

# Human-Agent Teaming for Effective Multirobot Management: Effects of Agent Transparency

Michael J. Barnes<sup>1</sup>, Jessie Y.C. Chen<sup>1(✉)</sup>, Julia L. Wright<sup>1</sup>,  
and Kimberly Stowers<sup>2</sup>

<sup>1</sup> U.S. Army Research Laboratory,  
Aberdeen Proving Ground, Aberdeen, MD, USA  
{michael.j.barnes.civ, yun-sheng.c.chen.civ,  
julia.l.wright8.civ}@mail.mil

<sup>2</sup> University of Central Florida, Orlando, FL, USA  
kstowers@ist.ucf.edu

**Abstract.** The U.S. Army Research Laboratory is engaged in a multi-year program focusing on the human role in supervising autonomous vehicles. We discuss this research with regard to patterns of human/intelligent agent (IA) interrelationships, and explore the dynamics of these patterns in terms of supervising multiple autonomous vehicles. The first design pattern focuses on a human operator controlling multiple autonomous vehicles via a single IA. The second design pattern involves multiple intelligent systems including (a) human operator, (b) IA-asset manager, (c) IA-planning manager, (d) IA-mission monitor, and (e) multiple autonomous vehicles. Both scenarios require a single operator to control multiple heterogeneous autonomous vehicles, and yet the complexity of both the mission variables and the relations among the autonomous vehicles makes efficient operations by a single operator difficult at best. Key findings of two recent research programs are summarized with an emphasis on their implications for developing future systems with similar design patterns. Our conclusions stress the importance of operator situation awareness, not only of the immediate environment, but also of the IA's intent, reasoning and predicted outcomes.

**Keywords:** Intelligent agents · Transparency · Patterns of human-agent interaction · Human factors · Supervisory control

## 1 Introduction

Autonomous systems are becoming part of the framework of American life. They are also changing the dynamics of future combat in ways that are not altogether predictable. Although the advantages are obvious (e.g., reduction in casualties, force multiplication and increased capabilities) [1, 2] their disadvantages (e.g., possible fratricides, civilian casualties, and mission disruptions) can only be estimated [3, 4]. It is important not only to minimize the potential dangers of autonomy but also to increase tactical flexibility by developing user interfaces that ensure that humans have ultimate decision authority [5].

When human operators must supervise or control many systems, the problems are multiplicative [6]. The operator, who is bounded by short-term memory limitations, must maintain situation awareness (SA) of a dynamic and potentially volatile environment while supervising multiple heterogeneous autonomous vehicles (AVs) [5]. To address these challenges, we discuss the implications of patterns of interaction between software-based intelligent agents (IA) and humans, and give examples of experiments that verify their usefulness (see [7] for an overview of design patterns for human-cognitive agent teaming). The definition of an IA varies [8]. In this context, an IA acts autonomously in the sense that it processes environmental information, has clear-cut objectives, and develops courses of actions (COAs) to achieve its objectives. Patterns of interaction refer to reusable human-agent architectures which are adapted to multiple problem spaces that share generic features. Below, we discuss two patterns which manage the type (or level) of information necessary for the human to supervise the IA while simultaneously maintaining SA of both the unfolding environment and the IA's intent, reasoning, and perceived outcomes. The first pattern is relatively simple and involves the human interacting with a single supervisory agent that serves as an intermediary and supervisor for multiple AVs. The second pattern requires the operator to interact with multiple intermediate agents. These agents can interact with each other but each is dedicated to a specific task: choosing the heterogeneous AVs, developing a route plan, and monitoring the AVs' progress. The purpose of identifying effective patterns is to pinpoint essential elements of human-agent systems for possible generic solutions in complex environments. Specifically, our emphasis in this effort is on determining the effectiveness of IA transparency in promoting operator awareness of both the physical environment and the IA's assumptions about the environment [8]. The construct of transparency and its application to patterns of human-IA interaction is discussed below.

## 2 The Situation Awareness-Based Agent Transparency (SAT) Model

Operator trust is an important research topic for both automated and autonomous systems, as it is a key determinant in the calibration of reliance on such systems. Trust can be measured either as an attitude (subjective measure) or as a behavior [9]. Subjective scales have been shown to correlate with automation reliability, the perception of the IA's capabilities, and task difficulties as well as individual differences [10]. Lee and See [11] defined appropriate trust as human reliance on automation that minimizes disuse (failure to rely on reliable automation) and misuse (over-relying on unreliable automation) [12]. Lee [13] suggested that to make the underpinnings of the automation algorithm transparent, it is necessary to display the purpose, process and performance (3-Ps) of the automation, with the caveat that too much information is counterproductive. Based on these and related concepts, U.S. Army Research Laboratory researchers developed a Situation awareness-based Agent Transparency (SAT) model (Fig. 1) to elucidate aspects of SA affecting trust [8]. SAT posits three transparency levels (L) of information to support the operator's SA of the IA's decision process: (a.) L1 – IA's actions and intent (b.) L2 – IA's reasoning process, (c.) L3 – IA's

predicted outcomes and uncertainty [14, 15]. The purpose of the SAT model is to define the type of information necessary to give the operator insight into the IA's basic plan, its intent, reasoning process and its objective end-state. Our hypothesis is that each level of transparency contributes to improving operators' appropriate trust calibration. However, an obvious problem of managing n-systems is that the SAT information for multiple systems could easily overwhelm the operator.

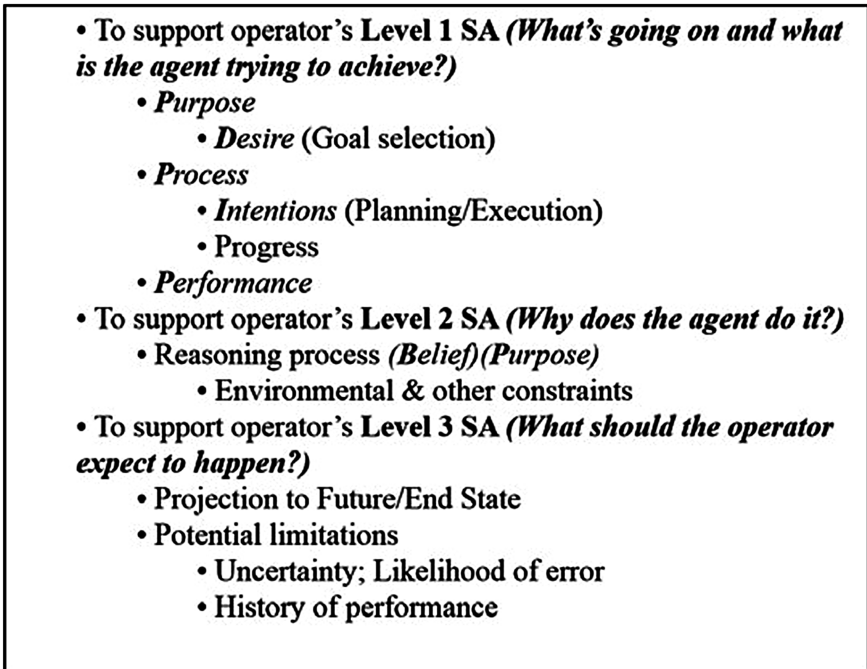
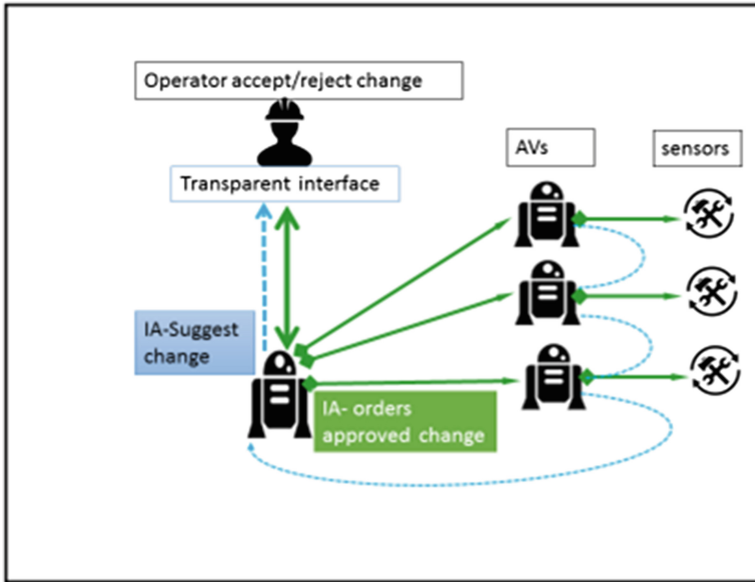


Fig. 1. The SAT model [8]

### 3 Pattern 1: Supervisory Agent

Autonomy and the necessity of human supervision appear to be contradictory. Especially when the operator is managing multiple systems, supervising many autonomous systems would seem to defeat the purpose of making them autonomous in the first place. The first pattern (Fig. 2) uses the power of an IA to reduce operator workload while maintaining the human's decision prerogatives. The IA (RoboLeader) acts as an intermediate supervisor monitoring the lower level systems and reporting back to the human supervisor through a chat window. Furthermore, the IA suggests an algorithmic solution to the operator who can either implement the IA's solution or use the operator's own solution [16].

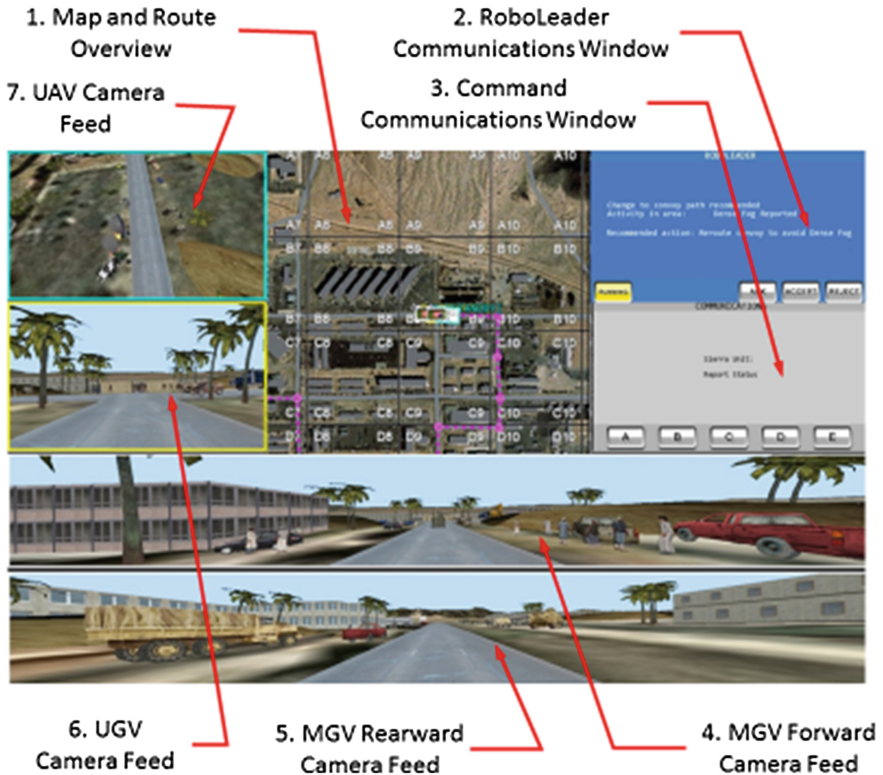
The advantage of this pattern is that the human operator can multitask, monitor the AVs when alerted by the IA, and also determine the utility of the agent's solution to the



**Fig. 2.** Pattern 1, the IA (RoboLeader) supervises AVs and suggests re-planning when the AVs original plan is no longer feasible, requiring the human operator to either accept the IA’s re-plan or over-ride the IA.

alerted problem. Multiple research projects have used the RoboLeader paradigm to explore a number of pertinent human-agent teaming issues, including manipulating the number of AVs supervised, IA reliability, the degree of autonomy, the type of autonomy, and IA transparency [8, 16–19]. The paradigm proved to be effective in numerous simulated combat-related scenarios, resulting in a variety of useful human-agent design guidelines [1, 8]. In general, results showed that the RoboLeader agent benefitted the operators’ concurrent task performance and reduced their workload while allowing operators to maintain their SA. Importantly, individual differences in gaming experience, spatial abilities, and perceived attentional control proved to be crucial factors in human-agent interactions—implying that training and decision support should be geared to individual aptitudes and experience—not “one size fits all” solutions [8].

Pertaining to our current emphasis on the role of transparency in trust calibration, our lab recently examined the efficacy of SAT level-2 information (reasoning explanation) during convoy operations [19]. In this effort, operators monitored an unmanned aerial vehicle (UAV), a manned ground vehicle (MGV), and an unmanned ground vehicle (UGV). We focused on IA reasoning transparency by manipulating the amount of explanatory information conveyed to the operator during high workload mission segments: the operator was monitoring three assets while also detecting threats to their own vehicle using the 360° SA displays (Fig. 3). The IA advisories (which made suggestions concerning when to re-route the convoy) were only accurate 66 % of the time, requiring the operator to monitor the situation continually.



**Fig. 3.** Operator control unit for convoy operations requiring monitoring multiple elements, identifying threats and making convoy route changes based on command updates.

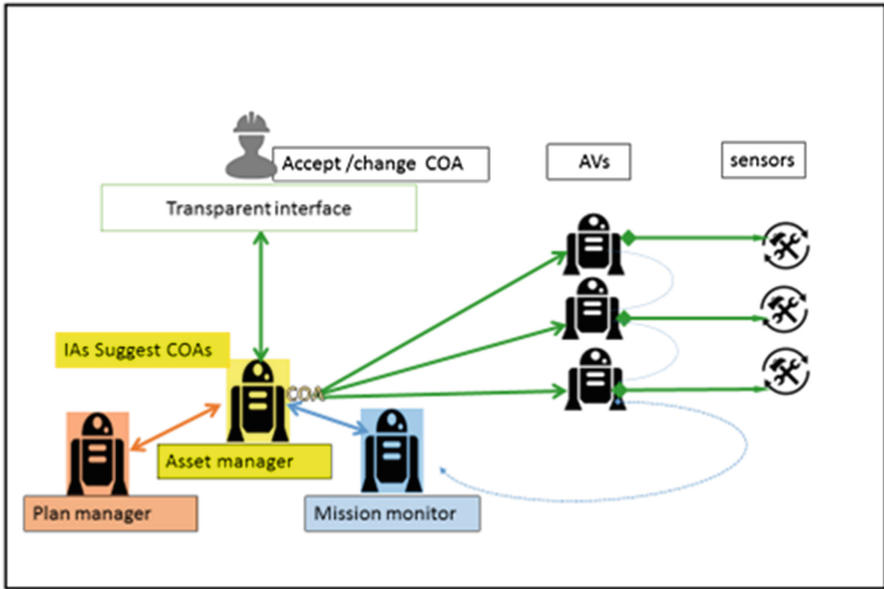
The results supported our hypothesis that explanatory information (e.g., “Change to convoy path recommended: Activity in area: Dense Fog”) would reduce misuse of automation. Operators rejected incorrect advisories significantly more often when given the rationale for the advisory. However, adding non-essential information to the explanation (a time stamp which was potentially useful to determine if the explanation is current) contributed to misuse of IA by decreasing correct rejections. This finding supports Lee’s [13] argument that when adding information to increase the transparency of a display, the information must adhere to the environment (or the task) for which it was intended. In Wright et al.’s case, a parsimonious explanation improved the ability of the operator to override incorrect advisories. In summary, having an IA interface between the AVs and the operator decision-maker proved to be a useful pattern of AV control under a variety of experimental conditions [8]. The caveat is that successful interaction between the IA and the operator depended on concise display formats that provided insight into the IA’s reasoning.

## 4 Multiple Supervisory Agents

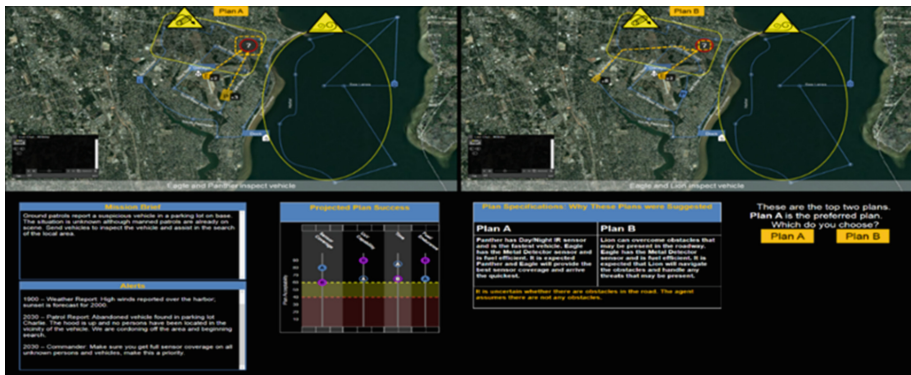
Figure 4 shows a more complex pattern representative of the Intelligent Multi-UxV Planner with Adaptive Collaborative/Control Technologies (IMPACT) program funded by the U.S. Department of Defense's Autonomy Research Pilot Initiative program [20, 21]. IMPACT researchers are investigating issues associated with multiple intelligent systems interacting both with one another and with the human operator to execute numerous combat missions. They are independent in the sense that each IA can generate a COA, and they are interdependent in the sense that to successfully complete a mission, they must interact with one another while taking their final decision cues from the human operator. Pattern 2 involves an Asset Manager agent which receives mission objectives from the operator and decides which AVs and sensors are best suited to the mission. Next, the Plan Manager compiles plans for routing the vehicles to arrive at the objective at a particular time whereas the Mission Monitor agent monitors the AVs in-route progress for plan deviations. The process is adaptive because either the human operator or the Mission Monitor can interrupt the mission and start the process over by interacting with the Asset Manager and Plan Manager to develop a revised plan to present to the operator. Each system tries to optimize mission objectives and the integrated plan requires permission from the operator before the Asset Manager gives execution instructions to the AVs. In Fig. 4, the integrator is the Asset Manager whose tasks include generating the information for the transparency display as well as sending the agreed upon instructions to the AVs.

The research addressed transparency displays containing integrated information from multiple sources. The human operator had decision authority that was manifested in two ways—identifying the initial framework for the mission and verifying the plans supplied by the Asset Manager. In our experiments, information integrated into the Asset Manager was used to generate three levels of SAT including: L1 – plan elements, L2 – rationale, and L3 – IA's predicted outcomes and uncertainty. We conducted two studies to investigate the effects of agent transparency on operators' task performance and trust. Because of the inherent complexity of the experimental conditions, the displays (Fig. 5) used in the experiments contained SAT information on the map, in text, as well as on graphics developed by U.S. Air Force researchers [22]. In addition, the user interface had tactical alerts from the command module showing the commander's intent, environmental, and situational changes. We explored interface options over two experiments in an attempt to isolate the effects of SAT levels and interface features. In particular, in the second experiment, we changed the L3 interface to better display the projected outcomes. L3 information was conveyed using the Air Force's revised graphics showing the relative weights of the predicted outcomes (e.g., time to target, fuel consumption, sensor coverage) on an integrated line graph [22].

In the first simulation experiment, we investigated the SAT model by incrementally varying the type of SAT level (L1 vs. L1 + 2 vs. L1 + 2 + 3) with scenarios counter-balanced over conditions. Each scenario developed a base defense vignette such as an unknown craft spotted near the shoreline adjacent to the naval base. The participants were given two options (A and B) that resulted in different asset compositions (naval, ground or aerial AVs), planned routes, and type of sensors. Updates from the command



**Fig. 4.** N- intelligent systems coordinating planning, executing, and monitoring a dynamic mission by sending mission instructions to N autonomous vehicles after receiving guidance from the operator.



**Fig. 5.** The improved transparency display used in the second experiment showing the relative outcomes predicted for options A and B.

modules supporting either plan A (IAs' favored COA) or B (IAs' less favored option) are shown in Fig. 5. The results indicated that trust calibration (correct rejections and correct positive responses-hits) improved for the latter two conditions compared with the L1 (baseline information) condition. Additionally, subjective trust scores also increased for the third condition ( $L1 < L1 + 2 + 3$ ) [23]. In the second experiment, aside from improving the user interface as discussed above, we separated the uncertainty

information (U) from other L3 information (projections) to create three experimental conditions (L1 + 2 vs. L1 + 2 + 3 vs. L1 + 2 + 3 + U). We found that participants performed their decision task significantly better with L1 + 2 + 3 and L1 + 2 + 3 + U compared with the L1 + L2 condition. Thus, we concluded that all three levels of SAT information incrementally improved operator performance, but that the impact of adding uncertainty information required further research [24]. The results indicate that the design pattern shown in Fig. 4 was useful in generating testable hypotheses. Specifically, SAT predictions were upheld in a complex environment where multiple intelligent systems contributed to transparency information during realistic tri-service scenarios.

## 5 Conclusions

In a world in which autonomy is being engineered into more and more systems, the role of the human is often overlooked. Particularly in military environments, humans face complex problems, multitasking, and increased volatility; situations in which autonomy offers the promise of reducing the problem space to manageable levels. To ensure safety and tactical flexibility, military systems will require human supervision [1–3, 8]. However, the human's span of control becomes limited as the complexity of the military environment and the number of autonomous assets increase beyond a certain limit [6]. Two patterns of autonomous control were discussed that reduced supervisory loading while ensuring human control of multiple heterogeneous autonomous assets. In both paradigms, humans interacted with IAs to control autonomous assets under various degrees of difficulty. Successful operations required the operator to have SA of both the changing military situation and the IAs' decision-making processes. The SAT model of transparency was tested for both patterns indicating the feasibility of effective supervisory control by enhancing the operator understanding of the IA's intent, reasoning, and projected outcomes [8, 14, 19, 23, 24].

**Acknowledgements.** This research was supported by the U.S. Department of Defense Autonomy Research Pilot Initiative, under the Intelligent Multi-UxV Planner with Adaptive Collaborative/Control Technologies (IMPACT) project. The authors wish to thank Olivia Newton, Ryan Wohleber, Nicholas Kasdaglis, Michael Rupp, Daniel Barber, Jonathan Harris, Gloria Calhoun, and Mark Draper for their contribution to this project.

## References

1. Barnes, M.J., Chen, J.Y., Jentsch, F., Redden, E.S.: Designing effective soldier-robot teams in complex environments: training, interfaces, and individual differences. In: Harris, D. (ed.) HCII 2011. LNCS, vol. 6781, pp. 484–493. Springer, Heidelberg (2011)
2. U.S. Department of Defense, Briefing on Autonomy Initiatives (2012)
3. U.S. Defense Science Board: Role of Autonomy in DoD Systems. Office of the Undersecretary of Defense, Washington, D.C. (2012)



4. Endsley, M.: *Autonomous Horizons: System Autonomy in the Air Force – A Path to the Future (Volume I: Human Autonomy Teaming)*. U.S. Department of the Air Force, Washington (2015)
5. Chen, J.Y.C., Barnes, M.J.: Human-agent teaming for multirobot control: a review of human factors issues. *IEEE Trans. Hum. Mach. Syst.* **44**, 13–29 (2014)
6. Lewis, M.: Human interaction with multiple remote robots. *Rev. Hum. Factors Ergon.* **9**, 131–174 (2013)
7. Schulte, A., Donath, D., Lange, D.: Design patterns for human-cognitive agent teaming. In: *Conference on Engineering Psychology & Cognitive Ergonomics. Lecture Notes on Computer Science*, Springer (2016)
8. Chen, J.Y.C., Procci, K., Boyce, M., Wright, J., Garcia, A., Barnes, M.J.: Situation awareness-based agent transparency. Technical report ARL-TR-6905, U.S. Army Research Laboratory, Aberdeen Proving Ground, Maryland (2014)
9. Meyer, J., Lee, J.: Trust, Reliance, Compliance. In: Lee, J., Kirlik, A. (eds.) *The Oxford Handbook of Cognitive Engineering*. Oxford University Press, Oxford (2013)
10. Schaefer, K.E., Billings, D.R., Szalma, J.L., Adams, J., Sanders, T.L., Chen, J.Y.C., Hancock, P.A.: A meta-analysis of factors influencing the development of trust in automation: implications for human-robot interaction. Technical report ARL-TR-6984, U.S. Army Research Laboratory, Aberdeen Proving Ground, Maryland (2014)
11. Lee, J.D., See, K.A.: Trust in automation: designing for appropriate reliance. *Hum. Factors* **46**, 50–80 (2004)
12. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: A model for types and levels of human interaction with automation. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **30**, 286–297 (2000)
13. Lee, J.D.: Trust, trustworthiness, and trustability. In: *The Workshop on Human-Machine Trust Robust Autonomous Systems*, 31 Jan. 2012, Ocala, FL, USA (2012)
14. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. *Hum. Factors* **37**, 32–64 (1995)
15. Endsley, M.R.: Situation awareness misconceptions and misunderstandings. *J. Cognit. Eng. Decis. Making* **9**, 4–32 (2015)
16. Chen, J.Y.C., Barnes, M.J.: Supervisory control of multiple robots: effects of imperfect automation and individual differences. *Hum. Factors* **54**, 157–174 (2012)
17. Chen, J.Y.C., Barnes, M.J.: Supervisory control of multiple robots in dynamic tasking environments. *Ergonomics* **55**, 1043–1058 (2012)
18. Wright, J.L., Chen, J.Y.C., Quinn, S.A., Barnes, M.J.: The effects of level of autonomy on human-agent teaming for multi-robot control and local security maintenance. Technical report, ARL-TR-6724, U.S. Army Research Laboratory, Aberdeen Proving Ground, Maryland (2013)
19. Wright, J.L., Chen, J.Y.C., Barnes, M.J., Hancock, P.A.: The effect of agent reasoning transparency on automation bias: an analysis of performance and decision time. Technical report, U.S. Army Research Laboratory, Aberdeen Proving Ground, Maryland (in press)
20. U.S. Department of Defense – Research & Engineering Enterprise.: *Autonomy Research Pilot Initiative*. <http://www.acq.osd.mil/chieftechologist/arpi.html>
21. Draper, M.: *Realizing Autonomy via Intelligent Adaptive Hybrid Control: Adaptable Autonomy for Achieving UxV RSTA Team Decision Superiority*, Yearly report. U.S. Air Force Research Laboratory, Dayton, OH (2013)
22. Behymer, K.J., Mersch, E.M., Ruff, H.A., Calhoun, G.L., Spriggs, S.E.: Unmanned vehicle plan comparison visualization for effective human-autonomy teaming. In: *Proceedings of the 6th Applied Human Factors and Ergonomics International Conference Las Vegas, NV* (2015)

23. Mercado, J.E., Rupp, M., Chen, J.Y.C., Barber, D., Procci, K., Barnes, M.J.: Intelligent agent transparency in human-agent teaming for Multi-UxV management. *Hum. Factors* **58**(3), 401–415 (2016)
24. Stowers K., Chen J.Y.C., Kasdaglis N., Newton O., Rupp, M., Barnes M.: Effects of situation awareness-based agent transparency information on human agent teaming for multi-UxV management. Technical report, U.S. Army Research Laboratory, Aberdeen Proving Ground, Maryland (2016)