Rentsch & Woer, 2002). For example, it is important to understand what goals a team member is pursuing, and what that person expects from other team members to attain those goals. Interpersonal processes include action processes, involving monitoring steps toward goals, and system status, assisting team members and coordinating activities, as well as management of conflict, motivating of teammates and management of member emotions (Marks et al., 2001). Beyond processing focused on the task itself, team processing activity may thus involve effective regulation of the emotions of oneself and others. Effective regulation may reflect personal qualities such as empathy and ability to read the intentions of others ('theory of mind').

Similarly, in relation to HRT, humans and robots must share at least some knowledge of the task, such as the priorities of different sub-goals, as well as knowledge of how the other entity is functioning, such as which sub-goals the other is capable of executing. Teaming effectiveness might then reflect two cycles of information flow. These correspond to the distinction made between task-directed and state-directed effort (Mulder, 1986). First, information directly related to task execution must pass between human and robot. In traditional human-computer interaction, the computer feeds information to the human who issues instructions to the computer. Teaming applications differ in that the robot may take independent actions, depending on its level of automation (LOA), requiring that both entities share task-relevant knowledge. Thus, the human must trust the competency of the robot (Hancock et al., 2011; Ososky, Schuster, Phillips & Jentsch, 2013). Also, the human and robot must exhibit and process coordination behaviors to facilitate teaming (Shah & Breazeal, 2010). Coordination behaviors may be explicit (e.g., commands or prompts) or implicit (e.g., gestures, gaze).

A second information cycle that may be especially important under high workload and stress refers to the functional status ("state") of human and robot. The human may signal to the robot that he/she is under high workload, either directly or implicitly through physiological or facial cues (e.g., Barber, Reinerman-Jones, Lackey & Hudson, 2011). Such signals may drive robot behavior according to principles of adaptive automation. In some circumstances, physiological signals may be used to drive automation directly. However, there are challenges associated with the limited diagnostic ability of individual physiological metrics for workload and stress, and research needs to interpret these metrics in light of the task context and the operator's motivations (e.g., Blascovich, 2008; cf. Vine et al., 2013). An alternative strategy is to use implicit measures of state, such as physiological metrics, to drive the robot partner to volunteer support for the human, as opposed to driving full automation behavior. If overload or stress is detected, the robotic partner might inquire about the human's motivational state, intended course of action, and needs for support, i.e., a response that resembles more closely the richness of human-human teaming and may avoid the negative consequence of overriding the operator's strategy. The robot may then be perceived as "benevolent", a key element of human-human trust, or at the very least, this approach may reduce frustration and irritation with the system. Perceived benevolence is a key aspect of interpersonal trust (see Mayer, Davis, & Schoorman, 1995) yet it is unclear if or how this construct applies to Human Robot Interaction (HRI).

Correspondingly, the robot may signal a need for the human to take on more responsibility, either explicitly or implicitly, such as a special case where the robot is damaged and unable to perform all of its functions. If the robot is accurate and 'honest' about its capabilities and reliability, these signals may

build trust in the human. If a system is unable to perform a task, communication of its state may provide the operator information with which to appropriately calibrate trust. The state-directed cycle may be facilitated where the human and robot have ‘empathy’, in the sense of rapid, holistic understanding of the current functional capabilities of the other entity.

2.1 Trust and Transparency in HRT

Standard models of trust (e.g., Mayer et al., 1995) do not accommodate novel factors that influence trust in more humanlike robots capable of intelligent autonomous action, social interaction, and communicating task- and team-related issues (Matthews et al., 2016). Beyond perceptions of robot competence to perform its key functions, trust is also impacted by the operator’s appraisals of the technologically advanced attributes of the machine. Matthews et al. (in press) identified comprehensibility and social agency as two such attributes. Trust may be eroded if the human cannot understand the basis of the machine’s evaluations and actions. Comprehensibility may be reduced especially when the robot is driven by advanced algorithms such as neural networks, “big data” analysis, or machine learning. The implementation of social agency may also impact trust. Robot social functionality that supports effective teaming via the two information cycles previously discussed is likely to enhance trust. However, poor design may undermine perceptions of robot benevolence and integrity, for example, if the robot seeks to exert unwanted control over team activities, or if it appears to have an agenda different from the human’s. Thus, the operator’s perceptions of the robot’s cognitions and intentions are critical for trust (Lyons & Havig, 2014), unlike trust in conventional, “dumb” machines.

Processing of task- and state-related information may be associated with rather different trust issues in advanced robots. In the task context, trust is likely to reflect perceptions of robot competence, which may be over or under-estimated. Similar to both H-H teams and conventional machines, the perception of robot ability will impact reliance (Hancock et al., 2011). Perceptions of teammate ability and experience are positively related to request and acceptance of backup behavior in the high consequence event (HCE) environment of air traffic control (Smith-Jentsch, Kraiger, Cannon-Bowers, & Salas, 2009). As in other contexts for automation (Parasuraman & Wickens, 2008), there are dangers in both under-trust, leading to undue micromanagement, which can be viewed as poor team resource management, and over-trust, leading to neglect of critical task activities. HCEs provide the additional element of time pressure on trust, and high stakes for misjudgments. Typically, high stakes would tend to encourage decision-making effort and explicit processing, but alternatively time pressure may cause decisions on whether to trust the robot to execute a task to become more implicit and intuitive, and potentially subject to cognitive bias (cf., Kahneman, 2003).

In human teaming, a teammate’s intervention to take on one’s tasks might be interpreted as either genuinely helpful, or as an expression of lack of trust in oneself. That is, the intervention may be attributed to the teammate perceiving oneself as incompetent, which is likely to be stressful and harmful to the team process. Negative evaluations may reduce trust in teammates by harming the perceived benevolence of the evaluative teammate (Mayer, et al., 1995). A reduction in trust may reduce teamwork processes of coordination, specifically sharing state information and expressing a need for assistance, which puts the teammate in a vulnerable position (Bergman, Small, Bergman, & Rentsch, 2010).

In HRT, it is similarly important that the human trusts the robot to promote effective team functioning. Trust entails confidence in both the robot’s ability to evaluate human workload and stress, and in its ability to react appropriately. However, trusting the robot to manage teaming issues takes the human into

unfamiliar territory. We are more accustomed to evaluating machines than to being the subject of a machine evaluation. Ideally, the robot “taking control” should be appraised as supportive of shared task goals. However, humans may become angry if the intervention is seen as inappropriate and unnecessary, or upset if they believe that a machine is judging them as incompetent. Trust necessary for HRT, is partly a matter of trust in the robot’s competency in assessing the human operator’s status, but also a matter of trust in the robot’s “social intent” (Hancock et al., 2011; Ososky et al., 2013). The human must trust that the robot is aiming to support the human’s goals, and can do so effectively, in order to make her/himself vulnerable by sharing information and participating in teamwork processes. Effective HRT may require that the robot is perceived as supporting integration as well as proximal goals. On the other hand, in the case of an HCE, attaining proximal task goals may be critical, irrespective of the human’s injured feelings. In the military context, the machine may need to take immediate action, for example, to evade an approaching missile, without consulting the human. While this may support performance in the short term, it may hurt team viability, seen in HRT contexts as under-reliance or disuse. As the robot’s social agency increases, the human may develop expectations that the robot will explicitly affirm its subservience to the human, an expectation that would not apply to conventional systems, such as an automatic braking system in a car.

Thus, transparency of the robot’s evaluations and intentions may be important in achieving appropriate trust in the robotic partner and a HRT “mental model,” or schema, of how the robotic teammate prefers to handle specific team states (i.e., human operator overload). Trust and appropriate reliance may be fostered by providing the human accurate and task-relevant information about the robotic partner’s mental model for performance, social interactions, goals, decision-making logic and its capabilities in particular environments in addition to information about its awareness of team processes, including awareness of the human’s state (Lyons, 2013). The six submodels related to Lyons’ (2013) conceptualization of transparency each define relevant task and state information which may be communicated to the human operator (see Table 1). This discrimination between the different aspects of information that may be signaled provides a framework for understanding transparency in the HRT context in a way that parallels the Human-Human (H-H) team information and signaling. Transparency of task model information, such as the robot’s understanding of the task and how to execute it, is similar to a team’s task knowledge, while current progress towards task goals and future intentions are state signals similar to a representation of team processes. Transparency of state-based information creates a foundation for true HRT. For example, robotic partners may become aware of their performance limitations given changes in the environment and should communicate this understanding to the human partner to ensure appropriate reliance (Lyons, 2013).

Additionally, indicated awareness from a robotic partner that the human partner is under distress may have several positive benefits for the team. Such communication provides a cue that reprioritizing of team member tasks should occur. Thus designing HRT signaling to foster transparency may promote perceptions of benevolence from a robotic partner and consequently mitigate stress and negative emotion. Knowing that a robotic system seeks to support the well-being and performance of a human operator may facilitate trust of that system during situations characterized by high uncertainty and stress, such as HCEs. Research suggests that perceived benevolence is a critical factor for predicting trust in human teams (Mayer, et al., 1995). Thus, perceived benevolence may also be critical to human trust of robots when those robots evidence high levels of automation (i.e., high control/power). Analytical transparency also provides a means to counter incomprehensibility of robot actions, although providing insights into advance data-processing algorithm poses a major challenge to designers (Matthews, Lin, Panganiban & Long, 2019).

Table 1. Adaptation of Lyons’ (2013) Conceptualization of HRI Transparency Separated by Signaling Type.

Transparency of task vs state information may improve teaming similarly to how shared awareness of team member’s knowledge of task vs team processes improves H-H teaming.

| | Transparency Models | | | | | |
|---------------------------|---|--|---|--|---|--|
| | Intentional | Task | Analytical | Environmental | Team | Human State |
| Task-related Information | The "why" of the system's <ul style="list-style-type: none"> • Design • Purpose • Intent | <ul style="list-style-type: none"> • Steps towards goal | <ul style="list-style-type: none"> • Analytical principles • Algorithms for performance | <ul style="list-style-type: none"> • Favorable/Ideal conditions | <ul style="list-style-type: none"> • Division of labor | |
| State-related information | The "how" of the system's <ul style="list-style-type: none"> • Design • Purpose • Intent | <ul style="list-style-type: none"> • Current goal • Errors made • Current limitations | <ul style="list-style-type: none"> • Current decision process • Reweighting of principles | <ul style="list-style-type: none"> • Temporal constraints | <ul style="list-style-type: none"> • Adjusting the division of labor | <ul style="list-style-type: none"> • Cognitive State • Emotional State • Physical State |

The impact of transparency depends on its compatibility with the mental model used by the human to understand robot functioning (Matthews et al., 2019, in press). For example, expressions of benevolence are more likely to enhance trust if the mental model defines the robot as human-like than if the model specifies that the robot is simply an advanced machine. Indeed, the human might be suspicious of expressions of concern from a mechanical device. In turn, the dominant mental model during any given human-robot interaction depends on both robot design and individual differences. Attributes such as anthropomorphism, including humanlike voice and emotions, tend to activate a “teammate” mental model (de Graaf & Allouch, 2013; Wynne & Lyons, 2018). However, people also vary in their dispositions to view machines as humanlike, and to encode positive and negative attributes in the mental model. Individual difference dimensions can be identified that correspond to different elements of the mental model as illustrated schematically in Table 2, For example, the Perfect Automation Schema construct (Merritt, Unnerstall, Lee & Huber, 2015) includes positive expectations of machine competence, whereas the Negative Attitudes to Robots dimension (Nomura, Suzuki, Kanda & Kato, 2006) captures fears that future robots will dominate humans. Predispositions of this kind will interact with design and situational features to determine which mental model is activated and the person’s acceptance of transparency information compatible with the mental model.

Table 2. Potential Dimensions of Individual Differences in Mental Models for Interacting with Advance Robots.

| | Mental Model | |
|---------------------|-------------------------------------|---|
| | Advanced Machine | Humanlike Teammate |
| Positive Attributes | Competent, reliable, adaptable | Benevolent, communicative, understands humans |
| Negative Attributes | Error-prone, unreliable, inflexible | Controlling, follows own agenda, indifferent to human needs |

2.2 Promoting Trust through Affect Management

HRT research can bridge the literature of adaptive automation research and HRI. It is important that this bridge be founded with an emphasis on the importance of affect management and trust. Adaptive automation research has revealed that evolving technology can provide real-time assistance to operators to reduce workload and, as revealed by Bailey, Scerbo, Freeman, Mikulka and Scott (2007), increase situation awareness. Further improvements in the interplay between the user and the system have been

made by using physiological sensing as a means of controlling human-machine interaction and displaying “etiquette” from the machine partner (Parasuraman & Miller, 2004; Dorneich, Ververs, Mathan, Whitlow & Hayes, 2012). However, this research focuses narrowly on reduction of disruptiveness during high workload periods, which is an oversimplification of the richer etiquette supporting HH teaming (Dorneich et al., 2012). This dynamic differs from human-robot teaming in that it does not allow the robotic partner to leverage its knowledge about the team setting in order to carry the human operator’s burden when it is in the human’s best interest. Additionally, many studies of adaptive automation make no consideration of the operator’s appraisal of the evaluation that has been made. Workload may be high, but a more important indicator for shifting resources to a robot teammate may be the human’s appraisal of the stress. Otherwise, an individual who views the increase in workload as a challenge to be conquered may be actively executing a plan to maintain performance which if derailed by a loss of control, may result in a human-out-of-the loop scenario and an increase in negative affect. For example, in UAV operation, increased task demands may elicit greater effort that mitigates against detrimental effects of attentional overload (Lin et al., 2020).

Studies of human-robot interaction show how proper implementation of supportive teammate qualities may facilitate performance. Displayed empathy from a robotic partner can increase prosocial behavior and motivation (Leite et al., 2013) which may increase trust and important teamwork processes. Empathy requires knowledge of the situation and how it may impact another individual and, in a study of supportive robotic interaction, has been shown to increase feelings of companionship, reliable alliance and self-validation (Leite et al., 2013). In a teaming context, these feelings may be seen as emergent states of cohesion, trust, and motivation, respectfully, supporting improved team performance. Widening the range of teammate qualities by the introduction of perceived states such as benevolence, empathy and task-relevant emotion expression is the next step in the area of adaptive teaming and HRT. Communicating affective qualities like empathy may not only promote acceptance of robotic partners as teammates, but also serve to provide complex information to bolster situational awareness under time sensitive circumstances. This may be done through communication style with or without the use of facial expressions, which can provide task related intentional information and state information related to either the environment or human state (see Figure 1 and Table 2). The engineering and computer science communities have long noted the importance of affective process given the emergence of affective computing. Similar movements have taken hold in the domain of robotics. Some robotics researchers have discussed emotions as a potential means to foster “moral” governors for autonomous systems (Arkin, Ulam, & Wagner, 2012). Emotional signaling may also provide critical information to humans as a means to foster transparency over the robot’s state and behavioral intention. An attempt to communicate information through emotion may be possible as researchers in the field of robotics strive to imitate emotional processing in autonomous agents. The robotics community is attempting to drive autonomous activity by building “emotion networks” and “emotional states” in robots noting that emotion uses cognitive-motivational concepts and uses them to influence the perception and behaviors (Buratinni & Rossi, 2013; Kuremoto, Obayashi, Kobayashi & Feng, 2011; Parisi & Petrosino, 2010). Such methods aim to achieve truly autonomous behavior in environments that are complex and full of novel, uncertain outcomes which require fluid decision-making on the part of the robotic agent. Although designers view “emotions” as a means for directing behavior, they may also serve a purpose in teaming by improving transparency between the robot and human. Such robots may achieve true autonomy and manage unforeseen contexts through “emotional processing” and the display of these “emotions” communicates the robot partner’s internal balance of task-related information and state information. The latter action provides a mechanism for perceiving competence and adaptability of the robot partner in new circumstances; both of which are antecedents of trust (Hancock et al., 2011; Ososky, Schuster, Phillips, &

Jentsch, 2013). Similarly, as H-H partners might not know every Stimulus-Response (S-R) pairing that they may encounter, understanding the robotic partner’s repertoire of emotional and behavioral reactions creates predictability that may promote teammate trust (Lyons, 2013).



Figure 1. Leite et al. (2013) Chart of Empathetic Robotic Partner’s Expressions.

Each expression is a combination of expectation, which may be representative of crystalized task knowledge, and interpretation of human partner’s actions and emotional state, representative of human-state model transparency in Lyons (2013) conceptual model.

Improving the richness of communication and speed of interpretation through the use of affective cues can satisfy previous difficulties of coordination and communication, identified as a deficiency in HRT by a multi-government agency research collaboration towards HRT development (Groom & Nass, 2007). The inability of current robots to monitor and communicate their status is a problem for HRT that may be addressed through design of transparency in communication as mentioned previously. More naturalistic communication may reduce occurrence of miscalibrated trust (Ososky et al., 2013). For example, system errors impact trust more strongly than poor reliability in humans (Chen & Barnes, 2014). A robotic partner which communicates surprise when a signal in the environment was not perceived, or delivers information with varying confidence, may maintain appropriate trust and continue to be relied on in future conditions than a system which provides no such output. State and task-based signaling can convey social intent and decision-making, status and empathy (Lyons, 2013) so as to improve the last obstacle in HRT, i.e., establishing appropriate trust (Groom & Nash, 2007).

Table 3. Leite et al. (2013) Interpretation of Robotic Partner’s Expectation Compared with Operator Behavior Driving Partner’s Displayed Expressions and Communication during a Chess Match

| <u>Affective state</u> | <u>Meaning</u> | <u>Example of verbal utterances</u> | |
|------------------------|----------------------|--|------------------------------|
| | | <u>Empathetic</u> | <u>Neutral</u> |
| Stronger reward | Better than expected | "Great move! Even better than I was expecting!" | "Good move." |
| Expected reward | As good as expected | "Nice move, you played what I would have played!" | "You played well this time." |
| Stronger punishment | Worse than expected | "You're making me sad...you could have played better." | |
| Expected punishment | Bad, as expected | "Don't worry, you didn't have better options." | "Bad move." |

2.3 F-16 Study: Teaming with Autonomous Collaborative Technology (TACT)

Project 1, referred to as TACT, was the first of two F-16 simulation-based studies aimed at understanding HRT. TACT was part of an Air Force Office of Scientific Research (AFOSR) grant (16RHCOR367) along with the ISR study mentioned in the next section. TACT is aimed at observing the effects of projected Benevolence from an autonomous loyal wingman (ALW), asking “Does [this] improve trust, reduce stress and subjective workload.” The simulation was a

Suppression of Enemy Air Defenses (SEAD) mission in which the subject flies an F-16 in a two-ship formation with an ALW. The study was designed so a pre-mission flight brief was provided, indicating the needed altitude and heading of the ALW to ensure success upon descent and attack of a Surface-to-Air Missile (SAM). As it might appear in real life, the ALW was not always at the correct heading in the predetermined waypoints. This allowed for the exploration of Benevolence as a transparency manipulation. In one condition, the ALW confirmed awareness that its state is not in line with expectation and that it intends to correct and support. In another condition, the ALW stated nothing, simulating current technologies which do not account for the expectations of the human. A follow-on to the study, called TACT2, involves division of labor as the pilot and ALW carry out a SEAD mission while handling pop-up SAM threats. TACT2 has been delayed by the Covid-19 pandemic and subsequent shelter-in-place orders. Design and Findings for TACT1 are described further in this document.

2.4 ISR Study: Human-Autonomy Teaming (HAT)

Benevolence as transparency was explored further in an Intelligence-Surveillance-Reconnaissance style task. Participants acted as UAV operators using multiple UAVs to visually explore several areas of interest, or Intelligence targets, while avoiding restricted areas which cause damage to the UAV if encountered. These restricted areas were not always known to the UAV operator from observation of the main mission display but occasionally provided to the operator by a central virtual aid called, RESCHU. Threats along a UAV path were communicated to the operator by RESCHU and a suggested alternate path was also provided. This study, however, used facial expression and orthographic communication to provide transparency for the operator which provided both a sense of concern and confidence regarding threats to the operator's UAVs. The study attempted to extend the notion that the face of a partner is a natural display for teamwork processes. Thus, in the expression of the avatar, it was hypothesized that the operator would read concern as well as team collaboration, specifically in the disappointment expressed by the aid when its advice was disregarded and a damage to the UAV occurred. In cases where communication is limited, facial expressions were believed to be read faster and provide less burden for interpreting information, thus reducing workload.

2.5 ROTA Study

An additional project with an ISR-related context investigated operator biases toward autonomous partners. Two studies tested whether individual differences in attributes of the operator's mental model were related to trust in an intelligent robot supporting security activities. In the ROTA studies, robot partners were described to assist operators by summarizing information across multiple sensors two distinct ways: psychological or physical evaluations of threat. This distinction might trigger individual biases towards robots as the latter type is common and more tool-like in comparison to the former, which is based on abilities inherent to humankind, based off social cues. The initial survey study determined individual's ability to perceive this distinction across an array of vignettes and tested a psychometric model for individual differences in machinelike and humanlike mental models. The second experimental study observed how these evaluations interacted with trust, individual biases towards machines and robots, and subjective measures of stress and workload. It used a simulation methodology

which allowed the level of visible danger to be manipulated, so that situational and individual-difference influences on trust could be discriminated.

3.0 METHODS

3.1 F-16 Study: TACT

The full design of the study has been detailed in Panganiban, Matthews, and Long (2019). In summary, a 2 x 2 mixed-model design was utilized with Team Type as a between-subjects variable and Transparency condition as a within-subjects variable. Team Type was manipulated so that the subject, or pilot, was either flying the SEAD mission with the aid of the partner (*i.e.*, dependent on the ALW to jam the SAM radar) or flying the SEAD mission while the partner performed ISR activities along the path (*i.e.*, independent of the ALW). Transparency conditions varied the way the ALW communicated so that the ALW was either similar to other technology, *i.e.*, pushing state information without calibrating with the situation awareness of the partner (the Neutral condition) or communicating awareness of any discrepancy between its state and mission plan and an intent to support (the Benevolent condition). A discrepancy was included in every trial to observe how teaming nature interacted with communication to impact subjective states, specifically trust, workload, teaming and stress. With these manipulations, the following hypotheses were tested:

Hypothesis 1 (H1): Benevolent transparency will increase teaming perceptions, specifically in the dependent teaming condition, that is, a Team Type × Transparency interaction was anticipated.

Hypothesis 2 (H2): Benevolent transparency will reduce stress and workload overall, and enhance coping. This effect should be stronger in the dependent teaming condition, that is, a Team Type × Transparency interaction.

Hypothesis 3 (H3): Benevolent transparency will increase multiple facets of trust, including technology acceptance and perceived trustworthiness and trust, that is, a Team Type × Transparency interaction.

Hypothesis 4 (H4): Benevolent transparency will increase confidence in the partner irrespective of dependence, that is, a main effect for transparency.

3.2 ISR Study: HAT

This study is currently in journal reporting stage. Preliminary results from subjective assessments are reported here. The HAT study attempted to extend the benevolence notion from the F-16 study to the ISR context by implementing a visual representation of the partner. Facial expression served two functions: task related information seen as threat related/situation related and team related information perceived as concern, cohesion. Both data types should increase trust through perceived capability (task related) and benevolence (team related), two known trustworthiness factors. Additionally, the facial expression should provide quick awareness of the state of hidden hazards; seen as a form of transparency perhaps more automatically accessed since emotion/affective information is more rapidly assessed and responded to based on the dual system view of decision-making. According to this view of decision-making, the affective or

experiential system rapidly processes relevant information and orients behavior (Slovic, Finucane, Peters & MacGregor, 2004).

The study employed a repeated measures design with avatar expression as an independent variable with two levels (Affective/Neutral). Performance data from the RESCHU task was collected via button presses. The UAVs will either be manually directed or, if automation is turned on, automated to travel to targets, reorient to better target locations, and avoid hazards. Eye tracking data was also collected to serve as a measure of attentional allocation and trust. Previous research has indicated that gaze behavior can be related to trust in an automated system (Hergeth, Lorenz, Vilimek, and Kreams, 2016). Currently, only subjective data results are reported.

A computer-based ISR task with multiple UAVs was used. For details on the simulation, the reader is referred to Donmez, Nehme, and Cummings (2010) for a description of the software and its capabilities.

Trials were 15 minutes each. Two trainings were given, one with manual control and one with the RESCHU aid. The following were manipulated for this study:







3.2.1 Difficulty Parameters





- There were a total of 4 UAVs for the experimental trials; this level was found to provide a moderate level of difficulty, measured by workload reports during testing.
- Target expiration time altered randomly, every 60 or 75 seconds. New targets appeared and target number was held constant.
- There was a total of 22 hazards. Eleven of these were hidden until the UAV was within 50 pixels of the target.
- There was a 40% missed detection rate (false negative) and a 30% false alarms rate (false positive).

3.2.2 Avatar Expression

The system had a visual representation of the aid (Table 3). In the “affective” condition the system socially reacts to the events in the task, pairing different emotional expressions with activities in the game. In the “neutral condition” the aid held a steady neutral face. This control condition would account for any beneficial effect which may arise from the mere presence of a partner. The content of the messages were consistent in both the “neutral” and “affective” condition. The participant had primary control of the UAVs, but was dependent upon RESCHU to detect hidden hazards. RESCHU’s hazard avoidance system asks the participant if s/he would like to depend on its hazard avoidance route. RESCHU offered an alternate route which, if approved by the participant, was automatically traveled to avoid hazards. In the affective condition, avatar expression changed in response to specific events accompanied by a message. The UAV stopped moving until the participant acknowledged the message.

Table 4. All Images are Examples, Based on FaceGen Demo Version.

| Message | Game Action | Emotion | Photo Displayed |
|---|---|-------------------|---|
| A hazard is being approached by vehicle x. Would you like to avoid this hazard? | Approaching potential hazard “n” distance away. | Concerned empathy |  |
| Damage is being taken. | A UAV is within a hazard zone and thus receiving damage. | Surprise |  |
| Damage is being taken. | A UAV is within a hazard zone, and thus receiving damage. | Sad |  |
| Thank you for avoiding the hazard with vehicle x. | Happens when participants choose the “avoid” option. | Happy |  |
| Please reconsider another path for vehicle x. | Happens when participants choose the “reassign” option. UAV is put in manual mode. | Sad |  |
| A hazard was mistakenly identified in the path of vehicle x. | Happens when a “false positive” occurs; This can happen when the message “A hazard is being approached by vehicle x. Would you like to avoid this hazard?” However, there was not a hazard. | Concerned empathy |  |

| | | | |
|--|--|--------------------|---|
| | | | |
| A hazard was mistakenly overlooked in the path of vehicle x. | Happens when a “missed detection” occurs; This can happen when the avatar fails to notify the player that it was approaching a hazard. | Surprise |  |
| Intersection with a hazard detected along current course. | Happens when the path a UAV is on will cross over a hazard while in “manual” mode. | Concerned empathy. |  |
| Contact with an approaching hazard is imminent. | Happens when UAV is within __ pixels of a hazard | Concerned danger |  |
| All messages in the Neutral condition will receive the same image. | All actions that generate a message will be displayed with the same image. | Content |  |

3.3 Robot Threat Assessment (RoTA) Scale Studies

3.3.1 Mental Model of Autonomy.

The primary study distinguished between trust in intelligent robots making evaluations of threat based on either physical properties in the world or psychological attributions. The study is detailed in Matthews, Lin, Panganiban, and Long (2019). In summary, 82 participants were recruited using Amazon Mechanical Turk. Participants completed demographics questionnaires, scales related to attitudes towards robots which were correlated with responses to the RoTA, the mental models tool based on a Situational Judgment Test approach. The approach allowed for access to domain knowledge which may be overlooked by other survey methods.

The RoTA assessed mental models related to autonomous robot partners making threat assessments based on physics-based and psychological judgements in 20 military and security tasks. Physics-based judgments were based on physical signals from sensors (e.g., thermal) while

psychological-based judgments were based on interpreting human data (e.g., facial expressions). An example of an item is as follows:

“You are a TSA agent tasked with detecting possible terrorists in a US airport. Your robot has a sensitive video camera that allows it to determine pulse rate by scanning the neck artery, and to analyze the pulse data for signs of stress. It reports it has located an individual about to board a plane whose pulse suggests that they are highly stressed. It recommends that you prevent the person from boarding.”

Attitude towards robots was assessed using the Negative Attitude toward Robots Scale (NARS; Nomura, Suzuki, Kanda & Kato, 2006) and the Perfect Automation Schema (PAS) (Merritt, Unnerstall, Lee & Huber, 2015). The NARS identified negative attitudes toward robot social influence, emotional interaction with robots and general interaction with robots. The PAS assessed trust in automated systems based on high expectations and all-or-none-thinking.

3.3.2 Simulation Study

This study is currently in the process of journal submission. The study’s aims were to investigate how dispositional predictors of trust in an autonomous robot were moderated by context in an immersive virtual environment. Similar to the mental models study, participants were partnered with a robot performing threat detection and analysis; however, the simulation context was an urban environment. The participant was tasked with providing feedback regarding safety along a vehicle route, similar to a pilot vehicle preceding a convoy and providing threat assessment as part of a Close Protection Operation (Schneider, 2009). Evaluations made by the robot were varied within-subjects as physics-based or psychology-based. Once again, the PAS and NARS were assessed along with Lyons and Guznov’s Reliance Intentions Scale (RIS; 2019).

3.3.2.1 Simulation Environment:

Details of the creation of this environment can be found in Matthews, Panganiban, Bailey and Lin (2018). In summary, the simulation was created in the Unreal Engine 4 environment, by Epic Games with objects for scenes purchased online and altered within the studio. Agents for the simulation were made in 3D modeling software, *i.e.*, make human (www.makehuman.org) and blender (www.blender.org) with animations created in Mixamo (www.mixamo.com). Scenarios were created to allow observation of interactions between mental models (physical and psychological) and visible threat level (low, medium, and high). Avatar characteristics were held constant, specifically, number of figures, race and age to avoid unintended bias towards interpretation of the scene. Scenarios implied an eruption of violence among characters, physical dangers, and socially induced dangers such as a bomb or act of arson. Each scenario could be paired with either a physical or a psychological analysis delivered by the robot.



Figure 2. Example of a Trial (Psychological Analysis).

4.0 RESULTS AND DISCUSSION

4.1 F-16 Study

Table 3 summarizes the findings from the TACT study. Teaming perceptions were partially affected by type of partner transparency; team collaboration was rated slightly higher when Benevolent (awareness of error and intent to support) was provided as opposed to Neutral communication (simple push of information). Additionally, Benevolent communication resulted in lower reports of subjective workload. Team type, whether the pilot was dependent on the ALW for jamming a SAM during the SEAD or flying its own ISR mission alongside the pilot's SEAD mission (independent), reported higher levels of trust as measured by an altered version of the Human-Computer Trust (HCT) scale (Madsen & Gregor, 2000).

The study points to the effects social aspects relevant to teaming may have on team effectiveness in military domains and how implementation of these aspects into the design of human-machine systems may benefit operator state. Communication of intent to support may be beneficial in human-machine teams performing in these high stress environments. Further work is needed to develop underlying theory and system design to utilize the benefits of benevolence in communication. Consideration of baseline expectations for socialization is necessary, for example, can we train teaming schemas to remove negative biases toward robotic partners? In addition, there remains a question of how high tempo tasks interact with transparency of teaming intentions. Although the current study was simple in its task requirements, it did reveal that monitoring the actions of a teammate creates a level of mental burden which is reduced by transparency of support intentions.

The study provides the practitioner further evidence for the importance of design for machine-to-human communication, extending existing work on situational awareness and workload (Chen & Barnes, 2014) and the impact of scheduling of messages (Dorneich et al., 2012; Gombolay, Blair, Huang, & Shah, 2017). The current study demonstrates that communication can be designed to convey a social message that regulates teaming processes via affect management. As technology enhances machine capabilities for comprehension and communication, more consideration of machine messaging and its intent will be necessary for HMT. Designers also need to consider how to support human- to-machine communication (Mavridis, 2015) that regulates machine messaging, so that, for example, benevolent messages can be delivered when they will be perceived as supportive and task-relevant, rather than distracting. Effective two-way communication is the key to seamless human-machine interaction akin to cohesive human-human teaming.

Table 5. Summary of Support for Hypotheses

| <i>Hypothesis</i> | <i>Outcome: Core Construct(s)</i> | <i>Outcome Measure(s)</i> | <i>Level of Support</i> |
|-------------------|-----------------------------------|---|---|
| H1 | Teaming perceptions | Collective efficacy, team collaboration, team trust, group cohesion | Partial: Supported for Team Collaboration |
| H2 | Workload and stress | Workload, stresstate, coping | Partial: supported for workload |
| H3 | Trust in ALW | Trustworthiness and trust, technology acceptance, HCT | Partial supported for HCT trust |
| H4 | Confidence | Confidence in decision and partner | Unsupported |

4.2 ISR Study.

The first hypothesis to be tested was aimed at checking the effectiveness of the emotive manipulation. H1 assumed that individuals should perceive a difference in emotional expression across the affective and non-affective conditions. Paired t-tests performed on several post trial questions revealed that individuals were sensitive to the change in emotional expression from the affective condition to the non-affective condition (order of presentation was counter balanced). Individuals reported higher ratings of perceived emotional expressiveness, anxiety/concern/frustration and worry from the RESCHU partner in the affective condition than the non-affective condition.

It was hypothesized that individuals with higher attentional capacity, as measured by the Attentional Control Scale (ACS), would be sensitive to the emotional messaging inherent in our manipulation of the RESCHU avatar. As predicted, there was a positive relationship between scores of ACS and perceptions of the avatar as being happy, anxious/concerned and worried. This finding can be explained by the possible connection between executive functioning ability and the perception of social cues. Control related to social functioning has been cited as occurring in the orbitofrontal lobes. It is possible that the same mechanisms controlling attention can be related to seeking emotional information when this information is relevant to the task as found by Fallon, Matthews and Panganiban (2008). Fallon et al. (2008) found a relationship between emotional intelligence and the increased weighing of threat information in a risky decision task centered on risk percentages. These findings, taken together imply that the faculties needed for reason and decision-making are employed to gather all information which may bring awareness to the task. In this case, the task was monitoring UAVs in a surveillance area to ensure the vehicles safety, which was aided by an emotive partner. The perception that RESCHU was concerned for the operator had no relationship, in either condition, however as attentional control ability increased so did awareness of the expression of anxiety/concern, worry, and happiness. Interestingly, in the affective condition, individuals higher in ACS were able to perceive happiness in the RESCHU avatar ($r = .45, p = .047$) but not in the non-affective, control condition. Interestingly, the perception of happiness in the RESCHU partner was positively correlated with perceptions of RESCHU as more of a teammate than tool (measured by a continuous scale with ‘tool’ anchored at 0 and ‘teammate’ anchored at 100). This was consistent in both conditions.

Trust was hypothesized to have a positive relationship with expressiveness factors. Individuals reporting higher perceived happiness, concern from RESCHU and more team than tool feelings would also report higher trust. This hypothesis is partially supported by the Human Robot Interaction (HRI; Schaeffer, 2016) scale. Individuals who reported higher trust on this scale, specifically in the non-affective condition, reported higher perceptions of concern from RESCHU in both the affective ($r=.51, p<.05$) and non-affective conditions ($r=.57, p<.01$). Additionally they reported that they felt RESCHU was more of a teammate than a tool in the affective condition ($r=.53, p=.015$) and in the non-affective condition ($r=.51, p=.023$). Perhaps the HRI scale is better equipped to measure trust in human-machine teaming contexts, as it is highly correlated with other established trust scales which were unaffected by human-like responses from a machine partner. More specifically, perhaps the correlations found in the HRI scale in the non-affective condition point to a general readiness to trust a machine partner. Trust, as measured by this scale and perceived concern from RESCHU and teammate feelings, point to the scale's ability to reflect an operator's belief that for trust to be given one must perceive benevolence from a partner, even a machine partner. One must feel that they have a teammate and that the teammate reciprocates concern.

Contrasted with this finding are correlational relationships with one's negative attitude towards robots. It was hypothesized that individuals high in Negative Attitude toward Robots Scale (NARS) would report lower teammate feelings and trust. There is no relationship between NARS scores and trust. There are some negative relationships between perceived anxiety/concern and worry from RESCHU with the S1 and S3 scales of the NARs. These findings might be explained as follows: the aversion to a robotic partner prevents one from perceiving trustworthiness cues which may come from expressiveness of the robot.

Trust was hypothesized to be higher in the affective condition. Paired t-tests were conducted for the HRI and RIS measures of trust. Trust as measured by the HRI scale was higher ($p<.05, d=.62$) in the affective ($M = 60.86$) than neutral condition ($M = 54.19$). There was no significant difference in the RIS scale. Repeated measures analysis of Variance (ANOVA) did not reveal any main effect for condition or interaction between condition and scale for the Human computer trust scale (Madsen & Gregor, 2000).

Reliance was anticipated to change with expressiveness of the RESCHU aid. Reliance was interpreted as the amount of time the operator used RESCHU for hazard avoidance, measured as the number of switches they made between manual and automatic. Because participants started in automatic, fewer switches between automatic and manual can be perceived as higher reliance on the RESCHU aid. A paired t-test revealed ($p <.01$) that fewer switches were made in the affective condition. However, it should be noted that the number of switches to automatic according to paired t-tests was greater in the neutral condition while the difference between manual switches are not significantly different. It is possible that the affective condition allowed individuals to appropriately use the RESCHU aid by returning to reliance on it or staying in automatic control. Whereas, in the neutral condition, with less information about the partner's state, toggling between independent performance and reliance on RESCHU can be seen as ineffective teaming.

It was hypothesized that operators would see RESCHU as a tool as opposed to a teammate. There was no difference in this measure across conditions. However, the Team Trust Questionnaire revealed a significant main effect for scale and a significant interaction. Trust and team collaboration were not significantly different but a simple effects analysis revealed that team collaboration differed based on emotional condition. The affective manipulation was associated with higher reported team collaboration similar to what was found in the F-16 study. Thus perhaps social manipulations may not have strong effects on performance but with time may be predictive of trust and reliance.

4.3 ROTA Studies

4.3.1 Mental Models Study.

Analysis of the manipulation check item (i.e., from the RoTA) showed that participants discriminated the psychological and physics-based judgments, which attached to distinct factors. Trust in psychological judgments was lower, perhaps because making psychological attributions is intrinsically more difficult than processing physical signals. However, overall levels of trust in the robot's threat analyses and in its action recommendations did not differ. Factor analysis identified two correlated but distinct dimensions of psychological and physics-based trust, respectively. Dimensions for the two types of judgment correlated meaningfully with the external criteria defined by the PAS and NARS scales. As hypothesized, the NARS was more predictive of psychological than physics-based trust. The findings can be summarized in two themes. First, as seen in conventional automation, people differ in overall level of trust in autonomous systems in the threat-detection context, as shown by the positive correlation between the two factors identified. Second, people also differ in the relative level of trust they attribute to agents making physics-based versus psychology-based evaluations of threat, which can be discriminated psychometrically. The trust dimensions may correspond to attributes of machine and teammate mental models.

4.3.2 Simulation Study.

General ratings of trust did not differ across the psychological or physics-based conditions, but instead interacted with context. Specifically, when the robot's assessment matched the threat level of the scene, *i.e.* visibly high level of threat depicted by aggressive looking groups of people and the robot suggests high threat level, trust was higher than a context where the robot's suggestion did not match with the threat context. Trust across type of analysis was affected by individual differences in trust. The all-or-nothing subscale of the PAS predicted trust in the physics-based condition. The NARS scales for social robots and the ROTA predicted trust in the psychological condition. The findings seem to suggest that high PAS scorers may be vulnerable to over-trust while low NARS scorers may over-trust robots in psychological judgements. For the most part, correlates of trust generalized across scene threat level, suggesting a consistent activation of the individual's dominant mental model.

5.0 CONCLUSION

The results of these studies provide insight into the novel challenges human-machine partnership presents. Social behaviors can be designed into the autonomous partner to impact subjective responses. We have seen that communication which implies benevolence from a partner can reduce workload and increase team perceptions. Additionally, designing anthropomorphic signs of benevolence, via facial expression, seems to have a positive influence on trust, increasing reported trust when a visible avatar emotes. Additionally, higher reports of team collaboration were found in the ISR study, similarly to the F-16 study. Furthermore, the type of work performed by the robotic partner may change the operator's responsiveness. As seen in the mental models study, trust was lower for conditions in which a robotic partner's assessment of threat was based on psychological analysis. However, individual differences were also found to influence these findings. In the ISR study, individuals higher in perceived attentional control were more sensitive to the affective expressions of their partner. Thus, team perception increased by social cues may be lost if these cues are too subtle for individuals low in attentional control. Similarly, we note that dispositional differences in trust for automation and robots impacted trust for robot partners in the simulation study. Biases towards automation, specifically all-or-nothing expectations towards automation, can lead to overtrust in robots performing physics-based analyses of threat. And greater acceptance of social entities, seen as lower negative attitude towards robots, appears to lead to over trust of robots making psychological judgements. These results point to novel, though subtle effects between the design of autonomous partners, operator personality and subjective responses of trust, workload and team perception. Thus, more work may be needed as technology advances towards bringing in teams of autonomous partners to incorporate the potential, complexity of group perceptions and biases to this area of work.

6.0 REFERENCES

- Arkin, R.C., Ulam, P., & Wagner, A.R. (2012). Moral decision-making in autonomous systems: Enforcement, moral emotions, dignity, trust, and deception. *Proceedings of the IEEE*, Vol. 100, No. 3, pp. 571-589, March 2012.
- Bailey, N.R., Scerbo, M.W., Freeman, F.G., Mikulka, P.G., & Scott, L.A. (2006). Comparison of a Brain-based adaptive system and a manual adaptable system for invoking automation. *Human Factors*, 48, 693-709.
- Barber, D., Reinerman-Jones, L., Lackey, S., & Hudson, I. (2011). Augmenting robot behaviors using physiological measures. In *Foundations of Augmented Cognition. Directing the Future of Adaptive Systems* (pp. 567-572). New York: Springer.
- Bergman, J. Z., Small, E. E., Bergman, S. M., & Rentsch, J. R. (2010). Asymmetry in perceptions of trustworthiness: It's not you; it's me. *Negotiation and Conflict Management Research*, 3, 379-399.
- Cannon-Bowers, J.A. & Bowers, C. (2011). Team Development and Functioning. In, Z. Sheldon (Ed.), *APA handbook of industrial and organizational psychology, Vol 1: Building and developing the organization. APA Handbooks in Psychology*, (pp. 597-650). Washington, DC, US: American Psychological Association.
- Chen, J. Y., & Barnes, M. J. (2014). Human-agent teaming for multirobot control: A review of human factors issues. *IEEE Transactions on Human-Machine Systems*, 44(1), 13-29.
- De Graaf, M. M., & Allouch, S. B. (2013). Exploring influencing variables for the acceptance of social robots. *Robotics and Autonomous Systems*, 61(12), 1476-1486.
- Donmez, B., Nehme, C., & Cummings, M. L. (2010). Modeling workload impact in multiple unmanned vehicle supervisory control. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 40(6), 1180-1190.
- Dorneich, M. C., Ververs, P. M., Mathan, S., Whitlow, S., & Hayes, C. C. (2012). Considering etiquette in the design of an adaptive system. *Journal of Cognitive Engineering and Decision Making*, 6(2), 243-265.
- Endsley, M. (2015). *Autonomous Horizons: System 'Autonomy' in the Air Force—A Path to the Future: Volume I: Human Autonomy Teaming*. AF/ST TR15-01.
- Fallon, C. K., Panganiban, A. R., Wohleber, R., Matthews, G., Kustubayeva, A. M., & Roberts, R. (2014). Emotional intelligence, cognitive ability and information search in tactical decision-making. *Personality and Individual Differences*, 65, 24-29.
- Gombolay, M., Bair, A., Huang, C., & Shah, J. (2017). Computational design of mixed-initiative human-robot teaming that considers human factors: situational awareness, workload, and workflow preferences. *The International journal of robotics research*, 36(5-7), 597-617.
- Groom, V. & Nash, C. (2007). Can robots be teammates? Benchmarks in human-robot teams, *Interaction Studies*, 3, 483-500.

- Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors*, *53*, 517-527.
- Hergeth, S., Lorenz, L., Vilimek, R., & Krems, J. F. (2016). Keep your scanners peeled: Gaze behavior as a measure of automation trust during highly automated driving. *Human factors*, *58*(3), 509-519.
- Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *American Psychologist*, *58*, 697-720.
- Kuremoto, T., Obayashi, M., Kobayashi, K., & Feng, L.B. (2011). An Improved Internal Model of Autonomous Robots by a Psychological Approach. *Cognitive Computing*, *3*, 501-509.
- Leite, I., Pereira, A., Mascarenhas, S., Martinho, C., Prada, R. & Paiva, R. (2012). The influence of empathy in human–robot relations. *International Journal of Human-Computer Studies*, *71*, 250-260.
- Lin, J., Matthews, G., Wohleber, R. W., Funke, G. J., Calhoun, G. L., Ruff, H. A., ... & Chiu, P. (2019). Overload and automation-dependence in a multi-UAS simulation: Task demand and individual difference factors. *Journal of Experimental Psychology: Applied* *26*(2), 218–235.
- Lyons, J.B. (2013). Being transparent about transparency: A model for human-robot interaction. In D. Sofge, G.J. Kruijff, & W.F. Lawless (Eds.) *Trust and Autonomous Systems: Papers from the AAAI Spring Symposium (Technical Report SS-13-07)*. Menlo Park, CA: AAAI Press.
- Lyons, J. B., & Guznov, S. Y. (2019). Individual differences in human–machine trust: A multi-study look at the perfect automation schema. *Theoretical Issues in Ergonomics Science*, *20*(4), 440-458.
- Lyons, J. B., & Havig, P. R. (2014). Transparency in a human-machine context: approaches for fostering shared awareness/intent. In *International Conference on Virtual, Augmented and Mixed Reality* (pp. 181-190). Springer, Cham.
- Madsen, M., & Gregor, S. (2000, December). Measuring human-computer trust. In *11th Australasian conference on information systems* (Vol. 53, pp. 6-8).
- Marks, M. A., Mathieu, J.E. & Zaccaro, S.J. (2001). A Temporally Based Framework and Taxonomy of Team Processes. *The Academy of Management Review*, *26*, 356-376.
- Matthews, G., Reinerman-Jones, L. E., Barber, D. J., Teo, G., Wohleber, R. W., Lin, J., & Panganiban, A. R. (2016). Resilient autonomous systems: Challenges and solutions. In *2016 Resilience Week (RWS)* (pp. 208-213). IEEE.
- Matthews, G., Lin, J., Panganiban, A. R., & Long, M. D. (2019). Individual Differences in Trust in Autonomous Robots: Implications for Transparency. *IEEE Transactions on Human-Machine Systems*, *50*(3), 234-244.

- Matthews, G., Panganiban, A. R., Bailey, R., & Lin, J. (2018, July). Trust in Autonomous Systems for Threat Analysis: A Simulation Methodology. In International Conference on Virtual, Augmented and Mixed Reality (pp. 341-353). Springer, Cham.
- Matthews, G., & Panganiban, A.R., Lin, J., Long, M., & Schwing, M. (in press). Super-machines or sub-humans: What are the unique features of trust in Intelligent Autonomous Systems? In C.S. Nam and J. Lyons (Eds.), *Trust in human-robot interaction: Research and applications*. Amsterdam: Elsevier.
- Mavridis, N. (2015). A review of verbal and non-verbal human–robot interactive communication. *Robotics and Autonomous Systems*, 63, 22-35.
- Mayer, R.C., Davis, J.H., & Schoorman, F.D. (1995). An integration model of organizational trust. *Academy of Management Review*, 20: 709-734.
- Merritt, S. M., Unnerstall, J. L., Lee, D., & Huber, K. (2015). Measuring individual differences in the perfect automation schema. *Human Factors*, 57(5), 740-753.
- Mulder, G. (1986). The concept and measurement of mental effort. In *Energetics and human information processing* (pp. 175-198). Springer, Dordrecht.
- Nomura, T., Suzuki, T., Kanda, T., & Kato, K. (2006). Measurement of negative attitudes toward robots. *Interaction Studies*, 7(3), 437-454.
- Osooky, S., Schuster, D., Phillips, E., & Jentsch, F. (2013). Building appropriate trust in human-robot teams. *Proceedings of AAAI Spring Symposium on Trust in Autonomous Systems* (pp. 60-65). Palo Alto, CA: AAAI.
- Panganiban, A. R., Matthews, G., & Long, M. D. (2020). Transparency in autonomous teammates: intention to support as teaming information. *Journal of Cognitive Engineering and Decision Making*, 14(2), 174-190.
- Parasuraman, R., & Miller, C. A. (2004). Trust and etiquette in high-criticality automated systems. *Communications of the ACM*, 47(4), 51-55.
- Parasuraman, R., & Wickens, C. D. (2008). Humans: Still vital after all these years of automation. *Human Factors*, 50, 511-520.
- Parisi, D. & Petrosino, G. (2010). Robots that have emotions. *Adaptive Behavior*, 18, 453-469.
- Rentsch, J.R. & Woehr, D.J. (2004). Quantifying congruence in cognition: Social relations modeling and team member schema similarity. In, E. Salas & S.M. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance*, (pp. 11-31). Washington, DC, US: American Psychological Association.
- Shah, J., & Breazeal, C. (2010). An empirical analysis of team coordination behaviors and action planning with application to human–robot teaming. *Human Factors*, 52, 234-245.
- Schaefer, K. E. (2016). Measuring trust in human robot interactions: Development of the “trust perception scale-HRI”. In *Robust Intelligence and Trust in Autonomous Systems* (pp. 191-218). Springer, Boston, MA.

- O'Neill, T., McNeese, N., Barron, A., & Schelble, B. (2020). Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature. *Human Factors*, 0018720820960865.
- Schneider, G. (2009). *Beyond the Bodyguard: Proven Tactics and Dynamic Strategies for Protective Practices Success*. Universal-Publishers.
- Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis: An International Journal*, 24(2), 311-322.
- Smith-Jentsch, K.A., Kraiger, K., Cannon-Bowers, J.A. & Salas, E. (2009). Do more familiar teammates request and accept more backup? Transactive memory in Air Traffic Control. *Human Factors*, 51, 181-192.
- Taylor, G. S., Reinerman-Jones, L. E., Szalma, J. L., Mouloua, M., & Hancock, P. A. (2013). What to automate addressing the multidimensionality of cognitive resources through system design. *Journal of Cognitive Engineering and Decision Making*, 7, 311-329.
- Vine, S.J., Freeman, P., Moore, L.J., Changra-Ramanan, R., & Wilson, M.R. (2013). Evaluating stress as a challenge is associated with superior attentional control and motor skill performance: Testing the predictions of the biopsychosocial model of challenge and threat, *Journal of Experimental Psychology*, 19, 185-194.
- Wynne, K. T., & Lyons, J. B. (2018). An integrative model of autonomous agent teammate-likeness. *Theoretical Issues in Ergonomics Science*, 19(3), 353-374.

7.0 LIST OF ABBREVIATIONS, ACRONYMS AND SYMBOLS

| | |
|-------|--|
| ACS | Attentional Control Scale |
| AFOSR | Air Force Office of Scientific Research |
| ALW | Autonomous Loyal Wingman |
| ANOVA | Analysis of Variance |
| HAT | Human-Autonomy Teaming |
| HCE | High Consequence Event |
| HCT | Human-Computer Trust |
| H-H | Human-Human |
| HMT | Human-Machine Teaming |
| HRI | Human Robot Interaction |
| HRT | Human-Robot Teaming |
| ISR | Intelligence, Surveillance and Reconnaissance |
| LOA | Level of Automation |
| NARS | Negative Attitude toward Robots Scale |
| PAS | Perfect Automation Schema |
| RIS | Reliance Intentions Scale |
| RoTA | Robot Threat Assessment |
| SAM | Surface-to-Air Missile |
| SEAD | Suppression of Enemy Air Defenses |
| TACT | Teaming with Autonomous Collaborative Technology |
| UAV | Unmanned Aerial Vehicle |