Best of both worlds: Design and evaluation of an adaptive delegation interface

Article *in* Proceedings of the Human Factors and Ergonomics Society Annual Meeting · January 2013

DOI: 10.1177/15419312135

Ewart de Visser
George Mason University
110 PUBLICATIONS **4,639** CITATIONS

Nathan Brooks
20 PUBLICATIONS **129** CITATIONS

Timur Chabuk
University of Maryland, College Park
11 PUBLICATIONS **78** CITATIONS

Amos Freedy
Perceptronics Solutions.com
60 PUBLICATIONS **877** CITATIONS

Best of Both Worlds: Design and Evaluation of an Adaptive Delegation Interface

Ewart de Visser[1,2], Brian Kidwell[2], John Payne[1], Li Lu[2], James Parker[2], Nathan Brooks[3], Timur Chabuk[1],
Sarah Spriggs[4], Amos Freedy[1], Paul Scerri[3] & Raja Parasuraman[2]

[1]Perceptronics Solutions, Inc., Falls Church, VA
[2]George Mason University, Fairfax, VA

[3]Carnegie Mellon University, Pittsburgh, PA
[4]Air Force Research Laboratory,
711th HPW/RHCI, WPAFB, Dayton, OH

The proliferation of unmanned aerial vehicles (UAVs) in civil and military domains has spurred increasingly complex automation design for augmenting operator abilities, reducing workload, and increasing mission effectiveness. We describe the Adaptive Interface Management System (AIMS), an intelligent adaptive delegation interface for controlling and monitoring multiple unmanned vehicles, with a mixed-initiative team model language. A study was conducted to assess understanding of this model language and whether participants exhibited calibrated trust in the intelligent automation. Results showed that operators had accurate memory for role responsibility and were well calibrated to the automation. Adaptive automation design approaches like the one described in this paper can be useful to create mixed-initiative human-robot teams.

The proliferation of unmanned aerial vehicles (UAVs) in civil and military domains has spurred increasingly complex automation designs aimed at augmenting operator abilities, reducing workload, and increasing mission effectiveness. Such human-robot teams introduce unique challenges to the planning and coordination of team performance. A key issue among these is the ability of the human-automation *mixed-initiative* team to predict, collaborate, and coordinate its actions with complex systems that can potentially act in unstructured and unpredictable environments with varied levels of autonomy. Advanced design approaches are necessary to support trustworthy human-automation collaboration in such conditions.

Design challenges for developing advanced automation

Automation designs can be broadly classified into two categories: *adaptable* and *adaptive* automation (Opperman, 1994; Parasuraman, 2000; Miller & Parasuraman, 2007; Scerbo, 2007; Feigh, Dorneich, & Hayes, 2012). Adaptable automation is primarily *user-initiated,* whereas adaptive automation is primarily *system-initiated.* The relative merits of these two approaches are still under debate. However, it may be that an optimal adaptive strategy is one that falls somewhere between complete human control and complete automation control, reflecting a trade-off between workload, unpredictability, and competency (Miller & Parasuraman, 2007).

Previous research has pointed to the general utility of adaptable delegation interfaces in allowing for effective supervisory control of multiple numbers of UVs (Parasuraman et al., 2005). Studies examining the efficacy of adaptive automation have shown benefits in situation awareness and workload (Parasuraman, Barnes, & Cosenzo, 2007; Parasuraman, Cosenzo, & de Visser, 2009), improved planning generation times (Cummings et al., 2010), and calibrated trust (de Visser & Parasuraman, 2011). Disadvantages include the potential for skill degradation (Kaber et al., 2004) and decreased user acceptance (Miller & Hannen, 1999). Automation design should thus focus on balancing mental workload, calibrating trust, maintaining high situation awareness, and extending flexible delegation control.

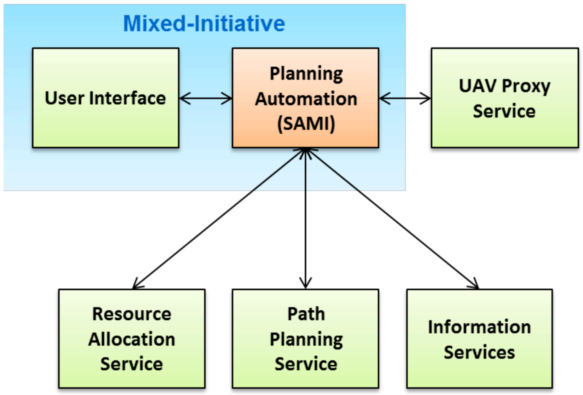


Figure 1. Adaptive Delegation Interface Architecture.

ADAPTIVE INTERFACE MANAGEMENT SYSTEM

What are the requirements needed to achieve a healthy balance between an adaptable and an adaptive interface? We propose a hybrid approach of an Adaptive Delegation Interface (ADI). An ADI is *adaptive* because it is responsive to context and user needs, and involves *delegation* in the same sense that a supervisor works with a human subordinate. Both the user and the automation can initiate a goal and propose a plan.

We developed such an ADI called the Adaptive Interface Management System (AIMS). AIMS is designed to create and execute complex mission models for multiple UAV operation with a comprehensive mission model language that includes situation awareness and mixed initiative (SAMI) markup (see Figure 1). To address issues identified in current design approaches, we focused on developing a shared task vocabulary, flexible automation control, and transparency.

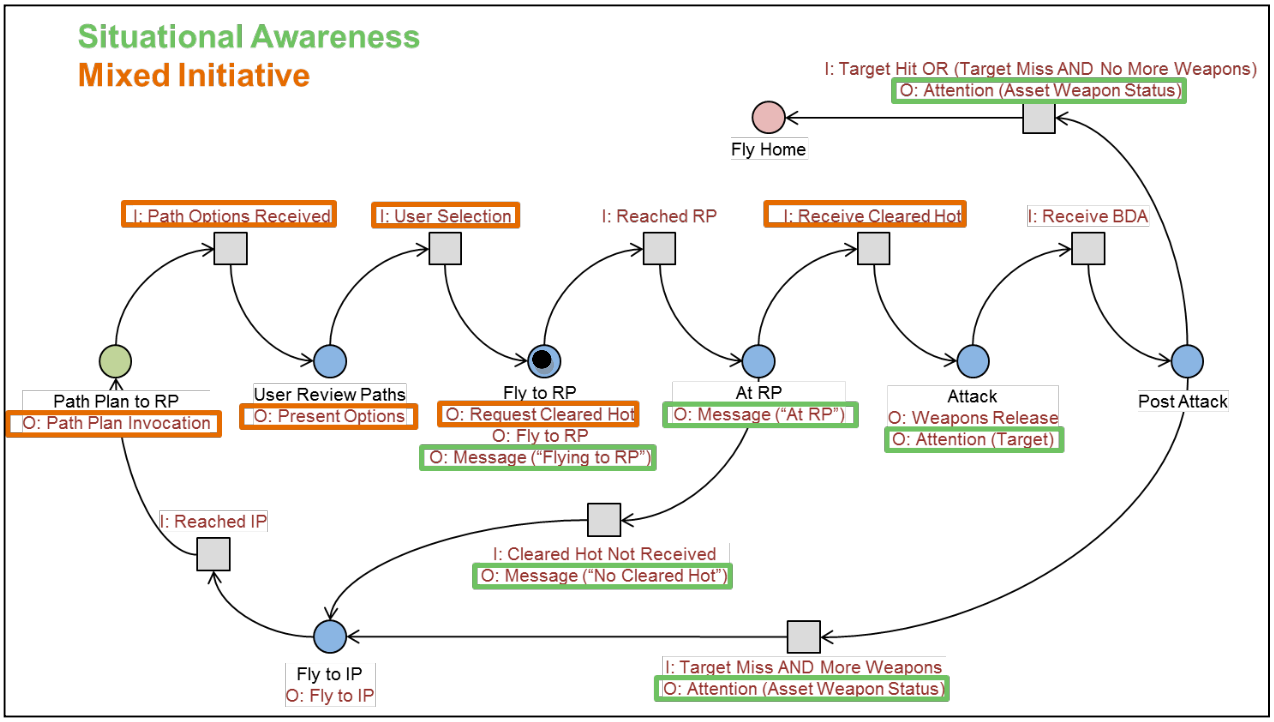


Figure 2. The Situation Awareness Mixed Initiative (SAMI) planning language.

AIMS Architecture

The AIMS architecture consists of planning automation and a user interface. As shown in Figure 1, the planning automation provides several relevant planning services that are typically harder for an operator to do quickly and effectively. These services include resource allocation, path planning, and aggregating relevant information services to the mission. This allows SAMI to propose pre-defined plans to the operator from a subordinate level, allowing for a true mixed-initiative system. The user interface design needs to communicate the shared task model between the user and the automation in a way that is transparent, understandable, and adjustable.

Situation Awareness Mixed Initiative (SAMI) Language

In order to foster close collaboration between the human and the automation, with a common task vocabulary, we developed a team plan specification language with SAMI markup (Brooks et al., 2013). Designers can specify these mission templates based on subject matter expertise prior to mission execution. An example collaborative attack mission model is shown in Figure 2. The mission is built using Petri-nets. Petri-nets consist of places (circles), transitions (squares), arcs (arrows), and tokens (black dot). In this case, tokens represent UAVs moving through transitions and places. Crucially, at each place or transition, four different types of information constitute the flexibility of the SAMI language. Plan information reveals the current status and progress of the plan. Situation awareness information includes directives to adjust the user interface to highlight an important aspect of the plan. An example of such a message would be to zoom in on the map or to switch to camera 1 on UAV 2. The mixed-initiative information includes the number of automation options shown to the user as well as criteria for autonomous decision-making. Finally, important information indicates the relative importance of a message compared to other messages in the system.

SAMI Agent Interaction

To the user we present SAMI as an agent that acts as a subordinate mission planner that guides the mission planning, execution, monitoring, and re-planning functions. At the top of the mission visualization we designed a cognitive agent to serve as a method to communicate with the operator at a high-level as a tactical subordinate. Previous research has shown that cognitive agents can foster relationships between humans and automation (de Visser et al., 2012). Operators are first presented with a library of missions from which they can select a mission of interest. When a mission is selected, a brief description of the mission is provided to explain the actions of each UAV and the goal they will execute.

SAMI creates a mission model template using the mission model visualization and the underlying team language. Operators can preview this model to acquire an initial sense of the mission phases and decisions before executing the mission.

Once an operator has decided to execute a mission, SAMI prompts the user to allocate UAVs to the mission model roles. In a collaborative attack, one UAV must act as the laser designator, and a second UAV acts as the attacker.

After UAVs have been allocated to their mission roles, the operator may select between several flight paths chosen by SAMI. Each path option can be seen on the mission map as the operator selects through the available options. Once path planning has been completed, SAMI orders the UAVs to embark on their mission. The mission progress can be monitored on the mission map and in the mission model visualization allowing the operator a precise moment-by-moment update of the mission status. In the collaborative attack scenario, the operator approves the final attack and as the attacker moves in on the target, the laser designator enables the camera to show the impact.

Shared Task Model User Interface Design

The mission model visualization module shows how SAMI executes a model step-by-step and provides valuable insight into the complex mission models created by the SAMI automation (see Figure 3). Conveying the mode of automation has been shown to increase understanding of the automation and subsequent trust. The mission model visualization module is comprised of several components as shown in Figure 3. These components are:

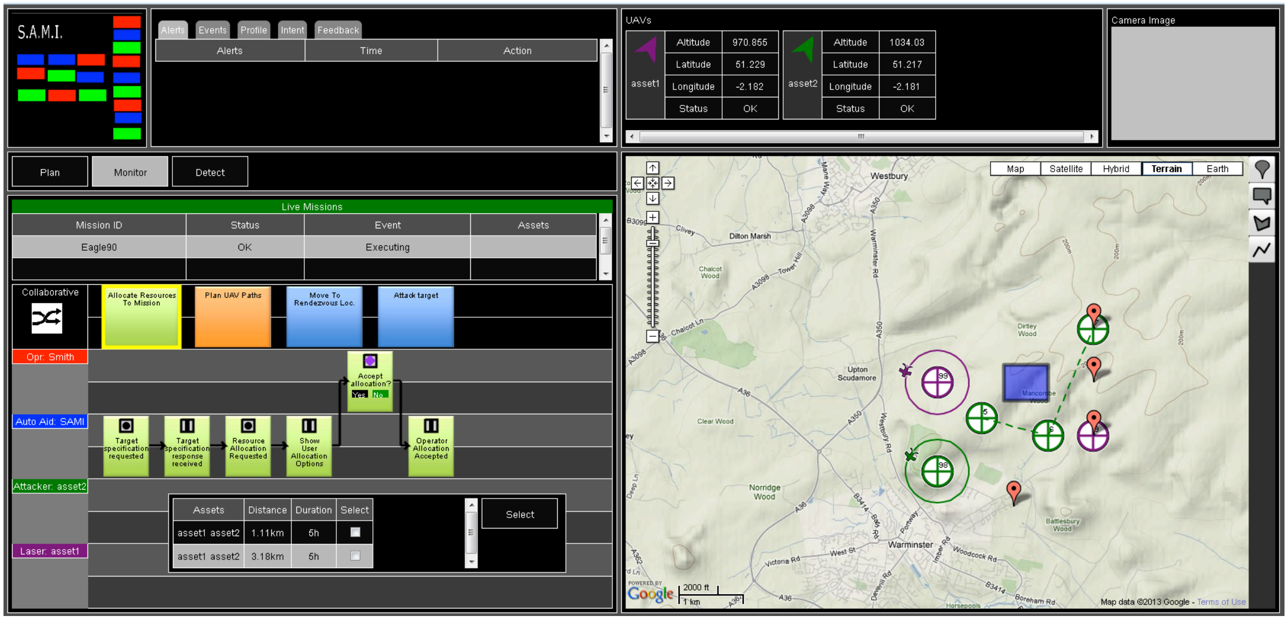
1. **Mission phases**. The top row of the mission model shows the phases or segments of each mission. Users can click on each phase to see the detailed task structure.
2. **Mission role timeline tracks**. The mission role column shows the roles for the mission and which actor fills this role. In this particular mission, there are 4 roles including the operator, the intelligent automation SAMI, a laser designator, and an attacker.
3. **Decision nodes.** Operator decisions are displayed using a special operator decision node (purple) for the mission visualization. The node can be set to either accept the automation recommendation or reject the automation recommendation.
4. **Task Status**. Petri-net transitions and arcs are shown on the timeline tracks. These elements indicate which part of the plan is currently being executed (orange), which parts have been successfully completed (green), and which parts still need to be completed (blue).

CALIBRATION STUDY

Purpose and approach

Since the AIMS system is a novel type of interface, a calibration study was conducted to assess its usability. The main goal for this study was to examine if participants could collaborate successfully with SAMI to plan, execute and monitor a mission. To test if participants were calibrated to SAMI, we varied executing behavior to be either congruent, or incongruent. In the congruent condition, SAMI showed map visualizations that were congruent with the mission task model. In the incongruent condition, the map visualizations did not match the mission task model. We designed the study in this manner to assess whether participants truly understood the planned behavior according to the mission model and the executed behavior on the map.

We used *concept maps* as a way to assess whether participants could remember which tasks were executed, which agent was responsible for those tasks, and how tasks were related. Concept maps are a flexible method for knowledge elicitation and have been used to test flight knowledge of novice pilots (Smith, 2008). This method was highly suitable for our purposes because we could easily capture our mission models with these maps and then test whether participants could produce a similar map. We predicted that participants would have accurate memory for role allocations and relationships. We further measured trust in SAMI. We predicted that trust in SAMI would be higher in the congruent condition compared to the incongruent condition.



Alerts	Time	Action

Live Missions

Mission ID	Status	Event	Assets
Eagle90	OK	Executing	

Assets	Distance	Duration	Select
asset1 asset2	1.11km	5h	☐
asset1 asset2	3.18km	5h	☐

asset1	
Altitude	970.855
Latitude	51.229
Longitude	-2.182
Status	OK

asset2	
Altitude	1034.03
Latitude	51.217
Longitude	-2.181
Status	OK

Figure 3. The AIMS Interface.

METHOD

Participants

Sixteen young adults (8 females), 18 – 24 years old (*M* =20.18, *SD*=0.42) were recruited and compensated with course credit. All participants signed consent forms allowing use of their data.

Design

This study was a repeated measures design with Execution behavior (congruent, incongruent) as the within-subjects factor. For congruent conditions, the automation system SAMI showed map visualizations including role allocations and path planning that were consistent with the mission task model and user selections. In the incongruent conditions, the map visualizations did not match the task model and the user selections. For instance, UAVs flew next to the designated path instead of right above the path or flew around a no-fly zone in a different direction then indicated.

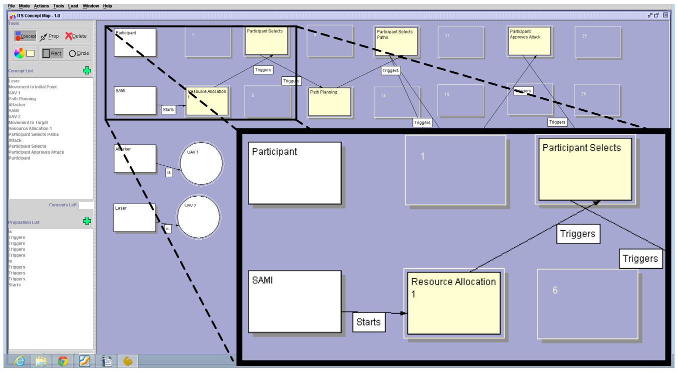


Figure 4. An example concept map.

Task and Scenarios

Four comparable mission scenarios were randomized and presented to participants during the course of this experiment. Each mission involved two UAVs coordinating a collaborative attack on a target. Participants were tasked with deciding the respective roles, planned route and final action for the two UAVs they were controlling. There were two possible roles for each UAV (designator and attacker), two possible routes around the no-fly zone, and one final action (attack; yes or no). For both resource allocation and route planning, participants were to select the most direct route for their UAV to travel. SAMI directed participants to make these decisions as the mission plan dictated. Participants were further instructed to monitor SAMI and the two UAV assets for possible erroneous decisions and use the deviation "detect" feature to record erroneous UAV behavior.

Measures

Mission model memory. Participants' mental models of the mission plans were evaluated using TPL-KATS, a concept mapping software (Hoeft, Jentsch, Harper, Evans, Bowers, & Salas, 2003). A representative concept map was created based on the mission model used for the scenarios. The concept map contained the mission tasks, roles, and relationships between these elements. Memory accuracy scores were derived by summing the total number of correctly identified elements, as compared to a concept map key, divided by the total number of elements in the concept map. Separate memory scores for the role and relationship elements were created.

Trust. We used a human-automation trust measure (Jian et al., 2000) to assess the SAMI automation.

Deviation detection. A detection button was included in the interface to detect deviations between the mission model and the mission map.

Workload. The NASA-TLX was used to assess subjective workload.

Procedure

Participants were welcomed into the experimental suite and began the procedure with a detailed training protocol explaining the AIMS interface. Participants then completed the experimental scenarios. After each scenario participants completed the NASA-TLX, the trust measures and a concept map of the specific scenario. Upon completing the study, each participant was debriefed and thanked for their participation.

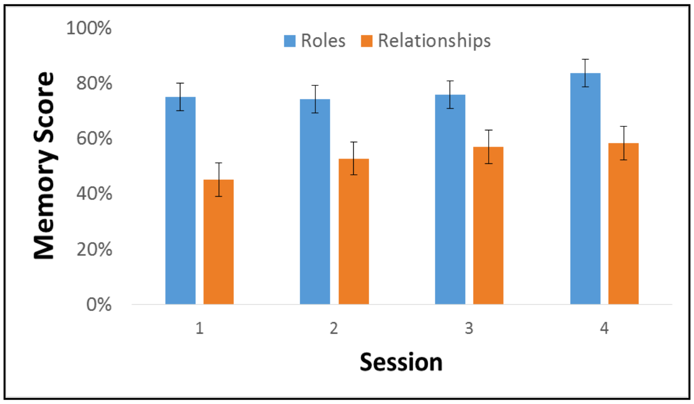


Figure 5. Mission model memory.

RESULTS

Mission Model Memory

Participants recalled the mission roles at around 80% accuracy for both the congruent and incongruent conditions (see Figure 5). Memory for relationships between the various mission components averaged to around 50%. Memory scores for relationships significantly improved in session 4 (*M*=0.58, *SEM*=0.07), compared to session 1 (*M*=0.45, *SEM*=0.06), $F(3,45) = 4.4$, $p < 0.05$.

Trust in SAMI

Participants did not significantly trust SAMI more in the congruent condition compared to the incongruent condition, $p = 0.14$. However, distrust was significantly higher in the incongruent condition (*M*=3.0, *SEM*=0.19) compared to the congruent condition (*M*=2.44, *SEM*=0.20), *F*(1,15) = 6.49, $p < 0.05$.

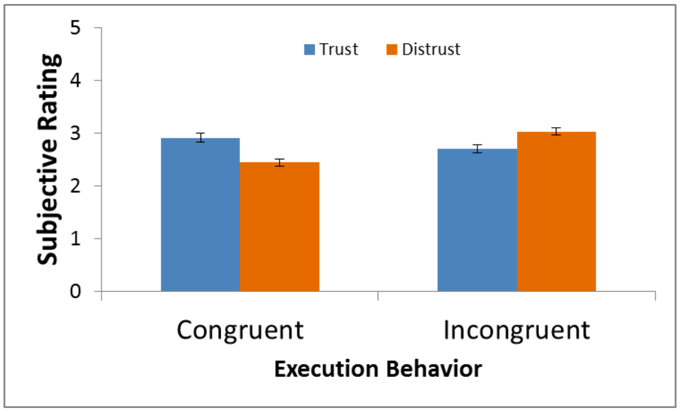


Figure 6. Trust and distrust in execution behavior.

Deviation Detection

Participants detected significantly more deviations in the incongruent condition (*M*=1.5, *SEM*=0.34) compared to the congruent condition (*M*=0.56, *SEM*=0.18), *F*(1,15) = 7.98, $p < 0.05$.

Workload

Subjective workload in the congruent trials (*M*=45.21, *SEM*=3.29) was not significantly different compared to workload in the incongruent trials (*M*=48.53, *SEM*=3.66), $p = 0.43$.

DISCUSSION

An adaptive delegation interface may be an intermediate solution between the extremes of fully adaptable or adaptive automation. We presented a design approach and a team language called SAMI as a method for fostering mixed-initiative team planning. We conducted a calibration study to assess user's trust in SAMI and their ability to understand and remember the mission models.

Our study showed participants were calibrated to SAMI and the mission task model. These results indicate that the shared vocabulary and SAMI were appropriate, usable, and understandable analogies and provides further evidence for the efficacy of these types of interfaces (Parasuraman et al., 2005). Distrust in the SAMI agent was higher when execution behavior was not appropriate. This result suggests that participants can anticipate when SAMI may or may not be correct, which is critical for human-automation team performance (de Visser & Parasuraman, 2011). The overall level of mental workload was moderate, which can be taken as evidence that the Adaptive Delegation Interface did not overburden the user, a key concern for adaptable interfaces (Miller & Parasuraman, 2007). Future studies with the ADI will focus on allowing users to adjust levels of automation and varying the levels of adaptive automation.

ACKNOWLEDGMENTS

This research was supported by SBIR Phase II contract No. FA8650-11-C-6224 funded by the Air Force Research Laboratories. The authors would like to thank Athena Johnson and Daniel Bernstein for providing software development support.

REFERENCES

Brooks, N., E. de Visser, T. Chabuk, E. Freedy, and P. Scerri (2013). An Approach to Team Programming with Markup for Operator Interaction. In *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems,* 2013.

Cummings, M. L., Buchin, M., Carrigan, G., & Donmez, B. (2010). Supporting intelligent and trustworthy maritime path planning decisions. *International Journal of Human-Computer Studies*, *68*(10), 616-626. doi: 10.1016/j.ijhcs.2010.05.002.

de Visser, E. J., Krueger, F., McKnight, P., Scheid, S., Smith, M., Chalk, S., & Parasuraman, R. (2012, September). The World is not Enough: Trust in Cognitive Agents. In *Proceedings of the Human Factors and Ergonomics Society* Annual Meeting (Vol. 56, No. 1, pp. 263-267). SAGE Publications.

de Visser, E., & Parasuraman, R. (2011) Adaptive aiding of human-robot teaming: Effects of imperfect automation on performance, trust, and workload. *Journal of Cognitive Engineering and Decision Making, 5*(2), 209-231.

Feigh, K. M., Dorneich, M. C., & Hayes, C. C. (2012). Toward a characterization of adaptive systems: A framework for researchers and system designers. *Human Factors, 544(6)*, 1008-1024.

Hoeft, R.M., Jentsch, F.G., Harper, M.E., Evans III, A.W., Bowers, A.C., Salas, E. (2003). TPL-KATS – concept map: A computerized knowledge assessment tool. *Computers in Human Behavior*, 19(6): 653-657

Jian, J., Bisantz, A., & Drury, C. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, *4*(1), 53-71.

Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science, 5*(2), 113-153.

Miller, C., & Hannen, M. (1999). The Rotorcraft Pilot's Associate: Design and evaluation of an intelligent user interface for cockpit information management. *Knowledge Based Systems, 12,* 443–456.

Miller, C., & Parasuraman, R. (2007). Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control. *Human Factors, 49*(1), 57-75.

Opperman, R. (1994). *Adaptive user support*. Hillsdale, NJ: Erlbaum.

Parasuraman, R. (2000). Designing automation for human use: empirical studies and quantitative models. *Ergonomics, 43*(7), 931 - 951.

Parasuraman, R., Galster, S., Squire, P., Furukawa, H., & Miller, C. (2005). A flexible delegation-type interface enhances system performance in human supervision of multiple robots: Empirical studies with RoboFlag. *IEEE Transactions on Systems, Man, and Cybernetics. Part A. Systems and Humans*, *35*, 481-493.

Parasuraman, R., Cosenzo, K. A., & de Visser, E. (2009). Adaptive automation for human supervision of multiple uninhabited vehicles: Effects on change detection, situation awareness, and mental workload. *Military Psychology*, *21*(2), 270–297. Routledge.

Scerbo, M. (2007). Adaptive automation. In R. Parasuraman and M. Rizzo (Eds.), *Neuroergonomics; The brain at work.* New York: Oxford University Press.

Smith, C. F. (2008). *The effect of functional display information on the acquisition and transfer of novice piloting knowledge*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses.